

EVALUACIÓN DE RESULTADOS EN SALUD: VALOR Y SOSTENIBILIDAD PARA EL SISTEMA NACIONAL DE SALUD

Javier Soto
José Manuel Martínez Sesmero
Miguel Ángel Casado
Miguel Ángel Calleja
Félix Lobo
(Directores)





EVALUACIÓN DE RESULTADOS EN SALUD: VALOR Y SOSTENIBILIDAD PARA EL SISTEMA NACIONAL DE SALUD

Javier Soto
José Manuel Martínez Sesmero
Miguel Ángel Casado
Miguel Ángel Calleja
Félix Lobo
(Directores)

Funcas

PATRONATO

ISIDRO FAINÉ CASAS
JOSÉ MARÍA MÉNDEZ ÁLVAREZ-CEDRÓN
FERNANDO CONLLEDO LANTERO
CARLOS EGEA KRAUEL
MIGUEL ÁNGEL ESCOTET ÁLVAREZ
AMADO FRANCO LAHOZ
MANUEL MENÉNDEZ MENÉNDEZ
PEDRO ANTONIO MERINO GARCÍA
ANTONIO PULIDO GUTIÉRREZ
VICTORIO VALLE SÁNCHEZ
GREGORIO VILLALABEITIA GALARRAGA

DIRECTOR GENERAL

CARLOS OCAÑA PÉREZ DE TUDELA

Impreso en España
Edita: Funcas
Caballero de Gracia, 28, 28013 - Madrid
© Funcas

Todos los derechos reservados. Queda prohibida la reproducción total o parcial de esta publicación, así como la edición de su contenido por medio de cualquier proceso reprográfico o fónico, electrónico o mecánico, especialmente imprenta, fotocopia, microfilm, *offset* o mimeógrafo, sin la previa autorización escrita del editor.

ISBN: 978-84-17609-49-8
Depósito legal: M-9082-2021
Maquetación: Funcas
Imprime: Cecabank

COLABORADORES	9
PRÓLOGO UN VISTAZO A LA INVESTIGACIÓN DE RESULTADOS EN TÉRMINOS DE SALUD	15
CAPÍTULO 1. EVALUACIÓN DE RESULTADOS EN TÉRMINOS DE SALUD: POR QUÉ, ¿CÓMO, CUÁNDO Y PARA QUÉ?	25
CAPÍTULO 2. RESULTADOS COMUNICADOS POR LOS CLÍNICOS	35
CAPÍTULO 3. RESULTADOS PERCIBIDOS POR LOS PACIENTES	55
CAPÍTULO 4. RESULTADOS EN SALUD PERCIBIDOS Y COMUNICADOS POR LOS OBSERVADORES: ¿QUÉ SON Y CÓMO SE MIDEN?	87
CAPÍTULO 5. ¿CÓMO INCORPORAR LAS MEDIDAS DE RESULTADOS EN SALUD A LA EVALUACIÓN ECONÓMICA?	95
CAPÍTULO 6. PRINCIPALES TÉCNICAS ESTADÍSTICAS Y PARÁMETROS PARA MEDIR RESULTADOS EN SALUD	115
CAPÍTULO 7. METODOLOGÍA DISPONIBLE PARA OBTENER RESULTADOS EN SALUD EN EL MUNDO REAL	141
CAPÍTULO 8. RESULTADOS EN SALUD A EVALUAR EN DIFERENTES PATOLOGÍAS: ENFERMEDADES INFECCIOSAS	179
CAPÍTULO 9. RESULTADOS EN SALUD EN LA INFECCIÓN POR HIV	199
CAPÍTULO 10. RESULTADOS EN SALUD EN PACIENTES CON HEPATITIS C CRÓNICA	225
CAPÍTULO 11. RESULTADOS EN SALUD EN REUMATOLOGÍA	245
CAPÍTULO 12. RESULTADOS EN SALUD EN DERMATOLOGÍA	271
CAPÍTULO 13. RESULTADOS EN SALUD EN APARATO DIGESTIVO: ENFERMEDAD INFLAMATORIA INTESTINAL E INTESTINO IRRITABLE	295
CAPÍTULO 14. RESULTADOS EN SALUD EN ESCLEROSIS MÚLTIPLE	313
CAPÍTULO 15. BENEFICIO CLÍNICO EN ONCOLOGÍA: UNA COMPLEJA PREGUNTA	345

CAPÍTULO 16. EVALUACIÓN DE RESULTADOS EN SALUD: HEMATOLOGÍA	369
CAPÍTULO 17. RESULTADOS EN SALUD: PATOLOGÍA CARDIOVASCULAR	389
CAPÍTULO 18. RESULTADOS EN SALUD A EVALUAR EN ENFERMEDADES RARAS	425
CAPÍTULO 19. INVESTIGACIÓN DE RESULTADOS EN SALUD EN NIÑOS Y ADOLESCENTES	453

COLABORADORES

Abad Sazatornil, María Reyes

Jefe de Servicio de Farmacia. Hospital Universitario Miguel Servet

Aibar Abad, María Pilar

Facultativo especialista de área. Hospital Obispo Polanco

Arenas Miras, María del Mar

Médico adjunta del Servicio de Enfermedades Infecciosas. Hospital Universitario del Mar (Barcelona)

Arévalo Ruales, Karla

Médico especialista en Reumatología. Médico adjunto. Servicio de Reumatología. Hospital Universitari i Politécnic La Fe (Valencia)

Bécares Martínez, Javier

Licenciado en Farmacia. Farmacéutico especialista en Farmacia Hospitalaria. Hospital Universitario Fundación Jiménez Díaz

Borrás Blasco, Joaquín

Farmacéutico. Hospital de Sagunto

Buti Ferret, María

Jefe clínico de la Unidad de Hepatología. Hospital Universitario Vall d'Hebron

Caballero Molina, María Teresa

Servicio de Alergia Hospital Universitario La Paz. Instituto de Investigación Sanitaria del Hospital La Paz (IdiPaz). Jefe de Grupo Unidad 754 CIBERER

Cajaraville Ordoñana, Gerardo

Facultativo especialista de Área de Farmacia Hospitalaria. Fundación Onkologikoa (San Sebastián)

Calleja, Miguel Ángel

Jefe del Servicio de Farmacia del Hospital Virgen Macarena de Sevilla

Calvo Rey, Cristina

Médico especialista en Pediatría. Servicio de Pediatría. Hospital Universitario La Paz (Madrid)

Casado, Miguel Ángel

Director de Pharmacoeconomics & Outcomes Research Iberia (PORIB)

Casas Deza, Diego

Graduado en Medicina y Cirugía. Servicio de Aparato Digestivo. Hospital Universitario Miguel Servet

Castrejón, Isabel

*Departamento de Medicina Interna-Reumatología.
Rush University Medical Center Chicago (Illinois)*

Chacón López-Muñiz, José Ignacio

Servicio de Oncología. Hospital Virgen de la Salud (Toledo)

Córdoba Mascuñano, Raúl

Jefe de Hematología, Instituto de Investigación Sanitaria IIS-FJD, Unidad de Innovación Clínica y Organizativa (UICO). Hospital Universitario Fundación Jiménez Díaz (Madrid)

Cueva Dobao, Pablo de la

Dermatólogo. Hospital Universitario Infanta Leonor

Fernández González, María Rosario

Enfermera. Complejo Asistencial Universitario de León. Sacyl

Forjaz, Maria João

Escuela Nacional de Sanidad y REDISSEC. Instituto de Salud Carlos III

Fraga, M^a. Dolores

*Especialista en Farmacia Hospitalaria. Facultativa especialista de Área.
Hospital General la Mancha Centro (Alcázar de San Juan, Ciudad Real)*

Fragío Gil, Jorge Juan

*Grado en Medicina. Residente de cuarto año. Servicio de Reumatología.
Hospital Universitari i Politècnic La Fe (Valencia)*

García López, Santiago

*Médico especialista en aparato digestivo, Unidad de Enfermedad Inflamatoria Intestinal,
Servicio de Aparato Digestivo. Hospital Universitario Miguel Servet*

García Palomo, Andrés

*Doctor en Medicina. Oncólogo médico. Jefe de servicio. Oncología médica.
Hospital Universitario de León*

Gastalver, Cristina

*Doctora en Farmacia y especialista en Farmacia Hospitalaria. BCPS. Farmacéutica adjunta
del Servicio de Farmacia del Hospital del Sureste (Arganda del Rey, Madrid)*

Ginestal López, Ricardo Constantino

Médico adjunto al Servicio de Neurología. Hospital Universitario Clínico San Carlos (Madrid)

Gómez Barrera, Manuel

*Licenciado en Ciencias Económicas y licenciado en Farmacia. Pharmacoeconomics
& Outcomes Research Iberia (PORIB). Universidad San Jorge*

González Mazarío, Roxana

Licenciada en Medicina. Especialista en Medicina Interna. Residente de tercer año. Servicio de Reumatología. Hospital Universitari i Politècnic La Fe (Valencia)

Horcajada Gallego, Juan Pablo

Médico adjunto del Servicio de Enfermedades Infecciosas. Hospital Universitario del Mar (Barcelona)

Lobo, Félix

Catedrático emérito de la Universidad Carlos III de Madrid y director de Economía y Políticas de Salud de Funcas

López González, Ana

Doctora en Medicina. Oncóloga Médica. Responsable de la Unidad de Investigación del Servicio de Oncología médica. Complejo Asistencial Universitario de León (CAULE)

Maderuelo, Mercedes

Gerente de la Federación Española de Diabetes (FEDE)

Manrique Rodríguez, Silvia

Farmacéutico especialista en Farmacia Hospitalaria. Servicio de Farmacia. Hospital General Universitario Gregorio Marañón (Madrid)

Martínez Fernández-Llamazares, Cecilia

Farmacéutico especialista en Farmacia Hospitalaria. Servicio de Farmacia. Hospital General Universitario Gregorio Marañón (Madrid)

Martínez Sesmero, José Manuel

Servicio de farmacia. Hospital Clínico de San Carlos. Madrid.

Monte Boquet, Emilio

Doctor en Farmacia. Especialista en Farmacia Hospitalaria. Jefe de Sección. Servicio de Farmacia. Hospital Universitari i Politècnic La Fe (Valencia)

Ortega Eslava, Ana

Especialista en Farmacia Hospitalaria. Doctora en Farmacia. Profesora titular. Servicio de Farmacia. Clínica Universidad de Navarra (Pamplona)

Pavez Perales, Cristóbal

Licenciado en Medicina. Residente de primer año. Servicio de Reumatología. Hospital Universitari i Politècnic La Fe (Valencia)

Queralt Gorgas, María

Jefe de Servicio de Farmacia. Hospital Universitario Vall d'Hebron

Revuelta Herrero, José Luis

Farmacéutico especialista en Farmacia Hospitalaria. Servicio de Farmacia. Hospital General Universitario Gregorio Marañón (Madrid)

Rodríguez Artalejo, Fernando

Doctor en Medicina. Catedrático Departamento de Medicina Preventiva y Salud Pública. Universidad Autónoma de Madrid

Rodríguez Sagrado, Miguel A.

Facultativo especialista de Área. Servicio de Farmacia. Hospital Universitario Ramón y Cajal (Madrid)

Rodríguez-Blázquez, Carmen

Centro Nacional de Epidemiología y CIBERNED. Instituto de Salud Carlos III

Roldán Rabadán, Inmaculada

Servicio de Cardiología. Hospital General Universitario La Paz, Madrid. IDIPAZ. CIBER-CV

Rubia Navarro, Marta de la

Licenciada en Medicina. Residente de segundo año. Servicio de Reumatología. Hospital Universitari i Politècnic La Fe (Valencia)

Rubio Salvador, Ana Rosa

Facultativo Especialista de Área de Farmacia Hospitalaria, Unidad de Farmacia de Onco-Hematología, Servicio de Farmacia, Complejo Hospitalario de Toledo (Toledo)

Ruiz, Miguel Ángel

Facultad de Psicología. Universidad Autónoma de Madrid

Saldaña, Mónica

Doctora en Medicina y Cirugía. Especialista en Farmacología Clínica. Jefe de Servicio de Farmacología Clínica. Hospital Universitario Puerta del Mar (Cádiz)

Sanmartín Fenollera, Patricia

Farmacéutico Adjunto al Servicio de Farmacia Hospitalaria. Hospital Universitario Fundación Alcorcón (Alcorcón, Madrid)

Santiago Pérez, Alejandro

Servicio de Farmacia. Hospital Universitario Infanta Leonor

Soto Álvarez, Javier

Dpto. de Farmacoeconomía y Resultados en Salud, Pfizer

Trapero-Bertran, Marta

Doctora. Profesora titular e investigadora de la UIC

Vallejo-Torres, Laura

Doctora. Profesora e investigadora doctora competitiva en la ULPGC

Vargas Castrillón, Emilio

Servicio de Farmacología Clínica. Hospital Clínico San Carlos (Madrid)

Ventayol, Pere

*Especialista en Farmacia Hospitalaria. Jefe de Sección del Servicio de Farmacia.
Hospital Universitario Son Espases (Palma de Mallorca)*

Vera García, Ruth

*Licenciada en Medicina y Cirugía por la Universidad de Valencia. Jefe de Servicio.
Responsable de la Estrategia del Cáncer en Navarra*

Villagarcía Casla, Teresa

*Doctora. Ingeniero Industrial. Profesora titular de Estadística, Departamento de Estadística,
Universidad Carlos III de Madrid*

Vivancos Gallego, María Jesús

Enfermedades Infecciosas. Hospital Ramón y Cajal (Madrid)



PRÓLOGO: UN VISTAZO A LA INVESTIGACIÓN DE RESULTADOS EN TÉRMINOS DE SALUD

MIGUEL ÁNGEL CALLEJA y FÉLIX LOBO

En 1913 Ernst **Codman**, un cirujano de Massachusetts, escribió que los hospitales informaban sobre el número de pacientes tratados, pero no indicaban si el tratamiento les había resultado beneficioso. Sostuvo que los hospitales deberían informar sobre los resultados de sus tratamientos de forma armonizada, para permitir comparaciones entre ellos. Incluso sugirió que todos los aspectos de la asistencia hospitalaria deberían ser evaluados para garantizar que producían efectos favorables (Jefford, Stockler y Tattersall, 2003). Codman fue una voz pionera de una visión que se consolidaría en el último tercio del siglo XX.

El término “resultado” (*outcome*, en inglés) fue consagrado por **Donabedian** en el marco de sus decisivas contribuciones al análisis y medida de la calidad de los servicios sanitarios. En particular, la aplicación con este fin de la trilogía “estructura”, “proceso”, “resultado”, supuso un nuevo paradigma para la evaluación de la calidad (Donabedian, 1966). Mientras que ciertos resultados, como la muerte, son fácilmente reconocibles y medibles, llamó la atención sobre la dificultad de definir y medir otros. Entre estos últimos incluía las actitudes y satisfacción de los pacientes, la reinserción social, la discapacidad física y la rehabilitación. Pero subrayó que los resultados en términos de salud de los pacientes deben ser siempre vistos como la dimensión fundamental para validar la efectividad y la calidad de la atención médica (Jefford, Stockler y Tattersall, 2003). Por eso, las medidas de estructura y de proceso resultan insuficientes.

En el Reino Unido, a mediados de los años setenta del pasado siglo, **Allan Williams** un profesor de Economía de la Universidad de York, que había empezado a investigar y trabajar en Economía de la Salud para el Ministerio de Sanidad británico y el NHS, buscaba procedimientos y medidas que permitieran evaluar las intervenciones sanitarias y establecer prioridades entre ellas para mejorar la eficiencia del sistema. En esa época conoció y empezó a trabajar con Rachel Rosser, una psiquiatra que se había planteado cómo medir el resultado de las hospitalizaciones en su especialidad. Esta fertilización cruzada entre profesionales de muy distintas disciplinas le llevó a “diseñar una medida de la efectividad del NHS centrada en el valor intrínseco de la salud como un bien por sus propios méritos, y no sólo porque permite a la gente mejorar su contribución al bienestar de la sociedad (especialmente al mantener su productividad), lo que suponía apartarse de la teoría del capital humano” (Williams, 2005; Sugden, 2007) fructífera en otras áreas. Así, se inscribía en una corriente de investigación –con Fanshel, Bush, Culyer, Torrance...– que desde principios de los setenta estaba avanzando vigorosamente en el diseño de medidas objetivas, basadas en datos y encuestas, de los estados de salud (Sugden,

2007). En 1985 publica su artículo seminal en el *British Medical Journal* sobre la economía de la revascularización de arteria coronaria, en el que utiliza por primera vez, de forma explícita y completa, el Año de Vida Ajustado por Calidad (AVAC o QALY por sus siglas en inglés) para evaluar la eficiencia de este procedimiento quirúrgico frente a otras intervenciones (Williams, 1985; MacKillop y Sheard, 2018). El AVAC, que combina tiempo y calidad de vida, es seguramente la medida genérica por excelencia de los resultados en salud.

En los EE. UU. los estudios sobre variaciones del ejercicio de la medicina en áreas pequeñas desempeñaron un papel fundamental en el movimiento hacia la evaluación de resultados en salud (Wennberg y Gittelsohn, 1973). Ante estas variaciones Wennberg se preguntó: ¿cuál es la tasa correcta de intervenciones médicas en un determinado segmento poblacional? ¿Cuál es la que lleva a mejores resultados para los pacientes? Para un determinado paciente aquejado de una enfermedad concreta algún tipo de operación tiene que ser más efectivo que otro. A partir de 1987 Wennberg fue becado por el National Center for Health Services Research and Health Care Technology Assessment (NCHSR) del Gobierno de los EE. UU. (Salive, Mayfield y Weissman, 1990) para estudiar estas cuestiones no resueltas sobre efectividad y resultados en salud. “Estos estudios demostraron la viabilidad y la productividad de conformar equipos multidisciplinares dirigidos a explorar territorios dominados por la incertidumbre utilizando la investigación de resultados en términos de salud” (Salive, Mayfield y Weissman, 1990).

La **Agency for Health Care Policy and Research (AHCPR)** sucedió en 1989 al NCHSR. La agencia desarrolló los *Patient Outcomes Research Teams (PORTs)*, encargados de estudiar la efectividad de los tratamientos, bajo la óptica de los resultados en términos de salud relevantes para los pacientes, en distintas enfermedades (inicialmente infarto agudo de miocardio, hiperplasia benigna de próstata, carcinoma invasivo local de próstata, dolor de espalda y cataratas). Los hallazgos de estos trabajos (por ejemplo, que ciertas operaciones de espalda o algunos medicamentos no eran efectivos) ocasionaron gran polémica. La agencia fue obligada a cambiar de orientación hacia derroteros menos conflictivos con los intereses creados. “La AHCPR cambió fundamentalmente el significado político de la Investigación de Servicios de Salud, transformándola desde una oscura actividad de interés sólo para un pequeño grupo, en una actividad visible y potencialmente importante que puede afectar a temas políticamente controvertidos” (Gray, Gusmano y Collins, 2003). Más en concreto sus directivos abogaron muy tempranamente por tener en cuenta las preferencias de los pacientes y la calidad de vida relacionada con la salud (Clancy y Eisenberg, 1998).

La relevancia de la investigación de resultados en salud se ha consagrado plenamente al ser adoptada como el enfoque más prometedor en el área de la gestión de servicios sanitarios por uno de los más destacados y conocidos profesores de Economía de la empresa y de las escuelas de negocios. Las siguientes palabras de **Porter** han sido citadas innumerables veces: “El objetivo general y predominante de los servicios sanitarios debería ser lograr un valor elevado para los pacientes, enten-

diendo por valor los resultados en salud obtenidos por unidad monetaria empleada. Este objetivo es el que importa a los pacientes y el que aúna los intereses de todos los actores del sistema. Si se incrementa el valor generado, pacientes, financiadores, clínicos y proveedores se benefician todos y mejora la sostenibilidad económica del sistema” (Porter y Teisberg, 2006; Porter, 2010).

Volviendo otra vez la vista hacia el Reino Unido, vemos cómo la institución más influyente en el mundo en el campo de la evaluación de tecnologías e intervenciones sanitarias, el National Institute for Health and Clinical Excellence (**NICE**), define hoy su misión en su página web por referencia a los resultados en términos de salud de interés para los pacientes: “¿Qué hacemos? La misión del NICE es mejorar los resultados para la gente, utilizando el NHS y otros servicios de salud pública y los servicios sociales” (National Institute for Health and Clinical Excellence, 2020). En la guía de métodos para la evaluación de tecnologías vigente, el segundo principio fundamental es que “una tecnología se considera clínicamente efectiva si en el ejercicio clínico normal proporciona un beneficio en salud, teniendo en cuenta cualquier efecto dañino y los costes de oportunidad”. En el caso base la perspectiva a considerar para los resultados es incluir “todos los efectos sobre la salud, de los pacientes o de los cuidadores”. “Se expresarán en AVACS y el EQ-5D será la medida preferida de la calidad de vida relacionada con la salud en adultos” (National Institute for Health and Clinical Excellence, 2018, pp. 14 y 32).

Estos desarrollos están directamente relacionados con un ámbito de investigación y estudio que se denomina **Investigación de los Resultados en términos de Salud (IRS)**. Es una metodología que identifica y mide la relación entre tratamientos o intervenciones y los resultados realmente obtenidos. De forma simple nos ayuda a determinar qué funciona y qué no funciona en el ámbito de la asistencia sanitaria (Ellis, 2015). El crecimiento de este enfoque responde a la tendencia hacia modelos de asistencia sanitaria y ejercicio de la medicina que desplazan su centro de gravedad desde los clínicos hacia el paciente y que permiten valorar mejor la responsabilidad de los servicios y los profesionales sanitarios frente a sus enfermos. Las tecnologías de la información, la disponibilidad de grandes bases de datos y la informatización de historias clínicas y registros administrativos han facilitado mucho el desarrollo de la IRS.

Es interesante señalar los paralelismos y diferencias entre IRS y **ensayos clínicos**. Estos consideran únicamente datos concretos, en un marco metodológico muy estricto, para determinar la eficacia de intervenciones específicas, mientras que la IRS adopta un punto de vista más amplio que incorpora resultados clínicos, impactos financieros y todo un abanico de medidas funcionales, incluyendo calidad de vida y satisfacción reportados por los pacientes. Sus datos provienen de una variedad de fuentes y metodologías: historias clínicas, bases de datos de entes financiadores, cuestionarios de pacientes, etc. Así, puede proporcionar orientaciones en relación con un conjunto de intervenciones y decisiones más amplio que los que consideran los ensayos clínicos. Estos primariamente tratan de fundamentar decisiones clínicas, mientras que la IRS se dirige a sustentar decisiones de política

de salud. La IRS tiende a centrarse en las condiciones de uso habitual, en el ejercicio diario de los cuidados de salud, con datos del mundo real, para determinar su efectividad clínica, por contraposición a los ensayos clínicos que se basan en diseños experimentales “de laboratorio”. Por tanto, son diferentes, pero no compiten sino que se complementan (Jefford, Stockler y Tattersall, 2003; Soto Álvarez, 2005; Ellis, 2015).

Después de la evolución que ha experimentado la Investigación de Servicios de Salud y de Resultados en términos de salud hoy nos parece obvio que se proclame que el objetivo de la asistencia sanitaria, los medicamentos, y otras tecnologías, los programas de prevención y, en general, las intervenciones sanitarias, es mejorar la salud de los pacientes o de las poblaciones, según corresponda una perspectiva individual, clínica, o una perspectiva poblacional, epidemiológica. Por otro lado, desde el punto de vista de la calidad de la asistencia sanitaria, también hemos aprendido que “las medidas de calidad basadas en variables de estructura o de proceso, se revelaron insuficientes. Por ejemplo, pueden incrementar procesos asistenciales que no mejoran los resultados últimos. “Si queremos mejorar los resultados de la asistencia sanitaria necesitamos medirlos directamente en términos de mortalidad, discapacidad y calidad de vida, mejor que medir las variables referidas a la estructura o al proceso de la asistencia” (Salive, Mayfield y Weissman, 1990).

Pero la salud no es una variable directamente observable y por eso se plantean los conocidos problemas de definir qué se entiende por salud y resultados en salud y decidir qué variables utilizar para medirlos. También, muchas veces, hay que preguntarse quiénes han de ser las personas destinatarias de tales intervenciones o tecnologías.

Para solucionar estos problemas es muy importante la distinción entre **medidas** de resultados o de efectos en términos de salud **intermedias** y **finales**. Las medidas intermedias (como nivel de glucosa, de colesterol o presión sanguínea), no reflejan resultados finales en términos de salud, sino que son parámetros fisiológicos que pueden tener relación con ella. Las variables intermedias sin duda son importantes para calificar las mejoras que una determinada intervención genera respecto de otras. Pero para conocer su impacto sobre la salud misma es necesario probar que están relacionadas con las variables del resultado final. Entre estas variables de resultado final es tradicional distinguir entre medidas **unidimensionales** (por ejemplo mortalidad, supervivencia, años de vida ganados, años de vida ajustados por calidad...) y medidas multidimensionales (que tratan de caracterizar y cuantificar un conjunto más o menos amplio de atributos de la salud). Estas últimas pueden ser **genéricas** (aplicables a todas las personas) o **específicas** de una enfermedad o afección determinada. Una medida genérica tiene la ventaja fundamental de que nos permite comparar entre distintas decisiones o cursos de acción, de modo que hace posible apreciar el coste de oportunidad de elegir una frente a otras posibles. Naturalmente las medidas genéricas, tienen limitaciones porque no dejan de ser una simplificación (Drummond *et al.*, 2015).

La Investigación de resultados en Salud abarca tanto los **resultados clínicos** (efectividad clínica), como los **humanísticos** (calidad de vida, nivel de satisfacción, estado de salud y preferencias de los pacientes), económicos (eficiencia e impacto presupuestario) y de **gestión sanitaria** (calidad asistencial, desempeño de los servicios y del sistema). Las dos primeras categorías serían las que más directamente atraen la atención de la perspectiva de los resultados en salud. Pero no cabe duda de que, indirectamente, unos resultados económicos y de gestión sanitaria positivos abren más oportunidades para conseguir mejores resultados directos en términos de salud y para más pacientes.

Desde hace unos años, estamos debatiendo sobre los resultados en salud y aunque se ha avanzado mucho, es aún mucho lo que tenemos que mejorar. Es habitual que ante la adquisición de cualquier tecnología o de un medicamento realicemos una evaluación para valorar su eficiencia y su adecuado posicionamiento terapéutico antes de su uso. Pero no es tan habitual que los resultados en los pacientes en los que los hemos aplicado sean reevaluados de una forma agregada y comparativa con los resultados de los ensayos clínicos que permitieron su aprobación. Los resultados en salud en vida real son fundamentales y nuestro sistema requiere un cambio y transformación completa para reorientarse hacia ellos. Necesitamos que se consolide la cultura de reevaluación de los resultados en salud pero, ¿qué resultados? ¿Cómo y cuando medirlos? ¿Solamente clínicos o también percibidos por el paciente? ¿Cuáles son los específicos de cada patología?

El objetivo de este libro es, precisamente, estudiar, con el enfoque de la IRS, todas estas interrogantes y darles respuesta. Para ello, el libro se estructura de la siguiente manera. Un primer capítulo introductorio expone la necesidad de la evaluación de resultados en términos de salud, sus tipos y los distintos diseños de los estudios, así como las tareas pendientes y las oportunidades que se plantean al respecto en España. Los tres capítulos siguientes examinan el tema desde las tres perspectivas de observación posibles: clínicos, pacientes y terceros observadores. El capítulo cinco se dedica a los problemas de la incorporación de medidas de resultados en salud en los estudios de evaluación económica y el sexto ofrece una síntesis de las técnicas estadísticas relevantes. El séptimo capítulo ofrece un resumen de los métodos disponibles para obtener resultados en salud en el mundo real. Desde el capítulo octavo hasta el diecisiete se detallan las medidas y métodos utilizados en distintos grupos de patologías: enfermedades infecciosas, infección por HIV, hepatitis C crónica, reumatología, dermatología, aparato digestivo, incluyendo enfermedad inflamatoria intestinal e intestino irritable, esclerosis múltiple, oncología con especial atención a la compleja pregunta del beneficio clínico en esta área, hematología y patología cardiovascular. Las enfermedades raras son objeto de atención en el capítulo dieciocho, y el diecinueve se dedica a la medición de resultados en términos de salud cuando los pacientes son niños y adolescentes.

Finalmente debemos felicitar a los autores por sus extraordinarias aportaciones, por haber sido capaces de sintetizar una temática compleja, por hacer accesibles los conceptos más importantes y por hacernos comprender qué son los

resultados en salud, cómo y cuándo se miden y cuáles deben medirse en cada patología concreta. Con este esfuerzo sin duda contribuirán a mejorar nuestro sistema sanitario. Esperamos que el libro resulte útil a los lectores y puedan aplicarlo en sus diferentes entornos asistenciales, docentes e investigadores.

■ REFERENCIAS

- CLANCY, C. C. y EISENBERG, J. M. (1998). Outcomes Research: measuring the end results of health care. *Science*, vol. 282, pp. 245-246. Disponible en: <https://www.semanticscholar.org/paper/Outcomes-Research%3A-Measuring-the-End-Results-of-Clancy-Eisenberg/4802aaae5ba43a4e7ac8a08f637b1b144f628e54>
- DONABEDIAN, A. (1966). Evaluating the quality of medical care. *Milbank Memorial Fund Quarterly*, 44, pp. 166-206.
- DRUMMOND, M. F., SCULPHER, M. J., CLAXTON, K., STODDART, G. L. y TORRANCE, G. W. (2015). *Methods for the economic evaluation of health care programs*, 4ª ed. Oxford: Oxford University Press.
- ELLIS, L. (2015). Using Health Outcomes Research to improve quality of care. Harvard T. H. Chan School of Public Health. Courses. Harvard University. Disponible en: <https://www.hsph.harvard.edu/ecpe/using-health-outcomes-research-to-improve-quality-of-care/>
- GRAY, B. H., GUSMANO, M. K. y COLLINS, S. R. (2003). AHCPR and the changing politics of Health Services Research. *Health Affairs*, 25, W3 283-307. Disponible en: <https://www.healthaffairs.org/doi/pdf/10.1377/hlthaff.W3.283>
- JEFFORD, M., STOCKLER, M. R. y TATTERSALL, M. H. N. (2003). Outcomes research: What is it and why does it matter? *Internal Medicine Journal*, 33, pp. 110-118.
- MACKILLOP, E. y SHEARD, S. (2018). Quantifying life: Understanding the history of Quality-Adjusted Life-Years (QALYs). *Social Science and Medicine*, 211, pp. 359-366. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6079188/pdf/main.pdf>
- NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE (NICE). (2018). *Guide to the methods of technology appraisal 2013. Process and methods*. Publicado el 13 de abril de 2013. Disponible en: <https://www.nice.org.uk/process/pmg9>. <https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781>
- (2020). *What we do?* Disponible en: <https://www.nice.org.uk/about/what-we-do>
- PORTER, M. E. (2010). What is value in health care? *New England Journal of Medicine*, 23. Disponible en: <https://www.nejm.org/doi/full/10.1056/nejmp1011024>
- PORTER, M. E. y TEISBERG, E. O. (2006). *Redefining healthcare: creating value-based competition on results*. Boston, EE.UU. : Harvard Business School Press.
- SALIVE, M. E., MAYFIELD, J. A. y WEISSMAN, N. W. (1990). Patient Outcomes Research Teams and the Agency for Health Care Policy and Research. *Health Services Research*, 25(5).
- SOTO ÁLVAREZ, J. (2005). Investigación de Resultados en Salud: El conocimiento del valor terapéutico añadido de los medicamentos. *Pharmacoeconomics – Spanish Research Articles*, 2(3), pp. 111-115.
- SUGDEN, R. (2007). Alan Harold Williams (Obituary). *Proceedings of the British Academy*, 150, pp. 313-331. Londres: The British Academy.

- WENNBERG, J. y GITTELSON, A. (1973). Small area variations in health care delivery. *Science*, 182, pp. 1102-1108.
- WILLIAMS, A. (1985). The Economics of coronary artery bypass grafting. *British Medical Journal*, 291, pp. 326-329.
- (2005). Discovering the QALY. Or how Rachel Rosser changed my life. En: A. OLIVER (ed.). *Personal histories in health research. The Nuffield Trust* (pp. 191-206). Londres. Disponible en: <https://www.york.ac.uk/media/che/documents/Williams%20on%20discovering%20the%20QALY.pdf>



1

EVALUACIÓN DE RESULTADOS EN TÉRMINOS DE SALUD: POR QUÉ, ¿CÓMO, CUÁNDO Y PARA QUÉ?

JAVIER SOTO ÁLVAREZ y EMILIO VARGAS CASTRILLÓN

■ 1.1. LA EVALUACIÓN DE RESULTADOS EN TÉRMINOS DE SALUD Y SU NECESIDAD

El Sistema Nacional de Salud (SNS) español, como los del resto de países desarrollados de nuestro entorno, se enfrenta en la actualidad a grandes retos demográficos y sanitarios, como el envejecimiento paulatino de la población, la cronificación de muchas enfermedades antes mortales y la continua llegada de nuevas intervenciones sanitarias (medicamentos y otras tecnologías) que en muchos casos aportan un grado importante de innovación en el tratamiento de las enfermedades, pero con un coste asociado elevado. Estos motivos sitúan al SNS en una encrucijada, ya que cada vez son mayores las necesidades sanitarias por parte de la sociedad y, sin embargo, los recursos disponibles cada vez son más limitados para la atención sanitaria de la población.

El fin último de nuestro SNS es incrementar la calidad asistencial y la calidad de vida de los pacientes, así como generar los mayores beneficios terapéuticos que aumenten su expectativa de vida y disminuyan al máximo la morbi-mortalidad. Por ello parece evidente que el SNS tendría que evolucionar y acometer reformas para poder garantizar en el futuro su calidad, sus prestaciones, y su sostenibilidad financiera. Igualmente, la industria farmacéutica deberá centrar sus esfuerzos en desarrollar intervenciones sanitarias realmente innovadoras, que tengan un impacto demostrable en el tratamiento de los pacientes.

En este contexto actual, una de las estrategias que el SNS debería implementar cuanto antes es maximizar el uso de aquellas intervenciones sanitarias que realmente aportan valor añadido a los individuos y a la sociedad, tanto valor clínico, como valor en eficiencia y valor para los pacientes y para la sociedad (incremento en la productividad) (Soto, 2005). A la hora de cuantificar el valor que aportan los nuevos tratamientos es decisivo el criterio de los resultados en términos de salud generados, que va más allá de lo ocurrido en los ensayos clínicos controlados (ECC), con los que tradicionalmente se evalúan. Esto es, se trata de saber cómo funcionan dichos nuevos tratamientos y qué beneficios terapéuticos producen en la práctica médica rutinaria del mundo real, cuando se utilizan en todo tipo de pacientes y en el marco del control y seguimiento ordinario de estos dentro del SNS.

Por tanto, cada vez va a ser más importante medir y evaluar los resultados en salud generados por las terapias administradas a los pacientes, de tal forma que se puedan monitorizar adecuadamente y adoptar las decisiones terapéuticas más

adecuadas en el día a día, en consonancia con los resultados obtenidos (Clauy y Eisenberg, 1998). De esta manera, los profesionales sanitarios, los pacientes, los financiadores y otros decisores del sistema podrían tomar decisiones fundamentadas en pruebas científicas de mayor calidad, lo que permitiría poner en práctica una medicina centrada realmente en el paciente, con beneficios para los ciudadanos, la sociedad y el SNS, tal y como se lleva haciendo en otros países desde hace tiempo (Frank, Basch y Selby, 2014).

Durante muchos años, y debido al movimiento de la “medicina basada en la evidencia”, se ha predicado que las decisiones terapéuticas de los profesionales sanitarios se debían basar, casi exclusivamente, en los datos proporcionados por los ECC y se ha atribuido un peso limitado a las evidencias aportadas por los estudios destinados a conocer los resultados en salud que producen las alternativas terapéuticas en condiciones de uso rutinario del mundo real. En estas condiciones los clínicos recetan según su criterio, los pacientes cumplen con el régimen posológico prescrito de manera irregular (con una mala adherencia y persistencia a los tratamientos), y los pacientes son seguidos y controlados por los profesionales sanitarios según la rutina del SNS (Sanson-Fisher *et al.*, 2007).

Por estos motivos, el interés por la evaluación de los resultados en salud dentro de las organizaciones sanitarias de nuestro SNS se ha incrementado de manera importante en los últimos años y va a ser mayor en un futuro cercano. Dicho interés se ha visto impulsado por la expansión de las nuevas tecnologías sanitarias, la preocupación por el rápido y constante aumento de los costes sanitarios, la variabilidad de la práctica médica y la atención sanitaria, la participación del paciente en la toma de decisiones y la utilización de nuevos sistemas de información en la práctica médica (Soto, 2007).

■ 1.2. TIPOS DE RESULTADOS EN SALUD

Los estudios que evalúan los resultados en salud se orientan y focalizan en cuantificar, analizar e interpretar los beneficios terapéuticos que producen los tratamientos administrados en el mundo real, una vez estos se han comercializado y se empiezan a prescribir a todo tipo de pacientes en los que están indicados según los criterios seguidos por los profesionales sanitarios (Garrison *et al.*, 2007). Por lo tanto, esta nueva área de conocimiento se centra en medir distintos resultados en salud, tales como resultados clínicos, económicos, humanísticos y de gestión sanitaria.

En relación con los resultados en salud clínicos, el área de mayor interés radica en conocer los beneficios clínicos de las alternativas terapéuticas en la práctica diaria comunicados por los profesionales sanitarios, esto es, conocer su grado de efectividad clínica y seguridad. Otros datos de gran interés son la evolución de los síntomas de la enfermedad, así como los informes de morbilidad (reingresos,

complicaciones, recidivas, exacerbaciones, etc.) y de mortalidad de las enfermedades tratadas a medio-largo plazo. También los resultados proporcionados por los observadores de los pacientes distintos de los profesionales sanitarios (familiares, amigos, cuidadores, etc.), junto a detalles de cumplimiento terapéutico y persistencia con los tratamientos, aspectos cada vez más cruciales para obtener los resultados en salud deseables (Walton *et al.*, 2015; Powers *et al.*, 2017; Sandy *et al.*, 2018; Osterberg y Blaschke, 2005).

En cuanto a los resultados en salud económicos, se evalúa el consumo de recursos derivados de emplear las distintas alternativas disponibles (diferentes tipos de costes generados), así como su nivel de eficiencia al relacionar los costes producidos con los resultados en salud generados, tales como años de vida ganados o años de vida ganados ajustados por calidad de vida (AVAC). Por otra parte, cada vez es más importante conocer cómo las intervenciones sanitarias modifican el coste del manejo de la enfermedad (a veces, ahorrando recursos de otras partidas presupuestarias), así como saber el impacto que su empleo va a suponer para el financiador.

Los estudios de resultados en salud humanísticos se centran en los resultados percibidos y comunicados por los pacientes (*Patient Reported Outcomes, PROs*), los cuales van a permitir que los pacientes puedan expresar de manera subjetiva como la enfermedad que padecen y los tratamientos administrados les afectan en su vida diaria. Por lo tanto, evalúan cómo las distintas opciones terapéuticas afectan a los pacientes en su calidad de vida relacionada con la salud, nivel de satisfacción, preferencias por los distintos tratamientos y nivel de utilidad asignado a los mismos y la discapacidad y el estado funcional que producen la enfermedad y los tratamientos (Casas, Repullo y Pereira, 2001; Villar *et al.*, 2009; AMCP Partnership Forum, 2018).

Los estudios que atienden a los resultados relacionados con la gestión sanitaria evalúan los servicios sanitarios del SNS y elaboran y calculan indicadores que puedan reflejar los resultados en términos de salud poblacional. Por ejemplo, las tasas anuales de infartos de miocardio o de ictus en una determinada región. Su conocimiento y evaluación sirve para conocer si las intervenciones sanitarias empleadas son las más adecuadas, o si deberían replantearse las políticas sanitarias a ese nivel.

■ 1.3. DISTINTOS DISEÑOS DE LOS ESTUDIOS DE RESULTADOS EN TÉRMINOS DE SALUD

Aunque durante mucho tiempo el patrón de oro en investigación han sido los ECC clásicos, por su elevada validez interna y su credibilidad dentro de la comunidad científica, cada vez se tienen más en cuenta sus importantes limitaciones. Sobre todo, los criterios restrictivos de selección de los pacientes participantes (sue-

len excluir ancianos, niños o embarazadas, pacientes polimedificados o pluripatológicos), hacen que sus resultados no se puedan extrapolar a todos los que padecen la enfermedad (Rothwell, 2005; Tannock *et al.*, 2016). Por lo tanto, hacen falta nuevos diseños orientados a obtener resultados en salud con elevada validez interna y externa, que permitan extrapolar los resultados a todo tipo de pacientes.

Uno de estos diseños es el de los ensayos clínicos pragmáticos que presentan una aceptable validez interna y externa de los resultados, ya que reflejan de manera bastante fidedigna lo que sucede en la práctica médica diaria. Sin embargo, requieren mucho tiempo hasta disponer de los resultados (varios años), son muy caros y necesitan de tamaños muestrales muy amplios (Lurie y Morgan, 2013).

Otro tipo de diseño cada vez más empleado para conocer resultados en salud es el de los estudios de cohortes observacionales prospectivos basados en registros médicos. En ellos se van incluyendo a los pacientes que padecen una enfermedad determinada a medida que son diagnosticados y se les sigue hasta el desenlace de la enfermedad, recogiendo información sobre los tratamientos administrados y los diferentes resultados en salud en los pacientes (eventos, complicaciones, progresión, supervivencia, efectos adversos, cumplimiento, consumo de recursos, calidad de vida, absentismo laboral, discapacidad, etc.). Actualmente en los países industrializados cada vez se elaboran más registros, en su mayoría auspiciados por las sociedades científicas y las autoridades sanitarias (Travers *et al.*, 2015).

Además, para realizar estudios de resultados en salud cada vez se emplean más los diseños observacionales retrospectivos basados en bases de datos administrativos y en la historia clínica electrónica de los pacientes, debido al auge de las tecnologías de la información en los últimos años. Otro posible diseño para obtener resultados de la vida real es el de las encuestas de salud, que preguntan directamente a los pacientes sobre diferentes aspectos de su estado de salud, consumo de recursos, patrón de tratamiento de la enfermedad, etc. (Hampson *et al.*, 2018; Lin *et al.*, 2013).

Los diseños observacionales (basados en registros, bases de datos, historia clínica electrónica, etc.), hoy predominantes para obtener resultados en salud, son métodos que están expuestos a sesgos importantes (selección, análisis y seguimiento) y a factores de confusión de muy difícil control, al no emplear la aleatorización. Por ello, es imprescindible que su diseño y su realización sigan los más altos estándares metodológicos y el máximo de rigor científico, de tal forma que los resultados obtenidos sean creíbles y relevantes para los profesionales sanitarios y otros decisores del SNS (Cox *et al.*, 2009; Berger, Dreyer y Anderson, 2012).

Por otra parte, la posibilidad de explotar, cruzar y analizar muchísima información, gracias al desarrollo exponencial de la digitalización de la sanidad y la revolución del *big data*, la inteligencia artificial y el *machine learning*, permitirá diagnosticar y tratar precozmente las enfermedades, alertar sobre posibles epidemias, adaptar de manera más eficaz y eficiente los tratamientos a los pacientes (según su

respuesta y la evolución de la enfermedad) y reducir visitas a los centros, con ahorro de recursos del SNS y mejores resultados para los pacientes. Van a ser técnicas muy importantes para obtener datos de resultados en salud en un futuro cercano, con un coste no muy elevado y con menores dilaciones en la disponibilidad de los datos (Crown, 2015; Jones *et al*, 2018).

■ 1.4. TAREAS PENDIENTES Y OPORTUNIDADES EN ESPAÑA

El SNS de nuestro país dispone de muchísima información terapéutica generada a lo largo de los años, pero no se evalúa de forma habitual y continua, sino que solamente se analiza en momentos puntuales, cuando es necesario disponer de datos por algún motivo concreto. Lo primero que hay que hacer es generar un cambio cultural en todos los agentes del SNS. Deben percibir que los programas dedicados a obtener datos de resultados en el mundo real pueden cubrir la brecha existente entre los ECC clásicos y lo que sucede en la práctica médica diaria y, por lo tanto, van a generar evidencias sustentadoras de mejores decisiones terapéuticas en beneficio de los pacientes. Además, hay que inculcar en el SNS la filosofía de que es necesario invertir recursos para que estos programas sean de suficiente calidad metodológica y creíbles para los profesionales sanitarios y otros agentes decisores.

Por otra parte, es necesario que en los servicios regionales de salud se realice una progresiva expansión de la historia clínica electrónica que integre atención primaria y asistencia hospitalaria y que sea interoperable entre todas las comunidades autónomas. En estos momentos ya hay buenos ejemplos de historia clínica electrónica integrada en algunas regiones, pero es necesario que se siga avanzando y se extienda a todo el país. Además, es deseable que dentro de las historias clínicas de cada paciente se estandarice la recogida de resultados en salud, no únicamente clínicos, sino también resultados relevantes para el paciente (PROs), y otro tipo de resultados claves para poder tomar decisiones terapéuticas acertadas por los profesionales sanitarios (Concannon, 2015).

El mayor reto de nuestro SNS es conseguir mantener una asistencia de calidad y equitativa, a la vez que establecer mecanismos para mejorar la eficiencia y sostenibilidad en respuesta al incremento de la demanda de la sociedad, el cambio en el perfil asistencial y la constante llegada de nuevas innovaciones. Para conseguir un sincretismo entre la inversión realizada y los resultados en salud alcanzados por la incorporación de intervenciones sanitarias innovadoras es clave disponer de datos de su desempeño en la práctica médica diaria y los resultados en salud que se consiguen en el mundo real.

Es necesario que dentro del SNS se considere una prioridad la obtención de datos de calidad y que se avance en la implantación de sistemas de información a través de los planes de salud de las comunidades autónomas (CC. AA.). Un ejemplo

es el plan oncológico de Euskadi 2018-2023 que prevé integrar una aplicación en la historia clínica electrónica e implantar un sistema de información específico para los procesos farmacológicos que van a permitir medir resultados en salud y calidad de vida de los pacientes. Además, estos datos van a ser claves a la hora de implantar en las CC. AA. los acuerdos de riesgo compartido o los esquemas de pagos por resultados en la adquisición de nuevos medicamentos, modalidades cada vez más empleados en nuestro país. En un futuro no muy lejano también podrían emplearse para vincular el presupuesto de los hospitales y centros de salud a los resultados en salud obtenidos.

Otra oportunidad que no hay que olvidar es la nueva ley de contratos del sector público, la cual obliga a todas las partes implicadas, administraciones y entidades públicas y las empresas proveedoras, a hacer un esfuerzo máximo de transparencia. Es de esperar que su desarrollo futuro evolucione hacia una compra estratégica, con una visión a medio-largo plazo, que tenga en cuenta los criterios de calidad, innovación y mejoría de resultados en salud, lo que permitiría avanzar hacia un contrato público basado en dichos resultados en salud.

Quedan aún muchos problemas por solucionar, como garantizar la seguridad y protección de los datos sanitarios a través de su anonimización, la explotación correcta de los resultados obtenidos, la colaboración idónea entre el SNS y las compañías farmacéuticas, y otros, pero es el momento de que todos los agentes implicados en el SNS trabajen de forma conjunta para lograr que el conocimiento de los datos de resultados en salud sea una realidad.

Por otra parte, sería muy interesante crear una agencia nacional coordinadora de la medición y evaluación de resultados en salud de los pacientes, empezando lógicamente por aquellas patologías más prevalentes, las que más morbi-mortalidad producen y las que mayor consumo de recursos generan al sistema. De esta forma, poco a poco, el SNS podría disponer de datos válidos y fiables de los resultados en salud generados por las alternativas de tratamiento disponibles en un número creciente de enfermedades. Este empeño puede obtener enseñanzas de experiencias internacionales, como la iniciativa *ICHOM (International Consortium for Health Outcomes Measurement)* (www.ichom.org), que pretende facilitar el progreso de los sistemas sanitarios mediante la medición y análisis de resultados en salud conforme a modelos estandarizados (Obbarius *et al*, 2017) o el *PCORI (Patient-Centered Outcomes Research Institute)* en USA, cuya finalidad es obtener resultados en salud para ayudar a los pacientes, profesionales sanitarios, financiadores y gestores del sistema sanitario a tomar decisiones más informadas y basadas en evidencias obtenidas del mundo real (Fleurence *et al.*, 2015).

■ REFERENCIAS

AMCP PARTNESHIP FORUM. (2018). Improving quality, value, and outcomes with patient-reported outcomes. *Journal of Management Care Special Pharmacy*, 24(3), pp. 304-310.

BERGER, M. L., DREYER, N. y ANDERSON, F. (2012). Prospective observational studies to assess comparative effectiveness: the ISPOR good research practices task force report. *Value in Health*, 15 (Issue 2), pp. 217-230.

CASAS, J., REPULLO, J. R. y PEREIRA, J. (2001). Medidas de calidad de vida relacionada con la salud. Conceptos básicos, construcción y adaptación cultural. *Med Clin (Barc)*, 116(20), pp. 789-796.

CLAUCY, C. M. y EISENBERG, J. M. (1998). Outcomes research: measuring the end results of healthcare. *Science*, 282(5387), pp. 245-246.

CONCANNON, T. W. (2015). Can patient centered outcomes research improve healthcare? *BMJ*, 351, h3859.

COX, E., MARTIN, B. C., VAN STAA, T. *et al.* (2009). Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effect using secondary data sources: the international Society for Pharmacoeconomics and Outcomes Research Good Research Practices for retrospective database analysis task force report. Part II. *Value in Health*, 12(8), pp. 1053-1061.

CROWN, W. H. (2015). Potential application of machine learning in health outcomes research and some statistical cautions. *Value in Health*, 18(2), pp. 137-140.

FLEURENCE, R., WHICHER, D., DUNHAM, K. *et al.* (2015). The patient-centered outcomes research institute's role in advancing methods for patient-centered outcomes research. *Med Care*, 53(8), pp. 2-8.

FRANK, L., BASCH, E. y SELBY, J. V. (2014). The PCORI perspective on patient-centered outcomes research. *JAMA*, 312(15), pp. 1513-1514.

GARRISON, L. P., NEUMANN, P. J., ERICKSON, P. *et al.* (2007). Using real-world data for coverage and payment decisions: the ISPOR real-work task force report. *Value in Health*, 10(5), pp. 326-335.

HAMPSON, G., TOWSE, A., DREITTEIN, W. B. *et al.* (2018). Real-world evidence for coverage decisions: opportunities and challenges. *J Comp Eff Res*, 7(12), pp. 1133-1143.

JONES, L. K., PULK, R., GIONFRIDDO, M. R. *et al.* (2018). Utilizing big data to provide better health at lower cost. *Am J Health Syst Pharm*, 75(7), pp. 427-435.

LIN, J., JIAO, T., BISKUPIAK, J. E. *et al.* (2013). Application of electronic medical record data for health outcomes research: a review of recent literature. *Expert Rev Pharmacoecon Outcomes Res*, 13(2), pp. 191-200.

LURIE, J. D. y MORGAN, T. S. (2013). Pros and cons of pragmatic clinical trials. *J Comp Eff Res*, 2(1), pp. 53-58.

OBARIUS, A., VAN MAASAKKERS, L., BAER, L. *et al.* (2017). Standardization of health outcomes assessment for depression and anxiety: recommendations from the ICHOM depression and anxiety working group. *Quality of Life Research*, 26(12), pp. 3211-3225.

OSTERBERG, L. y BLASCHKE, T. (2005). Adherence to medication. *NEJM*, 353(5), pp. 487-497.

POWERS, J. H., PATRICK, D. L., WALTON, M. K. *et al.* (2017). Clinician-reported outcome assessments of treatment benefit: report of the ISPOR clinical outcome assessment emerging good practices task force. *Value in Health*, 20(1), pp. 2-14.

ROTHWELL, P. M. (2005). External validity of randomized controlled trials: "to whom do the results of this trial apply? *Lancet*, 365(94539), pp. 82-93.

SANDY, L., DEMURO, C., BLOCK, S. L. *et al.* (2018). Development of a novel observer-reported outcome measure for the assessment of respiratory syncytial virus (RSV) infection symptoms in pediatric clinical trials. *Journal of Patient Reported Outcomes*, 2, p. 9.

SANSON-FISHER, R.W., BONEVSKI, B., GREEN, L.W. y D'ESTE, C. (2007). Limitations of the randomized controlled trial in evaluating population-based health interventions. *American Journal of Preventive Medicine*, 33(2), pp. 155-161.

SOTO J. (2005). Valor terapéutico añadido de los medicamentos: ¿qué es, cómo se evalúa y cuál debería ser su papel en política farmacéutica? *An Intern Med*, 22(1), pp.39-42. Madrid.

— (2007). Medicina basada en resultados en salud: la evolución lógica y deseable de la medicina basada en la evidencia. *Med. Clin (Barc)*, 128(7), pp. 254-255.

TANNOCK, I. F., AMIR, E., BOOTH, C. M. *et al.* (2016). Relevance of randomised controlled trials in oncology. *Lancet Oncology*, 17(12), e560-567.

TRAVERS, K., SALLUM, R. H., BURNS, M. D. *et al.* (2015). Characteristics and temporal trends in patient registries: focus on the life sciences industry, 1981-2012. *Pharmacoepidemiol Drug Saf*, 24(4), pp. 389-398.

VILLAR, J., LIZAN, L., SOTO, J. *et al.* (2009). La satisfacción con el tratamiento. *Atención Primaria*, 41(11), pp. 637-645.

WALTON, M. K., POWERS, J. H., HOBART, J. *et al.* (2015). Clinical outcome assessments: conceptual foundation-report of the ISPOR clinical outcomes assessment-emerging good practices for outcomes research task force. *Value in Health*, 18(6), pp. 741-752.



2

RESULTADOS COMUNICADOS POR LOS CLÍNICOS

ANDRÉS GARCÍA PALOMO, ANA LÓPEZ GONZÁLEZ y
MARÍA ROSARIO FERNÁNDEZ GONZÁLEZ

2.1. INTRODUCCIÓN

Invertir recursos en nuevas intervenciones sanitarias solo está justificado si la supuesta innovación llega cargada de valor. Según Porter, valor no es ni un ideal abstracto, ni una clave para reducir costes, sino el marco en el que se debe definir la mejora de la atención sanitaria y el impulso innovador del sistema. Valor es una razón con un numerador, que representa los resultados de la intervención, y un denominador con los costes que induce su implementación (Porter, 2010). (Figura 1).

Figura 1.

VALOR COMO RAZÓN ENTRE RESULTADOS (BENEFICIOS) Y COSTES (GASTO O INVERSIÓN)

El objetivo de un Sistema de Salud Público es devolver VALOR al paciente y que éste lo perciba.



Porter ME. N Engl J Med. 2010

Mientras que los costes se miden en unidades monetarias, los resultados son representados en términos de beneficio incremental. Se entiende como beneficio incremental cualquier efecto favorable en uno o varios aspectos significativos para el paciente, que pueden ser: vivir más, vivir con menos síntomas o retener una funcionalidad adecuada para sus necesidades diarias, comparado con lo que había o tenía. El efecto beneficioso debe ser un aspecto de la salud o la funcionalidad diaria afectado por la enfermedad, que al paciente le preocupa y sobre el que tiene expectativas de mejorar, no empeorar o que podrá ser prevenido por efecto de la intervención (Walton *et al.*, 2015).

El cálculo de los costes no es excesivamente confuso, aunque no está exento de dificultades técnicas y metodológicas. Sin embargo, la medición del beneficio, dovela del valor, se afronta con ambigüedad, es equívoco y ampliamente incomprendido. Muchas son las razones y no pocas las dificultades.

La primera y quizá la mas importante, se encuentre en la frase "...en uno o varios aspectos significativos para el paciente", porque no todos los resultados que resultan ser beneficiosos, son significativos para ellos. No cabe duda de que la involucración de los pacientes en el proceso de desarrollo de nuevos fármacos o intervenciones innovadoras, ayudará al diseño de una investigación con objetivos relevantes para ellos. Sin embargo, una interpretación literal, rigurosa y excluyente de la frase, puede representar una enorme barrera para el progreso terapéutico. La falta de "significado para los pacientes" no puede interpretarse radicalmente como una ausencia de beneficio, porque las expectativas de aquellos son heterogéneas, cambiantes, manipulables, frecuentemente inefables y algunas veces poco realistas (Coyne y Wyrwich, 2015).

La segunda se refiere a la jerarquía de los beneficios. La supervivencia, la funcionalidad y los síntomas pueden ser etiquetados en diferentes niveles según su relevancia para los pacientes. En el superior, y probablemente el más relevante e influyente sobre los inferiores, está la supervivencia (curación) y todos sus derivados "evento-tiempo". Es posible que se trate de un objetivo primordial para la gran mayoría de los pacientes, independientemente de la patología que sufran. Sin embargo, su maximización y sacralización puede no ser realista en algunas condiciones o no deseable como objetivo único para determinados grupos de pacientes. El segundo nivel es el grado de salud alcanzado, recuperado o retenido y que se mide mediante indicadores que analizan la funcionalidad para la actividad diaria básica, instrumental, avanzada o laboral. Finalmente el tercer nivel está representado por la sostenibilidad del estado de salud alcanzado y sus indicadores suelen ser referidos a recaídas y síntomas. Es posible que los niveles 2º y 3º puedan ser intercambiables en determinados marcos y condiciones (Porter y Teisberg, 2006).

La tercera está relacionada con las características de los resultados en salud. Estos suelen ser dependientes, inherentes y específicos de la condición médica que miden. Ningún resultado en particular captura todo el beneficio de la acción terapéutica. Es mas, el valor de la asistencia sanitaria, sobre todo en pacientes con enfermedades crónicas y/o pluripatología, se alcanza progresivamente en la evolución temporal y espacial del ciclo de atención y dependen del esfuerzo coordinado de todas y cada una de las intervenciones. Por tanto, la evaluación de resultados en salud debe interpretarse siempre como la medición interrelacionada de varios indicadores informados desde varias perspectivas, los cuales posteriormente serán integrados, proporcionalmente a su relevancia, en una única medida multidimensional, que representará el beneficio percibido y clínicamente significativo para el paciente (Porter, 2008).

2.2. INVESTIGACION EN RESULTADOS EN SALUD (IRS)

La IRS constituye la fuente mas importante de conocimiento sobre el valor de las intervenciones sanitarias, aportando información, fundamentalmente sobre los beneficios de estas, aunque también de los costes. Constituye la base conceptual de la Medicina Basada en la Evidencia (MBE) (Badia y Lizán, 2002).

Cuando se realiza en condiciones ideales, controladas y aleatorizadas, genera información de los resultados en clave de eficacia. Si se desarrolla sobre estudios observacionales posautorización no controlados y/o de fuentes de datos e información del mundo real (Datos de la Vida Real [DVR] o *Real Word Data* [RWD], en ingles), proporciona información sobre la efectividad, concepto este que complementa y da sentido a la eficacia, lo que nos acerca mucho más al valor (Badia y Del Llano, 2000). Dado que la estructura organizativa de los sistemas de salud y sus medios de información están, por ahora muy lejos de cubrir con calidad y fiabilidad la necesidad ética de entrega de información sobre efectividad, la IRS se nutre casi exclusivamente de resultados de eficacia, lo que da una idea solo aproximada de los beneficios.

Quizá la forma mas importante de IRS sobre eficacia sea la investigación experimental biomédica con fármacos, conocida como Ensayo Clínico Aleatorizado (ECA), en el que se realizan “evaluaciones de resultado”, entendidas como cualquier efecto inducido por un fármaco o fármacos en la modificación de un evento considerado y previamente definido. Las evaluaciones de resultados deben seguir una distribución de probabilidad conocida, que facilite su detección en la población estudiada. La base de cualquier ECA es rechazar la “hipótesis nula”, es decir, que el fármaco experimental A no mejora los efectos que consigue el control B. En función de que sea acertado o fallido el rechazo de la hipótesis nula se pueden cometer dos tipos de errores, el tipo I y el II (*Figura 2: Hipótesis nula y errores tipo I y II*). Para ello se usan herramientas estadísticas ad hoc que comparan los efectos sobre el evento medido en el grupo experimental frente al control (Laporte, 1993).

El resultado obtenido en un ECA y que se usa para interpretar si el efecto de la intervención ofrece beneficio incremental, se denomina *endpoint*. Se componen de una o varias evaluaciones de resultados preespecificadas en diferentes momentos del estudio, que se miden y analizan mediante herramientas estadísticas especifi-

Tabla 1.

HIPÓTESIS NULA Y TIPO DE ERRORES EN LA INVESTIGACIÓN BIOMÉDICA

¿Se rechaza la hipótesis NULA?	Hipótesis NULA verdadera	Hipótesis NULA falsa
Si	ERROR TIPO I (falso positivo)	CORRECTO (verdadero positivo)
No	CORRECTO (verdadero negativo)	ERROR TIPO II (Falso negativo)

Fuente: Modificado de: <https://www.eupati.eu/es/desarrollo-y-ensayos-clinicos/la-estadistica-en-los-ensayos-clinicos-conceptos-clave/>

cas. Pueden clasificarse en primarios o secundarios. Los *endpoint* primarios son los que directamente contestan a la pregunta del ECA sobre beneficio incremental. Por ello, condiciona el tamaño de la muestra y define la potencia del ECA para detectar diferencias que rechacen esa hipótesis nula ya citada. Los *endpoints* secundarios complementan y matizan esa respuesta. Este principio constituye la base para la gestión de las innovaciones farmacológicas por parte de las agencias evaluadoras y es el punto de partida para la puesta en el mercado de un fármaco. (Bakhai, Chhabra y Wang, 2006).

■ 2.2.1. Resultados para el desarrollo de fármacos

Las variables de resultados que definen beneficio pueden ser de evento-tiempo referidas a cuánto sobreviven los pacientes, evaluaciones clínicas que contestan a preguntas sobre funcionalidad y síntomas, o biomarcadores.

Se define biomarcador como una evaluación del paciente no influenciada por el juicio del evaluador, o la motivación o voluntad del paciente. Son ejemplos de biomarcadores las mediciones bioquímicas sanguíneas, muestras biológicas o las mediciones cuantitativas de las imágenes radiológicas. Siempre ofrecen una valoración indirecta y subrogada del beneficio. Las evaluaciones clínicas por el contrario son aquellas que se ven afectadas por la volición y/o motivación del paciente o por el juicio clínico del evaluador. Por tanto, son susceptibles de variación no relacionada con el efecto beneficioso o perjudicial de la intervención. Los COA (*Clinical Outcome Assessment*) son los instrumentos usados para medir estas evaluaciones clínicas en el contexto de los ECA¹

Es indudable el éxito alcanzado por la IRS basada en los ECA. Sin embargo, cada vez son más las voces que abogan por un cambio cuantitativo y cualitativo en la investigación biomédica.

La IRS basada en las ECA sufre de una profunda, crónica e inveterada falta de atención a la elección de *endpoints* primarios y secundarios causando un gran dispendio en tiempo, recursos y conocimientos. Debido a la rigidez y miopía de las normas regulatorias sobre las ECA, la mayoría de los diseños abusan de variables duras y unidimensionales como las “evento-tiempo”, que no capturan de manera integral el beneficio real y clínicamente significativo de la intervención. Es más, muchas de ellas vienen siendo sustituidas, por numerosas razones entre las que destacan los motivos económicos, por formas subrogadas de dichas variables o directamente por biomarcadores. La ya citada ausencia de pacientes en el diseño de la investigación biomédica, la falta de sensibilidad de los diseñadores, el interés económico de algunos promotores y la excesiva variabilidad, incoherencia, heterogeneidad y falta de estandarización de la mayoría de los COA, ha dado como

¹ <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM230597.pdf>. Último acceso: 27.07.2018.

resultado que lo estadísticamente significativo no deje paso a lo clínicamente relevante y que la medida del beneficio sea unidimensional. Como ya se ha señalado, el beneficio es inherentemente específico de la condición que se mide y es siempre multidimensionales, por lo que ningún resultado en particular captura toda la información sobre el beneficio.

Desde los años setenta se vienen produciendo numerosas iniciativas dirigidas a cambiar esta tendencia, mediante el desarrollo, mas allá de las evento-tiempo, de indicadores de resultados clínicos estandarizados y específicos para cada área de salud o problema sanitario. Junto a los COA (Benjamin, Vernon y Patrick, 2017) ya citados, iniciativa del ISPOR (International Society for Pharmacoeconomics and Outcomes Research), destacan los *Core Outcome Sets (COS)* (Tunis *et al.*, 2016) impulsados por el PCORI (*Patients-Centered Outcomes Research*). Aunque los datos de EMA no son demasiado claros en este sentido² FDA (US FDA, 1992) advierte de la necesidad de que los COA se incorporen de forma multidimensional a los ECA y sean parte inherente del proceso de análisis de resultados.

■ 2.2.1.1. Variables evento-tiempo y medidas de impacto

En la IRS basada en las ECA, las variables “evento-tiempo” se consideran variables “duras” para la definición de beneficio, porque informa sin ambages y directamente de él. La supervivencia se encuentra, casi siempre, en el primer nivel de relevancia de las expectativas de los pacientes. Esta consideración las hace insustituibles. Las variables “evento-tiempo” son reportadas por el clínico o el equipo investigador, pero no se afectan por el juicio e interpretación de éstos. No son por tanto, ni COA ni biomarcador.

Tanto EMA³ como FDA⁴ considera que la variable “evento-tiempo” de mayor relevancia es la *supervivencia global (OS)*, entendida como el tiempo desde la autorización (si es en el contexto de un ECA) o desde la inscripción en el estudio (fuera del contexto de los ECA) hasta la muerte por cualquier causa. Es la medida directa y universalmente aceptada de beneficio por su sencillez de interpretación. Es cualitativa, nominal y dicotómica. Sin embargo, tiene algunas dificultades. Encarece y enlentece el estudio al requerir una muestra poblacional mas grande y un seguimiento muy largo, sobre todo cuando las presumibles diferencias pueden ser muy pequeñas; es posible que incluya muertes no relacionadas con el proceso que se mide y, en caso de terapias encadenadas o secuenciales, el cruzamiento de brazos puede complicar la medición. Con la intención de reducir los tiempos de seguimiento, la OS está siendo expresada en periodos de tiempo, como la OS a corto

² http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2017/11/WC500238764.pdf. Último acceso: 17.07.2018.

³ http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137126.pdf

⁴ <https://www.fda.gov/downloads/drugsGuidanceComplianceRegulatoryInformation/Guidance/UCM071590.pdf>

(OS-1 año por ejemplo), medio (OS-3 años, por ejemplo) o largo plazo (OS-5 o 10 años, por ejemplo) o como un porcentaje de pacientes vivos cuando se alcanza un periodo de seguimiento dado (porcentaje de pacientes vivos a 1 año de seguimiento, por ejemplo). Los métodos estadísticos actuariales facilitan estos cálculos.

Los costes de los ECA han forzado a EMA⁵ y FDA⁶ a considerar la posibilidad de analizar los beneficios de un tratamiento mediante variables subrogadas. Tanto una como la otra definen variable subrogada como una medida de laboratorio, un signo físico o una variable evento-tiempo que sustituye a una variable clínicamente significativa y directa sobre lo que siente un paciente, sus funciones o su supervivencia y que se espera que prediga indirectamente el efecto de la terapia sobre la variable directa (US FDA, 1992). Existen dos variables evento-tiempo subrogadas, la *supervivencia libre de enfermedad o progresión* (SLE-SLP) y el *tiempo al fallo del tratamiento* (TTF). Se subrogan a OS porque se entiende que cuanto más se tarde en fracasar con el tratamiento, más OS se consigue a la postre. Las dos tienen en común el abaratamiento de costes y el acortamiento de las necesidades de seguimiento. La SLP o SLE, al igual que la OS, mide el tiempo hasta el fracaso del tratamiento interpretado como la aparición de un evento, previamente definido, que se interpreta como un fallo a aquel. La SLP o SLE incluye también la muerte por cualquier causa, si esta ocurre antes que dicho evento definido. Es también una variable cualitativa, nominal y dicotómica. La TTF es igual que la SLP-SLE, pero añade como evento de fallo la suspensión del tratamiento por toxicidad. Dado que la TTF mezcla eficacia y toxicidad, tanto EMA como FDA autorizan solo el uso de SLP como subrogado de OS. No obstante está sujeta a numerosos y peligrosos sesgos, no han demostrado ser universalmente subrogada de OS y requiere tiempos de evaluación equilibrados entre los brazos de tratamiento. Por ello la SLP parece convertirse poco a poco en una variable para aprobaciones condicionadas a demostrar OS.

Para evaluar las diferencias de efecto entre los brazos experimental y control se usan las denominadas medidas de impacto, herramientas estadísticas que permiten el contraste de hipótesis y el rechazo de la hipótesis nula, base de la investigación experimental. En el campo de los ECA, son medidas de impacto las mediciones de diferencias, la reducción de riesgos (relativo y absoluto) y el número necesario a tratar (NNT) (Molina Arias y Ochoa Sangrador, 2015).

Las *mediciones de diferencia* reflejan la diferencia existente entre las variables de resultados, medidas en el grupo experimental y las del control, e indican la contribución de la intervención. La cuantía de la diferencia es la magnitud del impacto. Las diferencias de medias son la herramienta para comparar variables

⁵ http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137126.pdf

⁶ <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidance/UCM071590.pdf>

cuantitativas, mientras que las diferencias de medianas se usa para el análisis de las variables cualitativas (Molina Arias y Ochoa Sangrador, 2015).

La *reducción de riesgos*, absoluta o RAR y relativa o RRR, establece la magnitud en la que se reduce el riesgo del evento medido. La absoluta o RAR es la resta directa de los riesgos en el brazo control y experimental, mientras que la RRR es la RAR dividida por el riesgo del grupo control. Aunque es moneda de cambio común usar la RRR, esta magnifica el efecto y no cuantifica claramente el esfuerzo para conseguir el beneficio. Con un ejemplo se entiende. Así si tenemos un riesgo de mortalidad del 20 % en el grupo control y del 10 % en el intervenido, la RAR será del 10 % ($0,20 - 0,10 = 0,10$), mientras que la RRR será del 50 % ($0,10/0,20 = 0,5$). Ahora si el control tiene un 0,2 % y el experimental es de 0,1 %, la RAR es del 0,1 % mientras que la RRR sigue siendo del 50 % ($0,01/0,02 = 0,5$) (Molina Arias y Ochoa Sangrador, 2015).

El *NNT* es la medida de impacto mas importante porque establece cuántos pacientes debemos someter a los efectos del fármaco para alcanzar beneficiar a uno. Se estima mediante el RAR y es su inversa, es decir, $1/RAR$. Con el ejemplo anterior, el esfuerzo en el primer caso con un RAR del 10 % es de 10 ($1/0,10$), mientras que en el segundo es de 1000 ($1/0,001$), con RRR semejantes. No todos los ECA aportan esta información tan relevante.

■ 2.2.1.2. Herramientas para evaluaciones clínicas (COA)

La FDA introdujo el término COA en octubre de 2011 para reflejar la importancia de la fuente informante en la medición de resultados clínicos. Los COA excluyen a las variables evento-tiempo porque su medición e información están influenciadas por elecciones humanas, juicios, sean estos profesionales o no, motivación y cooperación del informante. También excluyen a los biomarcadores porque estos dependen de algoritmos y procesos automatizados y no de informantes humanos. En definitiva, son mediciones basadas en evaluaciones humanas mediante observación o instrumentos de medición usados por un paciente, un clínico experto o un observador (Patrick y Burke, 2013).

Se describen cuatro tipos de COA. Las *evaluaciones de resultados de rendimiento* o *PerfO* (*Performance Outcome*) instruyen a un paciente para realizar una tarea definida que se cuantifica mediante una medición. Aunque no hay juicio clínico involucrado, es un COA porque precisa de la volición del paciente para realizar la prueba. El *resultado informado por un clínico* o *ClinRO* (*Clinician-reported Outcome*) es la evaluación del estado de salud o funcionalidad de un paciente mediante la observación, interpretación e información de la medición por un médico. El *resultado informado por un observador* u *ObsRO* (*Observer-reported Outcome*) es una evaluación que se usa cuando el paciente no puede autoreportar conceptos, observables o no. Los ObsRO proporcionan observaciones, pero no interpretaciones porque

el informante no es un profesional. Finalmente, los resultados observados por los pacientes o *PRO (Patient-reported Outcome)* son más desarrollados y conocidos. El PRO es un informe reportado por el paciente sobre su estado general, aspectos particulares o eventos relacionados con su patología o su salud, sin interpretación por parte de profesionales u observadores. Los PRO más significativos incluyen signos y síntomas, calidad de vida y funcionalidad, aunque también satisfacción y, en algunos casos, adherencia.

Existe un quinto tipo de COA, no globalmente aceptado que es el ProxyRO, un concepto mezcla de ObsRO y PRO, mucho más cercano a este último, porque es otra persona la que informa como si fuera el paciente.

Los COA proporcionan evidencia complementaria de los beneficios del tratamiento, pero en algunas ocasiones pueden considerarse como *endpoint* primarios en función de su relevancia. Los PerfO tienen una larga tradición en Medicina y en IRS, a las que últimamente se han incorporado los PRO, pero están sujetos a una rigurosa evaluación de sus aportaciones. Los ObsRO se han usado en condiciones muy específicas, como Psiquiatría o Pediatría, sobre todo, pero la consideración sistemática, sus fundamentos conceptuales y sus propiedades de medición son relativamente nuevas. Finalmente los ClinRO tienen una larga historia y son asumidos como válidos porque están cargados con el “juicio clínico”; sin embargo, carecen en muchos casos de evaluaciones rigurosas de su valor y la validez de su aportación.

Todas las iniciativas están de acuerdo en que constituyen un conjunto de información que deben integrarse y estandarizarse en los ECA y que la normativa sobre los mismos debe modificarse para recogerlos como parte fundamental y obligatoria para el investigador y los promotores. Este proceso es fundamental para incorporar los resultados centrados en el paciente (*Patient centered Outcomes* o *PCO*) en la IRS⁷.

■ 2.3. RESULTADOS REPORTADOS POR CLINICOS (CLINRO)

Según la definición de la FDA, se basa en un informe emitido por un profesional de la salud capacitado, después de la observación del paciente o la aplicación de un instrumento de medida o una escala. Implica siempre la emisión de un juicio clínico que interpreta signos observables, pruebas diagnósticas, comportamientos u otras manifestaciones físicas que se cree que están relacionadas con la enfermedad o la condición medida. Las medidas ClinRO no evalúan directamente síntomas, ya que estos solo pueden ser reportados por el paciente, salvo que ese síntoma tenga reflejo en un signo clínico (por ejemplo, palpitations que se reflejan en un ECG anormal, o fiebre que se termometra) (Powers *et al.*, 2017).

⁷ <http://wayback.archive-it.org/7993/20180424212148/https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm370262.htm#ClinRO&xcust=pdfmyurl>. Último acceso: 27.07.2018.

■ 2.3.1. Tipos de ClinRO

Según los informes del ISPOR⁸, las ClinRO se pueden dividir en tres tipos, lecturas e interpretaciones de pruebas (*ClinRO Reading*), calificaciones impulsadas por protocolo (*ClinRo rating*) y puntuaciones globales (*ClinRo global assessment*).

En cualquiera de las ClinRO citadas, la participación del paciente puede ser pasiva como en las ObsRO, cuando se le observa a él o a una prueba médica, o activa, como las PRO o las PerFO, según se le pregunte o se le haga realizar algún tipo de prueba. Tras ello, el clínico deberá obligatoriamente analizar los datos obtenidos, para en un segundo término informar aplicando su interpretación o “juicio clínico”, mediante una variable de resultado. Esta es la clave para diferenciar las ClinRO del resto de las COA, PerFO, PRO u ObsRO en las que no cabe el juicio clínico.

■ 2.3.1.1. ClinRO Readings

Son lecturas de signos clínicos o interpretación de pruebas. En ambos casos, y tras ser observados, caracterizados y analizados, serán informados en forma de variable dicotómica (sí/no; presente/ausente) tras el juicio clínico experto o el uso de un protocolo estandarizado de evaluación. La participación del paciente no es importante para el proceso de desarrollo, análisis e interpretación.

Pueden considerarse ClinRO la medición de constantes vitales, la detección de signos exploratorios o cualquier *end point* que sea mediado por el juicio clínico y conlleve una acción dicotómica. Por ejemplo, trasfundir o no, administrar oxígeno o no, realización de un procedimiento o prueba radiológica o la hospitalización. En todo caso, las *ClinRO Readings* no suelen ser consideradas *end points* primarios en los ECA.

■ 2.3.1.2. ClinRO rating

Las *ClinRO rating* o calificaciones impulsadas por tablas y protocolos, están representadas por escalas compuestas de variables categóricas, ordenadas o no, o bien medidas continuas en las que son reconocibles al menos, tres niveles posibles, que generan puntajes para representar el concepto de interés o variable de resultado.

Las *ClinRO rating* son las más abundantes. Son ejemplos en Oncología las escalas de toxicidad CTCAE⁹ o las escalas de medición de la respuesta como los

⁸ <https://www.fda.gov/downloads/drugsGuidanceComplianceRegulatoryInformation/Guidance/UCM071590.pdf>

⁹ <http://www.hrc.govt.nz/sites/default/files/CTCAE%20manual%20-%20DMCC.pdf>. Último acceso: 28.07.2018.

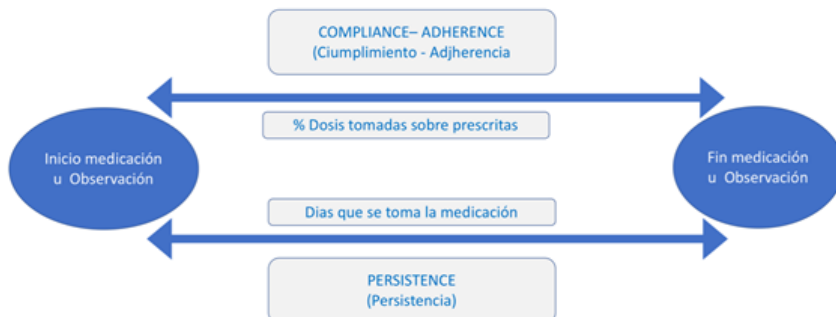
criterios RECIST¹⁰, ejemplo palmario con el que se construyen dos variables de resultados: las *tasas de respuesta* (ORR) y la *duración de la respuesta* (DoR). La ORR es una medida de proporción e incluye la cantidad de pacientes que reducen su carga de la enfermedad, medida por un criterio previamente definido y estandarizado como el RECIST, con respecto al total de pacientes incluidos en el brazo de estudio. Derivada de ella es la DoR, entendida como la duración de la ORR en el tiempo, desde que se detecta hasta que desaparece. Aunque ni la ORR ni la DoR deben considerarse una ClinRO propiamente dicha, son la consecuencia palmaria de la aplicación del juicio clínico del experto tras la aplicación de una escala protocolizada.

Algunas *ClinRO rating* pueden transformarse en escalas dicotomizadas, cuando las puntuaciones extremas se consideran rangos dicotómicos sobre la base de criterios de respuesta definidos. De hecho, las escalas RECIST se convierten en dicotómicas cuando se interpretan, porque identifica pacientes en progresión frente a pacientes sin progresión (estabilización, respuesta parcial o respuesta completa).

En determinadas ocasiones, las ClinRo Rating pueden ser consideradas *endpoint* primarios, como ocurre en algunos fases I en Oncología, como pruebas de concepto y eficacia directa del fármaco.

Figura 2.

CONCEPTOS DE COMPLIANCE Y PERSISTENCE EN EL CUMPLIMIENTO TERAPÉUTICO



¹⁰ <http://www.irrecist.com/recist/> Último acceso 28.07.2018. Último acceso: 27.07.2018.

■ 2.3.1.3. *ClinRO Global Assessment*

Las *ClinRo Global Assessment*, *CGA*, incluyen impresiones globales clínicas y de cambio evolutivo y están basadas en el juicio global del clínico sobre el estado de salud del paciente o un aspecto relevante del mismo para el que las variables examinadas no están bien definidas. Las *CGA* evalúan conceptos no explícitamente definidos, dan información suplementaria. Por ello, y aunque son de utilidad en la práctica clínica real, nunca deben ser considerados *endpoint* de un ensayo clínico.

Un ejemplo de este tipo de *ClinRO* son la escala ECOG-PSy la *Comprehensive Geriatric Assessment*¹¹.

■ 2.3.2. Otros *ClinRO* no agrupables en las categorías anteriores

El cumplimiento terapéutico afecta al beneficio terapéutico (Simpson, Eurich y Majumdar, 2006). Se define como el grado en que la conducta del paciente coincide con las recomendaciones médicas. El término ha sido sustituido recientemente por el de adherencia, para describir el comportamiento del paciente que toma los fármacos correctamente, a las dosis adecuadas y con la frecuencia precisa.

Para ISPOR (Lakdawalla *et al.*, 2018) cumplimiento (*compliance*) y adherencia (*adherence*) son sinónimos y significan “la medida en que un paciente actúa de acuerdo con las recomendaciones dadas por su equipo médico con respecto al cumplimiento de la dosis, intervalos y horarios del fármaco prescrito”. Define también el concepto de persistencia como la duración temporal del cumplimiento o adherencia (Figura 2).

Según ISPOR, tanto el cumplimiento como la persistencia deben ser variables de evaluación en los ECA, sobre todo cuando la medicación de estudio y/o de control es oral o autoadministrada sin la supervisión del equipo investigador. El cumplimiento se analiza mediante la razón entre la dosis tomada y la prescrita. Cuando se analiza en estudios retrospectivos se usa el concepto de tasa de posesión de la medicación (*medication possession ratio*) y es el “número de dosis dispensadas durante el periodo de dispensación”, asumiendo que el cumplimiento se produce automáticamente con la dispensación. La persistencia se define desde la fecha de inicio de tratamiento y la fecha final de la observación. Puede definirse como una variable continua o dicotómica.

Los métodos de medición de la adherencia y la persistencia en el tratamiento pueden ser los COA, según se utilice un método de evaluación u otros. Los métodos directos de evaluación son la observación médica de la toma o el

¹¹ <http://siog.org/content/comprehensive-geriatric-assessment-cga-older-patient-cancer>

uso de biomarcadores. La primera puede asimilarse como ClinRO de tipo reading porque hay un clínico o farmacéutico que observa, analiza, evalúa y obtiene un juicio clínico con una variable dicotómica (si la tomó/no la tomó). Los métodos indirectos son los COA de tipo ClinRO o PRO. Son los más sencillos y económicos, pero pecan de poca objetividad. Son ClinRO la relación entre citas y cumplimiento-persistencia (si no asiste es que tampoco cumple), la valoración de la eficacia terapéutica (si hay respuesta es porque hay cumplimiento-persistencia), el uso de marcadores fisiológicos (efectos secundarios), el recuento directo de los comprimidos que devuelve o la utilización de cuestionarios. Entre los métodos indirectos de tipo PRO se cuentan el test de Haynes & Sackett, la comprobación fingida o los diarios de paciente. En todo caso, ningún método ha demostrado ser claramente eficaz. Los ECA suelen usar el método indirecto de recuento de comprimidos.

■ 2.4. PROCESO DE CONSTRUCCIÓN DE UN COA¹²

La IRS ofrece numerosas fuentes de información de los beneficios y riesgos potenciales de las intervenciones sanitarias. Pero es un elemento fundamental garantizar la calidad de cualquier medición que pueda ser base de toma de decisiones. La calidad de cualquier medición tiene dos dominios, la validez, o grado en el que realmente se mide el fenómeno de interés, y la fiabilidad o reproducibilidad en múltiples mediciones.

Los resultados medidos en los ECA pueden ser variables evento-tiempo, biomarcadores y COA. La primera, entendida como la OS es una medida directa y objetiva cargada de validez y reproducibilidad. Los segundos son también objetivos, pero son mediciones indirectas o subrogadas, cuya reproducibilidad puede estar asegurada pero no así su validez. Finalmente los COA, pueden ser medidas directas o indirectas y generalmente subjetivas porque su información puede verse modificada por juicios, volición o motivación del informador.

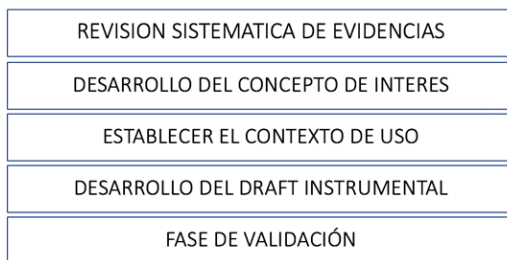
La FDA están llevando a cabo intensos y denodados esfuerzos para categorizar y estandarizar estas medidas de resultados. Fruto de ello ha sido la publicación de varios documentos tendentes a disminuir esa subjetividad, pero se han volcado en las medidas PRO¹³. Mientras, las ClinRO adolecen de guías que clarifiquen su desarrollo, implementación y regulación. Afortunadamente, el proceso de desarrollo de un ClinRO es prácticamente similar al que se ha diseñado para el desarrollo de los PRO, por lo que el desarrollo de constructo de los primeros favorecerá indudablemente a la eclosión de los segundos (Figura 3).

¹² Draft Guidance for Industry—Qualification Process for Drug Development Tools. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM230597.pdf>

¹³ Ibid.

Figura 3.

FASES DE DESARROLLO DE UN COA



El proceso de construcción de cualquier COA comienza con la búsqueda de evidencias desde un enfoque sistemático para determinar su marco conceptual y aproximarse a la validez de uso, estandarte de trabajo. Es de extrema importancia que un COA que vaya a ser usado como *end point* describa directamente o pueda demostrar que tiene relación con el beneficio del tratamiento. Las fuentes de evidencia podrán proceder de revisiones bibliográficas, consultas con expertos clínicos y de entrevistas con pacientes.

El segundo paso es el desarrollo del concepto de interés (COI). La mayoría de las veces, sobre todo cuando queremos medir funcionalidad, abarcar todas las fases o tareas de una actividad concreta y medible dificulta el desarrollo de una herramienta. Dicha actividad puede deconstruirse en varias tareas e identificar una que sea corolario fácilmente medible de la actividad global. Esa tarea es el concepto de interés, sobre el que se trabajará. La experiencia del dolor es paradigmática para entender lo que es el concepto de interés. Puede diferir en numerosas facetas, intensidad, duración, frecuencia, calidad. El concepto de interés es cuando elegimos una de ellas para medirlo y desarrollamos una escala para ello.

El tercer paso es la definición del contexto de uso¹⁴: enfermedad de estudio, subpoblación de pacientes en las que puede ser aplicado, factores culturales y lingüísticos que pueden afectar a su universalización, tipo de *end point*, entorno de medición y método de administración de la medición.

El cuarto paso es ya la elaboración y desarrollo del *draft* instrumental. Es extremadamente importante definir el contexto de uso, las definiciones y descripciones de los síntomas, a menudo discrepantes entre práctica clínica y contexto

¹⁴ <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm284077.htm>

ECA, y concretar las respuestas de cada ítem individual y sus escalas de relevancia y significación. Como el proceso es complicado e iterativo, con múltiples ciclos de revisión, es muy recomendable la formación de un grupo de expertos clínicos que reconozcan el concepto de interés medible, la definición del síntoma a medir, los dominios posibles, las escalas de valoración, el tipo de variable con el que se va a expresar la medición y el diseño del material de entrenamiento y capacitación para que la evaluación de resultados sea confiable.

Una vez diseñado, el COA pasa a la fase de validación, que depende de su fiabilidad y de su validez (Sackett, 1992). Mientras que esta expresa el grado en el que se mide el fenómeno de interés. La primera indica hasta que punto se obtienen los mismos resultados al efectuar la medición en más de una ocasión, bajo condiciones similares. Una medida precisa no tiene por qué ser válida. Así que las dos características tienen que ser analizadas. En los estudios que validan una medida se usa la comparación con un estándar ya establecido. Cuando el problema es la fiabilidad se valoran otros aspectos. La repetibilidad indica hasta que punto un instrumento proporciona resultados similares cuando se aplica a una persona de forma repetida y en idénticas condiciones. La concordancia intraobservador analiza el grado de consistencia al efectuar la medición por el mismo observador de forma repetida sobre el mismo objeto de medida y en las mismas condiciones; mientras, la consistencia interobservador mide la consistencia de la medición del instrumento cuando lo usan dos observadores distintos sobre el mismo objeto en las mismas condiciones. La fiabilidad y la validez son conceptos claves para un Clin RO exitosos.

■ 2.5. ATRIBUTOS DE UN COA¹⁵

Cualquier COA tiene tres atributos que deben definirse. El primero atributo es definir si el COA depende de la participación activa del paciente o del juicio del clínico. En el primero de los casos estamos ante un PRO y en el segundo ante un ClinRO. Es importante recordar que los ClinRO proporcionan la perspectiva del clínico sobre los PRO y, por tanto, a la hora de diseñar un ClinRO esta perspectiva no debe olvidarse. ¿Es necesario un ClinRO o sirve con un PRO? Si el paciente puede informar sobre su experiencia, sirve con un PRO, pero si el clínico puede “observar” el fenómeno, al PRO habrá que añadirle un ClinRO que complemente la experiencia.

El segundo atributo que debe aplicarse es definir quién es el evaluador. Hay tres categorías de evaluador, el paciente, el clínico o un observador no clínico. Debido a la diferente perspectiva, capacitación, habilidades y antecedentes, sus juicios pueden diferir y tener impacto en la calificación final.

Finalmente el tercer atributo es establecer si el COA puede medir directa o indirectamente la función o los síntomas. Los COA de medida directa proporcionan

¹⁵ Ver Walton *et al.* (2015).

información sobre un aspecto significativo de forma intrínsecamente interpretable. Es directamente información Mientras los COA indirectos requieren una elaboración posterior porque aportan solo datos que deben transmutarse en información, ya que la evaluación no es inherentemente clara. Son ejemplos de estos atributos, las mediciones de agudeza visual, en la que la información vertida es inherentemente clara y mide un concepto de interés muy cercano al de la vida diaria, la agudeza visual. Sin embargo un COA indirecto, por ejemplo, la fuerza isométrica del cuádriceps, no da información inherentemente clara porque no es un concepto de interés cercano a lo que el paciente hace en su vida diaria.

■ 2.6. CONCLUSIONES E IDEAS FINALES

- Valor es el marco en el que se define la mejora de la atención sanitaria y el impulso innovador del sistema. Es la relación entre los resultados de la intervención y el coste o inversión que induce su implementación.
- Los resultados se miden en términos de beneficio incremental, entendido como cualquier efecto favorable en uno o varios aspectos significativos para el paciente, que pueden ser vivir más, vivir con menos síntomas o retener una funcionalidad adecuada para sus necesidades diarias.
- No todo el beneficio incremental es “significativo para el paciente”. Por ello, es importante que el paciente se involucre en la definición e información de beneficios, sin que esto suponga una interpretación literal, rigurosa y excluyente de otros informantes.
- Aunque el beneficio es jerarquizable, con la supervivencia como objetivo prioritario, esta a veces puede quedar supeditada a otras ganancias más modestas objetivamente, pero más deseables subjetivamente. La retención de funcionalidad y la mejora de síntomas pueden ser más deseables en algunos contextos.
- Cualquier resultado en salud y, por tanto, su plasmación en beneficios, es dependiente, inherente y específico de la condición médica que mide. Ningún resultado en particular captura todo el beneficio global de la acción terapéutica.
- La Investigación en Resultados en Salud es la fuente principal de la Medicina Basada en la Evidencia porque informa de beneficios en términos de eficacia, cuando su información procede del Ensayo Clínico (ECA), o en términos de efectividad cuando procede del RWD o DVR.
- En el marco del ECA se entiende por “evaluaciones de resultado” a cualquier medición de los efectos inducidos por un fármaco en la modificación de un evento considerado y previamente definido, que permita rechazar la “hipótesis nula” y evitar el error tipo I.

- Las evaluaciones de resultados que se usan para interpretar beneficio, se denominan *endpoint*. Pueden clasificarse en primarios o secundarios. Los primarios contestan directamente a la pregunta sobre beneficio; los secundarios complementan y matizan esa respuesta.
- Las evaluaciones de resultados que definen beneficio pueden ser de evento-tiempo referidas a cuánto sobreviven los pacientes; evaluaciones clínicas que contestan a preguntas sobre funcionalidad y síntomas; o biomarcadores.
- Las evaluaciones clínicas se diferencian de las variables evento-tiempo y de los biomarcadores en que se ven afectadas por la volición y/o motivación del paciente o por el juicio clínico del evaluador, sea este un clínico o un observador no cualificado. Por tanto, son susceptibles de variación no relacionada con el efecto beneficioso o perjudicial de la intervención. Se denominan COA cuando se usan en el contexto de los ECA y vienen siendo impulsados por el ISPOR y la FDA.
- Se describen cuatro tipos de COA en función de quién sea el informante: PRO cuando es el paciente, ObsRO cuando es un observador no cualificado pero cercano al paciente, ClinRO cuando el informante es un clínico experto y PerFO cuando el informante no es una persona sino una escala no interpretable por el juicio clínico. Proporcionan evidencia complementaria de los beneficios del tratamiento, y algunos, en determinadas ocasiones pueden funcionar como *endpoint* primarios.
- El *resultado informado por un clínico* o ClinRO (*Clinician-Reported Outcome*) se basa en la emisión de un juicio clínico por un profesional sanitario experto y entrenado, después de la observación del paciente o la aplicación de un instrumento de medida o una escala sobre signos observables, pruebas diagnósticas, comportamientos u otras manifestaciones físicas que se cree que están relacionadas con la enfermedad o la condición medida. Tienen una larga historia en Medicina. Han sido asumidos como válidos porque están cargados con el “juicio clínico”, pero carecen, en muchos casos, de evaluaciones rigurosas de su valor y la validez de su aportación.
- ISPOR¹⁶ diferencia tres ClinRO, las *Reading* (lectura de signos o interpretación de pruebas), las *Rating* (calificaciones y juicios impulsados por escalas o protocolos) y las *Clinical Global Assessment* (CGA). En cualquiera de ellas la participación del paciente puede ser pasiva o activa.
- El cumplimiento terapéutico, en cuanto a adherencia y *compliance*, afecta al beneficio terapéutico y, por tanto, deben incorporarse como COA, bien ClinRO o PRO.

¹⁶ <https://www.fda.gov/downloads/drugsGuidanceComplianceRegulatoryInformation/Guidance/UCM071590.pdf>

- La construcción de un COA sigue las mismas directrices que cualquier proceso de construcción de un indicador que asegure su validez de constructo, su fiabilidad interna y externa, así como una adecuada definición de sus atributos. El desarrollo de un ClinRO debe iniciarse por la pregunta ¿es necesario?. Solo lo será si la experiencia del paciente que comunica un PRO puede ser observada e interpretada. Los ClinRO deben ser interpretados como complementarios de los PRO. Solo los sustituyen cuando lo que siente los pacientes o su funcionalidad es una experiencia inefable.

REFERENCIAS

- BADIA, X. y LIZÁN, L. (2002). Reflexiones sobre la investigación de resultados en salud. *Aten Primaria*, 30, pp. 388-91.
- BADIA, X. y DEL LLANO, J. (2000). La investigación de resultados en salud. *Med Clin (Barc)*, 114, pp. 1-7.
- BAKHAI, A., CHHABRA, A. y WANG, D. (2006). Endpoints. En: D. WANG y A. BAKHAI (Eds), *Clinical Trials. A practical guide to design, analysis and reporting* (pp. 37-46). Chicago: Remedica.
- BENJAMIN, K., VERNON, M. K. y PATRICK, D. L. (2017). Patient-Reported Outcome and Observer-Reported Outcome Assessment in Rare Disease Clinical Trials: An ISPOR COA Emerging Good Practices Task Force Report. *Value in Health*, 20(7), pp. 838-855.
- COYNE, K. S. y WYRWICH, K. W. (2015). ISPOR Task Force For Clinical Outcomes Assessment: Clinical Outcome Assessments: Conceptual Foundation—Report of The ISPOR Clinical Outcomes Assessment – Emerging Good Practices For Outcomes Research Task Force. *Value in Health*, 18(6), pp. 739-40.
- LAKDAWALLA, D. N., DOSHI, J. A., GARRISON, L. P. Jr. *et al.* (2018). Defining Elements of Value in Health Care-A Health Economics Approach: An ISPOR Special Task Force Report [3]. *Value Health*, (2), pp. 131-139.
- LAPORTE, J. R. (1993). Principios básicos de investigación clínica. *ZénecaFarma/Ergon*. Madrid.
- MOLINA ARIAS, M. y OCHOA SANGRADOR, C. (2015). Ensayo Clínico (IV). Resultados. Variables y medidas de efecto e impacto. *Evid Pediatr.*, 11, p. 33.
- PATRICK, D. L. y BURKE, L. (2013). Focusing on the patient in drug development and research. *ISPOR Connections*, 19(6), pp. 5-8.
- PORTER, M. E. (2008). *Defining and introducing value in health care. In: Evidence-based medicine and the changing nature of health care: 2007. IOM annual meeting summary* (pp. 161-172). Washington, DC: Institute of Medicine.
- . (2010). What is value in health care? *N Engl J Med.*, 363(26), pp. 2477-2481.
- PORTER, M. E. y TEISBERG, E. O. (2006). *Redefining health care: creating value-based competition on results*. Boston: Harvard Business School Press.
- POWERS, J. H. 3rd, PATRICK, D. L., WALTON, M. K. *et al.* (2017). Clinician-Reported Outcome Assessments of Treatment Benefit: Report of the ISPOR Clinical Outcome Assessment Emerging Good Practices Task Force. *Value Health*, 20(1), pp. 2-14.
- SACKETT, D. L. (1992). A primer on the precision and accuracy of the clinical examination. *JAMA*, 267, pp. 2638-2644.
- SIMPSON, S. H., EURICH, D. T., MAJUMDAR, S. R. *et al.* (2006). A meta-analysis of the association between adherence to drug therapy and mortality. *BMJ*, 333(7557), p.15.

- TUNIS, S. R., CLARKE, M., GORST, S. L. *et al.* (2016). Improving the relevance and consistency of outcomes in comparative effectiveness research. *J Comp Eff Res*, 5(2), pp.193-205.
- US FDA. (1992). New drugs, antibiotic and biological drug product regulations; accelerated approval. Proposed rule. 57 *Federal Register*, pp. 13234-13242.
- WALTON, M. K., POWERS, J. H. 3rd, HOBART, J. *et al.* (2015). Clinical Outcome Assessments: Conceptual Foundation-Report of the ISPOR Clinical Outcomes Assessment - Emerging Good Practices for Outcomes Research Task Force. International Society for Pharmacoeconomics and Outcomes Research Task Force for Clinical Outcomes Assessment. *Value in Health*, 18(6), pp.741-52.



3

RESULTADOS PERCIBIDOS POR LOS PACIENTES

ISABEL CASTREJÓN, MARIA JOÃO FORJAZ,
CARMEN RODRÍGUEZ-BLÁZQUEZ, MIGUEL ÁNGEL RUIZ

■ 3.1. INTRODUCCIÓN

Partiendo de un modelo biopsicosocial y centrado en el paciente, los resultados percibidos por los pacientes (*patient-reported outcomes, PRO*) recogen información directamente del paciente, sin interpretación del personal sanitario u otra persona (Doward y McKenna, 2004). Los PRO permiten la medición de constructos que no son directamente observables, como la fatiga o el dolor, así como conceptos multidimensionales como la calidad de vida o la discapacidad, que requieren la aplicación de varias preguntas o *ítems*, normalmente agrupadas en un cuestionario. Para que las medidas PRO puedan ser utilizadas como resultado de salud, es necesario que las respuestas a los *ítems* se traduzcan en valor numérico. Para eso, los cuestionarios deben tener buenas propiedades psicométricas, incluyendo: fiabilidad, validez, sensibilidad al cambio, y aceptabilidad, entre otros. Este capítulo se centrará en los siguientes PRO: calidad de vida relacionada con la salud (CVRS), satisfacción con el tratamiento, preferencias, comorbilidad, y discapacidad.

La calidad de vida relacionada con la salud (CVRS), como una PRO, se distingue de otros indicadores o índices estructurales objetivos, como los que valoran la calidad de vida de una ciudad o región, al centrarse en la percepción del individuo sobre aspectos importantes de su vida, especialmente la salud.

Otra PRO utilizada con frecuencia es la satisfacción con el tratamiento, que cuantifica el nivel en que se alcanzan las expectativas del paciente ante los distintos aspectos del tratamiento, valorando su propia experiencia.

La medida de las preferencias de los pacientes cuantifica la priorización que realizan estos de distintas alternativas terapéuticas disponibles para su tratamiento. Existen también medidas de las preferencias sociales por los diferentes estados de salud posibles en los que se puede encontrar un paciente, ordenándolos según la aversión por los estados más graves y la preferencia por los estados más benignos, a los que se denomina medidas de utilidad.

La comorbilidad se define como la coexistencia de varias enfermedades crónicas en un mismo paciente. Sin embargo, desde el punto de vista del paciente otros aspectos necesitan ser incorporados en la valoración de comorbilidad. Principalmente, el impacto en su calidad de vida y su capacidad funcional. Los pacientes tienen un conocimiento único sobre su situación clínica y cómo esta afecta a otras

esferas de su vida y por tanto pueden valorar de forma adecuada su grado de comorbilidad usando cuestionarios disponibles y validados para tal fin.

La discapacidad es una importante consecuencia de los problemas de salud. A pesar de que los avances sanitarios han reducido la mortalidad, el incremento asociado de la longevidad y de la supervivencia y el envejecimiento demográfico han provocado un aumento de las enfermedades crónicas (Vogeli *et al.*, 2007) que, con frecuencia, cursan con limitaciones y un mayor número de años vividos con discapacidad. La evaluación de la discapacidad se incluye en la actualidad en multitud de ensayos clínicos y en diversas iniciativas de uso rutinario de los PRO en la práctica clínica como PROMIS (*Patient-Reported Outcomes Measurement Information System*), un sistema de medidas desarrollado por el Departamento de Salud norteamericano centrado en el paciente para evaluar la salud física, psíquica y social en adultos y niños.

Además de las medidas PRO o PROM (*patient-reported outcomes measures*), que miden la situación clínica del paciente o el resultado de una intervención médica y engloban aspectos como la CVRS, la comorbilidad, la discapacidad, interferencia en las actividades de la vida cotidiana y el bienestar, hay otros términos relacionados. Entre ellos se encuentran las medidas de la experiencia informadas por el paciente o las PREM (*Patient reported experience measures*) que miden la valoración que hace el paciente del resultado de usar un tratamiento o servicio. Engloba las medidas de la satisfacción del paciente con la medicación, el tratamiento, o el servicio. Ambos tipos se construyen siguiendo los supuestos de la psicometría clásica o la teoría de respuesta al ítem. Por último, se encuentran las medidas de preferencia basadas en el paciente (*patient based preferences*). Estas medidas cuantifican la ordenación priorizada de un conjunto de estados de salud o tratamientos posibles, normalmente a partir de los resultados de un experimento de elección con métodos de escalamiento y medidas ipsativas. La finalidad de este capítulo es analizar las principales PRO y PREM y describir las escalas de evaluación más utilizadas en el entorno sociosanitario.

■ 3.2. CALIDAD DE VIDA RELACIONADA CON LA SALUD

■ 3.2.1. Definición y modelos

La introducción del concepto de calidad de vida en la medicina vino de la oncología, cuando los pacientes y médicos se dieron cuenta de que, además de aumentar la esperanza (o cantidad) de vida, los tratamientos debían enfocarse también a conservar cierto bienestar básico del individuo respecto a su salud. Este concepto, aplicado al contexto médico, se refiere como calidad de vida relacionada con la salud (CVRS). Hay muchas definiciones de CVRS, pero todas coinciden en que es

un concepto multidimensional (incluyendo por lo menos aspectos de salud física y mental y, frecuentemente, dimensiones sobre síntomas, función social y capacidad funcional), centrado en el paciente y, por lo tanto, subjetivo. El Centro de Control y Prevención de Enfermedades (CDC) americano la define como la percepción de un individuo o grupo sobre la salud física y mental a lo largo del tiempo (Zack y Centers for Disease Control and Prevention [CDC], 2013) mientras que la agencia norteamericana Food and Drugs Administration (FDA) concreta la CVRS como un concepto multidimensional que representa la percepción general del paciente sobre los efectos de la salud y la enfermedad en los aspectos físicos, psicológicos y sociales de la vida (U.S. Department of Health and Human Services FDA, 2006). Por su parte, la OMS define la CVRS como la percepción de un individuo sobre su salud y dominios del bienestar relacionados con la salud (The World Health Organization Quality of Life assessment [WHOQOL], 1995).

Uno de los modelos de CVRS más conocidos es el de Wilson y Cleary (1995), modificado por Ferrans *et al.* (2005). Dicho modelo integra los paradigmas biomédicos y de las ciencias sociales, al definir cinco dominios: biológico, síntomas, funcionamiento, percepción de la salud general y calidad de vida global. Además, los factores ambientales e individuales se asocian con resultados de salud, afectando así la CVRS total. En la revisión del modelo, Ferrans *et al.* (2005) aportan definiciones de las características individuales y ambientales, eliminan del modelo los factores no médicos, y explican que, aunque se presenten las principales relaciones causales entre los componentes del modelo, se admiten relaciones recíprocas.

■ 3.2.2. Uso en investigación y práctica clínica en el SNS

La medición de la CVRS tiene muchas ventajas y aplicaciones tanto en la investigación como en la práctica clínica. Algunas de sus aplicaciones son las siguientes: permite la identificación de problemas según una perspectiva centrada en el paciente, se usa en la investigación de intervenciones sanitarias, la evaluación de políticas de salud y de prácticas clínicas, como indicador de resultados en la investigación y facilita la comparación de enfermedades distintas. Además, también se usa en auditorías médicas, en los análisis coste-utilidad de los estudios de fármaco-economía y permite una mejor asignación de recursos y una mejor toma de decisiones. La FDA, agencia federal americana para el control de alimentos y medicamentos, indica la importancia de usar las PRO, incluyendo síntomas, función y calidad de vida, en ensayos clínicos y evaluación de tecnologías sanitarias (US Food and Drug Administration, 2016).

El *National Health Service-Patient Reported Outcome Measures (NHS-PROM)* es un programa del sistema nacional de salud del Reino Unido que ofrece un buen ejemplo del uso rutinario de medidas de CVRS en la práctica clínica, para comparación de resultados entre hospitales. Se aplica de forma habitual una medida genérica de CVRS, el EQ-5D, así como otros cuestionarios específicos, a pacientes que se someten a las siguientes cirugías electivas: prótesis de cadera y de rodilla,

hernia inguinal y venas varicosas. Los cuestionarios se administran antes de la intervención y después de un tiempo de seguimiento (3 o 6 meses, según la intervención). Además, el EQ-5D también se ha incluido en una encuesta a los usuarios de servicios de atención primaria, con cerca de novecientos mil pacientes al año.

En un esfuerzo por estandarizar y mejorar el uso de las PRO, el Departamento de salud y servicios humanos norteamericano ha desarrollado PROMIS, un sistema que incluye las PRO validadas para evaluar la salud física, psíquica y social (Fries, Bruce y Cella, 2005). No incluye medidas directas de CVRS, sino que parte del concepto de salud de la OMS, en tres áreas principales cuyos dominios ha identificado: mental (dominios de ansiedad y depresión), física (fatiga, dolor, sueño, función física) y social (rol social). En vez de crear *items* nuevos, se partió de cuestionarios ya existentes para medir síntomas y dominios de la CVRS. Así, se crearon bancos de *items*, que se adaptaron y estandarizaron para poder administrarlos de forma informatizada y dinámica (*computerized adaptive test*, CAT) o como versiones cortas.

En España, la *Encuesta Nacional de Salud* de 2011/2012 incluyó una medida de CVRS, el EQ-5D-5L para adultos y el KIDSCREEN-10 Index para niños y adolescentes de 8 a 14 años (Instituto Nacional de Estadística, 2013). La aplicación de estos instrumentos a nivel nacional, en una muestra representativa, permitió obtener datos de la CVRS de la población, lo que ofreció información importante para la planificación de políticas de salud pública. En la población adulta, se observaron diferencias por edad, clase social y actividad física, pero no por región, una vez se controló por estos factores (Martín-Fernández *et al.*, 2018).

■ 3.2.3. Evaluación: escalas genéricas y específicas

En la Tabla 1 se presentan algunos de los principales instrumentos de CVRS disponibles en España, y sus respectivas referencias. Mientras que los instrumentos genéricos permiten su aplicación en la población general y la comparación entre distintos problemas de salud, los instrumentos específicos captan aspectos particulares de la enfermedad en cuestión. Para la selección de un instrumento, hay que tener en cuenta varios aspectos, entre los cuales hay que destacar la complejidad del cuestionario, el objetivo de su aplicación, las dimensiones que se quieren medir, y sus propiedades psicométricas. Hay varios recursos que ayudan a la elección de un cuestionario de CVRS, como por ejemplo las revisiones sistemáticas de los cuestionarios de CVRS para un determinado grupo poblacional o problema de salud. Por otra parte, BiblioPro, una biblioteca virtual, ofrece información sobre varios cuestionarios PRO, permite su acceso o datos sobre cómo obtenerlos¹. Seguidamente se describen tres cuestionarios genéricos de CVRS para población general (EQ-5D-5L y SF-36), uno genérico para población pediátrica (KIDSCREEN), y dos sis-

¹ <https://www.bibliopro.org/>

temas modulares de cuestionarios, uno específico para oncología (EORTC) y otro para enfermedades crónicas. Todas estas medidas han sido sometidas a rigurosas validaciones psicométricas, con muy buenos resultados.

EQ-5D-5L (EuroQol Group, 1990). Desarrollado por el grupo internacional EuroQoL, es una medida de estado de salud y CVRS, que consta de dos partes. En la primera, se valoran cinco dimensiones o *ítems* (movilidad, cuidado personal, actividades cotidianas, dolor/malestar, ansiedad/depresión), en una escala de 1-5 puntos, de ausencia de problemas (Doward y McKenna, 2004) a muchos problemas (WHOQOL, 1995). La puntuación en las cinco dimensiones define un perfil o estado de salud (por ejemplo, 55555, describe la presencia de problemas extremos en todas las dimensiones), al que corresponde un valor índice entre 1 y 0, aunque pueda tomar valores negativos que indican estados de salud valorados como peores que la muerte. La segunda parte es una escala visual analógica, donde se pide que la persona valore su estado de salud hoy, en una regla vertical de 20 cm graduada desde 0 (peor salud imaginable) a 100 (mejor salud imaginable). El EQ-5D-3L corresponde a una versión anterior, en que cada dimensión se valoraba en tres niveles de respuesta. Tanto el EQ-5D-3L como el EQ-5D-5L disponen de valores de referencia para la población española, y los resultados de una versión se pueden convertir en la otra versión. Se puede acceder al EQ-5D y su manual a través de la página <https://euroqol.org/>.

MOS SF-36 (McHorney *et al.*, 1992). Es un instrumento de desarrollo norteamericano, pero ampliamente utilizado en todo el mundo. Está formado por 36 *ítems*, valorados en distintas escalas de respuesta, que se agrupan en ocho dimensiones (función física, rol físico, dolor, salud general, vitalidad, función social, rol emocional y salud mental) y dos componentes: físico y mental. Tanto las dimensiones como los componentes se expresan en puntuaciones de 0-100, en las que mayores valores se corresponden a una mejor CVRS. La población de referencia tiene una media de 50 y desviación estándar de 10. Este cuestionario consta de una versión actualizada (SF-36v2), y de varias versiones cortas: SF-8 y SF-12, con 8 y 12 *ítems*, respectivamente, que permiten el cálculo de los componentes, pero no de las dimensiones. Aunque existan varias versiones abiertas disponibles, el cálculo de las dimensiones y componentes está informatizado y requiere el pago de una licencia a Quality Metric.

KIDSCREEN (Aymerich *et al.*, 2005). Fue desarrollado simultáneamente en varios países para medir la CVRS en niños y adolescentes de 8 a 18 años. Puede ser contestado por padres y niños/adolescentes y está formado por 52 *ítems*, agrupados en diez dimensiones: bienestar físico, bienestar psicológico, estado de ánimo, autopercepción, autonomía, relación con los padres y vida familiar, amigos y vida social, entorno escolar y adaptación social (*bullying*). También están disponibles dos versiones cortas, una de 27 y otra de 10 *ítems*. El KIDSCREEN-10 solo permite el cálculo de un índice general de CVRS. Se puede acceder a las diversas versiones del cuestionario en español y al manual a través de BiblioPro.

Tabla 1.

CUESTIONARIOS DE CALIDAD DE VIDA RELACIONADA CON LA SALUD

Instrumento y referencia	Número de ítems y rango de puntuación	Contenido	Población diana
EQ-5D-5L Hernández <i>et al.</i> (2018) <i>QualLife Res</i>	5 + escala visual analógica. Dimensiones: 15; escala visual analógica: 0-100	Movilidad, cuidado personal, actividades cotidianas, dolor/ malestar, ansiedad/ depresión	Población general (hay una versión para niños, EQ-5D-5LY)
Short Form (36) Health Survey (SF-36) Alonso <i>et al.</i> (1995) <i>Med Clin</i>	36 ítems. 0-100	Componentes físico y mental. Dimensiones: función física, rol físico, dolor, salud general, vitalidad, función social, rol emocional y salud mental	Población general
KIDSCREEN Aymerich <i>et al.</i> (2005) <i>Gac Sanit</i>	52 ítems. 0-100	Bienestar físico, bienestar psicológico, estado de ánimo, autopercepción, autonomía, relación con padres y vida familiar, amigos y vida social, entorno escolar y adaptación social (<i>bullying</i>)	Niños y adolescentes
European Organisation for Research and Treatment of Cancer Quality of life Core 30 (EORTC QLQ-C30) Arrarás <i>et al.</i> (2002) <i>Psychooncology</i>	30 ítems. 0-100	5 escalas funcionales, 3 escalas de síntomas; 1 de estado de salud global/calidad de vida	Cáncer
Functional Assessment of Cancer Therapy-General (FACT-G) Cella <i>et al.</i> (1998) <i>Med Care</i>	27 ítems. 0-100	Bienestar físico, familiar/social, emocional, funcional	Cáncer

Fuente: Elaboración propia.

EORTC QLQ-C30 (Arrarás *et al.*, 2002). La Organización Europea para Investigación y Tratamiento de Cáncer (EORTC) desarrolló varios cuestionarios para la población oncológica. El EORTC QLQ-C30 es el cuestionario central, que se aplica juntamente con otros módulos específicos según el tipo de cáncer, grupo de pacientes (personas mayores), síntomas (fatiga), tratamiento (reconstrucción mamaria, cuidados paliativos) u otros aspectos (información recibida). El EORTC QLQ-C30

tiene 5 escalas funcionales, 3 escalas de síntomas y una de estado de salud global/calidad de vida. Los cuestionarios EORTC se aplican ampliamente en ensayos clínicos como medidas de resultado.

FACT-G (Cella *et al.*, 1998). Al igual que las escalas del grupo EORTC, Cella *et al.*, desarrollaron en Estados Unidos un sistema de escalas para valorar la calidad de vida de personas con enfermedades crónicas, incluyendo cáncer. Este sistema se compone de un cuestionario núcleo, el *Functional Assessment of Cancer Therapy-General (FACT-G)*, y de varias decenas de módulos específicos para tratamientos, tipos de cáncer, síntomas y enfermedades crónicas. El *FACT-G* consta de 27 *items*, valorados en una escala de respuesta de 0 a 4. Los cuestionarios están traducidos a una gran variedad de idiomas, a través de un protocolo de traducción muy riguroso.

■ 3.3. SATISFACCIÓN CON EL TRATAMIENTO

La necesidad de mejorar la atención sanitaria como servicio, la libertad de elección de tratamiento en algunos países, la puesta en marcha de iniciativas de copago de terapias de alto coste económico y el empoderamiento del paciente para su participación en la toma de decisiones clínicas, han hecho necesario evaluar las experiencias de los pacientes con el servicio prestado. Para ello, se han desarrollado un nuevo conjunto de medidas denominadas “Medidas de la experiencia del paciente” (PREMs) y en particular las medidas de la satisfacción del paciente.

Dependiendo de lo que se desee evaluar (un centro, una unidad, un programa, una intervención, un fármaco) se pueden considerar distintos niveles de evaluación y con ello distintos tipos de satisfacción: con el sistema sanitario, con el servicio, con el tratamiento, la medicación. Nosotros nos vamos a centrar en la evaluación de la satisfacción del paciente con la medicación, habitualmente denominada también satisfacción con el tratamiento.

En el sentido más amplio, la satisfacción del paciente con el tratamiento se relaciona con todos los aspectos de la atención sanitaria relevantes para su salud; incluye tanto la satisfacción con la atención médica como con los tratamientos específicos que se reciben. Se ha propuesto una estructura piramidal para organizar los aspectos relevantes de la satisfacción con el tratamiento. En la base se sitúa la satisfacción con la atención sanitaria, englobando todos los aspectos de la atención recibida e incluye la satisfacción del paciente con el acceso al cuidado médico, el comportamiento del médico y su competencia técnica, los servicios, los costes y el tratamiento elegido. En el nivel intermedio se sitúa el tratamiento médico en su conjunto, incluyendo la relación con el personal sanitario, los tratamientos alternativos disponibles, las orientaciones de educación para la salud y dietético-sanitarias recibidas del equipo médico, la información recibida sobre cambios de hábitos y de conducta, el estado de ánimo y la disponibilidad de terapias de apoyo. Por último,

en la punta de la pirámide se sitúa la satisfacción con la medicación recibida, que es la evaluación que hace el paciente del proceso de tomar una medicación y de los resultados asociados a ella: eficacia, comodidad, efectos no deseados, seguimiento, etc. (Shikiar y Rentz, 2004). El último escalón es el de mayor interés para evaluar el comportamiento de los fármacos, pero hay que tener en cuenta que este escalón no deja de estar influido por los otros en los que se apoya.

Otras PREM utilizadas habitualmente en la gestión sanitaria son: la satisfacción con la consulta médica (Maurice-Szamburski *et al.*, 2017), el tiempo de espera para ser atendido, la posibilidad de acceso a un recurso o traslado de servicio, la satisfacción con el cuidado médico, la implicación enfermo-cuidador en la toma de decisiones (contrato terapéutico), el conocimiento del plan de tratamiento y alternativas, la calidad de la comunicación con el paciente, el nivel de apoyo para gestionar estados de larga duración y la recomendación del servicio a la familia y amigos.

■ 3.3.1. Definición y modelos

La satisfacción con el tratamiento es “la evaluación del paciente del proceso de tomar una medicación y de los resultados asociados con la medicación” (Shikiar y Rentz, 2004). Además, “una medida de satisfacción es cualquier medida que incluya las evaluaciones de los pacientes o las respuestas afectivas a las distintas dimensiones que incluye la experiencia de los cuidados sanitarios” (Hudak y Wright, 2000). De estas definiciones se derivan varios aspectos importantes. En primer lugar, se trata de la evaluación de una experiencia, por lo que el paciente debe haber tenido que ser sometido un tiempo suficiente al tratamiento evaluado para poder valorar sus efectos y consecuencias. En segundo lugar, se trata de una valoración con un alto componente afectivo, al contrario que la CVRS, que tiene un componente únicamente cognitivo. En tercer lugar, como muchas otras PROM, es un concepto multidimensional, que agrupa muchas dimensiones (atributos), algunas de ellas asociadas al proveedor del tratamiento.

Pero ¿qué es lo que entendemos por experiencias del paciente y cuáles son las experiencias relevantes que debemos evaluar? Para explicarlo debemos revisar los modelos de satisfacción existentes. En este marco de referencia, entendemos la salud como un bien que nos es facilitado como si se tratara de un servicio recibido de un proveedor. Es en el ámbito del consumo y la prestación de servicios donde empezaron a desarrollarse las primeras mediciones de satisfacción de clientes. Básicamente, en todos los modelos utilizados actualmente, se define la satisfacción como la valoración afectiva realizada por el cliente de la discordancia existente entre el servicio obtenido (calidad recibida) y el servicio esperado (calidad ideal) (Grönroos, 1984; Parasuraman, Zeithaml y Berry, 1988). Esta valoración tendrá un signo positivo, neutro o negativo y, como consecuencia, promoverá el uso reiterado o lo evitará.

Cada instrumento de medición de la satisfacción con el tratamiento ha sido construido siguiendo las pautas de un modelo. Un buen referente es el modelo de equilibrado decisional (*Decisional Balance Model*) (Atkinson *et al.*, 2005). Este modelo es una extensión de la teoría de expectativa-valor (Ajzen y Fishbein, 1972) y propone que las conductas de un paciente, tales como las decisiones sobre adherencia al tratamiento, la recomendación a terceros o el cambio de tratamiento, vienen determinadas por el equilibrio entre los juicios de valor positivos (beneficios) y negativos (sacrificios) realizados sobre los atributos del tratamiento. Además, cada atributo tiene distinta importancia, que el paciente tendrá en cuenta a la hora de juzgar el grado de equilibrio. Básicamente, según el modelo, la satisfacción global con la medicación y las conductas resultantes (como la adherencia o la recomendación) están determinadas por el juicio del paciente comparando si los beneficios obtenidos con la eficacia medicación compensan los efectos secundarios y los inconvenientes.

Modelos como el anterior magnifican la necesidad de conocer cuáles son los atributos relevantes para el paciente y, a ser posible, tener en cuenta su importancia relativa. Una vez más se hace evidente que es imprescindible consultar con los usuarios del tratamiento para conocer las palancas relevantes de la satisfacción. Aun así, articular todo ello en un cuestionario puede ser complejo. Por ejemplo, la importancia dada por un paciente a los efectos negativos de un tratamiento será muy distinta cuando el tratamiento es eficaz en aliviar los síntomas de la patología (e.g. analgesia) a cuando un tratamiento es preventivo (e.g. control del colesterol) o cuando la enfermedad cursa sin sintomatología percibida (e.g. glaucoma).

■ 3.3.2. Dimensiones de la satisfacción

Los modelos de satisfacción se desarrollaron para diagnosticar las fortalezas y debilidades del servicio o el producto, con la intención de mejorarlo a partir de las necesidades y opiniones del destinatario del servicio (Oliver, 1997). Se ha investigado mucho sobre cuáles pueden ser las dimensiones, ejes o atributos comunes a todos los servicios (Parasuraman, Zeithaml y Berry, 1985), llegando a la conclusión de que se pueden distinguir tres dimensiones básicas de evaluación (Rust y Oliver, 1994):

- los tangibles o calidad técnica (las características del producto propiamente dicho),
- el trato o calidad funcional (la interacción con el cliente) y
- el contexto (el entorno en el que se presta el servicio).

Cada uno de estos ejes de evaluación se desdobra en una serie de atributos específicos. Dependerá de los intereses de cada instrumento particular medir o no todas las dimensiones y concretar los atributos. En el caso de la satisfacción con la

medicación es habitual medir la efectividad, los efectos indeseados, la comodidad, el trato y competencia del personal médico y el impacto sobre la vida cotidiana (Villar López *et al.*, 2009).

En las revisiones bibliográficas y metaanálisis que hemos encontrado se han recopilado las dimensiones utilizadas por distintos instrumentos genéricos y específicos (Ruiz *et al.*, 2008; Weaver *et al.*, 1997): (Doward y McKenna, 2004) satisfacción general con el tratamiento actual, (Vogeli *et al.*, 2007) eficacia del tratamiento, (Zack y CDC, 2013) información sobre la enfermedad, (U.S. Department of Health and Human Services, 2006) malestar con el tratamiento (incluyendo efectos secundarios), (WHOQOL, 1995) diseño/aspecto del producto, (Wilson y Cleary, 1995) conveniencia del tratamiento, (Ferrans *et al.*, 2005) intención de continuar con el tratamiento, (US Food and Drug Administration, 2016) facilidad/dificultad de la administración del tratamiento, (Fries, Bruce y Cella, 2005) información sobre el tratamiento, (Instituto Nacional de Estadística, 2013) flexibilidad del tratamiento (cuándo, dónde), (Martín-Fernández *et al.*, 2018) comodidad cuando no se usa (para transportar, guardar, etc.), (EuroQol Group, 1990) confianza del paciente en su habilidad para usarlo, (McHorney *et al.*, 1992) comparación con otro tratamiento, (Aymerich *et al.*, 2005) facilidad de compra del medicamento, (Arraras *et al.*, 2002) satisfacción con la planificación del tratamiento, (Cella *et al.*, 1998) rapidez del funcionamiento del tratamiento, (Shikiar y Rentz, 2004) seguridad, (Maurice-Szamburski *et al.*, 2017) duración del tratamiento, (Hudak y Wright, 2000) número de dosis, (Grönroos, 1984) grado en que el tratamiento cumple las expectativas del paciente, (Parasuraman, Zeithaml y Berry, 1988) proximidad al tratamiento ideal, (Atkinson *et al.*, 2005) recomendación a los amigos. Aunque se ha criticado incluir conductas tan distales como la recomendación a terceros o la facilidad de compra por obligar a hacer proyecciones temporales que complican la valoración del momento concreto en el que se realiza la medición.

Aparte de los atributos particulares, es muy frecuente incluir una pregunta o una dimensión que valore la satisfacción global con el tratamiento. Por un lado, por el hecho de que la valoración global combina los aspectos positivos y negativos de manera particular para el paciente que responde, por otro, porque se ha comprobado que esa valoración global es mejor predictora de la adherencia que las dimensiones particulares por separado (Atkinson *et al.*, 2005).

■ 3.3.3. Consecuencias de la satisfacción

La satisfacción con la medicación y con el tratamiento médico parece estar relacionada con la adherencia del paciente al tratamiento, es un indicador de calidad que se puede utilizar para mejorar la atención sanitaria y afecta a las preferencias de los pacientes (Barbosa *et al.*, 2012). La satisfacción con la medicación ha demostrado ser más sensible que la CVRS a los cambios de salud en los ensayos

clínicos con pacientes crónicos (Weinberger, Oddone y Henderson, 1996) y también es la mejor variable de resultado cuando se comparan tratamientos con el mismo mecanismo de actuación (Ruiz *et al.*, 2008).

Además, conocer en el grado de satisfacción con el tratamiento puede contribuir a predecir el cumplimiento terapéutico y ayudar a los clínicos a tomar decisiones, por lo que su medición es uno de los resultados en salud de obligada atención tanto en la práctica clínica diaria como en la investigación biomédica. Hasta hace relativamente poco tiempo, la mayoría de los instrumentos diseñados para medir la satisfacción del paciente con el tratamiento eran específicos de una determinada enfermedad o situación clínica, lo que obviamente no solamente limitaba su uso, sino que impedía poder comparar el nivel de satisfacción de los pacientes con el tratamiento entre distintas enfermedades o situaciones médicas.

Debemos tener en cuenta que la falta de adherencia también tiene consecuencias sociales como el absentismo, el presentismo, las complicaciones y las hospitalizaciones (Barbosa *et al.*, 2012). Por otra parte, adherencia, persistencia (continuación) y cumplimiento se suelen utilizar de manera intercambiable, lo que dificulta comparar los estudios.

■ 3.3.4. Instrumentos

Existen numerosos instrumentos específicos para distintas patologías (diabetes, artritis, glaucoma, etc.) pero no tantos instrumentos genéricos que puedan utilizarse con cualquier patología (Villar López *et al.*, 2009). Revisemos brevemente los instrumentos genéricos más utilizados.

Treatment Satisfaction Questionnaire for Medication (TSQM) (Atkinson *et al.*, 2005). Es un instrumento genérico diseñado para medir la satisfacción con el tratamiento farmacológico de cualquier enfermedad crónica, del que existen además dos versiones reducidas: la versión de 11 ítems y la de 9 ítems (Bharmal *et al.*, 2009). La versión inicial incluye cuatro dimensiones: efectos secundarios, eficacia de la medicación, comodidad de uso y satisfacción general. Se han contrastado sus buenas propiedades métricas y ha demostrado una correlación significativa ($r=0,17$) con la adherencia (Delestras *et al.*, 2013). El TSQM no incluye dimensiones como la satisfacción con el cuidado médico o el impacto de la medicación en las actividades de la vida diaria, aspectos relevantes dentro del constructo satisfacción con el tratamiento (particularmente en lo referente a su capacidad para pronosticar la adherencia al tratamiento) pues son considerados por los pacientes como atributos del tratamiento médico.

Satisfaction with Medication Questionnaire (SATMED-Q) (Rejas *et al.*, 2013). Es un instrumento genérico formado por 17 ítems que valoran seis dimensiones: efectos secundarios, eficacia de la medicación, comodidad de uso, impacto en la

vida cotidiana, atención médica y satisfacción general. Ha demostrado buenas propiedades métricas y una correlación con la adherencia ligeramente mayor que la del TSQM ($r=0,23$) y capacidad para discriminar entre grupos según su adherencia.

SAT-Q (Rebollo *et al.*, 2010). Con 34 *ítems* recoge la información de seis dimensiones, cinco sobre aspectos específicos del tratamiento farmacológico: satisfacción global con el tratamiento farmacológico, efectividad del tratamiento, efectos secundarios, conveniencia de la forma de administración/comodidad y olvidos en las tomas) y una sobre aspectos generales de satisfacción con el servicio sanitario (relación médico-paciente, tiempo dedicado al paciente e información al paciente).

■ 3.3.5. Covariables

Pero hay que tener en cuenta que existen características del paciente que pueden modular o incluso confundir los resultados obtenidos al medir la satisfacción. Un aspecto importante de la valoración de una experiencia, y más aún en el caso de las experiencias médicas, es que estará muy influenciada por los conocimientos previos de la enfermedad y del tratamiento, las creencias y las expectativas (Oliver, 1997). Aspectos todos ellos muy importantes cuando la evaluación la realiza un paciente que puede desconocer las características de un tratamiento, tener instauradas falsas creencias sobre el rigor en la pauta de ingesta o disponer de expectativas inalcanzables por el fármaco. La adecuada información del paciente debería contribuir a unas valoraciones más realistas. Por ende, las expectativas pueden variar con la edad, las comorbilidades o los efectos indeseados, muchas veces llevando a que el paciente se adapte y sea menos exigente con su salud.

La satisfacción con la medicación se relaciona con la efectividad autopercebida, el control del peso, el control de la glucosa y el bienestar. También se han observado correlaciones con la falta de dolor y la usencia de efectos secundarios. Está relacionada positivamente con la edad y la duración de la enfermedad y negativamente con la depresión y la ansiedad.

Conocer todas estas covariables es de interés cuando se hace necesario descontar su efecto modulador ya que pueden estar atenuando o amplificando las diferencias o relaciones observadas en nuestros estudios, distorsionando el tamaño real del efecto de las intervenciones médicas.

■ 3.4. PREFERENCIAS

La medición de las preferencias ha sido heredada del ámbito del *marketing* y la psicología del consumo debido a la necesidad de conocer la priorización realizada por los consumidores en la adquisición de bienes y servicios, cuando los presupuestos son limitados y es necesario elegir entre varias alternativas.

Para medir la preferencia hacia un conjunto de opciones se suele utilizar el método de “perfil completo” (*full profile*). El producto se descompone en los atributos que varían de uno a otro modelo del producto. Se crean tarjetas, una por cada modelo, en las que se describe el tipo o nivel de cada uno de los atributos que conforman el modelo mediante un perfil. Posteriormente se diseña un experimento de elección en el que una muestra representativa de sujetos debe priorizar las tarjetas según su agrado o disgusto por el modelo representado en cada tarjeta. Una vez ordenados los modelos, se estima la utilidad de cada uno de los niveles de cada atributo, con lo que se obtiene una función de utilidad multiatributo que permite obtener la utilidad global de cada modelo (o perfil) dados los niveles de los atributos de ese modelo. La utilidad final se expresa normalmente en una escala de 0 a 1 puntos.

Por tanto, los estudios de utilidad permiten obtener las preferencias sociales (de un grupo extenso de consumidores potenciales) hacia los distintos atributos del producto o servicio y la combinación de dichos atributos en un perfil concreto en una situación de elección.

Esta es la metodología que se ha utilizado para obtener la utilidad de los diferentes estados de salud ofrecidos por instrumentos de CVRS como el EQ-5D, el SF-6D y el HUI-3. En este caso particular, cada instrumento ofrece una serie de estados de salud en forma de perfiles y la corrección de un estado de salud particular permite obtener la utilidad social de ese perfil. Cuando administramos un instrumento como el EQ-5D, el paciente informa de su estado de salud sin ayuda del médico, pero su estado de salud se corrige reflejando la preferencia (o aversión) que tiene la sociedad por ese estado de salud particular (y no es la ordenación de los estados de salud según una muestra normativa de enfermos según la distribución de las puntuaciones observadas).

La utilidad es la medida de elección cuando se desea combinar la CVRS (como medida de morbilidad) con los años de vida ganados (la mortalidad) para obtener indicadores como los QALY (años de vida corregidos por calidad de vida).

Igual que se ha hecho con la CVRS, en la actualidad hay estudios que valoran la preferencia por distintos tratamientos, en función de atributos como la eficacia, los efectos secundarios, el coste o la vía de administración.

No debemos confundir el anterior tipo de medición de las preferencias con las preferencias personales de un paciente particular. En la actualidad también se están realizando esfuerzos para cuantificar las preferencias de cada paciente por las diversas alternativas terapéuticas disponibles para el tratamiento de su enfermedad. Pero en este último caso se trata de un ejercicio de toma de decisiones, en el que el paciente debe decidir entre varias alternativas (de las que normalmente le informará de manera detallada su médico), con distintas fortalezas y debilidades que el enfermo debe sopesar. El ejemplo más claro son los contratos terapéuticos utilizados actualmente en la práctica clínica.

■ 3.5. COMORBILIDAD

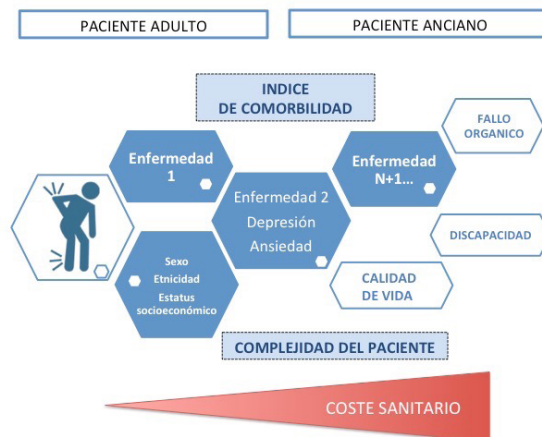
■ 3.5.1. Comorbilidad desde la perspectiva del paciente

La presencia de comorbilidad se asocia con una mayor complejidad en el manejo de los pacientes, peores resultados de salud y un incremento en los costes del servicio sanitario. Además, una mayor comorbilidad predice mayor mortalidad en la población general. Existen múltiples definiciones de comorbilidad dependiendo de si se pretende explicar en un contexto clínico o epidemiológico, o para planear costes sanitarios. La multimorbilidad es la coexistencia de dos o más enfermedades crónicas en el mismo paciente, sin tener en cuenta la relación temporal de las mismas. Desde la perspectiva del paciente, no solamente es importante el número de enfermedades, también es importante la carga psicológica que suponen y cómo afectan a su capacidad funcional.

En general, es fundamental usar índices ya validados para una correcta evaluación de la comorbilidad del paciente. A este respecto, han sido propuestos múltiples índices, siendo el más utilizado el índice de comorbilidad de Charlson (Charlson, *et al.*, 1987). La mayoría de estos índices tienen como objetivo englobar la complejidad de distintas enfermedades en el mismo paciente bien sea para evaluar el coste adicional, la utilización de recursos o para comparar pacientes desde el punto de vista clínico. Pero la comorbilidad no está únicamente influenciada por las características clínicas de las enfermedades asociadas, también está influenciada por

Figura 1.

RELACIÓN ENTRE COMORBILIDAD Y COMPLEJIDAD DEL PACIENTE



Fuente: Elaboración propia.

factores socioeconómicos, culturales, ambientales y factores de comportamiento del paciente. Capturar esta complejidad con índices compuestos es un verdadero reto (Figura 1). El paciente dispone de un conocimiento único sobre su propia enfermedad y cómo ésta afecta a las distintas esferas de su vida. Por otro lado, hoy en día se busca un modelo sanitario que englobe la participación del paciente y que tenga en cuenta sus necesidades, su calidad de vida, sus preferencias en cuanto a tratamiento y su nivel de satisfacción. La utilización de cuestionarios de comorbilidad validados es de gran utilidad para el manejo del paciente y su información correlaciona adecuadamente con los índices recogidos a través de la historia clínica o por el propio médico.

■ 3.5.2. Uso en investigación y en la práctica clínica rutinaria

En 2030 uno de cada cuatro habitantes de la Unión Europea tendrá más de 65 años². Con una población cada vez más envejecida, el nivel de comorbilidad en los pacientes será cada vez mayor. La mayoría de ensayos clínicos aleatorizados tienen unos criterios de inclusión muy estrictos y por lo tanto no representan el espectro de población real que se ve en la práctica clínica rutinaria como, por ejemplo, pacientes con múltiples comorbilidades y polimedicados. Sin embargo, en estudios observacionales y registros de pacientes estas variables pueden ser recogidas. Estos estudios necesitan integrar el impacto de las comorbilidades en el *screening*, diagnóstico, y seguimiento de distintas enfermedades. Una forma de evaluar su impacto es incorporar medidas validadas de comorbilidad y, a este respecto, las medidas informadas por el paciente pueden ser de gran utilidad. Integrar dicho conocimiento en la investigación puede ser de gran utilidad para generar recomendaciones en el manejo terapéutico de los pacientes con comorbilidad asociada. Incluso en la práctica clínica, los instrumentos para medir comorbilidad pueden ser importantes para planificar el manejo terapéutico.

■ 3.5.3. Evaluación de comorbilidad

La comorbilidad se puede medir de dos formas: 1) recogiendo un listado de comorbilidades en cada paciente, 2) utilizando un índice que da un peso específico a cada enfermedad y, al sumarlos, resultan en una puntuación final para cada paciente. La ventaja principal de los índices de comorbilidad es que engloban la severidad y la cantidad de enfermedades en una puntuación final, permitiendo el seguimiento de pacientes individuales y la comparación entre pacientes. Existen múltiples índices para evaluar la comorbilidad del paciente por lo que se debe seleccionar el índice más adecuado dependiendo de los datos disponibles, bien sean administrativos, de la historia clínica, o suministrados por el propio paciente. También es importante considerar la medida de desenlace que nos interesa medir.

² <http://www.consilium.europa.eu/media/30776/qc3210249enc.pdf>

A continuación, repasaremos brevemente algunos de los índices más conocidos centrándonos principalmente en los suministrados por el propio paciente.

Índice de comorbilidad de Charlson (ICC) (Charlson, *et al.*, 1987). Este índice fue descrito inicialmente en 1987 para evaluar el pronóstico basándose en la edad del paciente y el peso de determinadas comorbilidades. Se desarrolló con los datos de 607 pacientes ingresados en un servicio de medicina interna en un periodo de un mes. Se evaluaron los riesgos relativos de las morbilidades observadas, desestimándose aquellas con un riesgo relativo inferior a 1,2. Originalmente, se pretendía predecir la mortalidad a un año entre pacientes hospitalizados, pero ha sido adaptado para predecir también el número de visitas ambulatorias, la duración de los ingresos hospitalarios y los costes. Los datos se pueden obtener de la historia clínica y ha sido adaptado para su uso con el listado de clasificación internacional de enfermedades (ICD9), transformándolo en un cuestionario y combinándolo con la edad del paciente para crear así un índice edad-comorbilidad. Incluye un total de 16 comorbilidades, con una puntuación que oscila entre 1 y 6, dependiendo del grado de severidad (cuadro 2). La puntuación total tiene un rango de 0-36, considerando grave una puntuación >5. Múltiples aspectos de su validez han sido explorados, presenta una buena correlación con otros índices de comorbilidad, predice mortalidad, discapacidad, reingresos y la duración de la estancia hospitalaria, y muestra buena fiabilidad test-retest (de Groot *et al.*, 2003).

Self-administered Comorbidity Questionnaire (SCQ) (Sangha *et al.*, 2003). Este índice se diseñó para ser el equivalente del índice de Charlson cuando no se dispone de los datos de la historia clínica. Lo completa el propio paciente, indicando si sufre alguna de las 12 condiciones médicas seleccionadas por un panel de expertos basándose en el ICC. El paciente también indica si recibe tratamiento específico para cada una de estas enfermedades y si limitan su capacidad funcional. La lista permite añadir dos enfermedades adicionales. A cada respuesta afirmativa se le asigna un punto con una puntuación máxima de 45. Sirve para evaluar el impacto de las comorbilidades sobre la salud del paciente y el uso de recursos. Varios aspectos de su validación han sido examinados (Tabla 2). Existe una versión específica para pacientes con artritis inflamatoria precoz (EIA-SCQ) (Baron *et al.*, 2009). Esta versión se evaluó en una cohorte de pacientes con artritis precoz (una o más articulaciones con sinovitis desde hace más de seis semanas y menos de un año, con diagnóstico de artritis reumatoide o de artritis indiferenciada, sin otra causa de sinovitis aparente). Se evaluó la validez de constructo calculando la correlación con la edad (moderada), la calidad de vida HRQoL (débil), capacidad funcional HAQ-DI (débil) y dolor según el McGill Pain Questionnaire (débil).

Duke Severity of Illness (DUSOI) (Huntley *et al.*, 2012). Este índice fue desarrollado para evaluar pacientes ambulatorios en atención primaria. Basado en registros de pacientes y adaptado para su utilización por medio de una entrevista médico-paciente. Es un índice que permite cuantificar la gravedad de la enfermedad desde el punto de vista del médico para un paciente específico. Primero, se identifican

todos los problemas de salud. Cada comorbilidad va a tener cuatro dominios (los síntomas, las complicaciones, el pronóstico sin tratamiento, y con tratamiento). A cada uno se le da un valor de 0-4, con un valor final entre 0-100. Los datos analizados en diferentes revisiones apoyan su validez concurrente, predictiva y de constructo. La fiabilidad test-retest es moderada, y la fiabilidad entre observadores oscila entre moderada y buena (de Groot *et al.*, 2003).

Índice de comorbilidad de Elixhauser (ICE) (Elixhauser *et al.*, 1998). Este índice consiste en un método para categorizar las comorbilidades del paciente basado en la clasificación internacional de enfermedades (ICD9). Basándose en los datos administrativos de los pacientes de California, se identificó un total de treinta enfermedades presentes en pacientes hospitalizados y con un mayor impacto a corto plazo en las medidas de desenlace del paciente. Posteriormente, se estudiaron diferencias en la duración de la estancia hospitalaria, la mortalidad, y el gasto que presentaron los pacientes con estas comorbilidades, en comparación con los que no las presentaron. El índice considera solo comorbilidades crónicas y excluye aquellas que pueden estar relacionadas con la atención médica recibida durante la hospitalización. Para cumplimentarlo se recogen datos de los registros administrativos. En la versión original, Elixhauser consideraba las comorbilidades por separado, pero, posteriormente, se ha modificado para permitir realizar una medida ponderada. Las comorbilidades ejercen un efecto en los resultados de los tres indicadores estudiados (estancia, mortalidad y coste), tanto cuando se realiza sobre un grupo heterogéneo de pacientes con diversos diagnósticos principales, como cuando se aplican a grupos homogéneos de pacientes basados en un diagnóstico principal específico (misma enfermedad), si bien el valor predictivo no es igual en todos ellos, variando según los tipos pacientes, en cada una de las comorbilidades y en cada uno de los tres indicadores. Existen determinados riesgos y limitaciones en la aplicación de este índice, destacando los derivados de la calidad de los registros administrativos que se utilizan para la obtención de la información y, sobre todo, en lo que se refiere a la exhaustividad y fiabilidad de los datos de los diagnósticos principal y secundario: errores en la selección del diagnóstico principal, y errores de codificación.

Functional Comorbidity Index (Groll *et al.*, 2005). Se trata de un índice de comorbilidad cuyo objetivo es medir la función física en la población general. El índice se compone de 18 ítems con respuesta sí/no. Se calcula sumando el total de respuestas afirmativas, con una puntuación total que oscila de 0-18. Es un cuestionario autoadministrado. El índice no toma en consideración la gravedad de los diagnósticos. El índice se desarrolló en dos pasos: en el primero, se identificaron un amplio conjunto de ítems asociados con la función física a partir de una revisión sistemática de la literatura y seis grupos de trabajo seleccionaron conjunto extenso de ítems. En el segundo paso, los ítems se agruparon basándose en el diagnóstico. Los 40 diagnósticos resultantes fueron analizados utilizando los datos recogidos en dos bases de datos de pacientes: *Canadian Multi Centre Osteoporosis Study* (CaMos) y *National Spine Network* (NSN). Mediante análisis de regresión lineal múltiple, se seleccionaron 17 variables de la CaMos, que estaban significativamente

asociadas con el SF-36, y de la NSN se seleccionó una única variable estadísticamente significativa, que no estaba presente en la CaMos.

Index of Co-existent Disease (ICED) (Miskulin *et al.*, 2001). El índice de enfermedades coexistentes (ICED) combina dos dimensiones o subíndices, una medida de la gravedad (*Index of Disease Severity-IDS*) y otra del deterioro funcional (*Functional Severity-FS*) que producen la enfermedad asociada. El IDS incluye 13 ítems (o condiciones médicas coexistentes) que se valoran en una escala de cinco puntos dependiendo del nivel de severidad. El FS contiene diez áreas funcionales: deambulación, respiración, neurológico, mental, urinario, fecal, alimentación, visión, oído y comunicación. Para cada uno de estos ítems se le da al paciente una valoración de 1 a 3 en su capacidad funcional. IDS y FS se combinan para formar una puntuación final de cuatro niveles: no, leve, moderada o enfermedad grave según el impacto global (Sarfati, 2012). Se precisa acceso a la historia clínica para calcular la puntuación.

Índice de comorbilidad en enfermedades reumáticas (ICER) (Wolfe *et al.*, 2010). Las enfermedades musculoesqueléticas son la mayor causa de discapacidad en Europa y Estados Unidos. Se espera un incremento importante en la población con enfermedades reumáticas debido al envejecimiento poblacional. Por este motivo, aunque este índice es específico de enfermedades reumáticas, puede ser empleado en una gran proporción de pacientes. Este índice fue desarrollado para predecir la mortalidad, la hospitalización, las discapacidades funcional y laboral, y los gastos médicos causados por la comorbilidad en pacientes con enfermedades reumáticas. Fue creado a partir de cuestionarios de pacientes. Se compone de 11 comorbilidades (presentes o pasadas) en un cuestionario a rellenar por el paciente. Los resultados tienen un rango de 0-9. Tiene en cuenta el componente temporal de la comorbilidad recogida en el índice. Se puede utilizar usando datos admirativos o puede ser completado por el propio paciente.

Los índices utilizados con más frecuencia para evaluar comorbilidad requieren un entrenamiento médico del personal que se va a encargar de extraer los datos de las historias clínicas. Además del entrenamiento, requieren tiempo y están limitados por la disponibilidad y calidad de los datos que precisan. Una alternativa puede ser el uso de cuestionarios administrados al paciente. Diferentes estudios han mostrado cómo los pacientes pueden evaluar de forma adecuada su historia médica pasada y actual completando un listado de comorbilidades. Estos cuestionarios de comorbilidad informados por el paciente pueden ser de gran utilidad, especialmente para estudios epidemiológicos, ya que pueden ser una herramienta muy útil para ajustar posibles factores de confusión. Por otro lado, una ventaja de los índices de comorbilidad es que sintetizan todas las enfermedades coexistentes y su severidad en un único valor numérico, con lo que facilitan la comparación entre pacientes. Sin embargo, es necesario reconocer que pueden existir ciertos errores a la hora de clasificar las comorbilidades por parte del paciente. Por ejemplo, algunos pacientes pueden tener dificultades a la hora de diferenciar ciertas patologías: un ejemplo

Tabla 2.

CUESTIONARIOS DE COMORBILIDAD

Instrumento y referencia	Recogida de datos	Número de ítems y rango de puntuación	Contenido	Dificultad y tiempo	Medida de resultado de interés
<p>Índice de comorbilidad de Charlson (ICC) Charlson <i>et al.</i> (1987) <i>J. Chronic Dis.</i></p>	Datos administrativos	16 ítems 0-36	<p>1 punto: IAM, insuficiencia cardiaca crónica, enfermedad cerebrovascular y vascular periférica, demencia, EPOC, ETC, ECG, DM, hepatopatía.</p> <p>2 puntos: DM con daño orgánico, IRC, hemiplejía, leucemia, linfoma, cáncer.</p> <p>3 puntos: hepatopatía moderada o severa.</p>	<p>Dificultad: Baja-media</p> <p>Tiempo: 0-60 min</p>	<p>Mortalidad a 1 año entre los pacientes ingresados en el hospital.</p> <p>Posteriormente se adaptó para predecir costes.</p>
<p>Self-administered C morbidity Questionnaire (SCQ) Early Inflammatory Arthritis –SCQ Sangha <i>et al.</i> (2003) <i>Arthritis Rheum.</i></p>	Adaptado del ICC para el paciente Versión para pacientes con artritis precoz	13 ítems 0-45	<p>6 puntos: metástasis y SIDA</p> <p>Cada ítem tiene 3 dominios (respuesta sí/no)</p> <ol style="list-style-type: none"> ¿Tiene este problema? ¿Recibe tratamiento por él? ¿Limita sus actividades? <p>Versión para artritis precoz: 12 ítems</p>	<p>Dificultad: baja</p> <p>Tiempo: 5-15 min</p>	<p>Impacto de las comorbilidades sobre la salud del paciente y sobre el uso de recursos.</p> <p>Autoadministrado.</p>
<p>Duke Severity of Illness (DUSOI) Huntley <i>et al.</i> (2012) <i>Annals Fam Med.</i></p>	Entrevista médico-paciente	4 ítems 0-100	<p>Para cada comorbilidad se evalúan</p> <ol style="list-style-type: none"> 1) gravedad del síntoma 2) gravedad de la complicación, 3) pronóstico sin tratamiento 4) pronóstico con tratamiento 	<p>Dificultad: Media-alta</p> <p>Tiempo: >60 min</p>	<p>Cuantificar la carga de la enfermedad medida por el médico.</p>

Tabla 2. (continuación)

CUESTIONARIOS DE COMORBILIDAD

Instrumento y referencia	Recogida de datos	Número de ítems y rango de puntuación	Contenido	Dificultad y tiempo	Medida de resultado de interés
Índice de comorbilidad de Elixhauser (ICE) Elixhauser et al. (1998) <i>Medical care</i> .	Datos administrativos	30 ítems 0-18	Comorbilidades no agudas (17 del ICC+ 13 adicionales), excluyendo las que puedan estar relacionadas con la atención médica recibida durante la hospitalización.	Dificultad: Media-alta Tiempo: >60 min	Predice el uso de recursos hospitalarios y mortalidad.
Funcional Comorbidity Index (FCI) Groll et al. (2005) <i>J Clin Epidemiol</i> .	Datos administrativos o completado por el paciente.	18 ítems 0-18	Comorbilidades con respuesta si/no. Diferente para cada comorbilidad.	Dificultad: Baja Tiempo: 5-15 min	Impacto de las comorbilidades sobre la función física del paciente.
Index of Co-existent Disease (ICD) Miskulin et al. (2001) <i>Kidney Internal</i> .	Datos administrativos	24 ítems 0-100	Combina 2 dimensiones, la gravedad de la comorbilidad de 14 sistemas (1-5) y la discapacidad en 10 áreas funcionales (1-3).	Dificultad: Media Tiempo: >60 min	Impacto de comorbilidades sobre una medida de desenlace de interés en el periodo de observación.
Índice de comorbilidad en Enfermedades Reumáticas (ICER) Wolfe et al. (2010) <i>J Rheumatol</i> .	Se puede usar con datos administrativos o completado por el paciente	11 ítems 0-9	2 puntos: Enfermedad pulmonar, enfermedad cardiovascular, IAM, HTA. 1 punto: diabetes, úlcera o EGI, depresión, y cáncer.	Dificultad: Baja Tiempo: 5-15 min	Impacto de la comorbilidad en mortalidad, hospitalización, discapacidad funcional y laboral y gastos médicos.

Nota: IAM, infarto agudo de miocardio; EPOC, enfermedad pulmonar obstructiva crónica; ETC, enfermedad del tejido conectivo; IRC, insuficiencia renal crónica; EGI, enfermedad gastrointestinal; DM, diabetes mellitus; SIDA, síndrome de inmunodeficiencia adquirida; HTA, hipertensión arterial
Fuente: Elaboración propia.

claro es la diferenciación entre artrosis (degenerativa) y artritis reumatoide (inflamatoria). Además, en algunos cuestionarios las categorías no se excluyen mutuamente como es el caso de cáncer de pulmón que puede ser evaluado como enfermedad pulmonar y/o cáncer. A pesar de estas posibles limitaciones, si no se dispone de las historias clínicas o los recursos necesarios para recoger la información, los cuestionarios auto-administrados pueden ser una alternativa válida y práctica para evaluar la comorbilidad en los estudios prospectivos.

■ 3.6. DISCAPACIDAD

■ 3.6.1. Definición y modelos de discapacidad

El concepto de discapacidad es clave en la atención sociosanitaria, ya que concreta uno de los principales componentes de la salud. Está estrechamente relacionado con los conceptos de deficiencia, funcionamiento y actividad, hasta el punto de que la mayoría de escalas de evaluación de la discapacidad incorporan algunos de estos conceptos, a pesar de que no son equivalentes entre sí.

Se estima que aproximadamente el 15 % de la población mundial vive con algún tipo de discapacidad. En España, según datos de la *Encuesta de Discapacidad, Autonomía personal y situaciones de Dependencia 2008* (Instituto Nacional de Estadística, 2008), hay 3,5 millones de personas con alguna discapacidad, de las que casi el 60 % son mujeres y el 63 % son personas mayores de 65 años. La mayor demanda de servicios sociosanitarios de las personas con discapacidad y la necesidad de recursos específicos para esta población suponen un desafío para la sociedad y el Sistema Nacional de Salud (SNS). Las políticas sanitarias sobre discapacidad enfatizan las estrategias de prevención de enfermedades, las intervenciones rehabilitadoras y la aportación de ayudas técnicas y la atención integral mediante la elaboración de planes terapéuticos individualizados para cada paciente. La traslación de estas políticas requiere partir de una valoración previa del déficit funcional y de la capacidad del individuo para desempeñar sus actividades y roles.

Durante gran parte del siglo XX, la forma predominante de entender la discapacidad ha sido el modelo médico-rehabilitador. Según este modelo, la discapacidad se define como una consecuencia de una enfermedad, accidente o condición negativa de la salud, que requiere de cuidados médicos y rehabilitadores para conseguir la mejoría total o una adecuada adaptación de la persona a la sociedad. Uno de los principales efectos de este enfoque fue la puesta en marcha de leyes y la creación de servicios de atención y rehabilitación en la mayor parte de los países occidentales, pero a la vez favoreció una visión pasiva y subordinada de las personas con discapacidad. Como reacción, a partir de la década de los sesenta cobró fuerza el modelo social de la discapacidad, vinculado con los movimientos de vida indepen-

diente y de derechos humanos, por el que se apela a la responsabilidad de los factores sociales que restringen, limitan o impiden vivir una vida plena a las personas con discapacidad (Velarde Lezama, 2012). Este modelo es el que se recogió en la Convención Internacional sobre los Derechos de las Personas con Discapacidad de la Organización de las Naciones Unidas (ONU) en 2006.

En la actualidad, la Organización Mundial de la Salud (OMS) define la discapacidad como un fenómeno complejo que refleja los aspectos negativos de la interacción entre personas con un problema de salud (como parálisis cerebral, síndrome de Down o depresión) y factores personales y ambientales (como actitudes negativas, transporte y edificios públicos inaccesibles, y falta de apoyo social) (Organización Mundial de la Salud y Banco Mundial, 2011). Según el modelo de la OMS, publicado en la Clasificación Internacional del Funcionamiento y la Discapacidad de 2002 (World Health Organization, 2001), la discapacidad es un término genérico que engloba deficiencias, limitaciones de actividad y restricciones para la participación. Las deficiencias son los problemas, desviaciones o pérdidas en las funciones o estructuras corporales, la actividad es el desempeño/realización de una tarea o acción por parte de un individuo y la participación es el acto de involucrarse en una situación vital.

■ 3.6.2. Evaluación de la discapacidad

La evaluación de la discapacidad se realiza habitualmente mediante pruebas que pueden incluir la evaluación de las capacidades físicas (equilibrio, rapidez o amplitud de movimiento), las habilidades de autocuidado (actividades básicas de la vida diaria, ABVD) o la participación en diversas actividades (actividades instrumentales o avanzadas de la vida diaria, AIVD). La evaluación de las capacidades físicas se realiza mediante observación y medición objetiva del desempeño en pruebas como el *"timed up & go"* y la velocidad de la marcha, o en baterías como la *Short Physical Performance Battery* (SPPB) de Guralnik *et al.* (1994), por lo que no son objeto de este capítulo. Para la valoración de las ABVD y AIVD existe gran variedad de escalas auto-administradas. La Tabla 3 ofrece un listado de algunas de ellas, de las que se describen a continuación las más usadas.

Índice de Barthel (Mahoney y Barthel, 1965). Entre las escalas de evaluación de las ABVD, destaca este índice, que valora diez actividades: comer, vestirse, arreglarse, incontinencia fecal o urinaria, ir al retrete, traslado sillón-cama, deambulación y subir y bajar escaleras. La puntuación total de la escala oscila entre 0 (completamente dependiente) y 100 (completamente independiente), con puntos de corte para dependencia leve, moderada, grave o total. La escala cuenta con traducciones y estudios de validación en varios idiomas. Presenta una buena consistencia interna, fiabilidad interobservadores y correlación con medidas de calidad de vida relacionada con la salud como el EQ-5D. El índice es además predictor de mortalidad, institucionalización, utilización de servicios sociosanitarios y riesgo de caídas.

Sin embargo, presenta dificultades para detectar cambios en el estado funcional en las puntuaciones extremas.

Índice de Katz (Katz *et al.*, 1963). Evalúa el grado de dependencia en seis funciones básicas: baño, vestido, uso del retrete, movilidad, continencia y alimentación. La escala arroja ocho niveles de dependencia que van desde A (independiente en todas las funciones) a G (dependiente en las seis funciones). El índice de Katz presenta buena consistencia interna y fiabilidad test-retest. Además, es capaz de predecir mortalidad, institucionalización, hospitalización y expectativa de vida activa. Ha sido traducido y validado en diversos idiomas. Entre sus limitaciones destacan la baja correlación con medidas de gravedad de la enfermedad, la pobre sensibilidad al cambio, sobre todo en pacientes con menor grado de dependencia, y que se valoren principalmente tareas dependientes de las extremidades inferiores.

Escala de Lawton y Brody (Lawton y Brody, 1969). Es una de las más utilizadas internacionalmente. Incluye seis ítems de evaluación de ABVD, con contenido similar al índice de Katz, y ocho ítems para AIVD: capacidad para utilizar el teléfono, hacer compras, preparar la comida, cuidado de la casa, lavado de la ropa, uso de medios de transporte, responsabilidad en la toma de la medicación y manejo de sus asuntos económicos. Aunque cada ítem tiene varias opciones de respuesta, la versión auto-administrada de la subescala AIVD los puntúa como 0 (dependiente) o 1 (independiente), resultando en una puntuación total que va de 0 (máxima dependencia) a 8 (independencia total). Tiene una buena fiabilidad test-retest e interobservadores, aunque sus principales limitaciones son su baja sensibilidad al cambio y la influencia de los roles de género y los aspectos culturales en la puntuación de la escala.

Indicador General de Limitación de la Actividad (Global Activity Limitation Indicator, GALI) (Verbrugge, 1997). Es una medida sencilla compuesta de una sola pregunta: durante al menos los últimos 6 meses, ¿en qué medida se ha visto limitado/a debido a un problema de salud para realizar las actividades que la gente habitualmente hace? Las opciones de respuesta son tres: “gravemente limitado”, “limitado, pero no gravemente” y “nada limitado”. El GALI refleja la percepción subjetiva de las limitaciones de la actividad del modelo de la CIF, y se ha incluido en diversos estudios poblacionales, como la *Encuesta Europea de Salud* y la *Encuesta SHARE (Survey of Health, Ageing and Retirement in Europe)*. Correlaciona con medidas de capacidad física, ABVD y AIVD y es capaz de predecir resultados de salud, como mortalidad y morbilidad.

Escala de discapacidad de la OMS (World Health Organization Disability Assessment Schedule, WHO-DAS) (World Health Organization, 2009). Es una escala estandarizada para uso universal basada en el modelo de discapacidad de la CIF. Tiene dos versiones, de 12 y 36 ítems. La versión de 12 ítems es de utilidad para cribado de discapacidad, y ha sido recomendada por la OMS para estudios

Tabla 3.

ESCALAS DE EVALUACIÓN DE DISCAPACIDAD

Instrumento y referencia	Recogida de datos	Número de ítems y rango de puntuación
<p>Índice de Barthel Mahoney y Barthel (1965) <i>Md State Med J</i></p>	<p>10 ítems. 0 (completamente dependiente) - 100 (completamente independiente).</p>	<p>ABVD: comer, vestirse, arreglarse, incontinencia fecal o urinaria, ir al retrete, traslado sillón-cama, deambulación y subir y bajar escaleras.</p>
<p>Índice de Katz Katz et al. (1963) <i>JAMA</i></p>	<p>6 ítems. A (independiente en todas las funciones) - G (dependiente en las seis funciones).</p>	<p>ABVD: baño, vestido, uso del retrete, movilidad, continencia y alimentación.</p>
<p>Escala de Lawton y Brody Lawton y Brody (1969) <i>Gerontologist</i></p>	<p>6 ítems para ABVD, 8 para AIVD. 0 (máxima dependencia) - 8 (independencia total).</p>	<p>ABVD: usar el retrete, comer, vestirse, arreglarse, ambulación, baño. AIVD: capacidad para utilizar el teléfono, hacer compras, preparar la comida, cuidado de la casa, lavado de la ropa, uso de medios de transporte, responsabilidad en la toma de la medicación y manejo de sus asuntos económicos.</p>
<p>Indicador General de Limitación de la Actividad (Global Activity Limitation Indicator, GALI) Verbrugge (1997) <i>J Aging Stud.</i></p>	<p>1 ítems. 3 opciones de respuesta: gravemente limitado, limitado, pero no gravemente y nada limitado.</p>	<p>Limitaciones para realizar las actividades habituales.</p>
<p>Escala de discapacidad de la OMS (World Health Organization Disability Assessment Schedule, WHO-DAS) World Health Organization (2009)</p>	<p>23 ítems. 23-69 puntos (17-51 si nunca hizo tareas domésticas).</p>	<p>AIVD: cuidado personal, tareas domésticas, movilidad dentro de la casa y movilidad exterior.</p>

ESCALAS DE EVALUACIÓN DE DISCAPACIDAD

Instrumento y referencia	Recogida de datos	Número de ítems y rango de puntuación
<p>Escala de Incapacidad Física de la Cruz Roja Cruz Roja. Curso sobre el uso de escalas de valoración geriátrica</p> <p>Prous Science (2006)</p>	<p>1 ítems.</p> <p>0 (independencia) – 5 (incapacidad funcional total).</p>	<p>Incapacidad física.</p>
<p>Índice de Karnofsky</p> <p>Karnofsky et al (1948)</p> <p>Cancer</p>	<p>1 ítems.</p> <p>0 (fallecido) – 100 (normal).</p>	<p>Capacidad del paciente para poder realizar actividades cotidianas.</p>
<p>Cuestionario de Evaluación Funcional Multidimensional OARS</p> <p>Duke University (1981)</p> <p>J Gerontol</p>	<p>7 ítems para ABVD, 7 para AIVD.</p> <p>Ausencia de dependencia (puntuación 1-2) – dependencia grave o total (puntuación 5-6).</p>	<p>ABVD: comer, vestirse/desnudarse, andar sin ayuda, cuidar de su apariencia, acostarse/levantarse, bañarse/ducharse y uso del retrete.</p> <p>AIVD: usar el teléfono, llegar a los sitios (viajar), comprar, hacer la comida, hacer tareas de casa, tomar los medicamentos y administrar su dinero.</p>

Nota: ABVD: actividades básicas de la vida diaria; AIVD: actividades instrumentales de la vida diaria.

Fuente: Elaboración propia.

poblacionales. La versión de 36 ítems (WHO-DAS 2.0) evalúa el nivel de funcionamiento en seis dominios: cognición, movilidad, cuidado personal, relaciones, actividades cotidianas y participación. Las opciones de respuesta oscilan entre 1 (ninguna limitación) y 5 (limitación extrema), con una puntuación total que puede calcularse mediante la suma de los ítems individuales o estandarizarse en un rango de 0 a 100 (donde 0 %=sin discapacidad y 100 %=discapacidad completa), con los siguientes puntos de corte: sin discapacidad (0–4%), discapacidad leve (5–24 %), discapacidad moderada (25–49 %), discapacidad grave (50–95 %) y discapacidad extrema (95–100 %). La escala ha sido validada en diversos idiomas, con resultados satisfactorios en relación con la fiabilidad, validez y sensibilidad al cambio.

■ 3.7. CONCLUSIONES

- La CVRS es un constructo multidimensional y subjetivo, centrado en la percepción del paciente sobre su estado de salud, enfermedad y tratamiento. Su medición tiene aplicaciones a nivel clínico, de investigación y epidemiológico, para el desarrollo de políticas y programas sanitarios. Los cuestionarios genéricos SF-36 y EQ-5D-5L, permiten la comparación entre problemas de salud. El sistema EORTC, para cáncer, y FACIT, para enfermedades crónicas, son ejemplos interesantes de cuestionarios modulares. El KIDSCREEN es el más utilizado en niños y adolescentes. La investigación futura debería ir encaminada a evaluar los CVRS como herramienta de gestión en el Sistema Nacional de Salud y su incorporación en sistemas digitales de recogida de datos clínicos.
- La satisfacción con el tratamiento es una valoración afectiva de la discordancia entre la calidad de tratamiento esperada y la recibida. Aunque es un concepto multidimensional las valoraciones se agregan en una puntuación global. Es muy sensible a la eficacia del tratamiento y a sus efectos adversos, y predice la adherencia al tratamiento y el uso reiterado de un producto o servicio. Como inconveniente, con frecuencia suele ofrecer valores próximos al máximo posible lo que, paradójicamente, facilita identificar a los pacientes insatisfechos.
- Hay pocos estudios de preferencia por medicamentos, posiblemente debido a que, en nuestro ámbito, es el clínico quién prescribe y decide los cambios en el tratamiento. Algunos estudios de preferencias se han utilizado para determinar el valor social de los estados de salud y para determinar la utilidad de cada estado de salud, incorporándolo en el cálculo de los QALYs.
- No existe un índice de comorbilidad ideal para todas las situaciones clínicas ya que todos requieren cierta simplificación en el grado de comorbilidad del paciente, lo que lleva a una pérdida de información. La elección de una medida de comorbilidad depende del contexto en el que se quiera aplicar, la medida de desenlace de interés y la población bajo estudio pero, princi-

palmente, depende de los datos disponibles, siendo los cuestionarios informados por el paciente una opción válida y práctica para medir comorbilidad. Es una variable clínica muy importante ya que afecta a múltiples medidas de desenlace, principalmente mortalidad.

- La evaluación de la discapacidad es crucial para comprender el impacto de la enfermedad, determinar su progresión y poder predecir resultados en salud. Sin embargo, no hay consenso sobre la definición de la discapacidad ni sobre los límites que determinan la necesidad de intervención. Además, son necesarios instrumentos de evaluación adecuados, con buenas propiedades psicométricas y sensibilidad al cambio y libres de influencias culturales, roles de género y preferencias personales. La importancia de contar con medidas de evaluación global de discapacidad ha sido ampliamente reconocida y se ha propuesto la CIF como el marco conceptual a partir del que se puede desarrollar la evaluación estandarizada de la discapacidad integrada en iniciativas como el sistema PROMIS.
- En los últimos años ha aumentado de forma considerable el uso de las PRO en el ámbito sanitario. La selección del índice o cuestionario adecuado depende del constructo que se desea medir, ámbito en que se va a emplear y sus propiedades psicométricas. Existen diferentes recursos que pueden ayudar a seleccionar el PRO adecuado, como la biblioteca virtual BiblioPro³ o la desarrollada por la Liga Europea contra el Reumatismo (EULAR) y disponible de forma gratuita online⁴ que incluye PROs validados, algunos específicos para enfermedades reumatológicas y otros de uso genérico con versiones validadas en diferentes idiomas.

REFERENCIAS

- AJZEN, I. y FISHBEIN, M. (1972). Attitudes and normative beliefs as factors influencing behavioral intentions. *J Pers Soc Psychol*, 21(1), pp. 1-9.
- ALONSO, J., PRIETO, L. y ANTÓ, J. M. (1995). La versión española del SF-36 Health Survey (Cuestionario de Salud SF-36): un instrumento para la medida de los resultados clínicos. *Medicina clínica*, 104(20), pp. 771-776.
- ARRARAS, J. I., ARIAS, F., TEJEDOR, M. *et al.* (2002). The EORTC QLQ-C30 (version 3.0) Quality of Life questionnaire: validation study for Spain with head and neck cancer patients. *Psychooncology*, 11(3), pp. 249-256.
- ATKINSON, M. J., KUMAR, R., CAPPELLERI, J. C. *et al.* (2005). Hierarchical Construct Validity of the Treatment Satisfaction Questionnaire for Medication (TSQM Version II) among Outpatient Pharmacy Consumers. *Value Health*, 8(S1), pp. S9-S24.
- AYMERICH, M., BERRA, S., GUILLAMÓN, I., HERDMAN, M., ALONSO, J., RAVENS-SIEBERER, U. y RAJMIL, L. (2005). Development of the Spanish version of the KIDSCREEN: a health-related quality of life instrument for children and adolescents. *Gaceta Sanitaria*, 19(2), pp. 93-102.

³ <https://www.bibliopro.org/>

⁴ <http://oml.eular.org/>

- BARBOSA, C. D., BALP, M.-M., KULICH, K. *et al.* (2012). A literature review to explore the link between treatment satisfaction and adherence, compliance, and persistence. *Patient Prefer Adherence*, 6, pp. 39-48.
- BARON, M., SCHIEIR, O., HUDSON, M. *et al.* (2009). Evaluation of the clinimetric properties of the Early Inflammatory Arthritis--self-administered comorbidity questionnaire. *Rheumatol Oxf*, 48(4), pp. 390-394.
- BHARMAL, M., PAYNE, K., ATKINSON, M. J. *et al.* (2009). Validation of an Abbreviated Treatment Satisfaction Questionnaire for Medication (TSQM-9) among patients on antihypertensive medications. *Health Qual Life Outcomes*, 7, p. 36.
- CELLA, D., HERNANDEZ, L., BONOMI, A. E. *et al.* (1998). Spanish language translation and initial validation of the functional assessment of cancer therapy quality-of-life instrument. *Med Care*, 36(9), pp. 1407-1418.
- CHARLSON, M. E., POMPEI, P., ALES, K. L. *et al.* (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*, 40(5), pp. 373-383.
- DELESTRAS, S., ROUSTIT, M., BEDOUCH, P. *et al.* (2013). Comparison between two generic questionnaires to assess satisfaction with medication in chronic diseases. *PloS One*, 8(2), p. e56247.
- DOWARD, L. C. y MCKENNA, S. P. (2004). Defining patient-reported outcomes. *Value Health*, 7(Suppl 1), pp. S4-S8.
- ELIXHAUSER, A., STEINER, C., HARRIS, D. R. *et al.* (1998). Comorbidity measures for use with administrative data. *Med Care*, 36(1), pp. 8-27.
- EUROQOL GROUP. (1990). EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy*, 16(3), pp.199-208. The EuroQol Group.
- FERRANS, C. E., ZERWIC, J. J., WILBUR, J. E. *et al.* (2005). Conceptual model of health-related quality of life. *J Nurs Scholarsh*, 37(4), pp. 336-342.
- FRIES, J. F., BRUCE, B. y CELLA, D. (2005). The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. *Clin Exp Rheumatol*, 23(5 Suppl 39), pp. S53-S57.
- GROLL, D. L., TO, T., BOMBARDIER, C. *et al.* (2005). The development of a comorbidity index with physical function as the outcome. *J Clin Epidemiol*, 58(6), pp. 595-602.
- GRÖNROOS, C. (1984). A Service Quality Model and its Marketing Implications. *Eur J Mark*, 18(4), pp. 36-44.
- GURALNIK, J. M., SIMONSICK, E. M., FERRUCCI, L. *et al.* (1994). A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol*, 49 (2), pp. M85-M94.
- DE GROOT, V., BECKERMAN, H., LANKHORST, G. J. *et al.* (2003). How to measure comorbidity.a critical review of available methods. *J Clin Epidemiol*, 56(3), pp. 221-229.
- HERNANDEZ, G., GARIN, O., PARDO, Y., VILAGUT, G., PONT, À., SUÁREZ, M., NEIRA, M., RAJMIL, L., GOROSTIZA, I., RAMALLO-FARIÑA, Y., CABASES, J., ALONSO, J. y FERRER, M. (2018). Validity of the EQ-5D-5L and reference norms for the Spanish population. *Quality of life research*, 27(9), pp. 2337-2348. <https://doi.org/10.1007/s11136-018-1877-5>
- HUDAK, P. L. y WRIGHT, J. G. (2000). The characteristics of patient satisfaction measures. *Spine*, 25(24), pp. 3167-3177.
- HUNTLEY, A. L., JOHNSON, R., PURDY, S. *et al.* (2012). Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. *Ann Fam Med*, 10(2), pp. 134-141.
- INSTITUTO NACIONAL DE ESTADÍSTICA. (2008). *Encuesta sobre discapacidad, autonomía personal y situaciones de dependencia (EDAD)*, [citado 7 de julio de 2014]. Disponible en: www.ine.es/metodologia/t15/t1530418.pdf

- (2013). *Encuesta nacional de salud en España, 2011-2012*, [citado 20 de diciembre de 2017]. Disponible en: http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176783&menu=resultados&idp=1254735573175 (2013)
- KATZ, S., FORD, A. B., MOSKOWITZ, R. W. *et al.* (1963). Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychological function. *JAMA*, 185 (12), pp. 914-919.
- LAWTON, M. P. y BRODY, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist*, 9 (3), pp. 179-186.
- MAHONEY, F. I. y BARTHEL, D. W. (1965). Functional evaluation: the Barthel Index. *Md. State Med J*, 14, pp. 61-65.
- MARTÍN-FERNÁNDEZ, J., ARIZA-CARDIEL, G., POLENTINOS-CASTRO, E. *et al.* (2018). Explaining differences in perceived health-related quality of life: a study within the Spanish population. *Gac Sanit*, 32(5), pp. 447-453.
- MAURICE-SZAMBURSKI, A., MICHEL, P., LOUNDOU, A. *et al.* (2017). G-MISS Study Investigators. Validation of the generic medical interview satisfaction scale: the G-MISS questionnaire. *Health Qual Life Outcomes*, 15(1), p. 36.
- McHORNEY, C. A., WARE, J. E. Jr., ROGERS, W. *et al.* (1992). The validity and relative precision of MOS short- and long-form health status scales and Dartmouth COOP charts. Results from the Medical Outcomes Study. *Med Care*, 30(5 Suppl), pp. MS253-MS265.
- MISKULIN, D. C., ATHENITES, N. V., YAN, G. *et al.* (2001). Comorbidity assessment using the Index of Coexistent Diseases in a multicenter clinical trial. *Kidney Int*, 60(4), pp. 1498-510.
- OLIVER, R. L. (1997). *Satisfaction: A Behavioral Perspective on the Consumer*. McGraw Hill.
- ORGANIZACIÓN MUNDIAL DE LA SALUD y BANCO MUNDIAL. (2011). *Informe mundial sobre la discapacidad*. [citado 12 de julio de 2018]. Disponible en: http://www.who.int/disabilities/world_report/2011/es/
- PARASURAMAN, A., ZEITHAML, V. y BERRY, L. (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service. *Quality J Retail*, 64(1), pp. 12-40.
- PARASURAMAN, A. P., ZEITHAML, V. y BERRY, L. A. (1985). Conceptual Model of Service Quality, and its Implication for Future Research (SERVQUAL). *J Mark*, 49, pp. 41-50.
- REBOLLO, P., CUERVO, J., VILLA, G. *et al.* (2010). Desarrollo y validación de un cuestionario genérico para la evaluación de la satisfacción de los pacientes crónicos: Cuestionario SAT-Q. *Aten Primaria*, 42(12), pp. 612-619.
- REJAS, J., RUIZ, M., PARDO, A. *et al.* (2013). Detecting changes in patient treatment satisfaction with medicines: the SATMED-Q. *Value Health*, 16(1), pp. 88-96.
- RUIZ, M. A., PARDO, A., REJAS, J. *et al.* (2008). Development, and validation of the «Treatment Satisfaction with Medicines Questionnaire» (SATMED-Q). *Value Health*, 11(5), pp. 913-926.
- RUST, R.T. y OLIVER, R. W. (1994). The Death of Advertising. *J Advert*, 23(4), pp. 71-77 [citado 12 de julio de 2018]. Disponible en: <https://doi.org/10.1080/00913367.1943.10673460>
- SANGHA, O., STUCKI, G., LIANG, M. H. *et al.* (2003). The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. *Arthritis Rheum*, 49(2), pp. 156-163.
- SARFATI, D. (2012). Review of methods used to measure comorbidity in cancer populations: no gold standard exists. *J Clin Epidemiol*, 65(9), pp. 924-933.
- SHIKIAR, R. y RENTZ, A.M. (2004). Satisfaction with medication: an overview of conceptual, methodologic, and regulatory issues. *Value Health*, 7(2), pp. 204-215.
- THE WORLD HEALTH ORGANIZATION QUALITY OF LIFE ASSESSMENT (WHOQOL). (1995). Position paper from the World Health Organization. *Soc Sci Med*, 41(10), pp. 1403-1409.

- U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES FDA CENTER FOR DRUG EVALUATION AND RESEARCH, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES FDA CENTER FOR BIOLOGICS EVALUATION AND RESEARCH, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES FDA CENTER FOR DEVICES AND RADIOLOGICAL HEALTH. (2006). Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*, 4, p. 79.
- US FOOD AND DRUG ADMINISTRATION (2016). *Value and Use of Patient Reported Outcomes (PROs) in Assessing Effects of Medical Devices. CDRH Strategic Priorities 2016-2017* [Internet] [citado 12 de julio de 2018]. Disponible en: <https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHVisionandMission/UCM588576.pdf>
- VELARDE LEZAMA, V. (2012). Los modelos de la discapacidad: un recorrido histórico. *Rev Empresa Humanismo*, 15(1), pp. 115-136.
- VERBRUGGE, L. M. (1997). A global disability indicator. *J Aging Stud*, 11(4), pp. 337-362.
- VILLAR LÓPEZ, J., LIZÁN TUDELA, L., SOTO ÁLVAREZ, J. *et al.* (2009). La satisfacción con el tratamiento. *Aten Primaria*, 41(11), pp. 637-645.
- VOGELI, C., SHIELDS, A. E., LEE, T. A. *et al.* (2007). Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. *J Gen Intern Med*, 22(Suppl 3), pp. 391-955.
- WEAVER, M., PATRICK, D. L., MARKSON, L. E. *et al.* (1997). Issues in the measurement of satisfaction with treatment. *Am J Manag Care*, 3(4), pp. 579-594.
- WEINBERGER, M., ODDONE, E. Z. y HENDERSON, W. G. (1996). Does increased access to primary care reduce hospital readmissions? Veterans Affairs Cooperative Study Group on Primary Care and Hospital Readmission. *N Engl J Med*, 334(22), pp. 1441-1447.
- WILSON, I. B. y CLEARY, P. D. (1995). Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA*, 273(1), pp. 59-65.
- WOLFE, F., MICHAUD, K., LI, T. *et al.* (2010). Chronic conditions, and health problems in rheumatic diseases: comparisons with rheumatoid arthritis, noninflammatory rheumatic disorders, systemic lupus erythematosus, and fibromyalgia. *J Rheumatol*, 37(2), pp. 305-315.
- WORLD HEALTH ORGANIZATION. (2001). *International Classification of Functioning, Disability and Health*. Geneva: World Health Organization.
- . (2009). *World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0)*. Geneva: World Health Organization.
- ZACK, M. M. (2013). Centers for Disease Control and Prevention (CDC). Health-related quality of life - United States, 2006 and 2010. *MMWR*, 62(Suppl 3), pp. 105-111.



4

RESULTADOS EN SALUD PERCIBIDOS Y COMUNICADOS POR LOS OBSERVADORES: ¿QUÉ SON Y CÓMO SE MIDEN?

JAVIER SOTO ÁLVAREZ, MERCEDES MADERUELO

■ 4.1. CONCEPTO Y CIRCUNSTANCIAS EN LAS QUE DEBERÍAN EMPLEARSE

Dentro de la evaluación de los resultados clínicos (*clinical outcome assessment*) de los medicamentos y productos sanitarios, los resultados comunicados por los observadores (*observer-reported outcome assessment*) es una fuente adicional importante para poder conocer el beneficio terapéutico producido por un medicamento u otra tecnología sanitaria, en especial en situaciones donde el propio paciente no puede comunicar sus percepciones y vivencias por sí mismo, siendo complementaria a la evaluación de resultados en salud producidos por los clínicos (*clinician-reported outcome assessment*) y los propios pacientes (*patient-reported outcomes assessment*) (Powers *et al.*, 2017).

Por lo tanto, la medición de los resultados comunicados por un observador se basa en una medida de signos y síntomas que presentan los pacientes, así como actividades, comportamientos y verbalizaciones relacionadas con los pacientes, que son percibidas y valoradas por personas que no son el propio paciente, ni un profesional sanitario (médico, enfermera, farmacéutico, etc.). Habitualmente, el observador que percibe y comunica estos resultados relacionados con el paciente es un familiar (padres, hermanos, abuelos, etc.), un cuidador (en patologías con elevada discapacidad o con problemas cognitivos importantes) u otra persona que observa al paciente de manera habitual (maestros, entrenadores, cuidadores, cancheros, etc.), y son personas que no necesitan una formación y especialización sanitaria. Este tipo de medición de resultados en salud es muy importante para aquellos pacientes que no pueden comunicar por sí mismos las percepciones que tienen de la enfermedad que padecen o cómo les afectan los tratamientos pautados para el manejo de su enfermedad, tales como los neonatos, niños pequeños, sordomudos, pacientes con funciones superiores mermadas, pacientes con patología degenerativa del sistema nervioso central (SNC), pacientes con un grado importante de discapacidad, etc. (Walton *et al.*, 2015).

Hay tres circunstancias en las que debería emplearse la medición de los resultados comunicados por un observador:

1. Cuando un paciente no puede comunicar aspectos de su salud relacionados con la enfermedad que padece y los resultados percibidos de los tratamientos administrados, o si puede, los datos que van a proporcionar no

van a ser fiables y creíbles para tomar decisiones terapéuticas, como por ejemplo población pediátrica por debajo de 10 años, pacientes con enfermedades mentales o con afectación cognitiva importante, como por ejemplo, pacientes con enfermedad de Alzheimer avanzada, pacientes con otras enfermedades degenerativas del SNC, pacientes con patología psiquiátrica no bien controlada, etc.

2. Cuando los datos que van a recoger y medir los observadores van a ser de fácil recogida y no van a requerir una formación importante en temas sanitarios, pero van a ayudar de manera muy importante al profesional sanitario a valorar el control de la enfermedad de forma rutinaria y a tomar decisiones terapéuticas importantes, ya que esta información va a ser totalmente complementaria con la que éstos van a recopilar.
3. Cuando el resultado a evaluar va a ser recogido de manera más fiable por un observado que tenga una relación fluida y continua con el paciente más que por un profesional sanitario, como, por ejemplo, actividades y comportamientos que van a ser distintos día a día, o actividades que se van a producir solo durante algún momento a lo largo del día (en la escuela, mientras duerme, por la mañana a primera hora, etc.).

■ 4.2. FASES DEL DESARROLLO DE INSTRUMENTOS PARA MEDIR RESULTADOS EN SALUD COMUNICADOS POR UN OBSERVADOR

Este tipo de cuestionarios se componen de instrucciones, una serie de preguntas y unas opciones de respuesta, con una explicación sencilla para contestar. Los *ítems* se agrupan, habitualmente, en dimensiones que miden diferentes aspectos de la enfermedad que padece el paciente evaluado. El número y tipo de dimensiones e *ítems* incluidos dependen del propósito del instrumento, el modelo conceptual en el que se ha basado y de algunas consideraciones prácticas, tales como el tiempo disponible para su administración y lo que va a tardar en rellenarlo el observador.

El desarrollo de un instrumento nuevo para medir los resultados comunicados por un observador, comprende cinco fases diferentes, totalmente complementarias entre ellas (Benjamin *et al.*, 2017; Deshpande *et al.*, 2011):

- Identificación de los conceptos a incluir en el cuestionario, que va a constar de las siguientes fases: 1) revisión de la literatura; 2) entrevista a profesionales sanitarios expertos en la enfermedad y a posibles observadores del paciente que padece esta enfermedad (padres, hermanos, cuidadores, maestros, etc.); 3) análisis cualitativo de toda la información recopilada, lo que va a ayudar a generar las dimensiones e *ítems* que tendrá el nuevo instrumento.
- Generación de los *ítems* a incorporar en el cuestionario, para lo que será necesario elaborar un marco conceptual preliminar, así como definir un elen-

co provisional de *ítems* a incorporar en el instrumento, lo que va a permitir disponer al final de una primera versión del cuestionario, donde se habrán incluido todas aquellas dimensiones e *ítems* que se consideren importantes para valorar los resultados que se produzcan.

- Análisis de los *ítems* incorporados en el instrumento con el fin de valorar las posibles duplicidades y repeticiones de conceptos en distintos *ítems*, para lo que será necesario que esta primera versión del cuestionario sea cumplimentada por grupo de observadores, de tal manera que al final solo queden en la versión definitiva del instrumento aquellas preguntas que sean relevantes, claras y que sean útiles para el objetivo final del cuestionario.
- Refinamiento del nuevo instrumento una vez se hayan eliminado aquellos *ítems* que miden cosas similares y son redundantes, y se esté seguro de que encaja con el marco conceptual definido al inicio del proceso; en este paso será necesario que un grupo de clínicos expertos en la enfermedad lo revisen por si hubiera preguntas dudosas y, posteriormente, que un grupo de observadores lo revisen y estén de acuerdo.
- Validación de las propiedades psicométricas del nuevo instrumento, con el fin de estar seguros que mide lo que tiene que medir, para lo que es necesario valorar las propiedades de medición del cuestionario, tales como: 1) viabilidad, que se refiere a que el tiempo empleado en la cumplimentación del mismo no sea excesivo, que el cuestionario es sencillo, entendible y fácil de rellenar y que, además, su corrección e interpretación de los resultados es fácil; 2) fiabilidad, esto es que el cuestionario es preciso y estable, y que sus resultados se pueden reproducir si se realizasen mediciones en condiciones similares; esta propiedad va a indicar si la medición está libre de errores aleatorios, ya que la fiabilidad es una medida de la capacidad del cuestionario para conocer si una variable fluctúa como resultado de un error en la medida o cambio real. Para su valoración, se emplean habitualmente tres métodos diferentes: fiabilidad test-retest (va a explorar la estabilidad temporal o reproducibilidad), fiabilidad intra e inter observador (mide la concordancia de un observador consigo mismo o la concordancia entre 2 observadores independientes) y fiabilidad interna (o consistencia interna), que hace referencia a la estabilidad de las puntuaciones entre los *ítems* del instrumento; 3) validez, que nos dice el grado en que un cuestionario mide aquello para lo que está destinado, por lo que un nuevo instrumento será válido según su capacidad para medir realmente los resultados que percibe el observador. Hay varios tipos de validez, como la validez de contenido (valora si el cuestionario abarca todas las dimensiones de la enfermedad que se quiere medir), la validez de constructo (grado en que una medida particular se relaciona con otra de manera consistente) y la validez de criterio (correlación del instrumento con alguna otra medida de la variable a estudiar que se considera estándar); 4) sensibilidad al cambio, que es la capacidad del cuestionario para detectar cambios en la percepción del observador cuando mejora o empeora la

enfermedad evaluada (o alguno de sus síntomas y signos) después de un tratamiento, o en la evolución de la misma.

■ 4.3. PROPIEDADES IDEALES DE LOS INSTRUMENTOS PARA MEDIR RESULTADOS PERCIBIDOS Y COMUNICADOS POR LOS OBSERVADORES

Las propiedades ideales de un instrumento para medir resultados percibidos y comunicados por los observadores son las siguientes:

- Ser específico del concepto que se quiere medir.
- Estar basado en un modelo de punto final.
- Debería tener equivalencia conceptual.
- Debería estar basado en un modelo conceptual bien definido.
- Debería contener el número óptimo de *ítems* y dimensiones.
- Debería ser fácil de rellenar y no llevar mucho tiempo, así como disponer de propiedades de medida adecuadas.
- Debería tener evidencias apropiadas del modelo conceptual y excelentes propiedades psicométricas.
- Debería mantener la confidencialidad del observador.
- Debería ser reproducible.

■ 4.4. PRINCIPALES CARACTERÍSTICAS DE LOS INSTRUMENTOS EXISTENTES PARA VALORAR LOS RESULTADOS COMUNICADOS POR LOS OBSERVADORES

La mayoría de los cuestionarios existentes para evaluar los resultados comunicados por los observadores se ha desarrollado en enfermedades raras y en la edad pediátrica, con el fin de emplearse en los programas de investigación clínica de estas enfermedades (Geister *et al.*, 2017; Edwards *et al.*, 2018; Stacpoole *et al.*, 2018; Lewis *et al.*, 2018; Pascoal *et al.*, 2018).

En la génesis de todos estos instrumentos, se han seguido las siguientes fases:

- Entrevistas a padres y niños que padecen la enfermedad diana para obtener información sobre la enfermedad y los *ítems* y dimensiones más importantes a tener en cuenta en su evaluación, lo que va a ser clave a la hora de elaborar en contenido del cuestionario.

- Una vez elaborada la primera versión del instrumento, los padres y un grupo de clínicos expertos en la enfermedad (que han participado en el proyecto) revisan esta primera versión para valorar la comprensión y relevancia de los *ítems* y dimensiones incluidas, así como la factibilidad del mismo; de esta manera, se van a identificar *ítems* que presentan un lenguaje problemático que podría generar confusión cuando se contesten por los observadores, lo que va a motivar que se tengan que modificar estos *ítems*.
- Desarrollar un estudio piloto con un grupo de observadores para evaluar las propiedades de medición del nuevo instrumento, lo que va a garantizar que el cuestionario mide lo que tiene que medir y que presenta unas buenas propiedades psicométricas.

■ 4.5. CONCLUSIONES

Hasta ahora, no ha sido habitual medir los resultados en salud percibidos por los observadores de un paciente que padece una enfermedad y que no puede expresar directamente los resultados que el percibe y siente de su enfermedad; por este motivo, habitualmente esta información no se ha tenido en cuenta a la hora de decidir modificaciones terapéuticas en el manejo de su enfermedad. Sin embargo, medir los resultados en salud a través de observadores va a ser clave para que se pueda tener en cuenta las percepciones de los pacientes sobre su enfermedad y los tratamientos recibidos, en aquellas situaciones en que los pacientes no pueden transmitir directamente estos resultados en salud, bien por su edad (niños pequeños) o por padecer problemas cognitivos que impidan que los puedan comunicar adecuadamente, como, por ejemplo, pacientes con enfermedades neurodegenerativas.

Un resultado en salud medido y comunicado por un observador es una medida basada en la observación de alguien que no es un profesional sanitario (puede ser uno de los padres, el marido, el hermano, un cuidador, un maestro, etc.), que normalmente va a observar al paciente de manera regular y a diario, por lo que puede aportar información muy valiosa de como el propio paciente percibe su enfermedad y los tratamientos prescritos (un buen ejemplo es un niño pequeño que cuando tiene dolor es incapaz de transmitirlo, pero un observador puede ver cuando llora, y lo puede transmitir como un resultado derivado del dolor que padece).

Aunque pueda parecer que los resultados en salud comunicados por los observadores puede ser similar a los resultados obtenidos a través de un *proxi*, la diferencia está en que cuando es un *proxi* el que comunica los resultados, este se pone en lugar del paciente y describe los resultados como si él fuera el paciente que sufre la enfermedad (es algo totalmente subjetivo), mientras cuando es un observador el que comunica los resultados, simplemente observa lo que le sucede al paciente y como este se comporta (es una actividad totalmente objetiva, sin ningún componente subjetivo).

Dado que cada vez es más importante disponer de resultados en salud comunicados por los pacientes y, teniendo en cuenta que cada vez va a haber más pacientes que no van a poder transmitir sus resultados percibidos con garantías, en el futuro es de esperar que cada vez se empleen más los resultados en salud obtenidos a través de observadores, en un intento por disponer de datos que representen al paciente.

■ REFERENCIAS

- BENJAMIN, K., VERNON, M. K., PATRICK, D. L. *et al.* (2017). Patient-reported outcomes and observer-reported outcome assessment in rare disease clinical trials. An ISPOR COA emerging good practices task force report. *Value Health*, 20(7), pp. 838-855.
- DESHPANDE, P. R., RAJAN, S., SUDEEPTHI, B. L. *et al.* (2011). Patient-reported outcomes: A new era in clinical research. *Perspect Clin Res*, 2(4), pp. 137-144.
- EDWARDS, T. C., EMERSON, J., GENATOSSIO, A. *et al.* (2018). Initial development, and pilot testing of observer-reported outcomes (ObsROs) for children with cystic fibrosis ages 0-11 years. *J Cyst Fibros*, 17(5), pp. 680-686.
- GEISTER, T. L., BRUSNELL, D. M., YANG, J. *et al.* (2017). Initial psychometric validation of the questionnaire on pain caused by spasticity (QPS). *Health and Quality of Life outcomes*, 15, p. 229.
- LEWIS, S., DEL MURO, C., BLOCK, S. L. *et al.* (2018). Development of a novel observer-reported outcome measure for the assessment of respiratory syncytial virus (RSU) infection symptoms in pediatric clinical trials. *J Patient Rep Outcomes*, 2, p. 9.
- PASCOAL, C., BRASIL, S., FRANCISCO, R. *et al.* (2018). Patient and observer reported outcome measures to evaluate health-related quality of life in inherited metabolic diseases: a scoping review. *Orphanet J Rare Dis*, 13(1), p. 215.
- POWERS, J. H., PATRICK, D. L., WALTON, M. K. *et al.* (2017). Clinician-reported outcome assessment of treatment benefit: report of the ISPOR clinical outcome assessment emerging good practices task force. *Value Health*, 20 (1), pp. 2-14.
- STACPOOLE, P. W., SHUSTER, J., THOMPSON, J. L. P. S. *et al.* (2018). Development of a novel observer reported outcome tool as the primary efficacy outcome measure for a rare disease randomized controlled trial. *Mitochondrion*, 42, pp. 59-63.
- WALTON, M. K., POWERS, J. H., HOBART J. *et al.* (2015). Clinical outcome assessments: conceptual foundation-report of the ISPOR clinical outcomes assessment emerging good practices for outcomes research task force. *Value Health*, 18(6), pp. 741-752.



5

¿CÓMO INCORPORAR LAS MEDIDAS DE RESULTADOS EN SALUD A LA EVALUACIÓN ECONÓMICA?

LAURA VALLEJO-TORRES, MARTA TRAPERO-BERTRAN

■ 5.1. INTRODUCCIÓN

Las medidas de resultado en salud son un elemento clave en la evaluación de la eficiencia de las intervenciones sanitarias. La elección de la medida a emplear depende de la pregunta de investigación y viene determinada por el tipo de evaluación que prevea realizarse. En este capítulo se introducen los principales conceptos para evaluar la eficiencia de intervenciones sanitarias y se intenta delimitar cuáles son los pasos a seguir, y la información que se necesita recoger, cuando se quiere incorporar progresivamente la evaluación económica en el análisis de resultados en salud. Para conseguir este objetivo, el capítulo responde a cinco cuestiones principales. La primera es definir el concepto de eficiencia y de coste de oportunidad y justificar la presencia de estos conceptos en el día a día de las decisiones sanitarias. La segunda cuestión gira en torno a la definición de evaluación económica y los diferentes tipos de evaluaciones económicas que existen. También se exponen otros análisis económicos que nos pueden ser útiles a la hora de informar las decisiones sanitarias. La tercera cuestión intenta abordar qué medidas de costes hay que tener en cuenta para poder realizar este tipo de análisis y se pone de relieve la importancia de la perspectiva del análisis y su influencia en la elección y evaluación de los costes del análisis. La cuarta cuestión responde a qué medidas de resultado en salud debemos incorporar en el análisis y la importancia de estas. La última cuestión intenta explicar cuáles son los pasos a seguir para incorporar el análisis económico en salud, qué medidas de resultado habría que medir y qué costes deberíamos recoger para poder realizar una evaluación económica.

■ 5.2. ¿QUÉ ES LA EFICIENCIA Y EL COSTE DE OPORTUNIDAD?

El mercado sanitario se caracteriza por tener una demanda de servicios sanitarios infinita y una oferta limitada. En muchos sistemas sanitarios, las intervenciones son facilitadas a coste cero o con un gran subsidio para el consumidor en el momento de su consumo, aunque no hay que olvidar que la sociedad paga estas intervenciones a través de sus impuestos generales. Se espera que los pacientes quieran consumir servicios sanitarios siempre que el beneficio de la intervención sea mayor que el coste para ellos. La oferta es limitada debido a las restricciones presupuestarias del sistema. El resultado de esta demanda infinita y esta oferta limi-

tada es un desequilibrio en el mercado sanitario. Por lo tanto, existe la necesidad de equilibrar la oferta con la demanda racionalizando o priorizando los recursos.

Esta necesidad de priorizar se hace visible cuando nos enfrentamos ante la disyuntiva sobre qué tecnologías e innovaciones sanitarias podemos incorporar en un sistema o centro sanitario con los recursos disponibles para su financiación. La innovación sanitaria se puede definir como la actividad dirigida a la generación e implementación de nuevos conocimientos para mejorar la salud de la población (Drucker, 2002). El deseo individual y social sería acceder a toda la innovación sanitaria, sin embargo, los recursos, tanto económicos como clínicos, son limitados y como consecuencia, los gobiernos, los proveedores de salud y las familias no pueden acceder a toda la innovación en tecnologías sanitarias disponibles. A la hora de priorizar, la perspectiva económica y el concepto de eficiencia juegan un papel cada vez más importante. El objetivo del análisis de la *eficiencia* es informar las decisiones sobre financiación sanitaria teniendo en cuenta tanto los beneficios en salud como los costes, de manera que se alcance el máximo beneficio en salud de la población con los recursos disponibles. Para medir la eficiencia se pueden utilizar datos de eficacia o de efectividad, aunque lo ideal es utilizar los segundos ya que son los que obtendremos en la práctica real. Sin embargo, muchas veces se emplearán los primeros, ya que cuando se deben tomar decisiones sobre la inclusión o no de una tecnología o medicamento en el catálogo de prestaciones de un Sistema Nacional de Salud y determinar el precio al que se va a financiar solo se dispone de resultados obtenidos de ensayos clínicos y, por tanto, solo se conoce la eficacia. En cualquier caso, la eficiencia tiene en cuenta los resultados obtenidos, en condiciones ideales o en condiciones habituales, y el consumo de recursos empleados para su consecución.

La asignación eficiente de recursos limitados entre intervenciones alternativas no puede, bajo ningún pretexto, ignorar la consideración del *coste de oportunidad*, que implica tener en cuenta la restricción presupuestaria o los recursos económicos disponibles, es decir, los beneficios que no se van a obtener a causa del desplazamiento de recursos hacia la intervención seleccionada o recomendada. El coste de una determinada elección no es únicamente el dinero que gastamos al elegir esa opción, sino el sacrificio que nos supone tal decisión, tanto si implica una transacción monetaria como si no. Por lo tanto, el coste de oportunidad de los recursos no es necesariamente lo que se pagó por ese recurso sino el valor en su mejor uso alternativo. Por ejemplo, si un gerente de hospital dispone de 200.000 euros y se encuentra ante la posibilidad de comprar una nueva tomografía computarizada durante un año o comprar material para el laboratorio, el coste de oportunidad de elegir comprar una nueva tecnología para diagnóstico con imágenes se define como el número de pruebas en el laboratorio que dejarán de hacerse al no haber adquirido este material.

La Economía de la Salud tiene como fin último mejorar la asignación de los recursos destinados a la salud, y ofrece una serie de herramientas para incorporar

el análisis de la eficiencia y la consideración del coste de oportunidad. Entre ellas destacamos en este capítulo la evaluación económica en salud, que de manera creciente se emplea como un criterio de priorización en la toma de decisiones sobre financiación sanitaria.

■ 5.3. EVALUACIÓN ECONÓMICA, TIPOS DE EVALUACIÓN ECONÓMICA Y OTROS TIPOS DE ANÁLISIS ECONÓMICOS

La evaluación económica (EE) de tecnologías sanitarias es un análisis comparativo de alternativas en términos de costes y resultados en salud (Drummond *et al.*, 2015). Entendemos tecnología sanitaria como el conjunto de medicamentos, dispositivos y procedimientos médicos o quirúrgicos usados en la atención sanitaria, así como los sistemas organizativos y de soporte dentro de los cuales se propicia dicha atención (Office of Health Technology Assessment, 1998). La EE debe cumplir con dos requisitos básicos: analizar tanto los costes como los resultados en salud y comparar dos o más alternativas. Por tanto, para que un estudio pueda ser considerado una evaluación económica completa debe incorporar tanto el análisis de los costes como el de resultados en salud, y comparar estos con los que se obtienen con cursos de tratamiento alternativos, incluyendo, en caso de ser relevante, la práctica o tecnología habitual.

■ 5.3.1. Tipos de evaluación económica

Existen diferentes tipos de evaluaciones económicas, siendo las más importantes el análisis de coste-efectividad, coste-utilidad y coste-beneficio. Los costes en cada uno de estos tipos de análisis se miden en unidades monetarias. En el contexto español lo haríamos en euros. La distinción entre estos tipos de estudios radica en la medida que se emplea para cuantificar los resultados en salud. Es importante, sin embargo, señalar que el concepto de análisis de coste-efectividad se emplea a menudo como sinónimo de evaluación económica, independientemente del tipo de medida empleada para cuantificar los resultados en salud. A continuación, ofrecemos una descripción de cada uno de los diferentes análisis.

Análisis coste-efectividad (ACE)

Un *análisis coste-efectividad (ACE)* es una evaluación económica que valora los resultados de salud en unidades físicas o naturales con una dimensión clínica (por ejemplo, tasa de colesterol en sangre, número de personas correctamente diagnosticadas, muertes evitadas). Cualquier unidad clínica puede ser empleada para medir los resultados en salud en este tipo de análisis, usándose frecuentemente la medida más estrechamente relacionada con el objetivo de la intervención. Los resultados de un ACE se expresan en términos de la razón o *ratio de coste-efectividad incremental (RCEI)*, que mide el coste por unidad de resultados asociado a una

intervención en comparación con la alternativa bajo estudio, por ejemplo, coste por mg/dl de reducción de colesterol en sangre, o coste por caso detectado.

La RCEI se calcula de la siguiente manera:

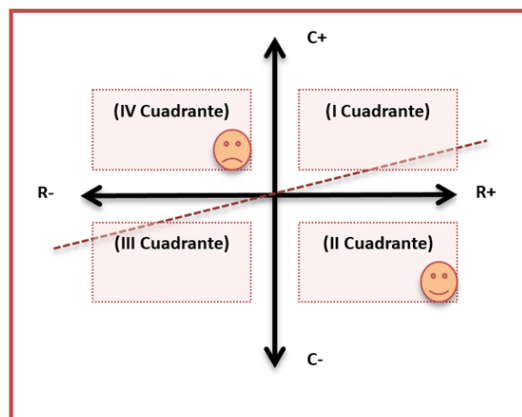
$$RCEI = \frac{C_A - C_B}{E_A - E_B}$$

donde C_A son los costes de la intervención A, C_B los costes de la intervención B, E_A los resultados (efectividad) de la intervención A y E_B los resultados de la intervención B. Por ejemplo, si se trataran de intervenciones destinadas a la detección temprana de una patología, siendo A una tecnología innovadora y B la práctica habitual con unos costes asociados de 30.000 euros y 20.000 euros, respectivamente, y un número de casos detectados de 20 y 10 casos, para la tecnología A y B, respectivamente, concluiríamos que la intervención A tiene un RCEI de 1.000 euros por caso detectado adicional en comparación con la práctica habitual.

Este ejemplo describe un caso muy común en la evaluación de innovaciones sanitarias, donde la nueva tecnología ofrece un beneficio adicional en salud, pero a un coste más elevado. Esta no es, sin embargo, la única posibilidad. La Figura 1 muestra las cuatro posibles combinaciones en relación con los costes y beneficios incrementales que podemos encontrar cuando se comparan dos intervenciones, independientemente de como midamos el resultado en salud. Este plano lo utilizamos tanto en el análisis coste-efectividad como en el coste-utilidad.

Figura 1.

PLANO DE COSTE-EFECTIVIDAD INCREMENTAL, QUE COMPARA LOS COSTES Y LOS RESULTADOS EN SALUD DE LAS DIFERENTES INTERVENCIONES



Fuente: Elaboración propia.

El eje vertical indica si la nueva intervención tiene un coste superior o inferior, mientras que el eje horizontal indica si el resultado en salud es mejor o peor, en comparación, en ambos casos, con la alternativa de tratamiento bajo estudio. El segundo y cuarto cuadrante de la gráfica muestran dos circunstancias ante las cuales la decisión es sencilla. Si existe una intervención más barata y que ofrece mejores resultados en salud que su alternativa, lo que llamamos alternativa dominante, esta debe ser la intervención incorporada al sistema o centro sanitario. Las dificultades se plantean cuando nos encontramos en el primer cuadrante (más costoso, mejor resultado) o en el tercer cuadrante (más barato, peor resultado). En ambos casos, la conclusión sobre qué intervención es coste-efectiva dependerá de la cantidad que se considere apropiada invertir por unidad de resultado en salud ganado. Este valor se denomina el *umbral de coste-efectividad* (línea discontinua en la Figura 1), pero no está definido para cada medida de resultado en salud posible. Por ello, una desventaja de los ACE es la complejidad que entraña a la hora de decidir si el tratamiento bajo estudio es considerado o no coste-efectivo en el marco de un sistema sanitario, especialmente cuando se deben comparar intervenciones que emplean métricas de efectividad diferentes para medir los resultados en salud. Esta es precisamente la principal limitación de los ACE: al trabajar con medidas de efectividad normalmente específicas al problema de salud bajo estudio, sus resultados no son comparables entre diferentes condiciones de salud o patologías. Estos problemas se resuelven parcialmente empleando una modalidad de evaluación económica denominada el análisis coste-utilidad (ACU).

Análisis coste-utilidad (ACU)

Los *análisis coste-utilidad (ACU)* son una forma de evaluación económica en la que la medida de resultados en salud contiene, no únicamente la dimensión clínica, sino que incorpora además las preferencias y los valores de la población en su medición. En lenguaje económico se emplea el término *utilidades* para definir los valores que representan las preferencias de la sociedad por diferentes estados de salud y que son empleadas en el ACU. Los Años de Vida Ajustado por Calidad (AVAC) son la unidad de resultado más comúnmente empleada en el ACU, aunque existen algunas variantes. Por ejemplo, los Años de Vida Ajustado por Discapacidad (AVAD) se utilizan para calcular los años de vida perdidos por la enfermedad y los años vividos con discapacidad para un gran número de enfermedades, especialmente en países no industrializados. Existen otras medidas de resultado que pueden ser consideradas para realizar un ACU, aunque su uso no está extendido.

Las características principales de los AVAC es que ofrecen una medida genérica que incorpora en una única métrica dos dimensiones cruciales sobre la salud, tiempo de vida y calidad de vida, esta última medida de acuerdo con las preferencias de la población. El resultado final de un ACU se expresa en términos de la ratio coste-utilidad incremental (RCUI), como el coste por AVAC ganado de una tecnología frente a su comparador. La clara ventaja de utilizar los AVAC como unidad de medida de resultados en salud es que permite interpretar los resultados y compararlos con otros estudios, independientemente de la patología o condición de salud

de la que traten, ya que la mayor parte de condiciones de salud pueden expresarse con esta unidad. Además, se dispone de estimaciones del valor umbral de coste-efectividad que indicarían el coste por AVAC ganado por debajo del cual la tecnología bajo estudio se consideraría eficiente en un contexto concreto. Específicamente para España ese valor ha sido estimado en torno a los 20.000-25.000 euros/AVAC (Vallejo-Torres, García Lorenzo y Serrano Aguilar, 2018).

Análisis coste-beneficio (ACB)

Otra modalidad de la evaluación económica es el *análisis coste-beneficio (ACB)* que mide los resultados en salud de cada alternativa en unidades monetarias, lo que permite una comparación directa entre los costes y los resultados en salud. El resultado de un ACB se expresa como la diferencia entre los beneficios en salud, medidos monetariamente, y los costes; lo que se conoce como el beneficio neto.

Para cuantificar monetariamente las mejoras en salud, este tipo de estudios emplean diversos métodos. Entre ellos destacan: i) la teoría del capital humano, enfocada en medir las ganancias en productividad derivadas de mejoras en salud; ii) las encuestas sobre disposición a pagar, destinadas a captar el valor monetario que los individuos le otorgan a la intervención o mejora en salud bajo estudio; y iii) la observación de la actitud de los individuos en el mercado, que se emplea como medio de obtener las preferencias reveladas de la población. Cada uno de estos métodos presenta una serie de dificultades metodológicas, lo que limita su uso en la práctica pese a la ventaja teórica de ofrecer un análisis de comparación de costes y resultados en salud expresados en la misma unidad monetaria.

■ 5.3.2. Otros tipos de análisis económicos en salud

Otras modalidades de evaluación económica existentes son el análisis de minimización de costes (AMC) y el análisis coste-consecuencia (ACC).

Análisis de minimización costes (AMC)

El *análisis de minimización de costes* es un tipo de evaluación económica en la que los resultados en salud de las diferentes alternativas evaluadas son esencialmente idénticos. En estos casos no se incluye el análisis de los beneficios en salud en la evaluación, al haber sido demostrado previamente que no existen diferencias. Por lo tanto, la decisión dependerá únicamente de la diferencia en costes. Este tipo de análisis no es el más frecuente en el área de salud, y únicamente podría verse justificado, en el marco de la evaluación económica, si existe evidencia clara de la equivalencia en resultados en salud de los tratamientos analizados.

Análisis coste-consecuencia (ACC)

El *análisis coste-consecuencia (ACC)* realiza las comparaciones de costes y de resultados entre alternativas de tratamiento, pero no calcula la RCEI, sino

presenta de una manera desagregada los costes y los resultados de cada alternativa. Normalmente se realiza cuando existen varias medidas de efectividad con una importancia similar entre ellas, por lo que la elección de una única medida para expresar el resultado en salud no es clara. La principal limitación de este tipo de estudios es que no ofrece la posibilidad de definir el criterio de eficiencia que otorga la medida de la RCEI, necesaria para concluir acerca del coste-efectividad de una intervención sanitaria.

Existen otros tipos de análisis económicos que no se consideran dentro del marco de la evaluación económica al no cumplir con los dos requisitos que estas deben seguir, es decir, cuantificar tanto los costes como los resultados en salud y comparar al menos dos alternativas. Por ejemplo, algunos estudios realizan una *descripción de costes y resultados de una intervención*, sin comparador. En ocasiones esto se debe a que el diseño del estudio no incorporaba a un grupo control que pudiera ofrecer el comparador frente al cual analizar la intervención evaluada. En otros estudios se realiza un *análisis de costes incrementales* de una intervención comparado con su alternativa, pero sin analizar, ni justificar la equivalencia, de los resultados en salud. Estos tipos de estudios también tienen su valor informativo, pero es muy importante resaltar que no se tratan de evaluaciones económicas completas.

Finalmente, definimos dos tipos de análisis económicos que a menudo complementan a/o sirven de fuente de información para, las evaluaciones económicas. Estos son el análisis de impacto presupuestario y el estudio de la carga de la enfermedad.

En la práctica, los decisores necesitan no solamente identificar las intervenciones más eficientes sino además conocer el impacto que puedan tener en el presupuesto a corto y medio plazo las decisiones de financiar una nueva tecnología sanitaria o programa de salud. El *análisis de impacto presupuestario (AIP)* se ha definido como la estimación del impacto de un nuevo tratamiento o diagnóstico en los costes anuales, beneficio en salud anual y otros resultados de interés para los primeros y subsiguientes años después de la introducción de la nueva intervención en un sistema de salud (Mauskopf *et al.*, 2007).

Por tanto, a diferencia de la evaluación económica, que tiene en cuenta tanto los costes como los beneficios en salud incrementales del individuo medio en un horizonte temporal que a menudo abarca el largo plazo, el AIP desglosa año a año los resultados aplicándolos al conjunto de la población que se beneficia de la intervención y centrándose en el corto/medio plazo.

Un *estudio de carga económica de enfermedad* intenta identificar, medir y evaluar los recursos sanitarios y no sanitarios y las pérdidas de productividad que resultan de una muerte prematura, morbilidad, discapacidad o lesión debido a la enfermedad correspondiente y/o sus comorbilidades (Drummond *et al.*, 2015). Esta información puede ser útil como base para la evaluación económica o para ayudar

a los tomadores de decisiones a priorizar las políticas e intervenciones de salud cuando se considera la eficiencia como criterio para informar las decisiones.

En este apartado hemos definido el concepto de evaluación económica y los tipos de estudios que cumplen con los dos requisitos que estas exigen: el análisis tanto de los costes como de los resultados en salud de dos o más tecnologías alternativas. Señalamos que los tipos de EE se definen de acuerdo con la medida que se emplea para medir los resultados en salud, y que cada uno de estos estudios ofrece determinadas ventajas e inconvenientes, siendo el análisis coste-utilidad el estudio que facilita el análisis de la eficiencia en sentido más amplio. Además, el análisis de impacto presupuestario y el estudio de la carga de la enfermedad son estudios que a menudo complementan o informan a la EE en salud. A la hora de decidir el tipo de estudio económico a realizar en un determinado contexto lo más importante es definir el objetivo o la pregunta que queramos responder, para poder así identificar el análisis más apropiado. Es, además, crucial conocer los requisitos en términos de la información necesaria para llevarlos a cabo. A continuación, detallamos la información sobre costes y resultados en salud necesaria para la realización de una EE en salud.

■ 5.4. ¿QUÉ PARÁMETROS DE COSTES NECESITO MEDIR?

La determinación de los recursos empleados en los tratamientos o intervenciones evaluadas y la forma de medición de ese uso de recursos es un elemento esencial para realizar una evaluación económica. Es imprescindible identificar los costes relevantes para la evaluación, así como la medición, y valoración de estos (Trapero-Bertran, 2015; Oliva Moreno *et al.*, 2018). La inclusión en el estudio de unos u otros recursos dependerá del objetivo y de la perspectiva adoptada, teniendo en cuenta los recursos (humanos, económicos, técnicos y de tiempo) disponibles. Aunque existe una terminología heterogénea entre los diferentes países para poder clasificar los tipos de coste, a grandes rasgos podemos encontrar¹: *costes directos (sanitarios y no sanitarios)*, *costes indirectos* o *pérdidas de productividad*, y *costes intangibles*.

Los costes directos sanitarios se refieren a los relacionados con el manejo de la enfermedad y de sus complicaciones, es decir, aquellas partidas comunes de gasto sanitario que tratan de reducir en lo posible el impacto en la salud de las personas que sufren una enfermedad. Estas partidas quedan registradas contablemente y se reflejan en las Cuentas Nacionales. Algunos ejemplos serían los costes de hospitalización, las pruebas diagnósticas, intervenciones, medicamentos, tratamiento de los efectos adversos, fisioterapia, etc. También podemos encontrar otros costes que no se reflejan directamente en los sistemas contables. Estos serían los

¹ Otra manera de clasificar los costes sería separar los costes sociosanitarios formales, que comprenderían los costes sanitarios y los cuidados formales, y por otro lado los costes sociales, que comprenderían los cuidados informales, pérdidas laborales y costes intangibles.

costes directos no sanitarios. Estos costes se asocian a las consecuencias generadas por la enfermedad y, dependiendo de la perspectiva, podrán incluir los costes de los cuidados personales formales e informales, la adaptación de la vivienda o los desplazamientos al centro sanitario. Es importante que estos costes se identifiquen de manera detallada a través de registros o encuestas diseñadas con ese fin. Estos podrían contemplar conceptos como los costes de los servicios sociales (financiados pública o privadamente) o, por ejemplo, los costes de los cuidados prestados por personas no profesionales (llamados también cuidados informales) surgidos a consecuencia de una enfermedad.

Los costes indirectos o pérdidas de productividad laboral se relacionan con una reducción de esta a consecuencia de un problema de salud. Este coste debería ser valorado ya que puede ocasionar un fuerte impacto en la situación laboral de las personas que la soportan (relacionados con cambios en la capacidad productiva del individuo). La productividad laboral es fuente de riqueza no solo para una persona sino para toda la sociedad, por tanto, la valoración de estos costes puede ser muy relevante según la patología que se analice. Más adelante se detallan las diferentes técnicas que existen para valorar la pérdida de productividad del paciente debido a la enfermedad.

Por último, también existen los costes intangibles que se definen como los costes relacionados con el dolor o sufrimiento de los pacientes. Generalmente no suelen incluirse en los estudios de evaluación económica porque son difícil de cuantificar, aunque se recomienda identificarlos cuando se considere que son relevantes (Drummond *et al.*, 2015).

Cuando hacemos una evaluación económica, en la mayoría de los casos, nos preguntamos qué costes deberíamos incluir en el análisis. La verdad es que no existe un manual intra o interpaíses que nos ayude a tomar la decisión de qué coste o qué costes deberíamos incluir. Sin embargo, la decisión en la elección de costes irá guiada por la perspectiva del análisis. La perspectiva del análisis es el punto de vista desde el que se realiza tanto la valoración de los costes como la de los beneficios en salud y determina cuáles son los recursos que deben incluirse y excluirse del análisis. Esta perspectiva se define según el agente social que utilizará los resultados del estudio para tomar las decisiones. Dependiendo de la perspectiva elegida, habrá que incluir en el estudio un tipo u otro de costes y de resultados, con lo que el análisis dependerá estrechamente de la perspectiva escogida. Las perspectivas pueden ser la *social*, la del *financiador público*, *financiador sanitario*, *proveedor* y la del *paciente*.

La *perspectiva social* es aquella que recoge todos los resultados relevantes desde la perspectiva de la sociedad. Abarca, por tanto, todos los beneficios resultantes de la aplicación de una intervención (frente a sus alternativas) y los costes asociados con las mismas, con independencia de quienes reciban los beneficios y quienes soporten los costes. Se recomienda incluir esta perspectiva en todas las evaluaciones junto a la perspectiva del financiador público, dado que este último

es uno de los principales destinatarios de los análisis de EE. Es la perspectiva más general y la que más información aporta sobre la importancia relativa y absoluta de una enfermedad. Por tanto, debería ser la de mayor ayuda para la correcta priorización de los recursos sanitarios y sociales. Por otra parte, la *perspectiva del financiador*, público o sanitario se refiere a todos los beneficios terapéuticos resultantes de la aplicación de una intervención (frente a sus alternativas) y los costes en los que incurre el financiador al poner a disposición del ciudadano el tratamiento requerido. La *perspectiva del proveedor* es aquella que recoge todos los costes y resultados relevantes en una evaluación económica realizada en un entorno de un agente que es proveedor del sistema sanitario de salud, como, por ejemplo, un hospital. En cambio, la *perspectiva del paciente* deberá contemplar todos los costes y resultados en salud que recaen directamente en el paciente. Esta perspectiva incluirá, por ejemplo, gastos de transporte y manutención para que el paciente pueda ir a recibir un tratamiento al hospital, o los gastos de pagar a una persona para que cuide del paciente en casa en caso de que la enfermedad genere discapacidad.

Cada perspectiva puede tener diferentes necesidades de información y distintos criterios para valorar costes y resultados. Por este motivo, es muy importante definir desde el principio la perspectiva de la evaluación económica, ya que el tipo de costes a cuantificar, la manera de evaluar y agregar costes y resultados dependerá de la perspectiva de la evaluación. Muchos estudios se realizan en el contexto hospitalario, pero en la mayoría de las ocasiones se aconseja que la evaluación económica tome una perspectiva social. En tal caso, se deben considerar de antemano todos los aspectos que no son recogidos de forma rutinaria durante la elaboración del estudio, ya que recaen fuera del sector hospitalario. Para ello, se deben diseñar métodos de recogida de información alternativos como cuestionarios para los pacientes y familiares, con el fin de recoger información sobre los costes directos y costes de productividad que recaen sobre los pacientes y familiares, así como el uso de otros servicios sanitarios, más allá de la estancia hospitalaria y los reingresos, como son el número de visitas al médico de atención primaria y los productos farmacéuticos consumidos. En ciertas enfermedades que generan discapacidad de carácter crónico, como la esclerosis múltiple, es recomendable considerar la perspectiva conjunta del financiador y la perspectiva social. Si en el análisis se utilizara únicamente la perspectiva del financiador sanitario, se debería justificar que los costes sociales (fundamentalmente, los costes de servicios sociales, cuidados informales y las pérdidas laborales) no tienen un peso relevante o no se dispone de información razonablemente fiable sobre los mismos, declarando las fuentes consultadas.

Tras la identificación de los tipos de los recursos que se incluirán en el análisis de acuerdo con la perspectiva del estudio, y su medición en términos de unidades naturales del uso que se hace de los mismos (por ejemplo, el número de pruebas diagnósticas, los minutos de quirófano, o las horas de cuidados informales recibidos), la última etapa consiste en la valoración de dichos recursos, aplicándoles un

coste unitario en términos monetarios. Los métodos y fuentes de información más comunes empleadas para obtener esta información son detallados a continuación.

■ 5.4.1. Costes directos sanitarios

En un contexto de servicios sanitarios financiados públicamente, parece adecuado para la valoración de los recursos recurrir al concepto de coste de producción del servicio analizado. Para ello se puede recurrir a publicaciones oficiales que contengan esta información, datos de la contabilidad analítica de los centros y a tarifas aplicadas a los contratos de prestaciones de servicios de las diferentes comunidades autónomas. Por ejemplo, en la valoración de los medicamentos de uso hospitalario, el fármaco debe ser valorado mediante el precio de adquisición del hospital, si bien también deben ser contemplados los costes de gestión, almacenamiento y dispensación hospitalarios cuando sea el caso. Si los medicamentos fueran prescritos en oficinas de farmacia, el precio a considerar debe ser el precio efectivamente pagado por el financiador del sistema sanitario. En caso de que se identifiquen servicios sanitarios financiados privadamente por el usuario o la familia, el coste a considerar sería el precio de mercado pagado por los servicios recibidos.

■ 5.4.2. Costes directos no sanitarios

Los costes directos no sanitarios son aquellos que los sistemas contables no computan directamente, pero que están asociados a las consecuencias generadas por la enfermedad. Por lo tanto, se refieren a conceptos muy relacionados con la discapacidad, como pueden ser los cuidados a las personas con falta de autonomía personal, los servicios sociosanitarios, la adaptación de infraestructuras en la vivienda o los subsidios por incapacidad laboral. Asimismo, el transporte no medicalizado al centro sanitario también se incluiría aquí, aunque en la práctica este tipo de costes suele obviarse ya que resulta complicado de medir y su magnitud sobre el coste total es pequeña. Dentro de esta partida existen varias técnicas que pueden ser aplicadas para la *valoración del tiempo prestado de cuidado informal*, que es aquel prestado por una persona no profesional, generalmente la pareja o un familiar. Estas se pueden clasificar en dos grandes grupos: (a) métodos de preferencias reveladas y (b) métodos de preferencias declaradas. En este punto, es importante insistir en que es tan importante la transparencia con el método de valoración elegido como con el método de medición del tiempo de cuidado, a través de encuestas apropiadas para ello.

Dentro del primer grupo, preferencias reveladas, se encuentran las dos técnicas tradicionalmente empleadas para valorar los cuidados informales: (a1) método del coste de oportunidad, que valora el tiempo en función de aquella actividad-laboral, doméstica, ocio a la que se renuncia para prestar el cuidado y, (a2) el método

del coste de sustitución o reemplazo, el cual valora los servicios prestados por el cuidador informal teniendo en cuenta que si éste no estuviera disponible, habría que contratar a un cuidador profesional.

El segundo grupo, preferencias declaradas, comprende varios métodos, entre los cuáles los más comunes son: (b1) métodos de valoración contingente, los cuáles estiman la disposición a pagar o la disposición a aceptar/ser compensado por un programa que suponga cambios en el número de horas de cuidados o de su composición entre diferentes tareas para la persona cuidadora); y, (b2) experimentos de elección discreta, los cuáles estiman la disposición a pagar por una hora adicional de cuidados a partir de una serie de elecciones entre dos escenarios hipotéticos.

■ 5.4.3. Pérdidas de productividad

Las pérdidas laborales se definen como el tiempo que un individuo deja de realizar una tarea laboral remunerada, a consecuencia de una enfermedad o lesión. Los aspectos que se deberían considerar en la valoración de estas son la productividad laboral perdida por los pacientes, ya que la productividad laboral perdida por los cuidadores normalmente se mide como cuidado informal (coste directo no sanitario), la productividad doméstica perdida por los pacientes y los cuidadores, y el tiempo de ocio perdido por paciente y por los cuidadores. Sin embargo, ni la productividad doméstica ni el tiempo de ocio se acaban incluyendo en la valoración de los costes debido a la complicación que supone su valoración, con lo que normalmente se mide y valora la pérdida de productividad de los pacientes. Existen dos métodos para la *valoración de las pérdidas laborales*: (c) enfoque de capital humano, que valora el periodo de empleo perdido monetariamente según el salario de ese individuo; y (d) enfoque del coste de fricción, el cuál valora el tiempo de pérdida de productividad de la persona enferma o incluso el trabajo extra llevado a cabo por otros compañeros para cubrir esas tareas según el coste que se origina por pagar a una tercera persona que reemplace la persona enferma o lesionada.

En las guías internacionales de EE no existe consenso sobre el método de valoración de las pérdidas de productividad. En el caso del tiempo de trabajo remunerado, en algunas guías se recomienda la utilización del método de capital humano. En otras, se prefiere el método de los costes de fricción. Por tanto, será labor de los autores del trabajo el justificar, de manera transparente, el tipo de método empleado para la valoración tanto de las pérdidas laborales del paciente como de los cuidados informales por parte de los cuidadores.

También cabe mencionar que la afectación del tiempo de trabajo doméstico y su productividad y las de ocio son también elementos susceptibles de ser estimados como pérdidas sociales puesto que se refieren a recursos sociales modificados a consecuencia de un problema de salud. La valoración del trabajo doméstico y del ocio de una persona que padece una enfermedad se puede aproximar a través

de los "precios sombra" adecuados. Sin embargo, este tipo de valoración es infrecuente en la literatura debido a la falta de información y bases de datos adecuadas, pero desde el punto de vista teórico la afectación total del tiempo y de su uso y disfrute por parte de una persona a consecuencia de un problema de salud se debería valorar.

■ 5.5. ¿QUÉ VARIABLES DE RESULTADO NECESITO RECOGER?

Una de las cuestiones más relevantes cuando se decide hacer una evaluación económica es la elección de las medidas de resultado en salud, ya que estas son en las que basaremos la decisión de recomendar financiar una tecnología sanitaria o no. Como señalamos al definir los tipos de EE, existen tres medidas principales de resultados en salud: las de dimensión clínica (ej. mortalidad y morbilidad, años de vida ganados), las que incorporan la valoración de las preferencias de los pacientes (siendo los Años de Vida Ajustado por Calidad la medida más comúnmente empleada), y las que miden el impacto en salud en términos monetarios (como la disposición a pagar por las mejoras en la salud de un servicio o de un cambio en el estado de salud).

La forma más simple de medir los resultados en salud en un ensayo clínico es utilizando unidades naturales, aunque dificulta las comparaciones entre diferentes condiciones en salud. Esto se soluciona parcialmente cuando incorporamos la mortalidad como medida de resultado (años de vida ganados), si bien no deja de ser un recurso insatisfactorio ante intervenciones que afectan principalmente a la dimensión de la calidad de vida. Por este motivo, existen las medidas de resultado que incorporan la calidad de vida relacionada con la salud (CVRS), que van más allá de la medida clínica y la mortalidad. En concreto, el AVAC, índice cardinal que combina los años de vida con la utilidad o bienestar derivado de la calidad con que estos se disfrutan, es la medida de resultado recomendada en buena parte de las guías de EE de tecnologías sanitarias. También existen las medidas de resultado basadas en el paciente, definidas como medidas a corto plazo que comprenden instrumentos como cuestionarios, entrevistas y otros métodos relacionados con la medida de la salud, la enfermedad y los beneficios de las intervenciones en salud desde la perspectiva del paciente.

Como señalábamos, en el entorno clínico, las medidas más habituales para valorar el beneficio en salud han sido las biomédicas y clínicas o naturales. En el caso de patologías respiratorias podríamos hablar, por ejemplo, del número de exacerbaciones o el Volumen Espiratorio Forzado en el primer segundo (FEV1). La principal ventaja de estas medidas es que se centran en el aspecto clínico más relevante objeto de la intervención. Estas unidades son las que se utilizarían como medidas de resultado en un *análisis coste-efectividad*.

No obstante, la evolución epidemiológica de la población, con un crecimiento en la prevalencia de enfermedades crónicas y un mayor número de personas con limitaciones en su autonomía, ha favorecido el desarrollo de otro tipo de indicadores complementarios a los anteriores, donde no se tiene en cuenta únicamente la muerte prematura o de eventos agudos ligados a la enfermedad considerada, sino también los resultados en salud percibidos por el paciente. Las medidas de morbilidad solamente indican la presencia de enfermedad, pero no nos dicen nada acerca de la duración o el impacto en la vida de una persona. Por este motivo aparecen medidas de morbilidad que intentan incorporar la calidad de vida de las personas. Estas medidas tienen como objetivo reflejar cómo la salud de una persona afecta su capacidad de realizar actividades cotidianas según su opinión subjetiva. Este concepto equivale a lo que llamamos Calidad de Vida Relacionada con la Salud definida como la “percepción subjetiva que posee el paciente del impacto de su enfermedad y su tratamiento en su vida diaria, funcionamiento físico, psicológico y social, así como en su bienestar”. Este tipo de medidas son las que utilizamos en un *análisis coste-utilidad*.

La CVRS suele medirse mediante *instrumentos específicos* para una enfermedad o con *medidas genéricas* del estado de salud. Aunque en el entorno clínico es más común utilizar cuestionarios específicos al problema de salud bajo estudio, los cuestionarios genéricos son los que se utilizan para medir los resultados en salud en evaluaciones económicas, principalmente porque estos garantizan que los resultados sean comparables entre diferentes patologías o condiciones de salud. Si medimos los resultados en salud con cuestionarios específicos esas medidas solo nos servirán para comparar intervenciones de esa misma patología, pero no para comparar el resultado en salud que aportan diferentes intervenciones para diferentes condiciones de salud. Las medidas genéricas han sido diseñadas para ser aplicables ampliamente en los diferentes tipos y grados de severidad de la enfermedad, en los diferentes tratamientos médicos o intervenciones en salud, y en diferentes grupos de población. Ambos instrumentos, específicos y genéricos, utilizan cuestionarios para describir la calidad de vida de los pacientes. Preferentemente, la descripción de la CVRS se realizará mediante la auto clasificación del propio paciente, pero cuando esta no sea posible por causas debidamente justificadas podrá obtenerse de los juicios formulados por otras personas (por ejemplo, cuidadores o personal sanitario).

El siguiente paso para que el instrumento de CVRS incorpore información sobre las preferencias de los individuos es asignarle un valor numérico que refleje el grado de satisfacción del paciente con el estado de salud descrito, es decir la utilidad asociada a ese estado. Dicha utilidad es una valoración cardinal entre el valor 0 y 1, siendo cero la muerte y 1 el estado con la mejor salud posible². Existen diferentes técnicas que se emplean para la obtención de estas utilidades: el síndrome de la lotería, el intercambio de tiempos o personas, el método de lotería estándar,

² En algunos casos, se consideran valores negativos que representarían estados de salud peores que la muerte.

o métodos introspectivos, como la escala visual analógica. Hoy en día existe un amplio consenso metodológico en favor de los métodos basados en el intercambio de tiempos o personas ya que parecen más consistentes.

Los cuestionarios de medidas genéricas más recomendados para una evaluación económica son el EQ-5D y el SF-6D (SF-36). La elección entre uno u otro depende de cuál de los dos sea más idóneo para reflejar la CVRS del grupo de pacientes relevante. Dicha idoneidad depende básicamente de las propiedades psicométricas (validez, fiabilidad, factibilidad) reveladas por cada instrumento en dicho grupo de pacientes. Existen además otros instrumentos genéricos validados en España, por ejemplo, el HUI 3 (Feeny *et al.*, 1995). La principal ventaja de estos instrumentos es que en muchos países se han estimado las utilidades correspondientes para cada uno de los estados de salud definidos por estos cuestionarios empleando muestras representativas de la población. A estos valores predefinidos se les conoce como tarifas, y en el caso de España están disponibles para el EQ-5D, en su versión 3L (Badia *et al.*, 1999) y 5L (Ramos-Goñi *et al.*, 2018) y el SF-6D(SF-36) (Méndez *et al.*, 2011).

En el análisis coste-utilidad no se podrán utilizar directamente los datos de las utilidades obtenidas con estos cuestionarios, sino que se deberán calcular los AVAC generados por una intervención. Estos se calculan multiplicando el número de Años de Vida Ganados (AVG) por la 'utilidad' asociada a la CVRS durante cada periodo. Es recomendable siempre mostrar por separado la variación en AVG y en CVRS, indicando claramente cómo se han combinado ambos cambios en forma de los AVAC ganados.

Como resumen de este apartado enfatizamos que la elección de la medida de resultados en salud es una decisión crucial para la evaluación económica. En muchas ocasiones será recomendable presentar diferentes medidas de resultados en salud, desde medidas clínicas estrechamente relacionadas con la intervención bajo estudio (si se trata de variables intermedias, se mantienen como tales, justificando su relación con el posible efecto final sobre la salud), hasta medidas más ampliamente recomendadas y empleadas en la EE, como son los AVAC, que permiten la comparación entre diferentes patologías e incorporan las preferencias de la población.

■ 5.6. CÓMO INCORPORAR EL ANÁLISIS ECONÓMICO EN SALUD

La importancia creciente de la presentación de evidencia, no solo en términos de seguridad y efectividad, sino también en términos de eficiencia y coste-efectividad, ha impulsado que muchos estudios incorporen en su diseño la valoración de resultados económicos. Para ello, existe un grupo de parámetros que deben ser valorados y medidos de manera continua y siguiendo una metodología apropiada. De esta forma se podrá disponer de datos de eficiencia que complementen a los datos de resultados en salud conseguidos por las terapias empleadas. La medición

de estas variables económicas tiene la misma relevancia que la medición de resultados en salud, por lo que habría que poner el mismo énfasis y esfuerzo.

En cuanto a las medidas de resultados en salud más propicias para la evaluación económica, hemos resaltado el papel de los AVAC. Para hacer factible la medición continuada y rutinaria de la efectividad de las intervenciones sanitarias en términos de AVAC, sería recomendable sistematizar la implementación de cuestionarios genéricos sobre CVRS a los pacientes, tales como el EQ-5D-5L. Esta práctica ya existe en algunos centros de España, independientemente del tamaño del centro, donde se está apostando por la evaluación continuada de los resultados en salud obtenidos por las diferentes tecnologías sanitarias existentes en un mismo centro. Esta práctica ayuda a generar evidencia del resultado en salud real que aportan estas tecnologías, y a generar la costumbre de crear indicadores y medidas para evaluar continuamente el beneficio en salud que aportan las diferentes decisiones del centro, de financiar una tecnología u otra, según el valor en salud que aporten.

En cuanto a la información sobre costes, hemos señalado los diferentes tipos de costes (sanitarios, no sanitarios, indirectos) que deben tenerse en cuenta en una EE dependiendo de la perspectiva del estudio. La disponibilidad de información es, en muchas ocasiones, mayor en cuanto a los costes directos sanitarios, ya que los hospitales y centros sanitarios recogen diariamente datos del uso de recursos de los pacientes para poder, en teoría, gestionar y tomar las decisiones sanitarias de financiación de ese centro fundamentadas en evidencia. Por lo tanto, en todos los centros debe existir información de contabilidad analítica – y por lo tanto uso de recursos y costes, para cada una de las condiciones de la salud para las que da servicio el centro y para las que deben tomarse decisiones. El problema es, en ocasiones, el acceso a dichos datos. Aunque los datos son recogidos con el trabajo detallado y diario de los profesionales sanitarios, aún pocos, en España, tienen o solicitan acceso a su centro sanitario. Por otro lado, en el caso de que la EE tome una perspectiva más amplia que la del propio centro sanitario donde se desarrolle la evaluación de la tecnología bajo estudio, se deben considerar métodos de recogida de información alternativos como cuestionarios para los pacientes y familiares, que permitan incorporar los costes sanitarios que recaen fuera del centro, así como los costes no sanitarios e indirectos, si éstos fueran considerados relevantes.

Una vez recogida la información necesaria sobre los resultados en salud y los costes de las alternativas de tratamiento, se requiere sintetizar esta información en una medida que permita emitir conclusiones sobre la eficiencia de la tecnología bajo estudio frente a su/s comparador/es. Como definimos anteriormente, la RCEI o RCUI nos indica el coste por unidad de efectividad adicional asociado a una tecnología frente a su comparador. Cuando la efectividad/utilidad es medida en términos de Años de Vida Ajustado por Calidad, este coste por AVAC estimado nos permite concluir más firmemente sobre la eficiencia de la tecnología, al permitir la comparación más allá de un mismo problema de salud y al disponerse de indicadores predeterminados que señalan el coste por AVAC máximo para que una intervención

sea considerada coste-efectiva, conocido como el umbral de coste-utilidad. Por otro lado, si el resultado en salud se mide directamente en unidades monetarias, la estimación del beneficio neto simplifica aún más la comparación de costes y resultados en salud de diferentes tecnologías. Sin embargo, las metodologías empleadas para la medición de los resultados en salud en términos monetarios no dejan de ser controvertidas, por lo que están poco extendidas. El estudio de la eficiencia debe ser, además, complementado con el análisis del impacto que la incorporación de la tecnología tendrá sobre la capacidad de financiación del centro o sistema que evalúa su introducción. Para ello, el análisis de impacto presupuestario toma un papel muy relevante ya que permite medir los recursos que son necesarios para financiar la innovación o tecnología sanitaria al conjunto de la población que se beneficiaría de dicha intervención.

En resumen, el estudio de la eficiencia se impone como un aspecto crucial a evaluar cuando se valoran los resultados en salud de intervenciones sanitarias. En este capítulo hemos definido los tipos de análisis económicos existentes, los aspectos en la medición de costes y resultados en salud más relevantes y las características que deben ser consideradas para facilitar que la eficiencia sea evaluada de forma rutinaria. Es imperativo que se reconozca desde el planteamiento del diseño de estudios que evalúen tecnologías sanitarias los aspectos que deben incorporarse para facilitar el análisis de la eficiencia, hasta los parámetros que deben ser valorados y medidos de manera continua, y las metodologías apropiadas para su análisis. Solo de esta forma se podrá disponer de información sobre la eficiencia que complementen a los datos de resultados en salud de las intervenciones que se evalúen.

■ 5.7. CONCLUSIONES

- La Economía de la Salud tiene como fin último mejorar la asignación de los recursos destinados a la salud, y ofrece una serie de herramientas para incorporar el análisis de la eficiencia y la consideración del coste de oportunidad.
- La evaluación económica (EE) de tecnologías sanitarias es un análisis comparativo de alternativas en términos de costes y resultados en salud.
- Existen diferentes tipos de evaluaciones económicas, siendo la distinción entre estos tipos de estudios la medida que se emplea para cuantificar los resultados en salud.
- Es imprescindible identificar los costes relevantes para la evaluación, así como la medición, y valoración de estos. La inclusión en el estudio de unos u otros recursos dependerá del objetivo y de la perspectiva adoptada.
- Los tipos de coste que podemos encontrar son: costes directos (sanitarios y no sanitarios), costes indirectos o pérdidas de productividad, y costes intan-

gibles. La disponibilidad de información es extensa, aunque el problema es, en ocasiones, el acceso a dichos datos.

- Existen tres medidas principales de resultados en salud: las de dimensión clínica, las que incorporan la valoración de las preferencias de los pacientes (ej. calidad de vida relacionada con la salud (CVRS)), y las que miden el impacto en salud en términos monetarios.
- Los Años de Vida Ajustado por Calidad (AVAC) son una unidad de medida que recogen tanto la duración como la calidad de vida de los pacientes.
- Sería recomendable sistematizar la implementación de cuestionarios genéricos sobre CVRS a los pacientes, tales como el EQ-5D-5L.
- Se requiere sintetizar la información de una evaluación económica en la ratio coste-efectividad/utilidad incremental (RCEI/RCUI).
- El estudio de la eficiencia debe ser complementado con el análisis de impacto presupuestario.

■ REFERENCIAS

- BADIA, X., ROSET, M., MONTSERRAT, S. *et al.* (1999). The Spanish version of EuroQol: a description and its applications. *European Quality of Life scale. Med Clin (Barc)*, 112 (Suppl 1), pp.79-85.
- DRUCKER, P. (2002). The Discipline of Innovation. *Harv Bus Rev*, 80(8), pp. 95-102.
- DRUMMOND, M. F., SCULPHER, M. J., CLAXTON, K., STODDART, G. L. y TORRANCE, G. W. (2015). *Methods for the economic evaluation of health care programmes*. (Fourth Edition). London: Oxford University Press.
- FEENY, D., FURLONG, W., BOYLE, M. *et al.* (1995). Multi-attribute Health status classification Systems. Health Utilities Index. *Pharmacoeconomics*, 7(6), pp. 490-502.
- MAUSKOPF, J. A., SULLIVAN, D., ANNEMANS, L. *et al.* (2007). Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices--budget impact analysis. *Value Health*, 10(5), pp. 336-347.
- MÉNDEZ, I., ABELLÁN PERPIÑÁN, J. M., SÁNCHEZ MARTÍNEZ, F. I. *et al.* (2011). Inverse probability weighted estimation of social tariffs: an illustration using the SF-6D value sets. *J Health Econ*, 30(6), pp. 1280-1292.
- OFFICE OF HEALTH TECHNOLOGY ASSESSMENT. (1998). *Assessing the efficacy and safety of Medical technologies*. Publication num.OTA-75. Washington DC: US Government Printing Office.
- OLIVA MORENO, J., GONZALEZ LÓPEZ-VALCARCEL, B., TRAPERO-BERTRAN, M., HIDALGO VEGA, A. y DEL LLANO SEÑARIS, J. E. *Economía de la Salud*. Madrid: Ed. Pirámide. ISBN: 978-84-368-3977-7.
- RAMOS-GOÑI, J. M., CRAIG, B. M., OPPE, M. *et al.* (2018). Handling Data Quality Issues to Estimate the Spanish EQ-5D-5L Value Set Using a Hybrid Interval Regression Approach. *Value Health*, 21(5), pp. 596-604.
- TRAPERO-BERTRAN, M. (2015). *Integrando la eficiencia en la Esclerosis Múltiple*. Barcelona: CEGE. ISBN: 978-84-944935-0-8.
- VALLEJO-TORRES, L., GARCIA LORENZO, B. y SERRANO AGUILAR, P. (2018). Estimating a cost-effectiveness threshold for the Spanish NHS. *Health Economics*, 27(4), pp. 746-761.



6

PRINCIPALES TÉCNICAS ESTADÍSTICAS Y PARÁMETROS PARA MEDIR RESULTADOS EN SALUD

FERNANDO RODRÍGUEZ ARTALEJO, TERESA VILLAGARCÍA CASLA,
ANA ORTEGA ESLAVA

■ 6.1. MEDIDAS DE ASOCIACIÓN O EFECTO EN ESTUDIOS EPIDEMIOLÓGICOS

■ 6.1.1. Introducción

Habitualmente el efecto de un factor de riesgo sobre una enfermedad se evalúa en estudios epidemiológicos, generalmente de cohortes y de casos-controles, y el efecto de una intervención sanitaria sobre una enfermedad se suele evaluar mediante ensayos clínicos. Un estudio de cohortes es un estudio epidemiológico en el que se selecciona a sujetos libres de la enfermedad o fenómeno de salud que se investiga y se mide la exposición a sus potenciales determinantes. Se les sigue a lo largo del tiempo y se compara la frecuencia con que aparece (incidencia) el fenómeno de salud en estudio entre los expuestos y no expuestos a cada uno de los potenciales determinantes. Un ensayo clínico es un diseño similar en el que la exposición (e.g. una intervención sanitaria) se asigna (habitualmente al azar) por el investigador a los grupos del estudio.

Hay dos formas de medir la incidencia (aparición de casos nuevos) de la enfermedad o variable resultado en los estudios de cohortes y en los ensayos clínicos. La primera es la incidencia acumulada, en la que se divide el número de sujetos en los que aparece la enfermedad a lo largo del seguimiento por el número de sujetos libres de la misma al inicio del seguimiento; por tanto, la incidencia acumulada es una proporción. La segunda forma es la tasa de incidencia, en la que se divide el número de sujetos en los que aparece la enfermedad a lo largo del seguimiento por la suma de los tiempos en que cada sujeto ha sido seguido; por tanto, es una tasa (enfermedad/tiempo de seguimiento). El análisis de los datos consiste en calcular la incidencia de la variable resultado (enfermedad, muerte, etc.) y compararla entre los sujetos expuestos y los no expuestos al potencial factor de riesgo/intervención sanitaria. Cuando la mayoría de los sujetos se han seguido aproximadamente durante el mismo tiempo, la medida más adecuada de incidencia es la llamada incidencia acumulada (riesgo). Si no es así, debe calcularse la densidad de incidencia.

Las diferencias absolutas o relativas (razones) de las incidencias de la enfermedad en los dos grupos de comparación constituyen medidas de la asociación entre el factor de riesgo/intervención sanitaria y la enfermedad; si se considera que

la relación es causal, constituyen medidas del efecto del factor de riesgo/intervención sanitaria. Estas medidas se calculan de la misma forma y tienen el mismo significado para la incidencia acumulada y para la densidad de incidencia. Por ello, realizaremos los cálculos usando incidencias acumuladas.

■ 6.1.2. El riesgo relativo y la *odds ratio*

El riesgo relativo (RR) es la incidencia de enfermedad en los expuestos dividida por la incidencia en los no expuestos al factor de riesgo (Tabla 1). Es una medida de la fuerza de la asociación, y proporciona información útil para juzgar si una asociación, que no se debe al azar ni a errores sistemáticos, puede ser causal; por ello es la forma más habitual de presentar los resultados de los estudios etiológicos. Carece de unidades, su límite inferior es 0 y el superior es infinito. Un RR=1 significa que ambas incidencias son iguales y, por tanto, el potencial factor de

Tabla 1.

MEDIDAS DE ASOCIACIÓN Y EFECTO EN ESTUDIOS DE COHORTES

Medida	Significado	Cálculo
Riesgo relativo (razón de riesgos)	Cuantas veces es más frecuente la incidencia de la enfermedad en los expuestos que en los no expuestos al factor de riesgo.	$RR = I_e / I_o$
Riesgo atribuible (diferencia de riesgos)	Incidencia de la enfermedad que es atribuible a la exposición al factor de riesgo.	$RA = I_e - I_o$
Proporción de riesgo atribuible	Proporción de la enfermedad que es atribuible a la exposición al factor de riesgo.	$RA \% = (I_e - I_o) / I_e \times 100$
Índice de intervención	Número de sujetos en los que hay que impedir la exposición para evitar una enfermedad.	$II = 1 / RA$
Riesgo atribuible poblacional	Incidencia de la enfermedad que es atribuible a la exposición al factor de riesgo en la población.	$RAP = I_t - I_o = RA \times P_e$
Proporción de riesgo atribuible poblacional	Proporción de la enfermedad que es atribuible a la exposición al factor de riesgo en la población.	$RAP \% = (I_t - I_o) / I_t \times 100$

Nota: I_e : incidencia de enfermedad entre los expuestos. I_o : incidencia de enfermedad entre los no expuestos. I_t : incidencia de la enfermedad en el conjunto de la población. I_t : incidencia total de la enfermedad en el conjunto de la población. P_e : Prevalencia de la exposición al factor de riesgo en el conjunto de la población.

Fuente: Elaboración propia.

riesgo carece de efecto sobre la aparición de la enfermedad. Valores superiores a 1 significan que el factor aumenta el riesgo de enfermar. La mayoría de los factores de riesgo toman valores entre 1 y 2. Se trata de efectos pequeños o moderados que requieren muestras de gran tamaño para alcanzar la significación estadística, es decir, para demostrar que son diferentes a 1. En general, hay que desconfiar de valores muy altos del RR. Suelen deberse a errores de cálculo o son hallazgos obvios. Por ejemplo, obtener que el RR de muerte por desplazarse en automóvil a 200 km/h es 25 en relación a los que conducen a 90 km/h resulta una obviedad. Sin embargo, también hay ejemplos en los que no se cumple esta regla. Así el tabaco, que es el factor aislado responsable de más muertes en las sociedades occidentalizadas y que presenta una gran prevalencia, tiene RR superiores a 5 para el cáncer de pulmón. Otros factores también pueden presentarlos, como el asbesto en la producción de mesotelioma y cáncer de pulmón, pero su prevalencia en la población general es escasa. Un RR inferior a 1 significa que el factor de estudio reduce el riesgo de la variable resultado (enfermar, morir, etc.); no se trata de un factor de riesgo sino de un factor protector. Es el resultado que se espera obtener cuando se evalúa un medicamento o una vacuna.

Cuando la variable de resultado es la mortalidad general y se usa la incidencia acumulada, es necesario acompañar el valor del RR del tiempo de seguimiento de los participantes en el estudio, pues el RR se irá aproximando a 1 en el transcurso de los años según las incidencias se aproximen al 100 % en los dos grupos

Tabla 2.

MEDIDAS DE ASOCIACIÓN ENTRE EL CONSUMO DE CIGARRILLOS Y EL CÁNCER DE PULMÓN EN VARONES*

Mortalidad por cáncer de pulmón entre fumadores de cigarrillos:	209/100.000/año
Mortalidad por cáncer de pulmón entre no fumadores de cigarrillos:	14/100.000/año
Mortalidad por cáncer de pulmón en el conjunto de la cohorte (fumadores y no fumadores):	141/100.000/año
Prevalencia de consumo de tabaco en el conjunto de la cohorte (fumadores y no fumadores):	0,65 (65 %)
Medidas de asociación:	
Riesgo relativo: $(209/100.000) / (14/100.000) = 14,9$	
Riesgo atribuible: $(209/100.000) - (14/100.000) = 195/100.000$	
Proporción de riesgo atribuible: $(195/100.000) / (209/100.000) \times 100 = 93,3 \%$	
Índice de intervención: $1 / (195/100.000) = 512$	
Riesgo atribuible poblacional: $(141/100.000) - (14/100.000) = (195/100.000) \times 0,65 = 127/100.00$	
Proporción de riesgo atribuible poblacional: $(127/100.000) / (141/100.000) \times 100 = 90,1 \%$	

Nota: *Estimado a partir de Doll *et al.*, 1994.

Fuente: Baker (2016).

de comparación. En cualquier caso, siempre es una buena práctica, ya que el RR puede cambiar durante el seguimiento, según se va manifestando el efecto del factor riesgo al alcanzarse el tiempo de inducción o a medida que envejecen los miembros de la cohorte. Por ejemplo, algunos de los factores de riesgo clásicos de las enfermedades cardiovasculares, como la colesterolemia o el tabaco, reducen su valor en los ancianos.

La Tabla 2 presenta las medidas de asociación entre el consumo de tabaco y la aparición de cáncer de pulmón calculadas a partir del seguimiento de 40 años de la cohorte de médicos británicos (Baker, 2016). Es un estudio diseñado para establecer si el consumo de tabaco aumentaba el riesgo de cáncer de pulmón, y para ello se comparó la incidencia de cáncer entre fumadores y no fumadores. En este caso un RR de 14,9 significa que, en un período de 40 años, los fumadores tienen 14,9 veces más riesgo de morir por cáncer de pulmón que los no fumadores.

El RR también puede calcularse cuando la variable de exposición es continua. Generalmente se buscan puntos de corte que definen categorías de exposición, y se calcula el RR de cada categoría comparando la incidencia de enfermedad en cada una de ellas con la de referencia, que suele ser el nivel más bajo de exposición (Tabla 3). Los puntos de corte han de establecerse en el diseño del estudio y suelen atender a uno de los siguientes criterios: a) grupos de exposición de igual tamaño definidos, por ejemplo, por cuartiles o quintiles del factor de riesgo, y b) grupos definidos por la magnitud de la exposición y su significado biológico o conductual. En el caso del tabaco, los puntos de corte pueden ser no fumadores, fumadores de 1-14 cigarrillos, de 15-24, y de 25 o más. Una vez calculados los RR para cada nivel de exposición, puede estudiarse si existe una relación dosis-respuesta y la función matemática que mejor la describe.

Tabla 3.

CÁLCULO DE LOS RIESGOS RELATIVOS EN UN ESTUDIO DE COHORTES SOBRE EL CONSUMO DE TABACO Y EL CÁNCER DE PULMÓN CON VARIAS CATEGORÍAS DE EXPOSICIÓN

	Mortalidad por cáncer de pulmón
No fumadores	14/100.000/año
Fumadores de 1-14 cigarrillos	105/100.000/año
Fumadores de 15-24 cigarrillos	208/100.000/año
Fumadores de ≥ 25 cigarrillos	355/100.000/año
*RR de fumadores de 1-14 cigarrillos frente a no fumadores:	$(105/100.000) / (14/100.000) = 7,5$
RR de fumadores de 15-24 cigarrillos frente a no fumadores:	$(208/100.000) / (14/100.000) = 14,9$
RR de fumadores de ≥ 25 cigarrillos frente a no fumadores:	$(355/100.00) / (14/100.000) = 25,4$

Notas: *RR: Riesgo relativo.

Fuente: Datos tomados de Doll *et al.* (1994).

Otra medida de asociación utilizada frecuentemente es la *odds ratio*; aunque se calcula de forma algo diferente a un RR, lo importante es que se interpreta de forma similar. La *odds ratio* se puede obtener en estudios transversales, de casos-controles y de cohortes, y se ha popularizado mucho porque los modelos de regresión logística, que son muy usados en el análisis de estudios epidemiológicos, proporcionan *odds ratios* en lugar de RR.

■ 6.1.3. El riesgo atribuible y medidas relacionadas

El riesgo atribuible (RA) es la diferencia de la incidencia de enfermedad entre los expuestos y no expuestos al factor de riesgo (Tabla 1). Es una medida del efecto absoluto del factor de riesgo y representa el “exceso” de riesgo de enfermar, entre los expuestos, atribuible al factor de riesgo. Por tanto, a diferencia de lo que ocurre con el RR, el cálculo del RA implica la asunción de una relación causal entre el factor de riesgo y la enfermedad. El límite inferior del RA es 0 y el superior es 100 o infinito, según se manejen tasas o densidades de incidencia, respectivamente. Si la incidencia se mide como densidad, el RA se divide por unidades de tiempo. En el caso del tabaco y el cáncer de pulmón (Tabla 2) un RA = 195/100.000 significa que el tabaco produce casi dos muertes por cáncer de pulmón al año entre cada 1.000 fumadores; esta es la mortalidad por cáncer de pulmón que se evitaría erradicando el tabaco en este colectivo. Una medida derivada del RA es la proporción de riesgo atribuible (RA %), también llamada fracción etiológica. Es el RA dividido por la incidencia de enfermedad en los expuestos (Tabla 1). Es otra forma más de presentar el impacto del factor de riesgo entre los expuestos a él. Expresado en términos útiles para la prevención, representa la proporción de la incidencia de enfermedad que se evitaría entre los expuestos si se erradicara el factor de riesgo. Al ser una proporción carece de unidades, su límite inferior es 0 y el superior es 1 (100 si se expresa como porcentaje). En nuestro ejemplo (Tabla 2), un RA % = 93,3 % significa que el 93,3 % de los cánceres desarrollados por los fumadores durante los 40 años del seguimiento se deben al consumo de tabaco. Por lo tanto, ese mismo porcentaje se habría evitado si no hubieran fumado. También puede estimarse el número de sujetos cuyo hábito tabáquico habría que interrumpir para evitar una muerte o evento adverso; es el llamado índice de intervención (II). $II = 1/RA$. En el ejemplo, II es 512, y significa que hay que evitar el hábito tabáquico en 512 fumadores para evitar una muerte por cáncer de pulmón cada año. En el contexto clínico, el II también se conoce simplemente como el “número de sujetos a tratar” para evitar un evento indeseable.

Si se trata de un factor protector, puede calcularse una medida análoga a la RA %, la llamada fracción preventiva (FE), que representa la proporción de la incidencia de enfermedad evitada entre los expuestos por la acción del factor protector:

$$FE = (I_0 - I_e) / I_0 \times 100 = (1 - RR) \times 100$$

donde "I₀" es la incidencia en los no expuestos, y "I_e" la incidencia en los expuestos al factor de riesgo.

El riesgo atribuible en la población general (RAP) es la diferencia de la incidencia de la enfermedad entre la población general y los no expuestos al factor de riesgo. También cabe calcularlo multiplicando el RA por la prevalencia del factor de riesgo en la población general (Tabla 1). Tiene las mismas unidades y límites que el RA. Es una medida del impacto del factor de riesgo en la población general, que ha de ser menor que en los expuestos al factor de riesgo, al estar formada por expuestos y no expuestos a dicho factor. Representa el “exceso” de incidencia de enfermedad que se evitaría en la población general si se erradicara el factor de riesgo. Es, junto con la RAP %, la medida de efecto de mayor importancia en salud pública. En nuestro ejemplo (Tabla 2), un RAP de 127/100.000 mide la reducción de la mortalidad anual por cáncer de pulmón en el conjunto de la cohorte si se erradicara el consumo de tabaco. La proporción de riesgo atribuible en la población general (RAP %) es una medida derivada del RAP y se obtiene dividiendo este por la incidencia total de enfermedad en el conjunto de la población (Tabla 1). Representa la proporción de la incidencia de enfermedad que se evitaría en la población general si se erradicara la exposición al factor de riesgo. La RAP % carece de unidades, su límite inferior es 0 y el superior es 1 (100 si se expresa como porcentaje). En el caso del tabaco y el cáncer de pulmón, la RAP % = 90,1 % significa que la erradicación del consumo de tabaco reduciría la mortalidad por cáncer de pulmón en la cohorte de estudio en un 90,1 %. Es posible también calcular el número de muertes atribuibles al tabaco multiplicando la RAP % por el número de muertes por cáncer de pulmón.

Un factor de riesgo con un RR modesto puede tener un gran impacto sobre la salud pública, una elevada RAP %, si su prevalencia es muy alta. Es el caso de la tensión arterial y su impacto sobre la cardiopatía isquémica; su RR es aproximadamente 2, pero alrededor del 20 % de la población presenta este factor de riesgo. Por el contrario, el asbesto tiene un RR superior a 10 sobre la mortalidad por cáncer de pulmón, pero su impacto sobre la mortalidad por esta enfermedad es solamente moderado, porque su prevalencia en la población general es muy baja. Sin embargo, tiene un enorme impacto sobre los expuestos profesionalmente a este material.

■ 6.1.4. Variabilidad en medidas de asociación

Las medidas de asociación descritas tienen una variabilidad derivada de la selección muestral en los estudios, y es posible calcular pruebas de significación estadística e intervalos de confianza para ellas. Por otro lado, el análisis de los datos debe iniciarse comprobando la comparabilidad de los grupos de expuestos y no expuestos, es decir, asegurando que son homogéneos en la distribución de otros predictores de la variable resultado, distintos del factor que se ha de evaluar. Si no es así, hay que ajustar las medidas de asociación por la acción de dichos predictores mediante métodos estadísticos, como la regresión multivariante (ver el apartado de dificultades en estudios observacionales y técnicas para solventarlos).

■ 6.1.5. Un ejemplo

A modo de ejemplo, la Tabla 4 presenta los resultados del efecto de una intervención en un ensayo clínico. En concreto se ha usado el ensayo SPRINT que se diseñó para establecer si, en pacientes hipertensos de alto riesgo cardiovascular, reducir la tensión arterial sistólica por debajo de 120 mm Hg era beneficioso por comparación con llevarla simplemente por debajo de 140 mm Hg (el valor habitual por debajo del que se considera que la hipertensión está controlada) (Doll *et al.*, 1994). Se reclutó a 9.361 pacientes, de los que 4.678 fueron asignados aleatoriamente al grupo de intervención (tensión <120 mm Hg) y 4.683 al grupo de control (tensión <140 mm Hg). La duración del ensayo fue de 3,26 años, y en el grupo de intervención se recibieron como media tres fármacos antihipertensivos y en el de control solo dos fármacos. La Tabla 4 presenta los resultados para una de las variables de resultado del ensayo, en concreto la mortalidad general. El RR fue 0,73, que indica que la intervención redujo la mortalidad un 27 % respecto al control. El RA fue 1,2 %, que indica que la intervención redujo en términos absolutos la mortalidad un 1,2 % a lo largo de los 3,26 años de seguimiento; ello supuso que bajar la tensión arterial por debajo de 120 mm Hg evitó 55 muertes. Por último, el número de pacientes a los que hubo que reducir la tensión a <120 mm Hg durante 3,26 años para evitar una muerte fue de 83. En conjunto la intervención produjo unos beneficios sustanciales sobre la mortalidad.

Tabla 4.

MEDIDAS DE EFECTO EN EL ENSAYO CLÍNICO SPRINT SOBRE REDUCCIÓN DE LA PRESIÓN ARTERIAL EN HIPERTENSOS Y MORTALIDAD GENERAL A LO LARGO DE 3,26 AÑOS

Mortalidad en el grupo de intervención (tensión arterial <120 mmHg): 3,3 %

Mortalidad en el grupo de control (tensión arterial <140 mmHg): 4,5 %

Medidas de asociación:

Riesgo relativo: $3,3 \% / 4,5 \% = 0,73$

Riesgo atribuible: $3,3 \% - 4,5 \% = -1,2 \%$

Número de pacientes que hay que tratar para evitar una muerte: $1 / 1,2 \% = 83$

Fuente: Estimado a partir de SPRINT Research Group *et al.* (2015).

■ 6.2. CONTRASTES DE HIPÓTESIS Y REGRESIÓN LINEAL

■ 6.2.1. Introducción

El contraste de hipótesis (CH) es la técnica estadística que se utiliza para decidir entre dos posibilidades. Ejemplos de uso podrían ser:

1. Tras un ensayo clínico, decidir si el resultado obtenido en el grupo que ha utilizado el tratamiento test es igual al resultado obtenido para el grupo que ha utilizado el tratamiento convencional o un placebo.
2. Se conoce el porcentaje de supervivencia a determinada enfermedad (p_0). Se hace un estudio con un nuevo medicamento obteniendo un porcentaje de supervivencia (p_1). ¿Es mejor el nuevo medicamento?

Los dos ejemplos anteriores, habituales en ensayos clínicos, muestran el problema principal del contraste de hipótesis. Supongamos en el ejemplo 2 que la probabilidad de supervivencia actual es del 30,00 % y con el nuevo medicamento hemos obtenido una supervivencia de 30,0000001 %. Casi no es necesario saber estadística para afirmar que no hay diferencias entre ambos grupos. Del mismo modo si hubiéramos obtenido un valor de supervivencia para el grupo test de 90 % decidiríamos que el tratamiento test es diferente del control y además es magnífico.

El problema surge, en la práctica, porque los resultados no son tan extremos y necesitamos herramientas que nos permitan decidir si el nuevo medicamento es efectivo o no. El CH nos va a permitir formular estas preguntas de modo sistemático y obtener una respuesta en términos de probabilidad. Para ello es preciso plantear dos hipótesis:

H_0 Hipótesis nula. Es la hipótesis que se desea probar con los datos.

H_1 Hipótesis alternativa. Es, como su nombre indica, la alternativa a H_0 .

Para realizar un contraste de hipótesis es preciso seguir los siguientes pasos:

1. Saber qué tipo de contraste se necesita.
2. Plantear correctamente H_0 y H_1 .
3. Obtención del p-valor.
4. Decisión.

Las etapas 1 y 2 son las más complejas hoy día, ya que requieren conocer bien el funcionamiento del problema y plantear las hipótesis correctamente. Las etapas 3 y 4 antiguamente requerían cálculos farragosos y el uso de cuadros de probabilidades. Actualmente son inmediatas ya que disponemos de *software* que realiza los cálculos (etapa 3) y, una vez obtenido el p-valor, de criterios claros sobre cómo tomar las decisiones. Ya no es necesario conocer la formulación matemática del contraste.

La decisión se toma en función del p-valor que nos da información de cuán verosímil es la hipótesis nula con los datos obtenidos. Cuanto menor sea el p-valor, tendremos menos evidencia de que nuestros datos sean compatibles con H_0 . El criterio más utilizado es que si el p valor es menor que $\alpha=0,05$ (5 %) rechazamos la

hipótesis nula. Es decir, pensaremos que es cierta la hipótesis alternativa. Si, por el contrario, el p-valor es mayor que 0,05 no rechazaremos la hipótesis nula. Es importante resaltar que el p-valor *da información sobre lo verosímil que puede ser H_0 . No es la probabilidad de H_0* . Este tema se trata en el apartado 6.2.3.

A continuación, se van a plantear cuatro tipos de contraste y su resolución con un ejemplo en cada caso.

■ 6.2.2. Tipos de contraste y cómo se plantean H_0 y H_1

■ 6.2.2.1. Contraste para una proporción

El ejemplo tipo de este contraste sería:

Se sabe que la supervivencia a determinada enfermedad es de 0,3 (30 %). Se prueba una nueva medicación en una muestra $n=56$ pacientes. Sobreviven a la enfermedad 19 (probabilidad de sobrevivir $19/56=0,3393$). La pregunta claramente es: ¿Es mayor la tasa de supervivencia con la nueva medicación?

La hipótesis nula siempre es la información externa que se quiere probar. En este caso,

$$H_0 : p=0,3$$

La hipótesis alternativa es igual a la nula siempre, salvo por el signo. Y el signo se decide comparando los datos muestrales con la hipótesis nula. Por tanto,

$$H_1 : p \neq 0,3$$

En nuestro caso hemos obtenido una proporción mayor de supervivientes (0,3393), por tanto, la alternativa es que la probabilidad de sobrevivir es mayor que 0,3,

$$H_1 : p > 0,3$$

El contraste se plantea completo de la siguiente manera:

Contraste para una proporción, con

$$H_0 : p=0,3$$

$$H_1 : p > 0,3$$

Una vez planteado el contraste cualquier *software* estadístico nos va a pedir la p de H_0 ($p=0,3$), el signo de H_1 ($>$), el tamaño muestral ($n=56$), el número de super-

vivientes (19) o la probabilidad de supervivencia ($19/56=0,3393$ o porcentualmente 33,93 %) y el nivel de significación α , que suele ser 0,05. El resultado es el p-valor.

La web Mathcracker¹ permite hacer los cálculos para obtener el p-valor. Introduciendo los datos requeridos, obtenemos p-valor=0,2606 en el apartado 4 de respuestas.

Como el p valor es superior a 0,05 no podemos rechazar la hipótesis nula y por tanto seguimos pensando que $p=0,3$ con la nueva medicación. Es decir, no ha supuesto una mejora en la tasa de supervivencia.

■ 6.2.2.2. Contraste para una media

Un ejemplo de este contraste sería:

Se sabe que la duración media de una enfermedad es 15 días. Se prueba un nuevo tratamiento en $n=48$ pacientes. Se obtiene una duración media de 13,2 días y una desviación típica de 2,2 días. La pregunta es ¿Ha disminuido la duración de la enfermedad?

La hipótesis nula es la información externa que se quiere probar. En este caso que la media (μ) es 15 días:

$$H_0 : \mu=15$$

La hipótesis alternativa, como se ha dicho en el apartado anterior, es igual a la nula siempre salvo por el signo, que se decide comparando los datos muestrales con la hipótesis nula. En nuestro caso la media de la muestra es de 13,2 días. Menor que 15. Por tanto:

$$H_1 : \mu<15$$

El contraste completo sería:

Contraste para una media con

$$H_0 : \mu=15$$

$$H_1 : \mu<15$$

El *software* estadístico nos va a pedir la μ de H_0 ($\mu =15$), el signo de H_1 ($<$), el tamaño muestral ($n=48$), la media obtenida (13,2) y la desviación típica obtenida (2,2). El resultado es el p-valor.

La web Mathcracker² permite hacer los cálculos para obtener el p-valor. Introduciendo los datos requeridos, obtenemos un p-valor=0,0016 en el apartado 4 de los resultados.

¹ <https://mathcracker.com/z-test-for-one-proportion.php>

² <https://mathcracker.com/t-test-for-one-mean.php>

Como el p valor es inferior a 0,05 rechazamos la hipótesis nula y por tanto hay evidencia estadísticamente significativa de que la duración de la enfermedad se reduce con el nuevo medicamento.

■ 6.2.2.3. Contrastes para comparar dos medias

El contraste para comparar dos medias resuelve el tipo de problema siguiente:

Se realiza un ensayo clínico de un nuevo medicamento para combatir la epilepsia. El grupo de control está formado por 235 pacientes que toman la medicina tradicional. Han tenido en 180 días un promedio de 5,3 crisis con una desviación típica de 2,3. El grupo de test está formado por 243 pacientes. Han tenido una media de 5,8 crisis con desviación típica 2,5. La pregunta es ¿Son iguales ambos tratamientos?

En el contraste de diferencia de medias la hipótesis nula es, casi siempre, que ambas medias son iguales. La alternativa es que no son iguales:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Los datos necesarios, que nos pedirá el *software*, serán, el signo de H_0 (=), el de H_1 (\neq), los tamaños muestrales ($n_1=235$, $n_2=243$) las medias (5,3 y 5,8) y las desviaciones típicas (2,3 y 2,5).

La web Mathcracker³ pide los datos anteriores y se obtiene un p-valor de 0,1065. No podemos rechazar la hipótesis nula y, ambos medicamentos son equivalentes.

■ 6.2.2.4. Contraste para comparar dos proporciones

El ejemplo sería:

Se realiza un ensayo clínico de un nuevo medicamento para combatir la epilepsia. El grupo de control está formado por 235 pacientes que toman la medicina tradicional. En 180 días han tenido crisis 161 pacientes (68,51 %). El grupo de test está formado por 243 pacientes. Han tenido crisis 180 pacientes (74,07 %). La pregunta es ¿Son iguales ambos tratamientos?

En el contraste de diferencia de proporciones la hipótesis nula es, casi siempre, que ambas proporciones son iguales. La alternativa es que no son iguales:

$$H_0 : p_1 = p_2$$

$$H_1 : p_1 \neq p_2$$

³ <https://mathcracker.com/t-test-for-two-means.php>

Los datos necesarios, que nos pedirá el *software*, serán: el signo de H_0 (=), el de H_1 (\neq), los tamaños muestrales ($n_1=235$, $n_2=243$) y el número de pacientes con crisis o sus proporciones.

La web Mathcracker⁴ pide los datos anteriores y se obtiene un p-valor de 0,1787 en el apartado 4 de resultados. No podemos rechazar la hipótesis nula y ambos medicamentos son análogos.

■ 6.2.2.5. Otros contrastes

Existen muchos contrastes, pero la forma de trabajar con ellos es siempre la misma. Plantear las hipótesis nula y alternativa correctamente. Mediante un *software* adecuado obtener el p-valor y con él decidir qué hipótesis debemos rechazar.

Un contraste muy habitual es el ANOVA, Análisis de la varianza. En este caso se tienen datos de varios grupos y se quiere saber si los grupos son iguales o si al menos uno de ellos difiere del resto. Suele utilizarse en ensayos en los que se miden los efectos de varias dosis diferentes a varios grupos de pacientes. En este caso H_0 será que todas las medias son iguales y H_1 que al menos una de las medias es diferente. El p-valor se analiza exactamente igual que hasta ahora.

También es importante mencionar que hay muchos contrastes diferentes para resolver el mismo problema. Por ejemplo, para estudiar una diferencia de medias se puede utilizar también el contraste no paramétrico de Mann-Whitney y el p-valor se interpretaría igual.

■ 6.2.3. El mal uso del p-valor

Como se ha explicado, el p-valor nos da información de cuán verosímil es H_0 con los datos que tenemos, pero no es la probabilidad de que H_0 sea cierta.

Técnicamente el p-valor es la probabilidad de que si H_0 es cierta, se obtenga una muestra más desfavorable a H_0 que la que hemos obtenido.

En la práctica esto equivale a que podemos equivocarnos un 5 % de las veces (cuando $\alpha=0,05$) en el sentido de rechazar H_0 y pensar que hay diferencias entre los dos grupos. Es decir, si repetimos el mismo ensayo sin diferencias 100 veces, en cinco ocasiones en promedio obtendremos p-valores menores que 0,05 aunque no haya diferencias entre los grupos y, consecuentemente rechazaremos H_0 y crearemos que hay diferencias entre los grupos. Por ejemplo, supongamos que realizamos el siguiente experimento absurdo:

Se escogen 50 personas sanas de 20 años al azar. Se les divide en dos grupos aleatoriamente y se les mide la tensión. Al grupo 1 se le ha medido la tensión sentado en una silla azul. Al segundo en una silla verde. Si

⁴ <https://mathcracker.com/z-test-for-two-proportions.php>

nos planteamos si hay diferencias en la tensión media de ambos grupos, obviamente esperamos un p-valor mayor que 0,05 ya que el color de la silla no debería influir en la tensión de esas personas. Ahora bien, si repetimos ese experimento 100 veces, habrá aproximadamente 5 en las que el p-valor saldrá menor que 0,05. Y la conclusión, en esos casos, será que el color de la silla influye.

Eso no es un problema grave si hacemos un ensayo y medimos una variable, pues la probabilidad de detectar algo inexistente es del 5 % que es baja. Pero si hacemos un ensayo y con la misma muestra medimos 200 variables –por ejemplo, todo tipo de marcadores– y hacemos 200 contrastes de diferencia de medias entre el grupo de control y el de test, vamos a tener en promedio 10 test con p-valores menores que 0,05 que no son reales. Es decir, son falsos positivos. Y pensaremos que ese medicamento es efectivo para esas diez cosas, y ese es el problema del mal uso del p-valor (Baker, 2016). Hay dos formas claras de evitarlo: no utilizar la misma muestra para hacer muchos contrastes de hipótesis o bajar notablemente el nivel de significación α . El problema en este segundo caso es que, si bien bajaríamos el número de falsos positivos, aumentaríamos el de falsos negativos y no se detectarían bien tratamientos que si son efectivos. Consecuentemente la mejor forma de evitar problemas es utilizar cada ensayo para el tema elegido y no como cajón de sastre para probar muchas otras posibilidades (Doll *et al.*, 1994).

■ 6.2.4. Regresión lineal

La regresión lineal se utiliza para estudiar cómo depende una variable Y de otra variable X.



Ejemplos de uso serían:

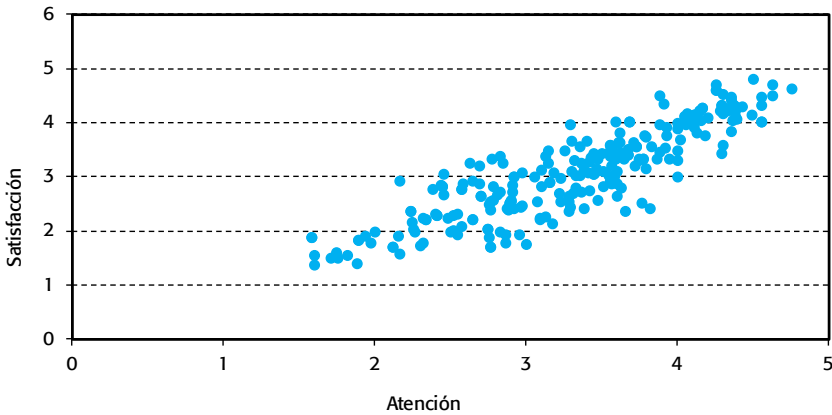
1. Calcular la concentración plasmática esperada (Y) de un medicamento en función de la dosis que se ha suministrado al paciente (X). Se tiene una muestra de pacientes y, para cada uno de ellos, se conoce la dosis recibida y la concentración plasmática.
2. Se realiza una encuesta de satisfacción de pacientes de un hospital. Se conocen, para cada paciente, los valores otorgados a la Satisfacción Global (Y) y a la Atención recibida (X).

El objetivo del análisis es estudiar dos aspectos:

- ¿Existe relación entre ambas variables?
- Cuantificar esa relación, es decir si se incrementa la dosis del medicamento en una unidad, ¿Qué incremento de la concentración plasmática cabe esperar?

Figura 1.

DISPERSIÓN DE LA SATISFACCIÓN VERSUS ATENCIÓN RECIBIDA POR EL PACIENTE



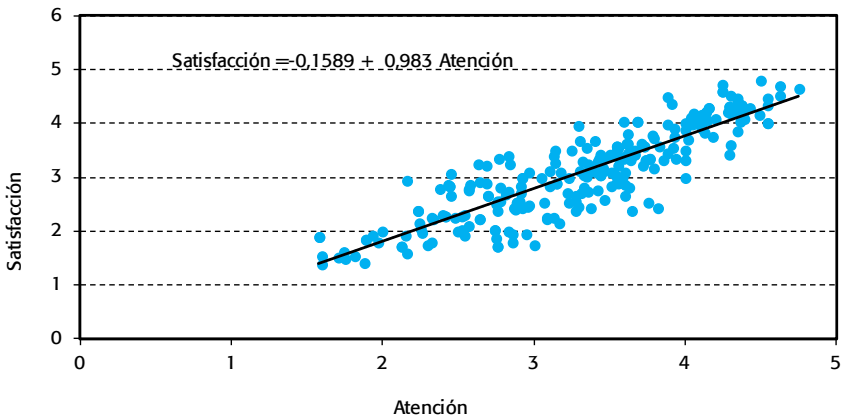
Fuente: Elaboración propia.

Se va a ilustrar el análisis con una muestra de datos del ejemplo 2. La Figura 1 muestra el gráfico de dispersión para los datos de satisfacción de los pacientes.

Como puede observarse en la figura, el gráfico presenta dos características importantes. En primer lugar, hay relación entre las dos variables, y en segundo

Figura 2.

REGRESIÓN DE SATISFACCIÓN DE LOS PACIENTES EXPLICADA POR LA ATENCIÓN RECIBIDA



Fuente: Elaboración propia.

lugar esta relación es de tipo lineal. En estas circunstancias es razonable ajustar una línea recta a los datos y utilizar su ecuación como resumen de la relación entre ambas variables. Esto es lo que hace una regresión simple o regresión lineal. Si la ajustamos a los datos de la Figura 1, obtenemos la regresión de la Figura 2.

La ecuación matemática de la recta que aparece en la Figura 2 es:

$$\text{Satisfacción} = -0,1589 + 0,983 \text{ Atención}$$

Como ya indicamos al hablar de contrastes de hipótesis, hoy día no es necesario conocer las fórmulas que permiten obtener este resultado, ya que esta ecuación puede obtenerse con cualquier *software* estadístico e incluso con Excel.

En general la recta de regresión se escribe:

$$Y = \beta_0 + \beta_1 X$$

y, como puede observarse, en nuestro caso $\beta_1=0,983$. Esto indica que si el valor de X se incrementa en una unidad, el valor de Y se incrementará en 0,983 unidades. Es decir, si un paciente da un punto más en la encuesta a Atención recibida en el Hospital su Satisfacción Global se incrementará en 0,983 puntos.

Si los datos fueran del ejemplo de las dosis, la conclusión hubiera sido que si un paciente recibe una dosis una unidad mayor, su concentración plasmática habría aumentado en 0,983 unidades.

Es muy importante resaltar que la regresión nos permite cuantificar el impacto que los cambios en el valor de X tendrán sobre el valor de Y.

■ 6.2.4.1. Contraste de significatividad de β_1

Cuando se ajusta una regresión es importante realizar el contraste de significatividad sobre el coeficiente β_1 . Este contraste sería:

$$H_0 : \beta_1=0$$

$$H_1 : \beta_1\neq 0$$

Si $\beta_1=0$, querría decir que la ecuación quedaría como:

$$Y=\beta_0+0 X=\beta_0.$$

Es decir que Y no depende de X. O dicho al revés, que X no influye sobre Y.



El contraste que resuelve esta cuestión se denomina Contraste t. Los programas que realizan regresiones proporcionan el p-valor del contraste t. Si el p-valor es menor que 0,05, concluiremos que X sí influye sobre Y y procederemos a cuantificar su efecto. En este caso hablamos de una variable *estadísticamente significativa*. Si, por el contrario, el p-valor es superior a 0,05 la variable no es significativa y podemos afirmar que X no influye sobre Y.

En nuestro ejemplo el p-valor=0,000013 que es menor que 0,05. Por ello concluimos que la variable X es significativa y sus cambios repercuten sobre Y.

■ 6.2.4.2. Coeficiente de Determinación R^2

Finalmente, en regresión se obtiene el valor R^2 , que indica cuánta variabilidad de Y es explicada por X. En el ejemplo que estamos utilizando $R^2=0,89$, es decir, 89 %. Podemos afirmar por tanto que X explica el 89 % de la variabilidad de Y. El valor de R^2 depende básicamente del tipo de fenómeno que se esté estudiando, ya que hay situaciones en las que un 20 % de capacidad explicativa es mucho y otras en que un 60 % es poco. El analista normalmente conoce los rangos de capacidad explicativa esperables.

■ 6.2.4.3. Otros modelos de regresión

Existen otros modelos de regresión que no se van a analizar en este capítulo como:

- La regresión lineal múltiple que es similar a la regresión lineal simple (visto en los apartados anteriores, con una variable Y y una variable X) pero en la que el modelo de regresión tiene más de una variable X, en ellas se analizaría el efecto de cada variable X sobre la variable Y controlado por las otras variables Xs.
- Regresión logística que se utiliza cuando la variable Y es una variable dicotómica cualitativa o categórica.
- Regresión de Cox que es un modelo utilizado cuando la variable de respuesta Y representa duraciones de procesos (por ejemplo, de enfermedades o tiempo de supervivencia de pacientes a una enfermedad mortal) y se desea conocer qué variables influyen en esa duración. El modelo se utiliza cuando parte de los datos de tiempos no se observan completos. Por ejemplo, porque un paciente sigue vivo en el momento del estudio. En ese caso, su tiempo de supervivencia real a la enfermedad será mayor que el que ha tenido hasta el momento de la toma de datos. Este tipo de datos se denominan censurados y el modelo de tasa de fallos proporcional de Cox es capaz

de trabajar con datos censurados proporcionando estimaciones centradas de los tiempos de supervivencia.

■ 6.3. LIMITACIONES DE ESTUDIOS OBSERVACIONALES Y ESTRATEGIAS DE ANÁLISIS

■ 6.3.1. Necesidad de los estudios observacionales

Los estudios observacionales, basados en la obtención de datos de la práctica clínica real con pacientes, son necesarios, en especial en la evaluación de resultados en salud, ya que proporcionan información que no se puede obtener de los ensayos clínicos. Informan sobre los resultados en la práctica clínica fuera del contexto ideal, restrictivo y controlado, de los ensayos clínicos, pueden aportar información respecto a efectos raros o a largo plazo, sobre factores que afectan al uso o efecto de los tratamientos, efectos en un mayor grupo de población o con unas características diferentes a las incluidas en los ensayos, etc.

La realización de estudios observacionales es una labor que se debe realizar de forma paralela a la práctica asistencial para así contribuir a un mejor conocimiento de los efectos de los tratamientos, etc. Los profesionales sanitarios podemos contribuir en gran medida a ampliar y aplicar este conocimiento con orientación a la mejora de la práctica asistencial y de los resultados en los pacientes. Además, para el análisis de los registros clínicos, se realizan, en general, estudios observacionales, en muchas ocasiones de cohortes, que están ganando importancia para monitorizar y mejorar la calidad del cuidado de los pacientes (Gil Martin-Thomas, Liew y Hopper, 2018).

■ 6.3.2. Principales limitaciones de los estudios observacionales

Sin embargo, los estudios observacionales tienen también limitaciones importantes que es necesario solventar para poder extraer conclusiones válidas. Entre estas limitaciones o dificultades se encuentran:

- La asignación no aleatoria de los individuos a las alternativas o factores de riesgo. En la práctica clínica es normal que los individuos hayan recibido una alternativa en lugar de otra por algunas características, disminuyendo la comparabilidad de los grupos.
- La falta de información: el limitado conocimiento sobre cómo se ha decidido la asignación a los tratamientos, la medida incompleta de los factores de riesgo que influyen en el tratamiento y en el efecto, etc. Es más, el carácter retrospectivo de muchos de estos estudios dificulta la recogida de cierta información. Además, la falta de datos puede no estar equilibrada entre los grupos

comparados, en general se recoge de forma menos completa en los individuos no tratados y, por ello, asumir que una condición no está presente cuando no está anotada puede ser problemático (Glynn, 2017).

Estas limitaciones pueden suponer sesgos e impedir establecer una relación causal entre una exposición o un factor de riesgo, ej un tratamiento, y un determinado efecto. El principal reto con los estudios observacionales para determinar relaciones causales válidas es solucionar en la medida de lo posible los posibles sesgos.

■ 6.3.3. Sesgos en los estudios observacionales

Sesgo es cualquier factor que pueda llevar a que los resultados se desvíen sistemáticamente de la realidad. Entre los sesgos podemos destacar los siguientes:

- Sesgo de selección: los grupos a comparar no han sido seleccionados de forma adecuada, por ello hay que tratar de evitar que se seleccionen individuos con una relación entre el factor y el efecto especial. Lo mejor es aleatorizar lo máximo posible. Ya que en estudios observacionales no se puede aleatorizar los individuos expuestos y no expuestos al factor, al menos se puede hacer una selección aleatoria de la muestra.
- Sesgo de información: se produce cuando no se recoge la información de todos los individuos o pacientes por igual o no se recoge la información de forma adecuada. Una buena estrategia es hacer que la persona que genera la información o la recoge no pueda manipularla, por ejemplo, sea “ciego” al grupo al que pertenecen los individuos. También es importante estandarizar la forma de recogida de los datos y las definiciones de estos, antes de recogerlos.
- Sesgo de indicación: se produce cuando existe una tendencia a que los individuos expuestos al factor, por ejemplo, que toman el medicamento evaluado, tengan unas características particulares, por las que se les ha dado el tratamiento. Este sesgo se evita con la aleatorización, pero, como hemos comentado, no es posible en los estudios observacionales. Es difícil de solucionar si claramente los expuestos al factor son diferentes a los no expuestos, porque nunca se conseguirá saber, por ejemplo, si se encuentra una asociación si ésta se debe al factor analizado o a otro. Sin embargo, si hay más individuos con unas características en un grupo que en otro, pero existen individuos similares en ambos grupos se puede corregir por el posible efecto de este sesgo. Una estrategia consiste en el uso de lo que se llama el índice de propensión, *propensity score* (PS) en inglés. Es un índice que indica la propensión a estar expuesto al factor analizado, por ejemplo, a reci-

bir un tratamiento, en base a unas características de los individuos. Cada vez se utiliza más esta técnica y por ello la desarrollaremos más adelante.

Por otro lado, existen lo que se conoce como factores de confusión y es necesario corregir su efecto para llegar a conclusiones válidas. Los factores de confusión son factores relacionados con el factor de exposición analizado, por ejemplo, el tratamiento, y relacionados con el efecto estudiado, pero no son la vía por la cual el factor produce el efecto. Por ejemplo, si se estudia la relación entre el ejercicio y el fallo cardíaco, fumar podría ser un factor de confusión, porque las personas que hacen ejercicio, en general, fuman menos, y fumar aumenta el riesgo de fallo cardíaco, y se podría concluir que los que no hacen ejercicio tienen más riesgo de fallo cardíaco; pero también podría ser porque fuman más, por lo que se tendría que corregir por este factor, fumar o no. La mejor forma para controlarlos es la aleatorización. Sin embargo, en los estudios observacionales, que no se puede aleatorizar, hay que corregir por los factores de confusión, aunque no se podrá hacer por los que no se conozcan. Existen diferentes estrategias para corregir el efecto de los factores de confusión, tanto en el diseño de los estudios (aleatorización, exclusión de pacientes con ciertas características, emparejamiento) como en el análisis de los resultados (estratificación o análisis de subgrupos, realización de modelos), cada una de estas estrategias tiene sus ventajas y sus inconvenientes (Tabla 5).

Tabla 5.

ESTRATEGIAS DE CONTROL DE LOS FACTORES DE CONFUSIÓN

Estrategia	Ventajas	Inconvenientes
<i>En el diseño del estudio</i>		
Randomización	<ul style="list-style-type: none"> • Garantiza la comparabilidad • Aumenta eficacia del estudio 	<ul style="list-style-type: none"> • No es posible en estudios observacionales
Emparejamiento	<ul style="list-style-type: none"> • Aumenta eficiencia, precisión 	<ul style="list-style-type: none"> • Alarga tiempo • Problemas factibilidad
Restricción o exclusión	<ul style="list-style-type: none"> • Simple • Aumenta homogeneidad 	<ul style="list-style-type: none"> • Alarga tiempo • Reduce validez externa
<i>En el análisis de los resultados</i>		
Estratificación	<ul style="list-style-type: none"> • Simple 	<ul style="list-style-type: none"> • Se puede perder información • Número limitado de categorías • Interpretación compleja
Modelos	<ul style="list-style-type: none"> • No se pierde información • No se categoriza • Puede dar estimaciones individuales de riesgo 	<ul style="list-style-type: none"> • Se asumen condiciones no siempre válidas • Resultados difíciles de interpretar

■ 6.3.4. Análisis aplicando el índice de propensión

Los pasos para la realización de análisis de los datos utilizando el índice de propensión son los siguientes:

1. Analizar las características de los pacientes en ambos grupos y ver cuáles son diferentes.
2. Crear un modelo que permita calcular el índice de propensión (PS) o la probabilidad o riesgo de que un individuo por sus características haya sido asignado a un grupo, por ejemplo, a recibir un fármaco u otro fármaco. Esto se realiza generalmente mediante un modelo de regresión logística ($Y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n$). En este modelo la variable dependiente (Y) será la asignación al grupo, por ej. a la intervención a evaluar sí=1 o no=0. Las variables independientes, $x_1, x_2, x_3, \dots, x_n$, serán variables que puedan estar relacionadas con la asignación a un grupo u otro y/o con el resultado que queramos medir. Como variables x se incluyen variables con sentido biológico, intentando no dejar fuera ninguna variable que pueda introducir un sesgo.
3. Una vez construido el modelo de regresión logística, calcular el valor del PS (Y) para cada individuo del estudio en función de los valores de las variables x_1, x_2, x_n de ese individuo en concreto (el PS será un valor entre 0 y 1). El PS nos indica la propensión de ese individuo, por sus características, a haber sido asignado al grupo de la alternativa/intervención evaluada.
4. Crear grupos comparables mediante emparejamiento utilizando los valores de PS o hacer modelos de regresión con el PS como una variable X. Al hacer emparejamiento, o grupos comparables en los valores de PS, se hacen los grupos comparables en las otras variables $x_1, x_2, x_3, \dots, x_n$. Para ello se pueden utilizar diferentes técnicas, las principales diferencias entre ellas están en el número de individuos que quedan tras el ajuste o emparejamiento y el peso relativo que se da a cada uno de los individuos. En ocasiones el PS se utiliza como una variable predictora o de confusión de un modelo (una variable independiente del modelo de análisis del efecto de la intervención en una determinada variable de resultado, ej $Y = a + b_1x_1 + b_2x_2$; donde sería Y la variable de resultado y x_1 la intervención y x_2 el PS, y se podrían añadir también otras variables).

El emparejamiento más común consiste en emparejar cada individuo del grupo con el factor analizado, ej. el tratamiento estudiado, con el individuo del grupo control más similar en PS en una relación k:1. El emparejamiento puede ser exacto incluyendo el mismo número de individuos con cada valor del índice de propensión en cada grupo, o emparejar en función del PS con un límite de tolerancia fijado (caliper) o incluir todos los individuos del grupo control con diferentes pesos en función de su PS (densidad de kernel) y así se pierden menos pacientes, pero en estos casos también conviene limitar

el análisis a los individuos con valores comunes de PS. Otro aspecto para considerar es el reemplazamiento, es decir, si un mismo individuo del grupo control puede ser usado como control de más de un individuo del grupo de intervención. El emparejamiento puede ser total o por estratos o subgrupos (Stuart, 2010).

5. Si se utilizan técnicas de emparejamiento, es necesario comprobar que las muestras son comparables en cuanto a sus características tras el emparejamiento. Si no se consigue, se debe descartar ese método de emparejamiento y realizar otro nuevo hasta que las muestras estén balanceadas en las covariables. Algunos autores recomiendan estudiar la diferencia estandarizada de medias (diferencia de medias dividido por la desviación estándar global) que sea $<0,25$ (Stuart, 2010), las gráficas de quintiles-quintiles y las probabilidades de las pruebas de comparación (ej. kolmogorv-Smirnov, bootstrap p value). Valorar si no va a ser posible separar las covariables de la asignación al tratamiento impidiendo el análisis.
6. Comparar el efecto en ambos grupos mediante las técnicas estadísticas convencionales. Se pueden usar modelos. Estimar la diferencia entre los tratamientos y su intervalo de confianza del 95 % (IC95 %). Para el cálculo de estos intervalos de confianza se pueden utilizar técnicas paramétricas, pero en muchas ocasiones es necesario recurrir a técnicas no paramétricas como el bootstrap (ver más abajo).

■ 6.3.5. Ventajas y desventajas del índice de propensión y otras estrategias

Las ventajas del uso de PS es que se simplifica la corrección del sesgo de indicación o la corrección del efecto de los factores de confusión. Por ejemplo, si se quiere hacer emparejamiento por todas las posibles variables de confusión serían muy complicado ya que habría que buscar en ambos grupos individuos con las mismas características. Sin embargo, hacer emparejamiento por el índice de propensión es más sencillo. Pero la desventaja es que solo permite corregir por las variables conocidas y medidas. Por lo tanto, es peor que la aleatorización.

Otras estrategias de corrección de estos posibles sesgos en los estudios observacionales son los métodos de *genetic matching* (GenMatch), que se puede realizar con el programa R. GenMatch es un algoritmo automático que maximiza el balance entre los grupos, busca el balance no sólo en el PS pero también en otras variables, y puede usar o no el PS. Requiere elegir las variables más importantes para buscar el balance entre los grupos, lo cual exige un criterio clínico y no solo estadístico, y el criterio para buscar la diferencia, para reemplazamiento, etcétera.

Tanto PS como GenMatch corrigen solo el efecto de las variables observadas, pero pueden existir variables no observadas que introduzcan un sesgo. En

ocasiones se realizan análisis de sensibilidad para analizar la posible confusión por factores no observados. Por ejemplo, analizar qué magnitud debería tener el efecto de las variables no observadas para cambiar las conclusiones y comparar esta magnitud con el efecto del factor de confusión observado más importante, si es menor asumir que es difícil que esto ocurra, es difícil que una variable no observada cambie las conclusiones y, por tanto, se estima que el resultado es robusto. Otros métodos incluyen técnicas de *machine learning*, corrección por variables instrumentales (variables relacionadas con la asignación al tratamiento, pero no con el resultado, por ejemplo, la distancia al hospital), etcétera.

■ 6.3.6. Bootstrap

Bootstrap es una técnica de remuestreo que genera múltiples muestras a partir de una muestra. Para generar cada una de estas muestras, se seleccionan aleatoriamente valores de la muestra de partida con posibilidad de repetirse algunos de ellos, ya que se generan muestras normalmente con el mismo número de individuos que la muestra inicial. En cada muestra que se va generando se toman individuos diferentes de forma aleatoria, pero todos incluidos en la muestra de partida. Mediante esta técnica se puede calcular la estimación del efecto (ej. diferencia entre grupos, *odds ratio* (OR), RR) en cada una de las muestras creadas y repetir el mismo proceso n veces, creando " n " valores de estimación del efecto, que seguirán una distribución normal y para el que podremos calcular la media y su intervalo de confianza del 95 % (IC95 %). En el caso de estimación de una OR, por ejemplo, en un modelo de regresión logística $Y = a + b_1x_1 + b_2x_2 + \dots$. Se puede estimar mediante bootstrap el valor de b_1 , y como e^{b_1} es la OR corregida por los otros factores ($x_2 \dots$), podemos estimar la OR mediante bootstrap y su IC95 %. Bootstrap es especialmente interesante para el cálculo del IC95 %.

■ 6.4. CONCLUSIONES

- Generalmente, el efecto de un factor de riesgo sobre una enfermedad se evalúa en estudios epidemiológicos y el efecto de una intervención sanitaria sobre una enfermedad mediante ensayos clínicos.
- Los estudios observacionales, basados en la obtención de datos de la práctica clínica real son necesarios en la evaluación de resultados en salud.
- Las medidas de asociación entre el factor de riesgo/intervención sanitaria y la enfermedad; si la relación es causal, constituyen medidas del efecto del factor de riesgo/intervención sanitaria. Las más utilizadas son el riesgo relativo, la *odds ratio* y el riesgo atribuible.
- Las técnicas estadísticas, requeridas en cualquier investigación, permiten determinar si las relaciones causales o de asociación existen.

- Los contrastes de hipótesis son técnicas estadísticas que se utilizan para decidir entre dos posibilidades, por ej entre el grupo con y sin intervención, y obtener respuestas en términos de probabilidad.
- Los estudios no pueden analizar muchas variables, para evitar falsos positivos se debe utilizar cada estudio para analizar solo el tema para el que ha sido diseñado.
- Las principales limitaciones de los estudios observacionales son la asignación no aleatoria, disminuyendo la comparabilidad de los grupos, y la falta de información, pudiendo llevar asociados sesgos que impidan determinar una relación causal entre el factor y el efecto.
- Es necesario corregir el efecto de posibles factores de confusión para llegar a conclusiones válidas, para ello se pueden utilizar estrategias como la exclusión, el emparejamiento, la estratificación, los modelos, el índice de propensión.
- Cualquier estudio científico debe apoyarse en la investigación observacional o experimental de datos y utilizar las técnicas de análisis adecuadas a los mismos.
- Un investigador o analista debe tener muy claro las limitaciones de los estudios y utilizar estrategias para controlar su efecto.

■ REFERENCIAS

- BAKER, M. (2016). Statisticians issue warning over misuse of P values. *Nature*, 531(7593), p. 151.
- DOLL, R., PETO, R., WHEATLEY, K. *et al.* (1994). Mortality in relation to smoking: 40 years' observations on male British doctors. *BMJ*, 309(6959), pp. 901-911.
- GIL MARTIN-THOMAS, J. F. M., LIEW, D., HOPPER, I. (2018). Observational studies and their utility for practice. *Aust Prescr*, 41(3), pp. 82-85.
- GLYNN, R.J. (2017). Use of propensity scores to design observational comparative effectiveness studies. *J Natl Cancer Inst*, 109(8), p. djw345.
- SPRINT RESEARCH GROUP, WRIGHT, J, T. JR., WILLIAMSON, J. D., WHELTON, P. K. *et al.* (2015). A Randomized Trial of Intensive versus Standard Blood-Pressure Control. *N Engl J Med*, 373(22), pp. 2103-2116.
- STUART, E.A. (2010). Matching methods for causal inference: a review and look forward. *Stat Sci*, 25(1), pp. 1-21.



7

METODOLOGÍA DISPONIBLE PARA OBTENER RESULTADOS EN SALUD EN EL MUNDO REAL

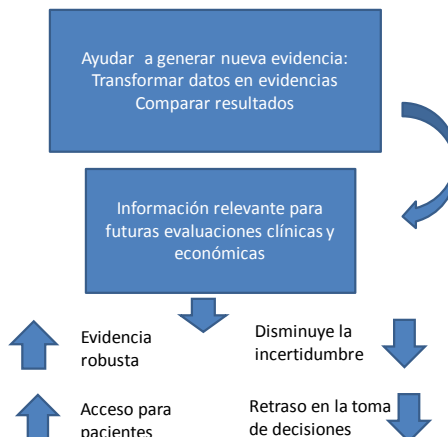
M^a. DOLORES FRAGA, CRISTINA GASTALVER, MÓNICA SALDAÑA,
PERE VENTAYOL

7.1. ENSAYO CLÍNICOS PRAGMÁTICOS

La evidencia sobre los beneficios y riesgos de los tratamientos en el cuidado sanitario puede ser obtenida de diferentes tipos de investigación que pueden agruparse, principalmente, en ensayos clínicos con asignación aleatoria (ECA) o estudios observacionales. Para la mayoría de los nuevos tratamientos, la evidencia en el momento de la aprobación por las agencias reguladoras no es completa y en muchos casos insuficiente para una adecuada toma de decisiones (Dickson *et al.*, 2018), por lo que los datos de la práctica clínica real emergen como una ayuda complementaria a los ECA para esta toma de decisiones (Calvert, Wood y Freemantle, 2011) (Figura 1).

Figura 1.

EXPLOTACIÓN DE LOS DATOS DE LA VIDA REAL PARA LA AYUDA EN LA TOMA DE DECISIONES Y GENERACIÓN DE NUEVAS EVIDENCIAS



Los datos de la práctica clínica real (RWD) pueden ser analizados para producir evidencias del mundo real, es decir, resultados en salud en el mundo real (RWE). Estas evidencias se pueden obtener con datos observacionales de la práctica clí-

nica diaria a través de la investigación de la efectividad comparativa, mediante, entre otros, grandes estudios de cohortes, estudios de registros o estudios retrospectivos en bases de datos (IOM [Institute of Medicine], 2011). Sin embargo, la información sin asignación aleatoria tiene importantes limitaciones y sesgos y, en este contexto, los ensayos controlados pragmáticos (también denominados *ensayos clínicos aleatorizados de efectividad comparativa*) ofrecen la oportunidad de combinar los datos de la vida real procedentes de un estudio observacional con el rigor científico de un ECA y de esta forma dar respuesta a preguntas relevantes en la práctica del día a día. En definitiva, los ensayos pragmáticos tienen como objetivo generar evidencias de la vida real sobre los efectos relativos de los tratamientos, generalizables a la práctica clínica real y ayudar a la toma de decisiones de clínicos y decisores.

■ 7.1.1. Diferencias entre ECA explicativos y ECA pragmático

Schwartz y Lelouch (1967) introdujeron el término pragmático para los ECA hace poco más de medio siglo y fueron los primeros en recomendar su uso, aunque reconocen que la mayoría de los ensayos contienen elementos explicativos y pragmáticos. Existe un *continuum* entre el ensayo explicativo o de eficacia (realizado en condiciones experimentales casi ideales) y el pragmático o de efectividad, más cercano a las condiciones de la vida real.

A diferencia del ECA explicativo que investiga sobre si una intervención produce un resultado de interés en circunstancias ideales, el ECA pragmático proporciona información sobre los resultados de las alternativas clínicas del mundo real en la práctica clínica diaria. Por otro lado, mientras que los ECA explicativos están diseñados para prevenir sesgos y asegurar la validez interna, los pragmáticos están centrados en maximizar la validez externa, su intención es la generalización de los resultados a todos los ámbitos. Estos últimos, sin embargo, pretenden preservar la validez interna; asignan aleatoriamente pacientes a dos o más intervenciones de uso en la práctica clínica, con el menor número de cambios en esta, para preservar la asignación aleatoria y asegurar representatividad de la práctica clínica. Los ECA explicativos son el *gold standard* para la aprobación de nuevos medicamentos, sin embargo el paradigma para la efectividad comparativa de los medicamentos en el mundo real es sobre todo pragmático.

El ensayo pragmático pretende demostrar que un medicamento es efectivo frente a la medicación estándar mediante su uso en una población heterogénea, muy poco seleccionada (pacientes de la práctica clínica rutinaria, con otras patologías, con medicación concomitante), que seguirá un protocolo sencillo (visitas, análisis, exploraciones según la práctica habitual) con flexibilidad en la administración según la práctica clínica habitual. Entre los diseños pragmáticos más habituales se pueden citar los aleatorizados en conglomerados o “cluster” (intervención a nivel grupal, la contaminación potencial hace que la aleatorización a nivel indi-

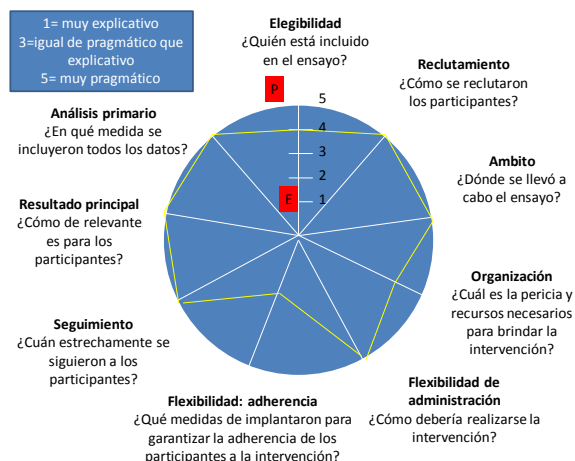
vidual sea problemática) y los diseños de cuña o *step-wedge* (todos los grupos reciben la intervención en diferentes momentos y actúan de grupos control cuando no la reciben).

7.1.2. Elementos claves de los ensayos clínicos pragmáticos

En 2008 se publicó la extensión de la declaración CONSORT para ensayos pragmáticos¹. En 2009 la herramienta PRECIS (*Pragmatic Explanatory Continuum Indicator Summary*) que intentó clarificar el concepto de pragmatismo y proporcionó una guía, un sistema de puntuación y una representación gráfica de las características pragmáticas de un ensayo. Estas características incluyen el reclutamiento de los investigadores y participantes, la intervención, seguimiento y determinación y análisis de resultados. Su actualización, la herramienta PRECIS-2 (Loudon *et al.*, 2015), consta de una rueda con nueve dominios que deben ser puntuados prospectivamente antes de la realización del ensayo y representan las decisiones sobre el diseño del ensayo (Figura 2). Muestra cuán pragmático o explicativo es un ensayo apreciando el tamaño y la forma de la figura que resulta de la conexión de las puntuaciones de cada dominio. Los ensayos explicativos producen ruedas más próximas al centro y los más pragmáticos están más cerca del contorno de la rueda.

Figura 2.

REPRESENTACIÓN ESQUEMÁTICA DE LA HERRAMIENTA PRECIS-2 (LOUDON ET AL., 2015)



Fuente: Adaptación y elaboración propia.

¹ <http://www.consort-statement.org/extensions/overview/pragmatic-trials>

Entre los elementos claves que impactan en el éxito de los ensayos pragmáticos se puede citar²:

- **Selección del lugar de realización.** Aspectos que considerar durante la selección del lugar de realización son el ámbito geográfico, las vías clínicas de tratamiento, la infraestructura y las características de los participantes. El cuidado de salud habitual puede no ser igual en diferentes ámbitos. Además de variaciones en el cuidado sanitario que puede encontrarse en las diferentes localizaciones geográficas, los pacientes pueden estar atendidos en diferentes ámbitos (primaria y especializada). Esto plantea el desafío de garantizar que se capturen todas las fuentes de datos relevantes para garantizar que los datos recopilados representen por completo la atención habitual del paciente.
- **Selección de los pacientes.** Todos los pacientes que son potenciales candidatos para tratamiento en la práctica clínica habitual deberían ser elegibles para un ensayo pragmático. Estos ensayos están focalizados en la inclusión y no en la exclusión de pacientes, ya que la homogeneidad puede entrar en conflicto con la relevancia para la práctica diaria. Si se espera una importante heterogeneidad entre subgrupos, se puede planificar un ensayo específico de tamaño suficiente para dirigirse al subgrupo de interés. Alternativamente, el número relativo de pacientes en estos subgrupos se puede aumentar en el ensayo heterogéneo y tenerlo en cuenta en un análisis agrupado (Knottnerus y Tugwell, 2017). Como en cualquier ensayo, un bajo reclutamiento y pérdida de seguimiento puede introducir selección y reducir la generalizabilidad, por lo que debe maximizarse la adherencia.
- **Consentimiento informado.** El Get Real consortium de la Innovative Medicines Initiative³ tiene el propósito de desarrollar estrategias para incorporar las evidencias del mundo real de manera temprana en el ciclo de vida de los medicamentos para informar mejor las decisiones sanitarias sobre los riesgos y beneficios comparativos de los nuevos medicamentos. Uno de los problemas para realizar ensayos clínicos pragmáticos es la necesidad legal de obtener el consentimiento informado. Dado que los modelos de consentimiento informado pueden llevar a dificultades en el reclutamiento, reducir la generalizabilidad de los resultados y producir sesgo de selección, proponen cuatro modelos alternativos: el consentimiento integrado (integra el consentimiento clínico y el de investigación), el consentimiento específico (se da información adicional, pero no se informa de proceso de asignación aleatoria), el enfoque de difusión y una renuncia al consentimiento.
- **Comparadores en la vida real.** En los ensayos pragmáticos el cuidado habitual o el tratamiento estándar es el comparador de elección, pero la imple-

² <http://www.imi-getreal.eu/>

³ <http://www.imi-getreal.eu/>

mentación puede ser difícil si el cuidado habitual difiere entre centros o países. Debería de representar el procedimiento de referencia más apropiado (Knottnerus y Tugwell, 2017). EunetHTA ha publicado recomendaciones para la elección de comparadores para la efectividad comparativa que que son relevantes para los ensayos pragmáticos⁴.

- **Medidas de resultados.** Los ensayos pragmáticos deben centrarse en resultados relevantes para los pacientes, para la práctica clínica y la elección de tratamientos. Las medidas de resultados no deben interferir con la práctica clínica habitual para asegurar la generalizabilidad de los resultados de los ensayos, y deben priorizarse los resultados recogidos rutinariamente. Las variables subrogadas deberían ser evitadas. Los típicos resultados incluyen mortalidad, morbilidad, estado funcional, bienestar y uso de recursos sobre un horizonte temporal relevante independientemente de si el protocolo del tratamiento es seguido. Los ensayos pragmáticos a menudo son abiertos, por lo que la inclusión de medidas de resultados objetivas puede reducir el riesgo de sesgo. Cuando se espera un riesgo de sesgo elevado o elevada variabilidad debería considerarse la estandarización y el cegamiento de la evaluación de resultados (Knottnerus y Tugwell, 2017). La selección de resultados en ensayos pragmáticos debería discutirse con todas las partes interesadas para asegurar la relevancia de los resultados del estudio para la toma de decisiones e la práctica clínica y la capacidad para realizar el estudio.
- **Seguridad, calidad y monitorización.** En los ensayos pragmáticos las vistas de seguimiento se minimizan lo que complica la recogida estructurada de datos de seguridad, es necesario establecer un proceso que no interfiera con la práctica clínica habitual. La principal consideración en el plan de monitorización es definir el manejo de los riesgos de manera que no interfiera con la práctica rutinaria, entre las opciones potenciales se pueden citar, entre otros, el registro en la historia clínica electrónica (HCE) y los resultados comunicados por los pacientes.
- **Recogida y manejo de datos.** Debe evitarse la interferencia con la práctica habitual o por lo menos limitarse, por lo que debe considerarse los registros en la HCE y el uso de sistemas electrónicos rutinarios para la recogida y manejo de los datos. El acceso a los datos y la privacidad, el momento para la recogida de datos, el nivel de detalle de estos y la falta de una clara comprensión del proceso de recogida de datos son unos de los principales desafíos para el uso de los datos recogidos de manera rutinaria. Iniciativas internacionales como el *TRANSFoRm Project*⁵ o el *Electronic Health Record sfor Clinical Research (EHR4CR)*⁶ están trabajando para mejorar la intero-

⁴ https://www.eunetha.eu/wp-content/uploads/2018/03/Criteria_WP7-SG3-GL-choice_of_comparator_amend2015.pdf

⁵ <https://www.i-hd.eu/index.cfm/resources/ec-projects-results/transform/>

⁶ <http://www.ehr4cr.eu/>

perabilidad de las fuentes de datos y desarrollando plataformas que combinen datos de los diferentes hospitales en distintos países. Otros como el PCORnet están centrados en los pacientes⁷ y el eClinicalconsortium están elaborando una herramienta para evaluar la calidad y seguridad de los registros de la HCE⁸.

Los ensayos clínicos pragmáticos puede ser una opción para dar respuesta a preguntas sobre efectividad comparativa, pero deben formar parte de la práctica clínica diaria intentando que los resultados puedan ser derivados de registros o historias clínicas electrónicas en las que estén recogidas las variables necesarias para obtener resultados y que no se sobrecargue el trabajo diario asistencial. Sería deseable que tras la evaluación crítica de las evidencias que motivaron la aprobación se diseñen formularios con variables a recoger en la práctica diaria y que su seguimiento nos oriente sobre los RWE y de respuesta a las incertidumbres presentes en el momento de la aprobación. Es una curva de aprendizaje en la que debe integrarse investigación y datos de la vida real para construir el conocimiento más robusto. Es necesario disponer de herramientas para valorar la calidad y fiabilidad de los ECA pragmáticos y unificar los criterios para obtener los mejores resultados con los diseños más apropiados. Además, los protocolos de los ECA pragmáticos con medicamentos mediante la HCE deben ser evaluados por los Comités Éticos de Investigación con medicamentos (CEIm) dentro de la categoría de ensayo clínico de bajo nivel de intervención (Real Decreto 1090/2015).

7.2. ESTUDIOS OBSERVACIONALES

Según el Real Decreto Legislativo 1/2015, de 24 de julio, por el que se aprueba el texto refundido de la Ley de garantías y uso racional de los medicamentos y productos sanitarios (Real Decreto Legislativo 1/2015) se denomina estudio observacional al “estudio en el que los medicamentos se prescriben de manera habitual, de acuerdo con las condiciones establecidas en la autorización. La asignación de un paciente a una estrategia terapéutica concreta no estará decidida de antemano por el protocolo del estudio, sino que estará determinada por la práctica habitual de la medicina, y la decisión de prescribir un medicamento concreto estará claramente dissociada a la de incluir al sujeto en el estudio. No se aplicará a los pacientes ninguna intervención, ya sea diagnóstica o de seguimiento, que no sea la habitual en la práctica clínica y se utilizarán métodos epidemiológicos para el análisis de los datos recogidos”. Los estudios observacionales, por tanto, difieren de los ensayos clínicos en que se llevan a cabo en condiciones de práctica clínica habitual y que carecen de intervención por parte de los investigadores del estudio. La realización de estudios observacionales permite conocer el efecto de fármacos, dispositivos y comportamientos sobre los que no sería ético aplicar la intervención del investiga-

⁷ <https://pcornet.org/>

⁸ <http://eclinicalforum.org/eSRA/esra-esource-readiness-assessment-handbook-and-assessment-template>

dor para realizar estudios de tipo ensayo clínico. Además, de ellos se obtiene información que los ensayos clínicos no aportan, como el comportamiento del fármaco en poblaciones especiales y durante un mayor periodo de tiempo que en los que se realizan los ensayos clínicos, por lo que es una herramienta básica en la farmacovigilancia de los fármacos y productos sanitarios. Esta información obtenida en estos estudios es esencial para orientar la práctica clínica y favorecer el uso racional de los medicamentos.

Dentro de los estudios observacionales hay que destacar la importancia de los estudios posautorización. En estos, el objeto del estudio es un medicamento ya comercializado que se emplea en las condiciones autorizadas en su ficha técnica. El objetivo principal de estos estudios es generar información adicional sobre el efecto de los medicamentos en condiciones normales de uso, con el fin de completar la información obtenida en los ensayos clínicos de fase I, II y III y contribuir a su mejor utilización. Para ello, estos estudios se centran en determinar la efectividad de los fármacos en condiciones normales de uso, en poblaciones especiales (ancianos, niños, embarazadas) y en presencia de factores que pueden modificarla (como enfermedades concomitantes o la presencia de otros fármacos); determinar los efectos adversos del fármaco, especialmente los no identificados en los ensayos clínicos y los que ocurren a largo plazo; evaluar la eficiencia de los medicamentos y establecer su lugar en la terapéutica atendiendo al uso racional de los medicamentos; obtener información sobre patrones de uso de los medicamentos y valorar los fármacos desde el punto de vista de los pacientes. Los estudios posautorización de tipo observacional tiene la obligación de estar inscritos en el Registro Español de Estudios Clínicos (REEC).

Al igual que los ensayos clínicos, los estudios observacionales requieren ser autorizados por las autoridades competentes antes de su inicio, lo que corresponde a la Agencia Española de Medicamentos y Productos Sanitarios (AEMPS) y los Comités Éticos de Investigación con medicamentos (CEIm) de los centros participantes (Orden SAS/3470/2009). Estos organismos evaluadores comprueban que el estudio sea ético y metodológicamente correcto (cumpliendo con la Declaración de Helsinki y Convenio del Consejo de Europa para la protección de los derechos humanos y la dignidad del ser humano con respecto a las aplicaciones de la biología y la medicina), que el estudio permita la obtención de resultados fiables y sólidos, que los beneficios para el sujetos de ensayo o para la salud pública supere con creces los posibles inconvenientes o perjuicios que puedan derivarse del estudio a los participantes del mismo, que se respeten los derechos del sujeto y su intimidad, que el sujeto esté debidamente informado y consienta libremente participar en el estudio (Reglamento (UE) 2016/679). Para la evaluación de todos estos aspectos, el promotor del estudio deberá remitir al CEIm y a la AEMPS la documentación del estudio.

La documentación del estudio está constituida por el protocolo del estudio, que incluye los siguientes elementos: identificación del promotor y del investigador coordinador del estudio, listado de centros participantes en el estudio, objetivos del

estudio, la población objeto del estudio, el número de sujetos a incluir, tiempo de duración del estudio, procedimiento metodológico que se va a emplear para la consecución de los objetivos, cuaderno de recogida de datos, cronograma del estudio, el análisis estadístico que se va a llevar a cabo para la obtención de los resultados, financiación del estudio y memoria económica, entre otras. Dos de las partes más importantes del protocolo son la Hoja de información al paciente y el Consentimiento Informado, que demuestra que cada uno de los sujetos participantes en el estudio ha sido debidamente informado y participa de forma libre y voluntaria en el mismo.

La AEMPS es la encargada de clasificar el estudio en los distintos tipos de estudios observacionales existentes:

- EPA-LA: aquellos estudios realizados por el laboratorio titular de un medicamento como condición impuesta por la AEMPS en el momento de la autorización de dicho medicamento.
- EPA-AS: estudios promovidos por las administraciones sanitarias o financiados con fondos públicos.
- EPA-SP: estudios observacionales prospectivos con medicamentos que no cumplan criterios de EPA-LA ni EPA-AS.
- EPA-OD: otros estudios posautorización de tipo observacional que tenga un diseño diferente al de seguimiento prospectivo.
- No EPA: estudios observacionales que no sean posautorización. El principal objeto de estudio no son los medicamentos (como, por ejemplo, los estudios de adherencia al tratamiento)

Una vez obtenida la autorización –si procede– y la clasificación por parte de la AEMPS, o de forma paralela a esta, el estudio debe ser evaluado por los CEIm de los centros en los que se va a llevar a cabo el estudio, que emitirán su correspondiente dictamen. En los estudios en donde participa más de un centro, únicamente es necesaria la evaluación del estudio por uno de los CEIm (que se denominará CEIm de referencia), cuyo dictamen será único y vinculante en todo el territorio nacional. No obstante, los CEIm de los demás centros participantes podrán requerir la documentación del estudio para comprobar que el desarrollo de este es posible en su centro.

En los estudios clasificados como EPA-SP se requiere, además, de la autorización de cada una de las comunidades autónomas en donde va a realizarse el estudio.

Una vez autorizado el estudio por todos los organismos competentes, el promotor deberá obtener la conformidad del centro participante y, de forma general,

aunque no necesario en todos los casos, firmar un contrato con el centro para poder comenzar el estudio en el mismo.

Una vez iniciado el estudio, el promotor deberá remitir a las agencias evaluadoras informes de seguimiento anuales que expongan cómo se está desarrollando el estudio y un informe final tras la finalización de este. En todos los estudios observacionales es necesario conservar la documentación generada durante un periodo de 5 años tras la finalización del estudio.

■ 7.3. REGISTRO DE PACIENTES

Generamos más conocimiento del que podemos absorber, esto que puede parecer una amenaza usado en nuestro favor puede convertirse en una fortaleza. El crecimiento de la tecnología la obtención de información es exponencial. Uno de los problemas es que la proyección de nuestra mente es lineal y no exponencial con lo que es una cuestión principalmente de aceleración y no de velocidad: “las cosas van despacio hoy en relación con como irán el año que viene”. *Big data* y RWD van a revolucionar nuestro sistema sanitario del futuro, y no solo en la forma en como vivimos sino también en como trabajamos y nos interrelacionamos (U.S. Department of Health and Human Services Food and Drug Administration, 2015).

Los grandes cambios que acontecen tras la aparición de una innovación no la percibimos inmediatamente, pero se incorpora de una forma abrupta tras un tiempo de letargo.

Volúmenes de datos tan grandes y heterogéneos no pueden ser manejados con el *software* tradicional, ni fácilmente analizarse con las herramientas convencionales de gestión de datos

Los registros de pacientes reflejan la atención real que reciben los pacientes en cada contexto y los resultados clínicos que realmente obtienen. Se corresponden con estudios observacionales prospectivos, de cohortes de pacientes que presentan una patología concreta o están recibiendo un determinado tratamiento o intervención. Entre otras utilidades permiten un mayor conocimiento de la historia natural de la enfermedad, evaluar la seguridad en condiciones reales, conocer la calidad de la atención recibida y/o para poder realizar estudios económicos (fundamentalmente de coste-efectividad). Los registros de pacientes suponen el tener que recoger datos de la información clínica y económica de forma prospectiva e idealmente en tiempo real. Su interés radica en que incluyen un grupo mayor y diverso de pacientes que los ECA en fase III, por lo que reflejan mejor los datos reales relacionados con los pacientes, permiten evidenciar distintas maneras de gestión y evaluar sus resultados. Ello supone que los datos clínicos de los pacientes sean registrados durante un periodo largo de tiempo, permitiendo una evaluación de los resultados a largo plazo. La mayoría de los registros requieren de la creación de un registro en bases de datos que prácticamente no precisan mantenimiento, lo que abarata sus costes; pero presentan limitaciones en cuanto a la necesidad de un registro manual periódico, la

cantidad/calidad de datos que se registran son limitados y los intervalos de recogida de estos pueden variar sensiblemente. Idealmente estos registros deberían establecerse de forma *online* a medida que se trabaja con el registro habitual en la historia clínica del paciente, sin la necesidad de incorporarlos a posteriori o en un sistema de registro paralelo al establecido como HCE (European Commission, 2014).

Figura 3.

CLASIFICACIÓN DE LAS BASES DE DATOS EN FUNCIÓN DE LA FUENTE Y SU ESTRUCTURACIÓN

Fuentes de datos	Externas	Los datos externos estructurados pueden proceder de otras empresas, censo poblacional, banco, aseguradoras,....	Es el área más difícil de utilizar por ser ajena como por la dispersión y variedad de los datos Ej: mensajes en redes sociales
	Internas	Registros de pacientes en HCE general o específicos (registros oncológicos, nutrición parenteral, pacientes externos, etc...	Datos clínicos en papel, notas manuscritas, pruebas funcionales, imágenes
		Estructuradas	No estructuradas
Tipos de datos			

Existen tres grandes tipos de registros de pacientes, los estructurados, los desestructurados y los mixtos. A su vez estos se catalogan entre registros procedentes de fuentes de datos internas y externas. Ver ejemplos en la Figura 3.

Características como la velocidad, variedad y volumen de datos pueden afectar a la cantidad de datos registrados, mientras que la veracidad, la validez y el valor de estos datos afectan a la calidad de estos. Es necesario disponer de recursos que faciliten por una parte la exhaustividad y por otra una supervisión adecuada que permita asegurar la validez de los datos obtenidos (IMS Institute for Healthcare Informatics, 2015).

7.3.1. Datos administrativos

Son habitualmente retrospectivos o a tiempo real. Pueden contener datos sobre diagnóstico, procedimiento, uso clínico, y costes permitiendo establecer al paciente como unidad de análisis, o bien por subgrupos o poblacional.

■ 7.3.2. La historia clínica electrónica

Desde la conversión del papel a su transformación digital mediante el escaneo de las historias clínicas en papel la digitalización de los datos ha supuesto un gran avance en la medicina: la información puede ser consultada a distancia y adicionalmente se evita el extravío y la posible manipulación a posteriori. Un paso más allá lo constituyeron las historias clínicas electrónicas (HCE) basadas en introducir toda la información clínica (registros de visitas clínicas, imágenes, analíticas, pruebas funcionales, ...), demográfica y administrativa relacionada con el paciente de la forma más estructurada posible en bases de datos, de modo que potencialmente pudiera ser explotada mediante herramientas informáticas.

La realidad en nuestro país es que en cada CC. AA. han ido desarrollando distintos sistemas de HCE, casi siempre incompatibles entre sí, algunas de ellas estructuradas, otras no estructuradas y la gran mayoría mixtas lo que dificulta disponer de datos globales cuyos resultados, aunque prometedores para la atención sanitaria, pierden valor por quedar confinados en una o pocas CC. AA.

■ 7.3.3. La receta electrónica

Es el servicio de prescripción y dispensación electrónica utilizado por los servicios de Salud de implantación en todas las autonomías en los últimos años y que permite a los pacientes obtener los medicamentos recetados simplemente presentando su tarjeta sanitaria en cualquier farmacia.

Para acceder a este servicio, entregando al médico la tarjeta sanitaria, este accede de forma informatizada a todos los datos del historial médico y actualiza la medicación. Este informa sobre todos los aspectos del tratamiento, incluyendo los cambios en el estilo de vida, los beneficios esperados, las posibles complicaciones... y sobre todo la importancia de tomar los medicamentos según la dosis, el horario y la duración que le indique. De esta manera, cuando el paciente acude a la farmacia solamente tiene que presentar la tarjeta sanitaria ya que desde allí se tiene la capacidad de acceder por vía telemática a la información farmacéutica prescrita en la HCE del enfermo, se le suministra la medicación y realiza la atención farmacéutica pertinente.

Sin lugar a dudas la receta electrónica es una fuente de información para evaluarla adherencia de los pacientes a un determinado tratamiento. Conocer qué medicamentos han sido prescritos por el facultativo y comprobar si se han retirado de las farmacias y cuando indirectamente informa del grado de adherencia al mismo, complementando los datos de prescripción de la HCE (Del Llano Señarís *et al.*, 2016; García López *et al.*, 2014).

■ 7.3.4. Encuestas de salud

La información proporcionada por las encuestas de salud es de gran utilidad para la planificación, la toma de decisiones y la evaluación de políticas de salud. A nivel local pueden llegar a facilitar indicadores que son necesarios para el conocimiento del estado de salud autopercebido por cada individuo de una población, sus hábitos de vida y la utilización de los servicios sanitarios, que pueden ser utilizados en los planes de salud y la programación sanitaria.

Aparte de este conocimiento de la percepción del estado de salud de las personas y la distribución en la población de los diferentes problemas de salud así como la valoración de los estilos de vida de la población, y la existencia de posibles desigualdades, una encuesta de salud permite que la población objeto de estudio pueda expresar sus opiniones, actitudes, estilos de vida, creencias y conductas en relación a los factores implicados en la salud. Desde el sistema sanitario, directamente, podemos acceder a múltiples informaciones sobre morbilidad, frecuentación, pero es escasa la información que tenemos de la opinión sobre la salud, la enfermedad, el proceso asistencial de los ciudadanos. Tampoco conocemos con los registros asistenciales cuáles son las creencias y las actitudes que generan las mismas respecto a actividades de prevención, tratamiento, etc. Solo preguntando a la población sobre determinadas cuestiones llegaremos a percibir cuáles son los ámbitos de actuación para mejorar nuestra prevención y atención sanitaria.

Otro objetivo fundamental en la puesta en marcha de encuestas de salud, desde una óptica de salud pública es el intentar identificar grupos de riesgo y las desigualdades en salud, sea en cuanto morbilidad, factores de riesgo, estilos de vida con objeto de establecer actuaciones específicas para eliminar ese riesgo y disminuir si fuera posible las desigualdades.

Y finalmente, la encuesta como herramienta de planificación es un elemento de evaluación si somos capaces de tener una serie de encuestas donde podamos analizar la evolución de determinadas variables de la población recogida. Evolución que nos permitirá monitorizar actuaciones implantadas, así como vigilar la aparición de situaciones de riesgo o desigualdades de salud.

Generalmente solo se realizan preguntas, pero existen encuestas de salud que acompañan a la cumplimentación de un cuestionario cómo la realización de mediciones antropométricas (peso, talla, etc.), o biológicas (toma de muestras clínicas...) o realización de exploraciones médicas o clínicas... aunque no son habituales este tipo de encuestas, aportan mucha más información que la indicada por la persona entrevistada, si bien llegan a ser muy caras sobre todo si la población muestral es grande.

Una vez recogida toda la información es necesario depurar y completar, si fuera necesario, la misma, de manera que resulte homogénea y, a continuación, proceder a incluirla en una base de datos al objeto de comenzar su análisis. Este

se puede realizar para determinar la frecuencia de cada variable por separado o con objeto de realizar asociaciones entre determinadas circunstancias (por ejemplo, situación socioeconómica u otros determinantes) y los resultados de salud (por ejemplo, morbilidad, utilización de servicios sanitarios, etc.) o para el análisis específico de situaciones concretas.

Un elemento importante es el determinar evolución de los indicadores e información si se realizaron varias encuestas consecutivas en el tiempo. De ahí la necesidad de realizar encuestas de salud periódicamente. Con ello podemos evaluar la evolución de comportamientos y valorar las actuaciones realizadas respecto a los mismos.

La aportación del RWD en España respecto a las encuestas de salud se centra en la posibilidad de sustituir estas encuestas por los registros de los HCE, lo que evitaría el tener que conseguir presupuestos para la realización periódica de las mismas, pero también respecto a aspectos concretos mucho más técnicos como la inclusión en las encuestas a la población institucionalizada, población inmigrante, etcétera.

Las encuestas de salud son instrumentos válidos para el suministro de información de la que carecen las distintas instituciones para elaborar sus políticas e intervenciones relacionadas con la salud, en especial aquellas derivadas de los elementos donde la opinión, la creencia y la actitud de la población es escasamente conocida. Son instrumentos con determinados tipos de limitaciones que una vez conocidos pueden ser controlados y facilitar su uso más frecuente en las políticas de salud a nivel local, regional, nacional o internacional.

Como ventajas cabe destacar características como la cobertura poblacional, permiten obtener información sobre determinantes, proporcionan datos biológicos además de hábitos de vida, sociodemográficos y contextuales y como desventajas los costes asociados a su ejecución y la falta de flexibilidad como mecanismos de evaluación en salud comunitaria

■ 7.3.5. Tecnologías “wearables”

Hay además una manera de generar datos que va a ser bidireccional, los dispositivos portátiles y o nuevos portables y diferentes “gadgets” van a permitir trasladar información clínica, farmacéutica y demográfica de forma continua desde el punto de atención donde se encuentre en paciente. En los últimos años estas tecnologías “wearables” han experimentado un uso creciente. Son pequeños dispositivos electrónicos que se llevan encima (relojes, pulseras, gafas, lentillas, etc.), dentro (en forma de implante subcutáneo, sonda insertada o administrada que se deposita en un órgano, o formado parte del tejido de la ropa y que proporcionan información, mide parámetros o interactúa con otros aparatos midiendo entre otras cosas la tem-

peratura corporal, el ritmo cardiaco o permiten liberar un medicamento, como, por ejemplo, la insulina.

Tecnología como los drones van a permitir conocer el paradero del paciente y hacerle llegar la medicación o dispositivos necesarios en base a los cinco correctos: paciente correcto, hora correcta, medicamento correcto dosis correcta y vía de administración correcta, ello junto a la domótica y robotización permitirá su preparación y administración en base a criterios de seguridad y de calidad en el domicilio habitual. La trazabilidad en aspectos como adherencia, caducidad, lote y dosis administradas, así como la monitorización en tiempo real de la cinética y los resultados en salud serán automáticos y continuos.

La parametrización por bioidentificación con lectores de pupilas y/o de huella digital permitirá identificar al paciente correcto sin margen de error, así como se establecerá una parametrización de la dosis correcta en base a características clínicas, farmacogenéticas, farmacocinéticas y biológicas precisas, todo ello a través de una aproximación holística y en el que la biología de sistémicas u “ómicas” habrán explotado.

■ 7.3.6. Utilidad y debilidades de los registros de pacientes

En el futuro, el gran reto no estará en diseñar dispositivos, que habrá muchos y variados, sino en cómo procesar la gran cantidad de datos que estos generarán.

Recoger información (acumular datos), analizarla (transformar datos en conocimiento) y predecir/prescribir (detectar prioridades, evitar enfermedades y actuar de forma más precisa cuando estas se producen contribuirá sin duda a conseguir una medicina de la precisión, orientada al paciente tanto individualmente como parte constituyente de un colectivo o población)

Las utilidades de los registros de pacientes permitirán: predecir hospitalizaciones, en base a factores climáticos como la humedad o la polución en relación con la epidemiología de las enfermedades, o en base a factores poblacionales; identificar pacientes de alto riesgo (clasificación en niveles de los paciente en cuanto a riesgos de salud), toma de decisiones en la consulta a tiempo real analizando casos similares y proponiendo alternativas de manejo y reducir la variabilidad en la práctica médica; análisis del estado de salud de una población (desarrollar modelos predictivos para distintas enfermedades o para prevenir hábitos no saludables); seguimiento de tendencias (controlar factores de riesgo y/o mejorar la adherencia a los tratamientos) y finalmente comparar tratamientos para una misma condición y establecer el papel de cada uno en los diferentes pacientes, un aspecto con importantes implicaciones en efectividad, seguridad y costes.

Las debilidades están relacionadas con la privacidad, la propiedad (de quien son los datos), problemas técnicos: interoperabilidad de sistemas, y fragmentación

de los datos, precariedad y debilidad de estructuras investigadoras y finalmente cuestiones de transparencia y accesibilidad.

7.4. BASES DE DATOS

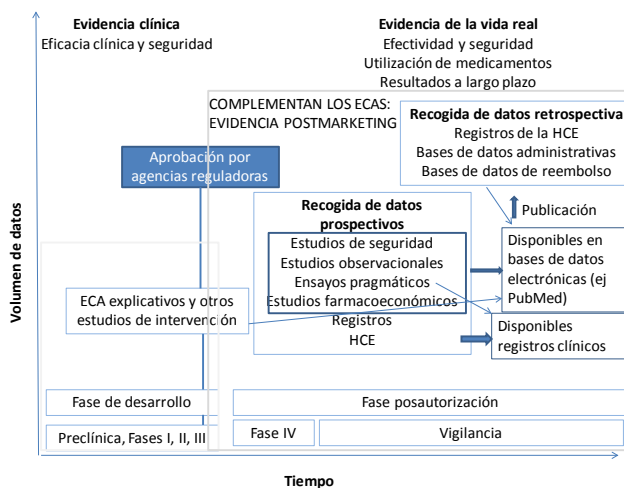
7.4.1. Fuentes de datos en el ciclo de vida de los medicamentos

Tal como ya se ha comentado en apartados anteriores, los RWD se refieren a los datos asociados con la salud de los pacientes recogidos de fuentes de datos diferentes a los ECA explicativos. Los RWD tiene su origen en una variedad de fuentes asociadas con, o utilizadas en, los diferentes ámbitos de la práctica clínica diaria y su análisis produce RWE (Katkade, Sanders y Zou, 2018).

Antes de hacer una revisión de las distintas bases de datos e iniciativas disponibles a nivel internacional y nacional, es necesario revisar las diferentes evidencias disponibles en el ciclo de vida de los medicamentos y así determinar la fuente de datos más adecuada para dar respuesta a la pregunta de investigación. Una aproximación sistemática, tanto en la búsqueda como en la selección de la base de datos es fundamental para conseguir los mejores resultados. La Figura 4 muestra la evidencia disponible durante el ciclo de vida de los medicamentos.

Figura 4.

EVIDENCIA DISPONIBLE DURANTE EL CICLO DE VIDA DE LOS MEDICAMENTOS



Fuente: Adaptación de Katkade, Sanders y Zou (2018).

La importancia e influencia de los RWD está ampliamente demostrada ya que en la mayoría de los países se está incrementando el desarrollo de infraestructura y tecnología para aumentar la capacidad de usar estos datos en la investigación de efectividad y seguridad comparativa, así como para valoración de la tecnología sanitaria (Wang *et al.*, 2017). Sólo abordando la búsqueda en una base de datos como PubMed centrado en los términos “real world data”, “real world evidence”, “real world data base”, RWE, RWD, databases, se localizan más de 400 referencias y si centramos en el título el número tampoco es depreciable, más de 90 referencias hasta julio de 2018.

Esta explosión de información debe ser analizada para conocer su reproductibilidad y validez. Wang *et al.* (2017) publicaron una lista de los parámetros específicos que deben comunicarse para incrementar la reproductibilidad de los estudios procedentes de la explotación de las bases de datos de salud. Más allá de las consideraciones generales para el estudio observacional, hay varios problemas únicos relacionados con la investigación de bases de datos que deberían abordarse, ya que en muchas ocasiones la comunicación de los resultados es muy conceptual y no hay suficiente información para los lectores y se desconoce cómo se incluyeron los pacientes, o como se definió la exposición, el resultado y las covariables. Los riesgos de resultados no interpretables o sesgados se pueden minimizar al proporcionar una descripción sólida de las tablas de datos utilizadas; centrándose en por qué y cómo fueron creados; medir e informar la precisión de la codificación de los diagnósticos y de los procedimientos utilizados; así como el registro adecuado de todas las variables. El sello distintivo de una buena investigación es un análisis e interpretación rigurosamente realizada, pero para ello es necesario que haya transparencia en la comunicación de los datos y que sean reproducibles para abordar nuevas investigaciones, sobre todo si va a ser la base para la toma de decisiones.

■ 7.4.2. Propuestas nacionales e internacionales de bases de datos con datos procedentes de la vida real

Los responsables de la toma de decisiones buscan cada vez más evidencia generada a partir de grandes bases de datos obtenidos de la práctica clínica habitual. Es necesario estandarizar la terminología y acordar qué parámetros son esenciales para comunicar la investigación y que esta sea reproducible y transparente, lo que permitiría evaluar su validez.

Tal como concluyen Wang *et al.* (2017) una mejora sustancial en la reproductibilidad, rigor y confianza en la evidencia del mundo real generada a partir de bases de datos de salud podría lograrse con una mayor transparencia de todo el proceso del estudio y la creación del conjunto de datos analíticos obtenidos a partir de las bases de datos de salud longitudinales.

Desde hace décadas se realizan estudios observacionales de efectividad (y seguridad) retrospectivos utilizando bases de datos unidas a los datos anonimiza-

Tabla 1.

BASES DE DATOS NACIONALES E INTERNACIONALES CON DATOS DE LA VIDA REAL (RWD)

Base de datos de RWD	Objetivo
Internacionales	
<p>Clinical Practice Research Datalink (GRPD) https://www.cprd.com/intro.asp</p>	<p>Facilitar la realización de estudios observacionales y ensayos clínicos disponiendo de los datos de 52 millones de británicos (HCE anonimizados).</p>
<p>The Learning Healthcare System http://www.learninghealthcare-project.org/section/background/learning-healthcare-system</p>	<p>Base americana. Busca integrar investigación, enseñanza y asistencia en un nuevo marco. Uno de los métodos es el <i>point-of-care</i> clinical trial, que pretende integrar los estudios de efectividad comparativa en el sistema de atención sanitaria. Proporciona enlaces a otros proyectos que utilizan uno o más de los seis casos de uso clave que cumplen con la definición de <i>The Learning Healthcare System</i> y representan áreas en las que se ha logrado un progreso significativo:</p> <p>Automatización inteligente, investigación comparativa de la efectividad, desviación positiva, vigilancia, modelo predictivo, apoyo a la decisión clínica.</p>
<p>TRANSFoRm Project https://www.i-hd.eu/index.cfm/resources/ec-projects-results/transform/</p>	<p>El proyecto nació con el propósito de desarrollar un sistema sanitario de aprendizaje rápido que pueda mejorar tanto la seguridad del paciente como la realización y el volumen de la investigación en Europa. Una interfaz dinámica está integrada con registros de salud electrónicos para identificar pacientes elegibles para la investigación y recogida de datos de resultados y seguridad.</p>
<p>Health information technology (IT) initiative https://www.ahrq.gov/professionals/prevention-chronic-care/improve/health-it/index.html</p>	<p>De la Agency for Healthcare Research and Quality de EE. UU., estrategia nacional que integra la tecnología de la información con el cuidado sanitario</p>
<p>NIHR Clinical Research Networks (NIHR-CRN) https://www.nihr.ac.uk/about-us/how-we-are-managed/managing-centres/crn/</p>	<p>El CRN hace posible que los pacientes y los profesionales de la salud en toda Inglaterra participen en estudios de investigación clínica dentro del sistema nacional de salud. Brinda oportunidades para que los pacientes obtengan un acceso más temprano a tratamientos nuevos y mejores a través de la participación en la investigación.</p>
<p>FDA Mini-Sentinel https://www.fda.gov/safety/fdassentinelinitiative/ucm2007250.htm</p>	<p>Mejorar la capacidad de la FDA para monitorizar proactivamente la seguridad de los productos médicos tras su comercialización y complementa el Sistema de Notificación de Eventos Adversos existente de la Agencia.</p>

Tabla 1. (continuación)

BASES DE DATOS NACIONALES E INTERNACIONALES CON DATOS DE LA VIDA REAL (RWD)

Base de datos de RWD	Objetivo
Nacionales	
<p>BIFAP http://www.bifap.org</p>	<p>Base de datos para la investigación farmacoepidemiológica en atención primaria. BIFAP es una base de datos informatizada de registros médicos de Atención Primaria para la realización de estudios farmacoepidemiológicos, perteneciente a la Agencia Española de Medicamentos y Productos Sanitarios, y cuenta con la colaboración de las comunidades autónomas participantes y el apoyo de las principales sociedades científicas implicadas.</p> <p>BIFAP incluye la información procedente de los pacientes atendidos por 4.910 médicos de familia y 842 pediatras de Atención Primaria del Sistema Nacional de Salud, integrando información de 7.890.485 historias clínicas anonimizadas.</p>
<p>SIDIAP http://www.sidiap.org</p>	<p>Promovida por el Instituto Catalán de la Salud (ICS) y el Instituto de Investigación de Atención Primaria. Su objetivo es generar bases de datos fiables para la investigación a partir de los registros de la historia clínica electrónica de atención primaria del ICS y otras bases de datos complementarias. Dispone de datos de más de 5,8 millones de ciudadanos.</p>

dos de la HCE en diferentes ámbitos, algunos ejemplos relevantes se muestran en la Tabla 1.

La promesa de valor de la evidencia del mundo real utilizando bases de datos en la toma de decisiones médicas debe equilibrarse con las preocupaciones relacionadas con las limitaciones de los datos. Los investigadores deberían tratar de evitar el sesgo en el diseño de un estudio, ajustar los factores de confusión y discutir los efectos del sesgo residual en los resultados, y comunicarlos adecuadamente. Otros autores proponen aproximaciones adaptativas a los datos mediante ajustes semiautomatizados y optimizados por factores de confusión para evitar los sesgos, una de las propuestas es utilizar algoritmos como el *high-dimensional propensity score* (Schneeweiss, 2018), estas técnicas se abordarán más detalladamente en el capítulo tres de este libro.

Poco a poco se está extendiendo el uso de datos de la vida real como complemento a la evidencia clínica regulatoria, por ejemplo, en algunos países como Francia, Italia, Holanda Bélgica y los Países Nórdicos es mandatorio disponer de evidencia positiva para mantener el acceso a los nuevos medicamentos (Katkade, Sanders y Zou, 2018). Aunque por el momento el uso de RWD por agencias evaluadoras de tecnologías sanitarias en Europa es dispar y parece que los RWD se utilizan más en valoraciones de coste-efectividad que para la valoración relativa

de efectividad (Makady *et al.*, 2018). Está claro que existe una necesidad de mejorar la metodología, así está disponible una guía metodológica elaborada por The European Network Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP), que recientemente actualizó el apartado sobre el uso y efectos de los medicamentos en la vida real⁹.

Estandarización, unificación de criterios, transparencia, ajuste de factores de confusión y listas guías para diseñar y comunicar los resultados en salud, entre otros, parecen una buena fórmula para sustentar la toma de decisiones sobre efectividad comparativa en el mundo real y complementar los datos procedentes de los ECAS explicativos. Las iniciativas tanto internacionales como nacionales para avanzar en el diseño de bases de datos con datos de la vida real dan idea del interés en el tema.

■ 7.5. HISTORIA CLÍNICA ELECTRÓNICA

Según la Ley 41/2002 de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica, se define historia clínica como el conjunto de documentos que contienen datos, valoraciones e informaciones de cualquier índole sobre la situación y evolución clínica de un paciente a lo largo del proceso asistencial. Por tanto, la historia clínica es el compendio de datos clínicos de un paciente, englobando informes médicos de cada uno de los procesos asistenciales por los que ha pasado el paciente y las pruebas de imagen y de laboratorio que se le han realizado (Ley 41/2002 de 14 de noviembre). La aplicación de la tecnología de la información y la comunicación (TIC) ha permitido la digitalización de esta información, dando lugar a la historia clínica electrónica.

La historia clínica electrónica (HCE) es un archivo dinámico que permite a todos los profesionales que asisten a un paciente consultar los datos de este en tiempo real, ya sean datos del proceso de atención actual o de procesos anteriores. La informatización de las historias clínicas permite la consulta de informes, pruebas de imagen, analíticas y otros procedimientos en periodos de tiempo muy inferiores a cuando estos datos no estaban informatizados, mejorando la accesibilidad a la información y evitando traslados y pérdidas de documentación. La HCE también permite la integración de todos los procesos asistenciales por los que puede pasar un paciente: hospitalización, urgencias y consultas ambulatorias. Entre las ventajas de la HCE frente a la historia clínica tradicional (en papel) hay que destacar lo siguiente:

- Contiene información estructurada y coordinada entre los diferentes profesionales sanitarios que tratan a un paciente, favoreciendo el manejo multidisciplinar de las diferentes patologías.

⁹ http://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml

- Permite el seguimiento del estado de salud del paciente a tiempo real, agilizando la consulta de información y favoreciendo su entendimiento, ya que omite los problemas de tipo caligráfico.
- El acceso de los usuarios del sistema es seguro, identificándose cada uno de ellos mediante nombre y contraseña, lo que permite conocer quién y cuándo se modifica la información contenida en la HCE. Además, permite la firma electrónica de los informes médicos, dotando a estos documentos de carácter legal.
- La accesibilidad de la HCE puede realizarse desde cualquier lugar y en cualquier momento, permitiendo la atención del paciente a pie de cama.
- Contiene un sistema de control de citas y agendas clínicas, lo que favorece la optimización y el ahorro de tiempo y costes de personal.
- La información contenida en la HCE puede utilizarse con fines docentes e investigadores.

La aplicación de las TIC permite la agregación de datos para la elaboración de informes y su posterior análisis. Para ello se requiere que el personal sanitario que maneja estas historias clínicas esté debidamente entrenado y que la introducción de los datos se haga conforme a unos códigos y normas que posteriormente permitan la explotación de la información. Con todo ello, la HCE se convierte en una fuente de conocimiento.

- Sirve como medida de calidad asistencial.
- Reduce en gran medida los errores de transcripción, gracias a la prescripción electrónica. Además, durante los procesos de hospitalización, las HCE permiten la prescripción electrónica de la medicación con la que se está tratando al paciente, lo que supone un aumento en la seguridad de los tratamientos médicos. La tecnología de la información permite que esta información sea visible a cualquier otro profesional sanitario que atienda a ese paciente, por lo que el Servicio de Farmacia puede validar esta información y dispensar la medicación prescrita, y la enfermería puede administrar aquella medicación incluida por el médico en el sistema, evitando traslado de papeles de una parte del hospital a otra, disminuyendo los errores de medicación y aumentando la seguridad del paciente en la medicación que recibe.

Los sistemas de historia clínica electrónica se completan con herramientas de asistencia a la prescripción y a la validación, lo que facilita a los usuarios el manejo del sistema informático mediante acceso rápido a funciones de uso frecuente, búsqueda de tareas, avisos y alertas a través de simbología o ventanas emergentes, etcétera. Por otra parte, también permite la solicitud de pruebas de imagen y pruebas de laboratorio y la comunicación entre los diferentes profesionales que tratan a un paciente, gracias a un sistema de petición de interconsultas entre los diferentes

servicios médicos y notas dirigidas que pueden ser contestadas. Todo ello aumenta la eficiencia de los procesos y favorece el manejo multidisciplinar del paciente. La aplicación de las TIC a los tratamientos de los pacientes ha permitido la automatización de los servicios de Farmacia en su tarea de dispensación intrahospitalaria de medicamentos.

Como desventajas de la HCE hay que destacar la necesidad de que el personal esté entrenado en el manejo de los programas informáticos utilizados, su dependencia de una fuente de energía y la susceptibilidad a las caídas del sistema.

En España, la HCE está implantada de forma general en todo el territorio nacional, gracias a la creación, en el año 2006, de la Historia Clínica Digital del Sistema Nacional de Salud (HCDSNS) (Ortega *et al.*, 2011). Esta nació como un proyecto del Ministerio de Sanidad con el objetivo de que un paciente pudiera ser atendido en cualquier comunidad autónoma con la garantía de disponer de su información clínica preexistente. El programa surge de la necesidad de compatibilizar los sistemas de información relacionados con la sanidad de cada una de las comunidades autónomas, ya que la competencia estas en materia de sanidad y la aplicación de las TIC en este ámbito habían dado lugar a que cada una de las CC. AA. hubiera establecido un sistema informático propio para el manejo de información sanitaria de sus pacientes. Por este motivo, se creó la HCDSNS, en donde los profesionales sanitarios y los ciudadanos tienen acceso telemático a un conjunto definido de datos que contienen información relevante sobre su estado de salud o el de sus pacientes. Como valor añadido, el proyecto permite el acceso del ciudadano a sus informes de salud y pone en su conocimiento qué profesional sanitario ha consultado su información clínica. La seguridad del proyecto constituye su principal punto crítico, por lo que se ha certificado la plataforma HCDSNS con la norma ISO 27001. La interoperabilidad de los diferentes sistemas informáticos de las CC. AA. y los informes de los distintos profesionales sanitarios está garantizada por el Real Decreto 1093/2010 de 3 de septiembre, por el que se aprueba el conjunto mínimo de datos de los informes clínicos en el Sistema Nacional de Salud.

■ 7.6. REVISIONES SISTEMÁTICAS Y METAANÁLISIS

La revisión de los estudios publicados acerca de cada tema clásicamente ha constituido la forma de profundizar en el conocimiento científico del mismo. Asimismo, desde hace décadas, ha sido la manera de resolver problemas y preguntas de interés clínico que se puedan plantear en la práctica. Se pueden distinguir dos grandes tipos de revisiones: las *narrativas* –tradicionales o de autor– y las denominadas más recientemente, *revisiones sistemáticas*.

■ 7.6.1. Revisiones narrativas

Las revisiones narrativas consisten en la búsqueda y lectura de las publicaciones acerca de un tema determinado, ello de forma más o menos exhaustiva. Por

lo habitual, es obra de un experto destacado y reconocido en el campo del tema en cuestión. En cierto modo, cada capítulo de una obra científico-médica viene a suponer una revisión de esta naturaleza, de manera que los textos de referencia en la literatura médica suponen *compendios* de revisiones narrativas

Estas revisiones pueden abordar diversos aspectos del tema de estudio. En efecto, pueden tratar el tema desde las diferentes perspectivas posibles, a saber, epidemiología, etiología, fisiopatología, diagnóstico, pronóstico o terapéutica (Letelier, Manríquez y Rade, 2005).

Por lo general, el autor redacta la revisión como una sucesión de resúmenes de los diferentes estudios revisados. Ello suele incluir el diseño, los criterios de selección y los resultados. La procedencia de los estudios incluidos en la revisión narrativa no suele especificarse necesariamente, como tampoco se suelen especificar los criterios empleados en la búsqueda bibliográfica. La selección de los estudios incluidos en este tipo de revisiones no obedece a unos criterios pre-establecidos, sino que, a criterio del autor, se consideran los trabajos por su relevancia en el ámbito científico. Finalmente, las conclusiones son formuladas mediante criterios cualitativos con cierto componente subjetivo basado en la evidencia.

Las características de las revisiones narrativas se resumen en la Tabla 2.

Tabla 2.

CARACTERÍSTICAS DE LAS REVISIONES NARRATIVAS (TRADICIONALES)

Características	Revisión narrativa
Pregunta de investigación	Amplia
Selección de fuentes de información	Habitualmente no especificada Potencial sesgo
Criterios de búsqueda bibliográfica	Habitualmente no especificados
Criterios de selección de los estudios	Habitualmente no especificados Potencial sesgo de selección
Evaluación	Variable
Resumen de los resultados	Habitualmente cualitativo y subjetivo
Inferencias	A veces, basadas en la evidencia

Fuente: Modificada de Cook, Mulrow y Haynes (1997).

7.6.2. Revisiones sistemáticas

Las revisiones sistemáticas, por su parte, son aquellas que resumen y analizan la evidencia existente sobre un determinado tema de una forma estructurada y sistemática. Estas evidencias proceden de estudios “originales” o “primarios” previamente publicados.

Existen dos tipos de revisiones sistemáticas:

- a. Revisiones sistemáticas *cualitativas* o revisiones *sistemáticas* propiamente dichas. Son aquellas que resumen los resultados de los estudios originales o primarios de forma descriptiva pero no se combinan ni analizan necesariamente mediante métodos o procedimientos estadísticos.
- b. Revisiones sistemáticas cuantitativas o metaanálisis: son aquellas revisiones sistemáticas en las que se aplican métodos estadísticos para analizar los resultados de los estudios primarios que, de esta forma, se combinan de forma cuantitativa. Tal como se establece en esta definición, todo metaanálisis debe partir de una revisión sistemática. El término metaanálisis fue acuñado por el estadístico norteamericano Gene V. Glass en 1976 para referirse al “análisis de análisis” o análisis estadístico de los resultados de una serie de ensayos clínicos individuales con el propósito de integrar sus hallazgos (Glass, 1976). El objetivo principal de un metaanálisis consiste básicamente en calcular una medida promedio de los resultados recogidos en los estudios analizados (Moliner, 2008).

Las revisiones sistemáticas, por tanto, pueden considerarse investigaciones científicas en sí mismas en las que los sujetos en estudio son los propios estudios y en las que se aplican métodos preestablecidos cuyo objetivo principal es reducir sesgos y errores aleatorios (Cook, Mulrow y Haynes, 1997). De esta forma se pretende eliminar cualquier componente subjetivo, tanto en la búsqueda y selección de los estudios incluidos como en la interpretación de los resultados.

Por otra parte, la especificación *a priori* de los métodos que se deben seguir en una revisión sistemática cumple el objetivo de garantizar la reproducibilidad con idénticos resultados.

Los métodos empleados en las revisiones sistemática cualitativas y en los metaanálisis consisten en (Ferreira González, Urrutia y Alonso-Coello, 2011):

- Enumerar y definir *a priori* las fuentes de procedencia de los estudios primarios.
- Establecer *a priori* los criterios de búsqueda de los estudios que se incluirán en la revisión sistemática.
- Establecer *a priori* los criterios de selección de los estudios que se incluirán en la revisión sistemática.
- Describir adecuadamente el diseño y las principales características de los estudios seleccionados para la revisión, así como sus resultados y la interpretación de los mismos.

La Tabla 3 muestra las características de las revisiones sistemáticas y metaanálisis en contraposición a las características de las revisiones narrativas (Tabla 3).

Tabla 3.

CARACTERÍSTICAS DIFERENCIALES DE LAS REVISIONES NARRATIVAS O TRADICIONALES Y LAS REVISIONES SISTEMÁTICAS

Características	Revisión narrativa	Revisión sistemática
Pregunta de investigación	Amplia	Centrada en un problema clínico concreto
Selección de fuentes de información	Habitualmente no especificada Potencial sesgo	Fuentes completas
Criterios de búsqueda bibliográfica	Habitualmente no especificados	Incluye estrategia de búsqueda sistemática y detallada
Criterios de selección de los estudios	Habitualmente no especificados Potencial sesgo de selección	Criterios definidos y aplicados sistemáticamente
Evaluación	Variable	Crítica y rigurosa
Resumen de los resultados	Habitualmente cualitativo y subjetivo	Resumen cuantitativo siempre. Si incluye análisis estadístico, se trata de un metaanálisis
Inferencias	A veces, basadas en la evidencia	Habitualmente, basadas en la evidencia

Fuente: Modificada de Cook, Mulrow y Haynes, (1997).

■ 7.6.3. Fases de la elaboración de una revisión sistemática (o metaanálisis)

Formulación de la pregunta clínica a la que la revisión sistemática debe dar respuesta

Esta etapa debe considerar y caracterizar detalladamente tres componentes fundamentales: los participantes, la intervención que se estudia (o experimental) y el comparador y las variables o resultados.

Selección de los criterios de elegibilidad de los estudios primarios

Esta etapa está íntimamente vinculada a la primera, ya que una vez definida y formulada la pregunta en investigación, es posible decidir y establecer los criterios de elegibilidad de los estudios primarios que se incluirán en la revisión en cuanto a diseño y características. Esta decisión está vinculada con la pregunta en investigación, ya que, si el propósito de la revisión sistemática es evaluar la eficacia de una intervención farmacológica o de otra naturaleza, lo más adecuado será seleccionar ensayos clínicos aleatorizados y controlados, siempre que estén disponibles. En ocasiones puede ser necesario seleccionar estudios observacionales e incluso series de casos o casos aislados o estudios centrados en Farmacoeconomía para la realización de revisiones sistemáticas de evaluaciones económicas.

Estrategias de búsqueda de estudios

Esta etapa debe identificar los estudios primarios aplicando los criterios de selección adoptados en la etapa anterior. Las revisiones sistemáticas y metaanálisis, en tanto que constituyen una forma de investigación, no están exentas de sesgos errores sistemáticos y aleatorios. También es posible cometer sesgo de selección de los estudios primarios que se incluyen en la revisión sistemática, bien porque no se han predefinido correctamente los criterios de selección o porque existe una tendencia a no publicar los resultados negativos. La ausencia de sesgos de selección en las revisiones sistemática es fundamental y está ligada a una adecuada selección de los criterios de elegibilidad y su aplicación en esta fase.

Existen aspectos que pueden parecer triviales pero que son relevantes como la aplicación de restricciones sobre el idioma de publicación o sobre la inclusión de estudios no publicados, por lo general, muy difíciles de obtener.

Otro aspecto que es necesario definir adecuadamente es la fuente de obtención de los estudios primarios, entre las que se encuentran las siguientes:

Bases de datos de publicaciones científicas

Las bases de datos bibliográficas pueden ser electrónicas y manuales. En la actualidad, las numerosas bases de datos electrónicas de publicaciones científicas han desbancado a las bases de datos bibliográficas manuales. Algunas de las bases de datos bibliográficas electrónicas más ampliamente usadas son:

- **MEDLINE.** Es una base de datos científica producida por la *US National Library of Medicine* (NLM) que contiene referencias bibliográficas y resúmenes de más de 4.000 revistas biomédicas publicadas en los Estados Unidos y en otros 70 países. Contiene millones de referencias desde 1966 en áreas de medicina, enfermería, odontología y veterinaria.
- **CENTRAL** (*Cochrane Central Register of Controlled Trials*). Este registro consiste en una recopilación de informes de ensayos clínicos controlados aleatorizados y cuasialeatorizados. La mayoría de los registros incluidos en CENTRAL proceden de bases de datos bibliográficas (principalmente MEDLINE y Embase), aunque también incluye otras fuentes publicadas y no publicadas.
- **EMBASE.** Según la propia web de Embase, se trata "de una base de datos biomédica muy versátil, polivalente y actualizada. Abarca la bibliografía biomédica internacional más importante desde 1947 (Excerpta Medica) hasta nuestros días".
- **Science Citation Index (SCI)**, también conocido como ISI, en alusión al Institute for Scientific Information, es una base de datos que incluye todas las contribuciones (artículos, editoriales, cartas, revisiones, discusiones, etc.) indexadas por Thomson Reuters.

- Bases de datos específicas, como, por ejemplo, la *NHS Economic Evaluation Database (NHS EED)*, que forma parte de la *Cochrane Library* o la base de datos CINAHL¹⁰, especializada en temas de enfermería.

La utilización de bases de datos manuales es poco frecuente hoy en día. Tanto si se emplean bases de datos bibliográficas como manuales es muy importante predefinir los términos de búsqueda e incluirlos en la revisión sistemática en aras de garantizar la reproducibilidad de la búsqueda.

Registros de ensayos clínicos en marcha

Dentro de estos registros debe mencionarse el registro de ensayos clínicos europeo *Clinical Trials Register*¹¹; el Registro Español de Estudios Clínicos¹² y el estadounidense¹³.

Todos estos registros son de carácter público y gratuito; su principal inconveniente es que no todos los ensayos incluyen los resultados intermedios o finales.

Otras fuentes

La literatura gris también puede contribuir con su aportación a la detección de estudios que pueden resultar de interés. Por literatura gris se conoce al conjunto de documentos que no se editan o que se publican por canales no convencionales. Entre ellos, se cuentan las tesis doctorales, informes de investigación, memorias, proyectos, patentes, traducciones científicas, así como las comunicaciones a congresos. Los resúmenes y comunicaciones enviados a congresos habitualmente se publican en numerosos especiales o suplementos de revistas con factor de impacto de la correspondiente especialidad; sin embargo en otros casos, se editan en publicaciones específicas. El principal problema del empleo de la literatura gris radica en la dificultad que puede suponer acceder al estudio, que a veces requiere contactar directamente con el autor.

Otras fuentes que pueden ser útiles son los manuales de medicina y las propias referencias de los estudios localizados, así como las consultas a expertos de solvencia en el tema.

No existe ningún criterio con respecto al número de fuentes que es necesario consultar, aunque, probablemente ninguna revisión sistemática que se precie debe dejar de realizar su búsqueda en las bases de datos MEDLINE y CENTRAL, así como en EMBASE. El inconveniente fundamental de estas bases de datos es que ofrecen acceso gratuito a un número limitado de referencias, siendo necesaria suscripción para el acceso a la gran mayoría de publicaciones. También es importante

¹⁰ <https://www.ebscohost.com/nursing/products/cinahl-databases/cinahl-complete>

¹¹ <https://www.clinicaltrialsregister.eu/>

¹² <https://reec.aemps.es/reec/public/web.html>

¹³ <https://clinicaltrials.gov/>

tener en cuenta la posibilidad de que aparezcan repeticiones de los estudios, ya que en muchos casos, las bases de datos comparten información.

Una vez obtenidos los estudios primarios, deben aplicarse los criterios de inclusión y exclusión para filtrar los estudios y descartar repeticiones. Es conveniente que la selección de los estudios primarios sea realizada por, al menos dos revisores de forma independiente, para aumentar el nivel de fiabilidad de los criterios aplicados a los estudios y que determina si se seleccionan o se excluyen.

Probablemente el paso de decidir cuáles son los estudios que se incluyen en la revisión sistemática sea el más relevante para la validez de la revisión.

Obtención de los datos

La información que debe extraerse de cada estudio primario dependerá de las características de la revisión sistemática. Es conveniente diseñar una hoja de registro o formulario de extracción, equivalente al cuaderno de recogida de datos de los estudios con sujetos. Los contenidos de este formulario deben especificarse en la fase del diseño de la revisión sistemática o metaanálisis y suelen incluir información relativa a:

- fuente o identificación del estudio primario;
- criterio de elegibilidad, incluyendo los motivos para incluirlo o excluirlo;
- método o diseño del estudio primario;
- características de los participantes;
- descripción de las intervenciones experimental y de control;
- variables;
- resultados;
- otros aspectos de interés tales como la fuente de financiación, las conclusiones de los autores, o comentarios de los revisores.

Estos formularios deben detallar no solo el tipo de datos de interés para la revisión sino también la forma en la que se recopilarán los datos. Habitualmente, los datos recogidos son comprobados por otro revisor.

Análisis de los resultados y exposición de estos

El análisis de los resultados puede efectuarse como un resumen estructurado narrativo que constituye la esencia de una revisión sistemática cualitativa. La otra posibilidad consiste en integrar los resultados de los estudios primarios mediante métodos estadísticos, es decir, como metaanálisis.

Los métodos estadísticos aplicados en los metaanálisis calculan el resultado combinado como la media ponderada de los resultados de cada uno de los estu-

dios primarios. Cada estudio primario contribuye al resultado integrado de forma proporcional a su relevancia o “peso”, de forma que los estudios primarios que más aportan contribuyen en mayor medida al resultado integrado.

Se considera que los estudios de “*mayor peso*” son aquellos de mayor tamaño muestral, los de menor variabilidad o los que presentan un mayor número de eventos. Por contra, los estudios con un tamaño muestral más limitado o aquellos con mayor variabilidad en la respuesta contribuirán en menor medida al valor del resultado global.

Básicamente, existen dos modelos estadísticos empleados para la integración de los resultados de los metaanálisis en función de la heterogeneidad: el modelo de efectos fijos y el modelo de efectos aleatorios (Perlega Díaz y Pita Fernández, 2005).

El modelo de efectos fijos asume que no existe heterogeneidad en los estudios primarios incluidos en la revisión, decir, considera que todos estiman el mismo efecto y que las diferencias observadas se deben únicamente al azar. El modelo de efectos aleatorios asume, por el contrario, que los estudios incluidos en la revisión estiman efectos diferentes y constituyen una muestra aleatoria de todos los estudios existentes.

Si no existe heterogeneidad entre los estudios, los resultados obtenidos por ambos modelos serán idénticos. Si realmente existe heterogeneidad, los intervalos de confianza para el efecto serán más amplios si se emplea el modelo de efectos aleatorios que si se aplica el modelo de efectos fijos y la estimación resultará más conservadora. La heterogeneidad es un aspecto muy relevante en los metaanálisis y debe ser identificada y cuantificada y debe analizarse desde una doble perspectiva: heterogeneidad en el sentido de la consistencia clínica de los estudios incluidos y la heterogeneidad desde el punto de vista estadístico.

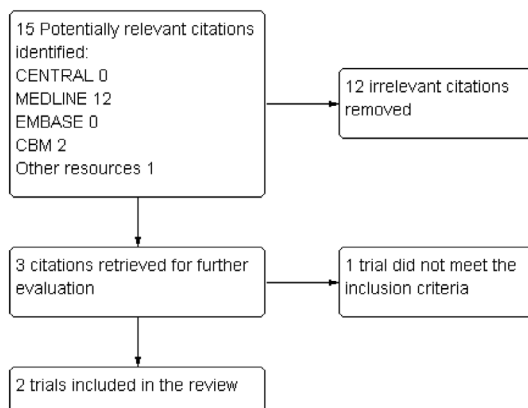
La heterogeneidad clínica se refiere a las posibles diferencias existentes en cuanto a las características de los estudios (criterios de inclusión, diseño del estudio, pérdidas de seguimiento, tipo y características de los grupos de intervención, entre otras). La heterogeneidad estadística consiste en una cuantificación matemática de la discrepancia de los resultados numéricos alcanzados por los diferentes estudios.

Cada uno de estos modelos emplea diferentes métodos de análisis estadísticos, tales como el inverso de la varianza, el método de Peto, el de Mantel-Haenzel, y el método de Dersimonian y Laird. La elección de un método u otro dependerá del tipo de resultado en estudio: continuos, dicotómicos, asimétricos, ordinales, tasas o tiempo hasta el evento y del modelo asumido.

La exposición de los resultados debe efectuarse de forma sistemática y ordenada. Es característico de las revisiones sistemáticas utilizar herramientas gráficas como tablas, figuras y gráficos. Habitualmente, los informes de resultados incluyen

Figura 5.

EJEMPLO DE DIAGRAMA DE FLUJO DEL METAANÁLISIS

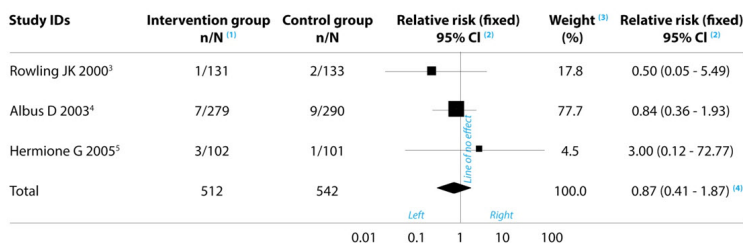


Fuente: Tomado de Xiao, Wang y Luo (2012).

Figura 6.

EJEMPLO DE FOREST PLOT

How to read a forest plot?



Test for heterogeneity Chi-square = 0.79, df = 2, p = 0.67, I² = 0.0% ⁽⁵⁾

Test for overall effect z = 0.35, p = 0.7 ⁽⁶⁾

- (1) N = total number in group, n = number in group with the outcome.
- (2) Outcome of interest in picture and in number. Fixed effect model used for meta-analysis.
- (3) Influence of studies on overall meta-analysis.
- (4) Overall effect.
- (5) Heterogeneity (I²) = 0%. So, we use fixed effect model.
- (6) p value indicating level of statistical significance

Fuente: <https://uk.cochrane.org/news/how-read-forest-plot>

una tabla con las características de los estudios incluidos, así como diagramas de flujo que ilustran los resultados de la búsqueda de los estudios primarios, la aplicación de los criterios de selección y las razones de la exclusión. La Figura 5 es un ejemplo real de diagrama de flujo.

Probablemente, el gráfico más característico y definitorio de los metaanálisis es el conocido como *forest plot* o “diagrama de bosque” (Figura 6). El *forest plot* combina elementos de una tabla (columnas) y elementos gráficos. Suele incluir seis columnas. La primera columna consiste en el listado de los estudios incluidos en el metaanálisis. La segunda columna expone las intervenciones experimentales y la tercera, la intervención empleada como control. Estas dos columnas pueden dividirse, a su vez, en columnas que reflejan el tamaño de la muestra (n) y el valor de los resultados de cada estudio primario, en términos de medias y desviaciones estándar o riesgos relativos.

La cuarta columna representa gráficamente los resultados de cada estudio. Incluye una línea vertical central que representa la ausencia de efecto, es decir, la ausencia de diferencias entre el grupo de intervención experimental y el grupo control. Si los resultados se expresan como ratios (riesgo relativo u *odds ratio*), la línea representa la unidad; si los resultados se expresan como diferencias de medias, la línea de “no efecto” representa el cero. Ambos lados de la línea representan en qué medida los resultados favorecen a la intervención experimental (habitualmente, pero no siempre, a la izquierda de la línea de no efecto) o a la intervención de control.

Correspondiendo a los resultados de cada estudio, se trazan líneas perpendiculares a la línea de no efecto, que representan los intervalos de confianza. Cuanto más larga sea la línea, más amplio será el intervalo y los resultados del estudio, menos precisos. Estas líneas contienen un cuadrado que representa el efecto logrado por cada estudio. El tamaño de este cuadrado es proporcional al peso de cada estudio en el metaanálisis, expresado numéricamente en la quinta columna. Los cuadrados, es decir, los efectos individuales de cada estudio se situarán a la izquierda o a la derecha de la línea de no efecto, dependiendo de si el resultado del estudio favorece a la intervención experimental o a la de control. Si no existen diferencias, el cuadrado se sitúa sobre la línea de no efecto. Si la mayoría de los efectos de cada uno de los estudios se sitúa en el mismo lado de la línea de referencia, se puede concluir que la heterogeneidad es baja; si la distribución se reparte a ambos lados de la línea de no efecto, puede interpretarse como que los resultados son discrepantes y la heterogeneidad elevada.

Finalmente, los “diamantes” representan los resultados integrados del metaanálisis (totales o subtotales por subgrupos). El centro del diamante representa el efecto de la intervención de forma integrada y su ancho es proporcional al intervalo de confianza global. La diferencia entre los grupos de intervención y control puede considerarse estadísticamente significativa si el diamante se sitúa a un lado de la línea de no efecto, sin aproximarse. Si el diamante rozara la línea de no efecto, no podrán extraerse conclusiones concluyentes.

La sexta columna refleja los resultados numéricos de cada estudio primario, bien en términos de riesgo relativo en el caso de las variables dicotómicas o bien como diferencias de medias en el caso de las variables continuas, junto con su correspondiente intervalo de confianza. En ocasiones, el orden de las columnas puede variar y el gráfico se ubica en la última columna.

Interpretación de los resultados

La revisión sistemática finaliza con la interpretación de los resultados que debe incluir un análisis de las limitaciones de la revisión sistemática o del metaanálisis, una evaluación de los potenciales sesgos tanto de los estudios primarios (sesgo de selección, sesgo de realización, sesgo de desgaste, sesgo de notificación) como de la revisión en sí misma.

Es recomendable que la interpretación de los resultados también considere la consistencia de los resultados, su aplicabilidad clínica y siempre que sea posible, recomendar actuaciones para el futuro.

■ 7.6.4. Ventajas y limitaciones de las revisiones sistemáticas y metaanálisis

Desde la óptica de la medicina basada en la evidencia, los metaanálisis de ensayos clínicos aleatorizados constituyen el más elevado nivel de evidencia científica, tal y como afirma la Agency for Healthcare Research and Quality (AHRQ) en USA¹⁴.

Los metaanálisis permiten la posibilidad de detectar efectos clínicos significativos a partir de estudios de tamaño muestral tan reducido que no pueden detectar efectos significativos de forma aislada. El agrupamiento de estudios permite también aumentar los tamaños muestrales para realizar análisis de subgrupos que no alcanzarían un tamaño estadísticamente significativo si se consideraran individualmente. Sin embargo, los resultados de una revisión sistemática o metaanálisis no pueden ser más fiables que los de los estudios en los que se basan.

Por todo ello, es conveniente ser prudente y crítico en la interpretación de las revisiones sistemáticas tanto si quien consulta la revisión sistemática es el clínico que debe resolver un problema terapéutico en un paciente en concreto, los integrantes de las comisiones de farmacia, las agencias de evaluación o los órganos responsables de las políticas sanitarias.

■ 7.7. COMPARACIONES INDIRECTAS

Uno de los principales objetivos cuando se evalúan medicamentos para su uso en los hospitales es conocer cuál es su posicionamiento terapéutico, lo que requiere

¹⁴ <https://digital.ahrq.gov/research-method/meta-analysis>

la comparación frente a las alternativas existentes. Los ensayos clínicos comparativos aleatorios y las revisiones sistemáticas y/o metaanálisis se consideran las mejores fuentes de evidencia, aunque presentan ciertas limitaciones. En muchas ocasiones es necesario comparar la evidencia existente en ausencia de estudios comparativos. El hecho de que se disponga alternativamente de estudios de cada uno de los fármacos frente a un comparador común pueden ser de gran utilidad lo que se conoce como comparaciones indirectas. Estas también son de utilidad, cuando hay comparaciones directas, pero no son concluyentes, resultan insuficientes o son de mala calidad.

Poder aplicar comparaciones indirectas requiere ciertas asunciones importantes: la similitud clínica y metodológica, la consistencia y la homogeneidad.

En cualquier caso, las comparaciones indirectas tienen un papel en la evaluación de medicamentos, aunque tienen muchas limitaciones y hay que ser muy conscientes de ellas para saber cuándo se pueden utilizar y en qué medida se puede uno apoyar en sus resultados. Es imprescindible analizar su validez interna considera en qué medida se pueden extrapolar los resultados a un ámbito particular. Para ello tienen que ser estudios frente a los comparadores adecuados y que se hayan elegido todos los estudios relevantes. Por ello tiene mucha importancia cómo se ha realizado la búsqueda de los estudios y si se han considerado los posibles sesgos, como el de publicación, etc., la calidad de los estudios combinados y evitar el sesgo de confusión por violación de la asunción de similitud y consistencia.

Al evaluar nuevos medicamentos para su uso en el hospital, van a darse con frecuencia situaciones en las que las comparaciones indirectas pueden resultar útiles y antes que hacer las comparaciones de forma subjetiva y arbitraria, es mejor realizarlas con mayor rigor metodológico. Es importante recordar que las comparaciones indirectas ayudan a tomar la decisión entre alternativas a comparar, pero no serán mejores que los estudios de los que parten y habrá que velar porque se hagan sin sesgos y sin violación de las asunciones para no llegar a conclusiones erróneas (Tonin *et al.*, 2017).

Es en este contexto los registros de RWD van a permitir comparar tratamientos para una misma condición y establecer el papel de cada uno en los diferentes pacientes, un aspecto con importantes implicaciones en efectividad, seguridad y costes y permitirán establecer estimaciones de la efectividad más que de la eficacia, en una variedad de escenarios clínicos diferentes.

- Comparación con múltiples alternativas de investigación o estrategias clínicas para informar de opciones terapéuticas óptimas, más allá del uso del placebo como comparador.
- Estimación de los riesgos y beneficios de una nueva intervención, incluyendo los beneficios y daños a largo plazo. Minisentinel ha permitido recoger datos de seguridad de medicamentos que han supuesto retirar fármacos o modificar fichas técnicas.

- Obtención de resultados clínicos en una población diversa que refleja el rango y la distribución de los pacientes observados en la práctica clínica.
- Resultados obtenidos desde una perspectiva más amplia que en los ECA tradicionales (resultados informados por los pacientes, calidad de vida y síntomas).
- Datos utilizables para cálculo de los costes de los servicios de salud y evaluación económica.
- Información de cómo un producto es aplicado y prescrito en la práctica clínica y cuál es la adherencia al mismo.
- Datos en situaciones donde no es posible llevar a cabo un ECA.
- Justificación de la recolección de datos en más de una sede.

■ 7.8. METAANÁLISIS EN RED

El apartado 7.6. se dedicó a las revisiones sistemáticas y a los metaanálisis. Este apartado abordará un tipo especial de metaanálisis denominados metaanálisis en red (*network meta-analysis*). Estos metaanálisis pretenden contrarrestar una de las clásicas desventajas de los metaanálisis convencionales que es la imposibilidad o dificultad para comparar más de dos intervenciones simultáneamente.

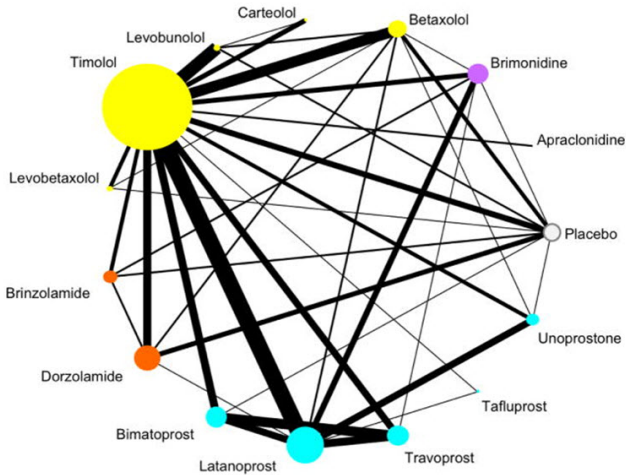
El meta-análisis de red permite comparar múltiples tratamientos simultáneamente en un solo análisis mediante la combinación de comparaciones directas e indirectas “*en red*” entre ensayos clínicos aleatorizados. Para ello, emplea técnicas estadísticas que incorporan la información procedente de las comparaciones directas e indirectas (Catalá-López, Tobias y Roqué, 2014).

De forma similar al metaanálisis *convencional* o *pair-wise*, el metaanálisis en red basa su validez en la calidad de los estudios y la variabilidad entre los mismos, considerando los sesgos de información. Estas circunstancias tienen en el metaanálisis en red especial importancia ya que se efectúan comparaciones indirectas. Por ello es muy importante que los estudios incluidos en el metaanálisis en red sean comparables en los factores modificadores del efecto. Las etapas de la realización de un meta-análisis en red son las mismas que las de una revisión sistemática convencional: desarrollo de una pregunta de investigación apropiada, selección de estudios y aplicación de estrategias de búsqueda bibliográfica, obtención de los datos, realización de la síntesis cualitativa y cuantitativa, presentación de resultados y conclusiones (Rouse, Chaimani y Li, 2017).

El metaanálisis en red realiza una estimación numérica de comparaciones indirectas o mixtas, es decir, aquella en la que se combinan los resultados o efectos de los tratamientos a partir de comparaciones directas e indirectas mediante métodos

Figura 7.

GRÁFICO EN RED



Nota: Cada círculo es un nodo que representa un agente farmacológico. El tamaño del nodo es proporcional al número de sujetos aleatorizados a recibir ese principio activo y el grosor de la línea que une dos nodos es proporcional al número en ensayos clínicos que compara ambos agentes.

Fuente: Tomado de Rouse, Chamaini y Li (2017).

estadísticos que parten de la aplicación de la propiedad transitiva (si el tratamiento A es mejor que B y B es mejor que C, entonces A es mejor que C). El *software* más utilizado para aplicar estos métodos en los metaanálisis en red son los programas Stata y WinBUGS (Tonin *et al.*, 2017).

Si el *forestplot* es la característica gráfica por antonomasia del metaanálisis, los metaanálisis en red emplean gráficos en forma "de red" que representa los diferentes tratamientos evaluados. Cada intervención o tratamiento se representa como un círculo o nodo y las comparaciones directas se representan como líneas que unen dos nodos (ver Figura 7).

El metaanálisis de red puede ayudar a evaluar la efectividad comparativa de diferentes tratamientos regularmente utilizados en la práctica clínica y probablemente su mayor ventaja es la capacidad de proponer una jerarquía de intervenciones. Sin embargo, no está exento de limitaciones. A las limitaciones propias de los metaanálisis convencionales (calidad de los estudios primarios, sesgos o errores sistemáticos, variabilidad o heterogeneidad) se une el hecho de que los meta-análisis en red realizan comparaciones indirectas por lo que, tal y como se ha mencionado con anterioridad, los estudios que se incluyen en estos metaanálisis deben ser comparables desde el punto de vista metodológico y los factores modificadores del

efecto no deben diferir entre estudios. Un metaanálisis en red debe partir siempre de una revisión sistemática rigurosa.

■ REFERENCIAS

- CALVERT, M., WOOD, J. y FREEMANTLE, N. (2011). Designing "Real-World" trials to meet the needs of health policy makers at marketing authorization. *J Clin Epidemiol*, 64(7), pp. 711-717.
- CATALÁ-LÓPEZ, F., TOBIAS, A. y ROQUÉ, M. (2014). Conceptos básicos del metaanálisis en red. *Aten Primaria*, 46, pp. 573-581.
- COOK, D. J., MULROW, C. D. y HAYNES, R. B. (1997). Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med*, 126, pp. 376-380.
- DICKSON, R., BOLAND, A., DUARTE, R. *et al.* (2018). EMA and NICE Appraisal Processes for Cancer Drugs: Current Status and Uncertainties. *Appl Health Econ Health Policy*, 16(4), pp. 429-432.
- EUROPEAN COMMISSION. (2014). *Green paper on mobile Health (mHealth)*. Disponible en: <http://ec.europa.eu/digital-agenda/en/news/green-paper-mobile-health-mhealth>
- Ferreira González, I., Urrutia, G. y Alonso-Coello, P. (2011). Revisiones sistemáticas y metaanálisis: bases conceptuales e interpretación. *Rev Esp Cardiol*, 64, pp. 688-696.
- GARCÍA LÓPEZ, J. L., DEL LLANO SEÑARIS, J. E., DEL DIEGO SALAS, J. y RECALDE MANRIQUE, J. M. (2014). *Aportación de los "Real World Data (RWD)" a la mejora de la práctica clínica y del consumo de recursos de los pacientes*. Ed. Fundación Gaspar Casal. Disponible en: http://fgcasal.org/publicaciones/Libro_RWD.pdf
- GLASS, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educ Res*, 5, pp. 3-8.
- IMS INSTITUTE FOR HEALTHCARE INFORMATICS (2015). *Patient Adoption of mHealth. Use, evidence and remaining barriers to mainstream acceptance*. Disponible en: <http://www.imshealth.com/en/thought-leadership/ims-institute/reports/patient-adoption-of-mhealth>
- IOM (INSTITUTE OF MEDICINE). (2011). *Learning what Works: infrastructure required for comparative effectiveness research. Workshop summary*. Washington, DC: The National Academies Press. [consultado: 228/07/18]. Disponible en: <https://www.nap.edu/catalog/12214/learning-what-works-infrastructure-required-for-comparative-effectiveness-research-workshop>
- KATKADE, V. B., SANDERS, K. N. y ZOU, K. H. (2018). Real world data: an opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. *J Multidiscip Healthc*, 11, pp. 295-304.
- KNOTTNERUS, J. A. y TUGWELL, P. (2017). Research methods must find ways of accommodating clinical reality, not ignoring it: the need for pragmatic trials. *J Clin Epidemiol*, 88, pp. 1-3.
- LETELIER, L. M., MANRIQUEZ, J. J. y RADE, G. (2005). Revisiones sistemáticas y metaanálisis: ¿son la mejor evidencia? *Rev Med Chile*, 133, pp. 246-249.
- Ley 41/2002 de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica.
- DEL LLANO SEÑARIS, J. E., DEL CASTILLO, J. A., DILLA QUINTERO, T. y GARCÍA LÓPEZ, J. L. (2016). *Datos de la vida real en el sistema sanitario español*. Ed. Fundación Gaspar Casal. Disponible en: http://www.fgcasal.org/publicaciones/Datos_de_la_vida_real_SSE_2016.pdf
- LOUDON, K., TREWEEK, S., SULLIVAN, F., DONNAN, P., THORPE, K. E. y ZWARENSTEIN, M. (2015). The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ*, 350, p. h2147.

- MAKADY, A., VAN VEELLEN, A., JONSSON, P., MOSELEY, O., D'ANDON, A., DE BOER, A., HILLEGE, H., KLUNDEL, O. y GOETTSCH, W. (2018). Using Real-World Data in Health Technology Assessment (HTA) Practice: A Comparative Study of Five HTA Agencies. *Pharmacoeconomics*, 36(3), pp. 359-368.
- MOLINERO, L. M. (2008). Metaanálisis, una guía rápida para lectores y usuarios. *Hipertension (Madr)*, 25, pp. 108-120.
- Orden SAS/3470/2009, de 16 de diciembre, por la que se publican las directrices sobre estudios posautorización de tipo observacional para medicamentos de uso humano. *BOE* 310, de 25 de diciembre de 2009, pp. 109761-109775. Disponible en: https://www.boe.es/diario_boe/txt.php?id=BOE-A-2009-20817
- ORTEGA, A., FRAGA, M. D., ALEGRE DEL REY, E. *et al.*; (2011). Comparaciones Indirectas. *Farmacia Hospitalaria*, pp. 173-175.
- PERLEGA DÍAZ, S. y PITA FERNÁNDEZ, S. (2005). Revisiones sistemáticas y metaanálisis. *CAP Aten Primaria*, 12, pp. 109-112.
- Real Decreto 1090/2015, de 4 de diciembre, por el que se regulan los ensayos clínicos con medicamentos, los Comités de Ética de la Investigación con medicamentos y el Registro Español de Estudios Clínicos. *BOE*, núm. 307, de 24 de diciembre de 2015, pp. 121923-121964. Disponible en: http://www.boe.es/diario_boe/txt.php?id=BOE-A-2015-14082
- Real Decreto Legislativo 1/2015, de 24 de julio, por el que se aprueba el texto refundido de la Ley de garantías y uso racional de los medicamentos y productos sanitarios. *BOE*, núm. 177, de 25 de julio de 2015, pp. 62935-63030. Disponible en: http://www.boe.es/diario_boe/txt.php?id=BOE-A-2015-8343
- Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 de Protección de Datos (RGPD).
- ROUSE, B., CHAIMANI, A. y LI, T. (2017). Network meta-analysis: an introduction for clinicians. *Intern Emerg Med*, 12, pp. 103-111.
- SCHNEEWEISS, S. (2018). Automated data-adaptive analytics for electronic healthcare data to study causal treatment effects. *Clin Epidemiol*, 10, pp. 771-788.
- SCHWARTZ, D. y LELLOUCH, J. (1967). Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis*, 20(8), pp. 637-648.
- Sistema de Historia Clínica Digital del Sistema Nacional de Salud. Disponible en: https://www.msssi.gob.es/organizacion/sns/planCalidadSNS/docs/HCDSNS_Castellano.pdf [último acceso 26/07/18].
- TONIN, F. S., ROTTA, I., MENDES, A. M. y PONTAROLO, R. (2017). Network mta—analysis: a technique to gather evidence from direct and indirect comparisons. *Pharm Pract*, 15, p. 943.
- U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES FOOD AND DRUG ADMINISTRATION (2015). Mobile Medical Applications. *Guidance for Industry and Food and Drug Administration Staff*. Disponible en: <https://www.fda.gov/downloads/MedicalDevices/.../UCM263366.pdf>
- WANG, S. V., SCHNEEWEISS, S., BERGER, M. L. *et al.* (2017). joint ISPE-ISPOR Special Task Force on Real World Evidence in Health Care Decision Making. Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0. *Pharmacoepidemiol Drug Saf*, 26(9), pp. 1018-1032. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5639362/pdf/PDS-26-1018.pdf>
- XIAO, WANG y LUO (2012). Sildenafil citrate for erectile dysfunction in patients with multiple sclerosis. *Cochrane Database of Systematic Reviews*, Issue 4. Art. No.: CD009427. DOI: 10.1002/14651858.CD009427.pub2



8

RESULTADOS EN SALUD A EVALUAR EN DIFERENTES PATOLOGÍAS: ENFERMEDADES INFECCIOSAS

MARÍA DEL MAR ARENAS MIRAS, JUAN PABLO HORCAJADA GALLEGO

■ 8.1. INTRODUCCIÓN

Las enfermedades infecciosas han sufrido a lo largo de la historia una evolución sin precedentes. A principios del siglo XX las enfermedades infecciosas seguían siendo una de las primeras causas de mortalidad a nivel global, asociadas a secuelas y a una disminución en la calidad de vida de los pacientes. De forma progresiva, y gracias a hechos tan relevantes como el desarrollo de los antibióticos, la mejoría de técnicas diagnósticas, el desarrollo de programas de control de salud, mejoras en los sistemas de higiene y saneamiento, la implantación de programas de vacunación, etc., se produjo un cambio que impactó directamente en la mortalidad poblacional. Sin embargo, cuando el problema parecía solucionado surgió un nuevo problema relacionado con el uso generalizado de antibióticos: la selección y diseminación de resistencias a los antibióticos. Tanto dentro como fuera del ámbito hospitalario, la emergencia de resistencias antibióticas es uno de los problemas más importantes de salud pública en el que deben tomar partido tanto los gobiernos como la sociedad en general según recomendaciones de la Organización Mundial de la Salud (OMS) (WHO, 2015).

En esta situación el abordaje de los pacientes se hace muy complicado, y requiere la participación de profesionales expertos en la materia que sean capaces de hacer frente al diagnóstico y tratamiento del problema, así como realizar el manejo clínico optimizado y seguimiento del paciente con una enfermedad infecciosa concreta.

Otros fenómenos que están influyendo en el manejo de los pacientes con enfermedades infecciosas en la actualidad son la globalización de las infecciones, la mejora del pronóstico de los pacientes VIH y de los infectados por el virus de la hepatitis C gracias a la mayor eficiencia de los tratamientos antivirales, el desarrollo de los programas de trasplante de órganos y la cronicidad y la mayor complejidad de los pacientes atendidos tanto en atención primaria como en los hospitales.

Es importante evaluar los procesos desarrollados en el manejo de las enfermedades infecciosas y los resultados en salud derivados de dicha actividad para mejorar la calidad de la actuación y la atención global al paciente, así como su impacto en los sistemas de salud.

■ 8.2. RESULTADOS EN SALUD CLÍNICOS

Con respecto a los resultados clínicos, la eficacia clínica y la seguridad de las acciones realizadas en el abordaje de las enfermedades infecciosas son los más relevantes. La incidencia de infecciones también es un indicador crucial en esta área.

Los indicadores más utilizados para medir resultados en salud clínicos son:

- Tasa de mortalidad.
- Estancia media hospitalaria.
- Tasa de reingresos.
- Tasa de complicaciones.
- Necesidad de nueva cirugía en caso de infección quirúrgica intervenida.
- La medida de años de calidad de vida.

Otros indicadores relacionados con el manejo integral de las infecciones son:

- La tasa de infección nosocomial.
- La tasa de bacteriemia asociada a catéter endovenoso.
- La tasa de infección de la herida quirúrgica.
- La tasa de neumonía asociada a ventilación mecánica.
- La tasa de infección urinaria asociada a sonda urinaria.
- La tasa de infección/colonización por microorganismos multirresistentes.
- La tasa de consumo de antibióticos.

Muchos de estos indicadores pueden compararse con indicadores regionales, nacionales e internacionales. De esta manera pueden detectarse desviaciones susceptibles de corrección. Algunos de los indicadores deberían utilizarse para evaluar la efectividad de los tratamientos administrados y las decisiones tomadas en relación con el problema infeccioso del paciente.

La evaluación sistemática y continua de los indicadores es necesaria para lograr los objetivos clínicos, así como a los relacionados con la calidad de vida relacionada con la salud, y las preferencias de los pacientes o satisfacción de estos con el sistema sanitario.

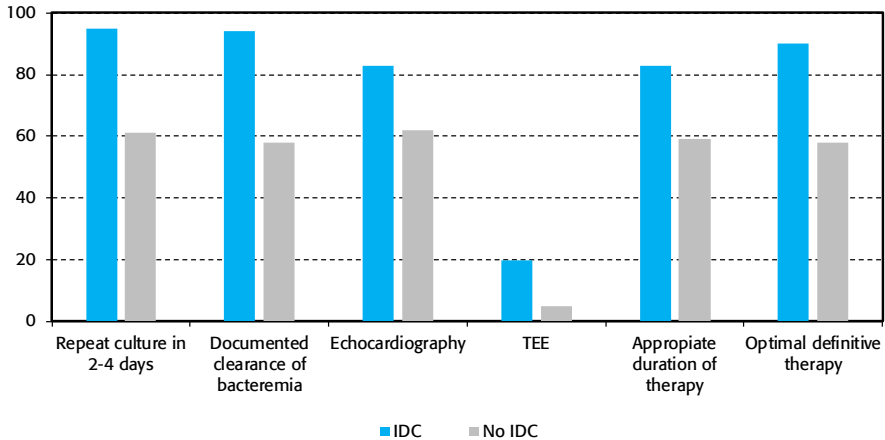
A modo de ejemplo, el estudio de la eficacia del control de las infecciones nosocomiales iniciado en los años setenta por los Centers for Disease Control (CDC), demostró cómo los programas de control de infecciones conseguían reducir las cifras de infección nosocomial en un 32 % (Haley *et al.*, 1985); hecho también corro-

borado en un estudio posterior que demostró un 44 % de reducción en el número de bacteriemias en las unidades de cuidados intensivos cuando se instauraban las medidas de control adecuadas (CDC, 2000).

Dada la creciente complejidad de las enfermedades infecciosas, no solamente debido al aumento de resistencias a los antibióticos sino también a la complejidad de los pacientes, la participación del médico especialista en enfermedades infecciosas en la atención infecciones complejas, graves o resistentes ha demostrado su utilidad en varios resultados en salud. Por ejemplo, es posible reducir el consumo de antibióticos sin afectar negativamente a la evolución del paciente y sin aumentar la mortalidad ni aumentar la tasa de reingresos (Nilholm *et al.*, 2015). Asimismo, la mejor prescripción de antibióticos produce un impacto positivo tanto sobre la estancia hospitalaria como en reducción de complicaciones secundarias al uso de antibióticos (como la diarrea por *clostridium difficile*) (Davey *et al.*, 2013). Los pacientes ancianos son los más vulnerables a los efectos colaterales de los antibióticos, tanto por el riesgo incrementado de infección por *C. difficile* como por un mayor riesgo de convertirse en portadores de bacterias multirresistentes por su edad, dependencia y fragilidad (CDC, 2012).

Otro ejemplo de la necesidad de optimizar del manejo clínico para lograr mejores resultados en salud es la bacteriemia de origen nosocomial. Se ha demostrado en esta patología que la mortalidad intrahospitalaria es mayor en pacientes que recibieron tratamiento empírico inadecuado frente a los que recibieron un tratamiento adecuado (26.1 % vs. 16.6 %; $P=0.015$) (Scharamm *et al.*, 2006). Dentro de las bacteriemias, la bacteriemia por *S. aureus* se ha asociado con una importante morbilidad con cifras de mortalidad relacionada del 20 % al 30 % (Lopez-Cortes *et al.*, 2013). La inclusión de un consultor de enfermedades infecciosas en el abordaje del paciente mejoró los resultados clínicos de las bacteriemias por *S. aureus* (Jenkins *et al.*, 2008). Entre los indicadores de calidad que debe controlar el infectólogo en el manejo de esta patología se encuentran: la identificación del foco de infección y retirada de material (vías intravenosas, por ejemplo) la realización de hemocultivos repetidos hasta la resolución de la bacteriemia, ecocardiograma, diferenciación entre bacteriemia complicada y no complicada, selección del tratamiento intravenoso adecuado y control de la duración adecuada del tratamiento antibiótico (Liu *et al.*, 2011). En un estudio retrospectivo multicéntrico realizado en cinco hospitales, en el cual se pretendía evaluar el impacto del consultor de enfermedades infecciosas en cuanto a la adherencia a los indicadores de calidad y resultados clínicos, en pacientes con bacteriemia por *S. aureus*, se evidenció que la consulta con enfermedades infecciosas se asoció con una menor incidencia de fallos de tratamiento y menor riesgo de muerte, existiendo una mayor adherencia a los indicadores de calidad en los pacientes en los que se realizó interconsulta con el infectólogo (Figura 1) (Turner *et al.*, 2016). Esto no se ha evidenciado únicamente en la bacteriemia por *S. aureus*, sino también en otras infecciones graves como las candidemias o las bacteriemias por bacterias multirresistentes. En ellas, la participación del especialista de

Figura 1.

INDICADORES DE CALIDAD PARA PACIENTES CON BACTERIEMIA POR *S. AUREUS* CON Y SIN CONSULTA DE ENFERMEDADES INFECCIOSAS

Nota: IDC: consulta de enfermedades infecciosas. No IDC: no consulta de enfermedades infecciosas.
Fuentes: Adaptado de Turner *et al.* (2016).

enfermedades infecciosas con apoyo en la prescripción antibiótica consiguió disminuir la incidencia de estas infecciones y la mortalidad asociada (Molina *et al.*, 2017).

El especialista de enfermedades infecciosas también puede actuar como colaborador en los programas de prevención de enfermedades transmisibles y manejo de brotes de infecciones nosocomiales. Asimismo, puede también participar en la elaboración, implementación y cumplimiento de programas de vacunación (McQuillen *et al.*, 2008).

El reciente desarrollo de los programas de optimización de antibióticos (PROA), de los que hablaremos más adelante, ha producido mejoras en resultados en salud como la reducción de la estancia hospitalaria y del gasto sanitario mediante la detección precoz de tratamientos antibióticos inadecuados y la disminución de la duración de los tratamientos sin suponer un aumento de mortalidad (Nilholm *et al.*, 2015; Güerri-Fernández *et al.*, 2016). Estos programas también han demostrado una reducción en la incidencia de candidemias y bacteriemias por bacterias multirresistentes y en la mortalidad relacionada con las mismas (Molina *et al.*, 2017).

Para facilitar la adecuación del tratamiento antibiótico de las infecciones es importante la elaboración de guías de terapia antimicrobiana tanto en atención primaria como en hospitales y otros centros sanitarios, que sean actualizadas, adaptadas al entorno ecológico y accesibles a todos los médicos del centro. Son numerosos los procesos infecciosos en los que se recomienda elaborar guías de tratamiento

Tabla 1.

PROCESOS INFECCIOSOS EN LOS QUE SE RECOMIENDA ELABORAR PROTOCOLOS DE TRATAMIENTO EMPÍRICO Y DIRIGIDO

Infecciones respiratorias

Neumonías adquiridas en la comunidad

Neumonías nosocomiales no asociadas y asociadas a ventilación mecánica

Infecciones de la piel, partes blandas y de las mucosas

Erisipela y celulitis

Fascitisnecrosante

Pie diabético

Infecciones cardiovasculares

Endocarditis (nativas y protésicas)

Infecciones de marcapasos y desfibriladores automáticos implantables

Infecciones relacionadas con los catéteres vasculares

Infecciones intraabdominales

Colecistitis y colangitis

Peritonitis (secundarias y terciarias)

Infecciones del tracto urinario

Pielonefritis y prostatitis

Bacteriuria asintomática (pacientes sondados y no sondados), cistitis

Infecciones del sistema nervioso central

Meningitis aguda

Encefalitis

Infecciones osteoarticulares

Artritis aguda

Osteomielitis aguda

Sepsis grave/shock séptico sin focalidad

Fuente: Rodríguez-Baño *et al.* (2012).

empírico y dirigido (Tabla 1) ya que mejoran el pronóstico de los pacientes (Frei *et al.*, 2010). Es conveniente los centros cuenten con expertos en enfermedades infecciosas para la elaboración de las guías y como consultores en su aplicación cuando sea preciso (Nilholm *et al.*, 2015). El papel del especialista en enfermedades infecciosas en este ámbito consiste en adaptar las guías a su entorno ecológico local interpretando los datos de sensibilidad junto al servicio de microbiología y las formulaciones y dosis de antibióticos junto al servicio de farmacia. La confección y revisión de las guías deben realizarse en común con los servicios clínicos implicados (McQuillen *et al.*, 2008).

■ 8.3. RESULTADOS EN SALUD ECONÓMICOS

Desde el punto de vista de las enfermedades infecciosas, en los últimos años se han ido implantando prácticas que han tenido repercusiones económicas favorables en el ámbito de la salud. Entre ellas podríamos destacar:

- El desarrollo progresivo y paulatina generalización, tanto a nivel intra como extrahospitalario, de los PROA, con aumento creciente de apoyo institucional.
- Apoyo al resto de servicios médicos y quirúrgicos a través de la interconsulta a enfermedades infecciosas.
- El desarrollo de equipos de trabajo multidisciplinares para el control de la infección nosocomial y relacionada con la asistencia sanitaria (IRAS).

Los PROA nacen como una de las medidas principales propuestas por las sociedades científicas y expertos para optimizar el uso de antibióticos y luchar contra las resistencias bacterianas. Están formados por equipos multidisciplinares con participación de farmacéuticos, infectólogos, microbiólogos e intensivistas en centros con UCI y con estrechas relaciones con los referentes de las demás unidades médicas y quirúrgicas (Dellit *et al.*, 2007; Rodríguez-Baño *et al.*, 2012). Es conocido que el uso inadecuado de antibióticos influye de forma directa en la selección y diseminación de las resistencias a los antibióticos (Laxminarayan *et al.*, 2013), por lo que los esfuerzos encaminados a controlar este problema deben ser una prioridad. Varios estudios han identificado que entre un tercio y la mitad de las prescripciones antibióticas en los hospitales son inadecuadas (Apisarnthanarak *et al.*, 2006; Hecker *et al.*, 2003). Entre las causas de estas prescripciones inadecuadas se encuentran: la falta de conocimientos específicos por parte de los prescriptores debido al creciente número de organismos multirresistentes, entorno cada vez más exigente y complejo, falta de recursos de ayuda a la prescripción, medicina defensiva, etc. (Rodríguez-Baño *et al.*, 2012). Como ya se ha adelantado, también se ha demostrado cómo estos programas contribuyen a disminuir el gasto en antimicrobianos, suponiendo ahorros de hasta 1.7 millones de dólares anuales según las series (Nowak *et al.*, 2012; Álvarez-Lerma *et al.*, 2018).

Respecto a las interconsultas de otros servicios, además de los beneficios clínicos previamente expuestos, existen evidencias de que la participación del especialista en enfermedades infecciosas como consultor puede mejorar los resultados económicos en el manejo de los pacientes con enfermedades infecciosas (Lanbeck, Tennvall y Resman, 2016).

En cuanto a las IRAS, el gasto hospitalario derivado de las complicaciones infecciosas es un hecho relevante a tener en cuenta. El coste medio directo de un episodio de infección nosocomial en EE. UU. es de unos 15.275 dólares, que al incluir los costes colaterales supondrían unos 38.600 dólares (Roberts *et al.*, 2003). En el caso de bacteriemia por *S. aureus* resistente a meticilina (MRSA), el coste medio por episodio de bacteriemia en pacientes que recibieron un tratamiento ade-

cuado fue de 13.688 dólares frente a los 19.427 dólares de los que lo recibieron inadecuado ($p=0.019$) (Shorr, Micek y Kollef, 2008). En un intento de abordar este problema, en un estudio realizado en pacientes ingresados en unidades de cuidados intensivos con bacteriemia asociada a catéter, se demostró como el implementar un programa de control de infecciones multidisciplinar conseguía reducir la cifra de bacteriemia asociada a catéter un 66 % tras 16-18 meses de implantación del programa (Provonost *et al.*, 2006). Las infecciones por bacterias multirresistentes también tienen un importante impacto económico como demuestran varios estudios (Esteve-Palau *et al.*, 2018; Morales *et al.*, 2012).

Consecuencia de esta reducción, se redujo significativamente la morbilidad hospitalaria, así como los costes derivados de este problema tan frecuente en los hospitales. Otras experiencias similares subrayan la necesidad de trabajar conjuntamente para reducir la incidencia de las IRAS y sus consecuencias clínicas y económicas.

Para adecuar las necesidades hospitalarias en cuanto al control de la infección nosocomial, es fundamental el apoyo institucional y destinar recursos económicos a los programas de control de infecciones y antibióticos. De cara al futuro y a la sostenibilidad del sistema sanitario, la disminución de las tasas de infección nosocomial y el acortar las estancias hospitalarias son aspectos críticos donde un especialista en enfermedades infecciosas cualificado es la mejor opción para alcanzar y cuantificar este ahorro hospitalario (McQuillen *et al.*, 2008).

■ 8.4. PRINCIPALES TÉCNICAS ESTADÍSTICAS Y PARÁMETROS PARA MEDIR RESULTADOS EN SALUD

Tanto la epidemiología como la estadística son necesarias en ciencias de la salud para proporcionar instrumentos que nos permitan diseñar, realizar, interpretar y publicar los resultados en salud. Las opciones de posibles diseños de estudios son numerosas por lo que la elección debe estar condicionada por la pregunta formulada, en forma de objetivos. Antes de iniciar cualquier recogida de información se debe elegir el diseño más adecuado para el estudio. Hablaremos en primer lugar de los distintos diseños de estudios de los que disponemos y posteriormente de las medidas de frecuencia y asociación utilizadas más frecuentemente para expresar los resultados en salud.

■ 8.4.1. Tipos de estudios epidemiológicos

De forma clásica, los estudios epidemiológicos se han dividido en experimentales y observacionales (no experimentales) (Tabla 2). El interés del estudio puede ir desde la simple descripción de una realidad (epidemiología descriptiva) hasta la confirmación de una hipótesis en relación con determinados factores de riesgo o

Tabla 2.

CLASIFICACIÓN DE LOS ESTUDIOS EPIDEMIOLÓGICOS

<i>Estudios experimentales</i>
Ensayos clínicos aleatorizados controlados
Estudio cuasiexperimental
<i>Estudios observacionales</i>
<i>Descriptivos</i>
En poblaciones: estudios ecológicos
En individuos: estudios transversales o de prevalencia, series de casos, a propósito de un caso
<i>Analíticos</i>
Estudios de casos y controles
Estudios de cohortes

Fuente: Elaboración propia.

pronóstico (epidemiología analítica). Desarrollaremos de forma breve los distintos estudios disponibles.

Estudios experimentales: en los que el investigador manipula las condiciones de la investigación. En este grupo están los ensayos clínicos controlados y los cuasiexperimentales. En los ensayos clínicos controlados hay aleatorización y el investigador produce una manipulación de una exposición determinada en un grupo de individuos que se compara con otro grupo en el que no se intervino, o que se expone a otra intervención. Si ha existido intervención, pero no aleatorización se habla de estudios cuasiexperimentales. Un ejemplo de ensayo clínico sería el realizado por Darouiche *et al.*, en 2010, en el que se aleatorizo a pacientes a usar una solución previa a la cirugía con clorhexidina frente a povidona yodada siendo el resultado del estudio la presencia de infección de la herida quirúrgica en los 30 días después de la cirugía (Darouiche *et al.*, 2010).

Estudios observacionales: en este tipo de estudios no hay posibilidad de asignación aleatoria de la exposición. Las personas presentan exposición o no a distintos factores, sin que intervengan los investigadores y limitándose éstos a observar lo que sucede. Los estudios observacionales se subdividen a su vez en: estudios descriptivos y estudios analíticos.

- a. *Los estudios descriptivos* describen la frecuencia y las características más importantes de un problema de salud. Los principales tipos de estudios descriptivos son los estudios ecológicos, los estudios transversales o de prevalencia y los estudios de series de casos:

Los *estudios ecológicos* no utilizan la información del individuo de una forma aislada, sino que utilizan datos agregados de toda la población. Estos

estudios son utilizados con cierta frecuencia en el campo de la salud pública como sería determinar la inmunidad de un grupo poblacional frente a una determinada enfermedad. La principal limitación es que no pueden determinar si existe una asociación entre una exposición y una enfermedad a nivel individual. En este contexto, la falacia ecológica consiste en obtener conclusiones inadecuadas a nivel individual basadas en datos poblacionales.

Los *estudios transversales o de prevalencia* estudian simultáneamente la exposición y la enfermedad en una población bien definida en un momento determinado. Un ejemplo de este tipo de estudios sería determinar la prevalencia en el primer trimestre del año de las infecciones nosocomiales en el hospital. Esta medición simultánea no permite conocer la secuencia temporal de los acontecimientos y, por tanto, no es posible determinar si la exposición precedió a la enfermedad o viceversa.

Las *series de casos* describen la experiencia de un paciente o un grupo de pacientes con un diagnóstico similar. Se describe una característica de una enfermedad o de un paciente, que sirven para generar nuevas hipótesis.

- b. Por otro lado, los *estudios analíticos* son una buena opción cuando se intentan verificar una hipótesis de trabajo. Los estudios analíticos más frecuentes son los estudios de cohortes y los estudios de casos y controles:

En los *estudios de cohortes*, los individuos son identificados en función de la presencia o ausencia de exposición a un determinado factor, estando al principio del estudio libres de enfermedad. Durante el estudio, son seguidos durante un periodo de tiempo para observar la frecuencia de aparición de la enfermedad de interés. Ejemplo de este tipo de estudio sería evaluar si el tratamiento con fluorquinolonas previas influye en la aparición de infecciones por MRSA. Los estudios de cohortes pueden ser prospectivos o retrospectivos, dependiendo de la relación temporal entre el inicio del estudio y la presencia de la enfermedad. La medida de asociación más utilizada es el Riesgo Relativo.

En los estudios de casos y controles, se identifican a personas con una enfermedad (casos) y se comparan con un grupo control (controles) que no tenga la enfermedad. Son por definición retrospectivos. Un ejemplo sería el estudio de los pacientes que se han colonizado por MRSA durante el ingreso en el mes previo e identificar los factores relacionados con la colonización. En casos y controles se mide la exposición a uno o varios factores relacionados con la enfermedad y se compara la frecuencia de exposición a este factor u otros factores entre los casos y los controles. Si la frecuencia de exposición es mayor en el grupo de casos que en los controles, hay una asociación entre la causa y el efecto. La medida de asociación más utilizada es la *Odds Ratio*. Dentro de los estudios de casos y controles tenemos estudios especiales:

- *Estudios de casos y controles anidados (nested case-control studies)*: A partir de una cohorte original, se seleccionan todos los casos que han aparecido y se identifican “n” controles por caso. Un ejemplo de este estudio sería el realizado por Landelle *et al.*, 2014, en el que se estudió si la colonización por SAMS protegía frente a la adquisición nosocomial de MRSA.
- *Diseño de caso-caso-control*: Para dos tipos de caso se utilizan los mismos controles. Este tipo de estudios es de especial interés para el análisis de los factores de riesgo en resistencias a antimicrobianos.

Para la elección de un determinado estudio, debemos poner en la balanza, por un lado, las ideas teóricas y éticas y por otro los medios prácticos de los que disponemos, para así intentar conseguir finalmente un estudio con alta validez interna evitando los sesgos. Los aspectos claves que se deben tener en consideración para el diseño más apropiado en enfermedades infecciosas se resumen en la Tabla 3 (McGregor y Furuno, 2014).

Tabla 3.

CONSIDERACIONES IMPORTANTES EN EL DISEÑO Y ANÁLISIS DE LOS ESTUDIOS QUE EVALÚAN EL AJUSTE DE LA ANTIBIOTERAPIA

<i>Diseño del estudio</i>	<i>Análisis de los datos</i>
Medida de resultados (ejem. clínicos, microbiológicos, de proceso)	Tipo y distribución de los datos
Medida de objetivos	Selección del modelo de regresión
Completar y validar la recolección de datos	Correlacionar la estructura de los datos
Duración de periodos de tiempo	Tendencia secular
Duración del estudio	Efactor retrasados
Selección del grupo control	Modelo de parameterización
Uso de variables dependientes no equivalentes	Representación gráfica de los datos
Influencias externas y fuentes de sesgos	Control de variables confusoras

Fuente: Adaptado de Mc Gregor y Furuno (2014).

■ 8.4.2. Medidas de frecuencia y asociación más utilizadas en enfermedades infecciosas

Las medidas de frecuencia permiten valorar lo común que es un fenómeno de interés. Las más utilizadas para la descripción y comparación de grupos son la prevalencia y la incidencia:

Prevalencia: cuantifica la proporción de individuos de una población que padecen una enfermedad en un momento o periodo de tiempo determinado. Podríamos así hablar de la prevalencia de infección nosocomial en el año 2017 en nuestro hospital. La prevalencia no tiene dimensiones y se suele expresar como un porcentaje. Es una medida característica en los estudios transversales.

Incidencia: se define como el número de casos nuevos de una enfermedad que se desarrollan en una población durante un período de tiempo determinado. Hay distintos tipos de medidas de incidencia:

- **Incidencia acumulada:** proporción de sujetos, libres de la enfermedad al comienzo del estudio, que desarrollan la enfermedad durante un período de tiempo de seguimiento. Por ejemplo, el número de infecciones quirúrgicas tras una cirugía de prótesis de rodilla en un año. La incidencia acumulada se expresa como una proporción y representa la probabilidad de que un sujeto libre de la enfermedad la desarrolle durante el período de seguimiento.
- **Densidad de incidencia:** cociente entre el número de casos nuevos de enfermedad ocurridos durante el tiempo de seguimiento y la suma de todos los tiempos a riesgo que aportan los sujetos de la población de estudio (personas-tiempo). Sería ejemplo el número de bacteriemias relacionadas con catéter en la población de estudio en los días totales en los que han llevado catéter.
- **Tasa de incidencia:** cociente entre el número de casos nuevos de enfermedad y tiempos no individuales (los datos son agregados de la población en riesgo a lo largo de todo el estudio). Un ejemplo de este caso sería el número de infecciones por MARSA entre el total de estancias hospitalarias.

Las medidas de asociación miden la fuerza con la que una determinada enfermedad o evento (resultado) está asociada o relacionada con un determinado factor o exposición (causa). Las principales medidas de la magnitud de una asociación son el Riesgo Relativo y la *Odds Ratio*:

El *Riesgo Relativo* compara la incidencia de la enfermedad en los sujetos expuestos con la incidencia de la enfermedad en los sujetos no expuestos; por ejemplo, el riesgo de infección por *Clostridium* tras exposición a antibióticos. El RR informa de cuántas veces se incrementa el riesgo de padecer la enfermedad o evento en los expuestos con respecto a los no expuestos, siendo mayor de 1 cuando el factor de exposición se comporte como factor de riesgo, menor de 1 cuando tenga efecto protector e igual a 1 cuando no exista asociación entre el factor de exposición y el evento. Es una medida de asociación característica de los estudios de cohortes.

La *Odds Ratio* compara si la frecuencia de exposición es mayor en el grupo de casos de la enfermedad que en los controles. Es un cociente entre una probabilidad (la de haber desarrollado la enfermedad) y su probabilidad complementaria

(no haberla desarrollado) Es la medida de asociación utilizada en los estudios de casos y controles.

■ 8.5. METODOLOGÍA DISPONIBLE PARA OBTENER RESULTADOS EN SALUD EN EL MUNDO REAL (RWE)

El *Real World Evidence (RWE)* o Evidencia en la Vida Real tienen su fuente de información en los datos observacionales del sistema de salud y cada vez más en los datos de los pacientes. Estos datos son reflejo de la experiencia del paciente que recibe un determinado tratamiento o bien que es usuario un dispositivo médico en entornos del mundo real, a diferencia de cuando lo usa en el contexto de un ensayo clínico controlado. Las fuentes tradicionales de RWE pueden incluir datos clínicos de registros médicos, la codificación de patologías según la clasificación internacional ICD, el Conjunto Mínimo Básico de Datos (CMBD) que se utiliza en gestión hospitalaria, encuestas a pacientes y a profesionales sanitarios, registros de pacientes, datos de resultados moleculares y de laboratorio, medios de comunicación social, tecnologías móviles...etc. Como ejemplos concretos podríamos citar la encuesta de satisfacción durante el ingreso en planta de hospitalización, las complicaciones de los pacientes durante el ingreso, los efectos secundarios de los tratamientos, etc. La información obtenida a través de estas fuentes puede ser útil para tomar decisiones estratégicas y organizativas y para analizar la eficiencia de los protocolos clínicos o de los fármacos o dispositivo en la vida real.

Es difícil realizar en todos los ámbitos de las enfermedades infecciosas ensayos clínicos controlados, tanto por la complejidad de la puesta en marcha como por los recursos que se precisan. Sin embargo, estudios realizados en un entorno real con pacientes accesibles para estudiar su evolución ante un determinado tratamiento antibiótico o proceso diagnóstico, puede ayudar a mejorar la práctica habitual e incluso en ocasiones a modificarla. Serían estudios que, aunque no cuenten con la complejidad del ensayo clínico, tienen la suficiente validez como para darnos resultados útiles y fiables. Ejemplos de este tipo de estudios serían los estudios ecológicos, los transversales o de prevalencia, las series de casos, los estudios de cohortes y los de casos y controles.

Teniendo en cuenta este tipo de aproximaciones a la vida real del paciente se generan nuevas hipótesis de trabajo y se identifican nuevos objetivos.

■ 8.6. TRANSFORMACIÓN DIGITAL Y *BIG DATA* EN LA OBTENCIÓN DE DATOS DE RESULTADOS EN SALUD EN EL SNS

La recogida de la información debe ser cuidadosa y completa y lo más estandarizada posible y con medios informáticos. Para avanzar en la medición de los resultados en salud resulta imprescindible contar con registros de pacientes y disponer

de la historia clínica informatizada. Esta se ha convertido en una de las herramientas más útiles que permiten al clínico la obtención de datos del paciente de forma rápida y fiable. En varias comunidades autónomas se han desarrollado sistemas de trabajo en red compartida interhospitalaria e intercentros, de forma que la información de un mismo paciente puede ser visualizada en distintos centros a los que acuda, permitiendo tener acceso rápido y veraz a información del paciente. En el campo de las enfermedades infecciosas la información útil es la referente a eventos clínicos, tratamientos antibióticos prescritos previamente, pruebas diagnósticas realizadas e información microbiológica y de laboratorio. El sistema de red compartida llega incluso a la medicina primaria, siendo posible un intercambio de información más fluido y bidireccional del especialista con el médico de familia.

El futuro es aspirar a tener una red global a nivel mundial de datos compartidos que sea analizable en cualquier parte del mundo. En este sentido disponer de bases de datos compartidas de microbiología (con información sobre resistencias a nivel mundial) o incluso de la secuenciación genómica de diversas bacterias permitiría estudiar tendencias a nivel mundial y detectar de forma precoz problemas de resistencias antibióticas, brotes y epidemias. Las grandes bases de datos también permitirían responder a algunas preguntas clínicas y de gestión hospitalaria sin tener que llevar a cabo estudios prospectivos, lentos y costosos.

Otra de las fuentes de información que nos permiten un acceso rápido y sencillo a información actualizada sobre enfermedades infecciosas son las aplicaciones móviles y páginas web específicas en este campo. A lo largo de los últimos años hemos visto como se han ido desarrollando diversas aplicaciones que nos permiten obtener información sobre tratamiento de antibióticos, indicaciones de tratamientos según la patología que presente el paciente, acceso a fórmulas matemáticas de uso diario para ajuste de tratamiento e incluso posibilidad de obtener información de interacciones medicamentosas. Como ejemplo de páginas web, existen varias que ayudan a mejorar la prescripción de antibióticos¹.

■ 8.7. PAPEL DE LA MEDICIÓN DE RESULTADOS EN SALUD EN LA GESTIÓN SANITARIA: PRESENTE Y FUTURO

La complejidad cada vez mayor de las enfermedades infecciosas obliga disponer de indicadores de resultados medibles y fiables que ayuden a gestionar mejor los recursos. En la actualidad entre los indicadores que pueden ser útiles en este campo se encuentran la tasa de mortalidad, la complejidad del paciente, la estancia hospitalaria, la tasa de reingresos y las complicaciones. Además, el consumo de antibióticos y las resistencias son excelentes indicadores del manejo de estas

¹ Prudent Antibiotic User Website (<http://www.pauseonline.org.uk>); Healthcare Infection Control Special Interest Group (http://www.asid.net.au/hicsigwiki/index.php?title=Antibioticstewardship_programs#Guides); Centers for Disease Control and Prevention: (<http://www.cdc.gov/drugresistance/healthcare/default.htm>) y Infectionnet (<http://infectionnet.org/>).

enfermedades. Las tasas de infección posquirúrgica o de las relacionadas con dispositivos son necesarias para conocer la calidad asistencial relacionada con dichos procedimientos. En el futuro los *big data* nos servirán como fuente de información que permitirán analizar nuestra situación con respecto a nosotros mismos en otros periodos de tiempo o con respecto a otras unidades o centros del ámbito nacional e internacional.

Con el fin de conseguir que los resultados en salud puedan ser útiles para medir la calidad de la actividad clínica, el International Consortium for Health Outcomes Measurement (ICHOM) (organización sin ánimo de lucro que promueve la definición de resultados específicos para cada proceso clínico) organiza grupos de trabajo con profesionales clínicos y pacientes para definir qué resultados son los más importantes durante el proceso de la enfermedad. Esta plataforma tiene publicados varios resultados referentes a distintos procesos médicos, estando pendiente aún de definición los referidos a enfermedades infecciosas para que puedan ser utilizados como guía de mejora de la calidad asistencial en nuestro campo.

Por otra parte, la Sociedad Española de Directivos de la Salud (SEDISA) ha dado a conocer recientemente (octubre de 2016) en sus IX Jornadas Nacionales, los diez puntos clave en el valor y aportación de los resultados de salud. Centrándonos en las enfermedades infecciosas, consideramos como más importantes los siguientes.

- Deben ayudar a la toma de decisiones para reorganizar los recursos.
- Son el lazo de unión entre la investigación básica y la clínica.
- Pueden ser parte clave en la materialización de la participación de los pacientes en la sanidad, siendo básicos para optimizar recursos y procesos asistenciales.
- Nos deben servir de unión entre los distintos servicios sanitarios regionales, con acceso a la innovación de forma equitativa para todos los pacientes siendo claves en la evolución del modelo sanitario.

■ 8.8. EVALUACIÓN DE RESULTADOS EN SALUD: PERSPECTIVA DESDE LA ADMINISTRACIÓN SANITARIA

Después de todo lo desarrollado anteriormente en cuanto a resultados en salud, y a modo de resumen se puede afirmar que:

- Los programas de control de infecciones consiguen reducir las cifras de infección nosocomial y los gastos derivados.
- La participación del consultor en enfermedades infecciosas mejora los resultados y manejo clínico del paciente en infecciones como la bacteriemia por

S. aureus (Jenkins *et al.*, 2008), con disminución de incidencia y mortalidad en otras como la candidemia o bacteriemias por patógenos multirresistentes (Molina *et al.*, 2017).

- La participación del médico especialista en enfermedades infecciosas reduce el consumo de antibióticos sin efectos negativos sobre la evolución, mortalidad ni reingresos (Nilholm *et al.*, 2015).
- La optimización del uso de los antibióticos tiene un impacto positivo sobre estancia hospitalaria y complicaciones secundarias al uso de antibióticos (Davey *et al.*, 2013).
- Los PROA suponen reducción de estancia hospitalaria, detección precoz de tratamientos inadecuados y disminución de la duración de estos sin aumento de mortalidad (Güerri-Fernández *et al.*, 2016).
- La participación del especialista en enfermedades infecciosas como consultor mejora los resultados económicos en los pacientes con enfermedades infecciosas (Lanbeck, Tennvall y Resman, 2016).

Vistos los resultados en salud positivos que tienen las acciones llevadas a cabo por los especialistas en enfermedades infecciosas, tanto en el ámbito clínico como económico, la administración debe responder de forma positiva para asegurar la sostenibilidad de la actividad sanitaria. Para esto son indispensables:

- La dotación de personal especializado.
- Ayudas en la mejora del proceso diagnóstico: asegurar la incorporación de nuevas técnicas que confieran mayor eficacia y rapidez al proceso diagnóstico.
- Fomentar la interacción entre distintas áreas. Permitiendo el trabajo del médico de enfermedades infecciosas integrado en distintas unidades médico-quirúrgicas.
- Fomentar el desarrollo de equipos multidisciplinares que incluyan a distintos especialistas (médicos, farmacéuticos, microbiólogos, enfermería...).

Las organizaciones sanitarias son en este sentido un motor importante que ayuda a impulsar y dar valor a todas estas acciones consideradas como indispensables en el campo de las enfermedades infecciosas. Así, la Sociedad Española de Enfermedades Infecciosas y Microbiología Clínica (SEIMC) junto con la Sociedad Española de Directivos de la Salud (SEDISA), realizaron unas jornadas en las que ambas sociedades destacaron la necesidad de que el Sistema Nacional de Salud mejore el nivel de integración de los protocolos de salud pública (tanto los elaborados por Ministerio de Sanidad como los adaptados a los distintos centros sanitarios), se garantice el liderazgo del médico con un proceso de formación adecuado y específico así como el garantizar un sistema de información fiable integrado en el ámbito nacional, que permita la comparación en tiempo real entre los distintos subsistemas de información autonómicos de salud y entre centros.

Ambas sociedades recalcan que el proceso de gestión de la asistencia sanitaria ante el reto de las enfermedades infecciosas debe partir de la elaboración de un plan de acción diseñado y liderado por profesionales expertos en la materia, con el apoyo institucional y el compromiso en su ejecución garantizado por parte de las administraciones públicas. Asimismo, se deberían establecer criterios de evaluación que faciliten la monitorización y actualización de los planes de acción, de forma periódica y transparente.

■ REFERENCIAS

- ÁLVAREZ-LERMA, F., GRAU, S., ECHEVARRIA-ESNAL, D. *et al.* (2018). A before-and-after study of the effectiveness of an antimicrobial stewardship program in critical care. *Antimicrob Agents Chemother*, 62(4), pp. e01825-7.
- APISARNTHANARAK, A., DANCHAIVIJITR, S., KHAWCHAROENPORN, T. *et al.* (2006). Effectiveness of education and an Antibiotic-Control Program in a Tertiary Care Hospital in Thailand. *Clin Infect Dis*, 42(6), pp. 768–775.
- CDC. (2000). Monitoring hospital-acquired infections to promote patient's safety-United States, 1990-1999. *MMWR Morb Mortal Wkly Rep.*, 49 (08), pp. 149-153.
- (2012). Vital sings: preventing Clostridium difficile infections. *MMWR Morb Mortal Wkly Rep*, 61(09), pp. 157-62.
- DAROUICHE, R. O., WALL, M. J., ITANI, K. M. F. *et al.* Chlorhexidine-Alcohol versus Povidone-Iodine for Surgical-Site Antisepsis. *N Engl J Med*, 362(1), pp. 18-26.
- DAVEY, P., BROWN, E., CHARANI, E. *et al.* (2013). Interventions to improve antibiotic prescribing practices for hospital inpatients. *Cochrane Database Syst Rev*, 30(4); CD003543.
- DELLIT, T. H., OWENS, R. C., MCGOWAN, J. E. *et al.* (2007). Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America guidelines for developing an institutional program to enhance antimicrobial stewardship. *Clin Infect Dis*, 44(2), pp. 159–177.
- ESTEVE-PALAU, P., GRAU, S., HERRERA, S. *et al.* (2018). Impact of an antimicrobial stewardship program on urinary tract infections caused by extended-spectrum β -lactamase-producing Escherichia coli. *Rev Esp Quimioter*, 31(2), pp. 110–117.
- FREI, C. R., ATTRIDGE, R. T., MORTENSEN, E. M. *et al.* (2010). Guideline-concordant antibiotic use and survival among patients with community-acquired pneumonia admitted to the intensive care unit. *Clin Ther*, 32(2), pp. 293–299.
- GÜERRI-FERNÁNDEZ, R., VILLAR-GARCÍA, J., HERRERA-FERNÁNDEZ, S. *et al.* (2016). An antimicrobial stewardship program reduces antimicrobial therapy duration and hospital stay in surgical wards. *Rev Esp Quimioter*, 29(3), pp. 119-121.
- HALEY, R. W., CULVER, D. H., WHITE, J. W. *et al.* (1985). The efficacy of infection surveillance and control programs in preventing nosocomial infections in US hospitals. *Am J Epidemiol*, 121 (2), pp. 182-205.
- HECKER, M. L., ARON, D. C., PATEL, N. P. *et al.* (2003). Unnecessary use of antimicrobials in hospitalized patients: current patterns of misuse with an emphasis on the antianaerobic spectrum of activity. *Arch Intern Med*, 163(8), pp. 972–978.
- JENKINS, T. C., PRICE, C. S., SABEL, A. L. *et al.* (2008). Impact of routine infectious diseases service consultation on the evaluation, management, and outcomes of Staphylococcus aureus bacteremia. *Clin Infect Dis*, 46(7), pp. 1000-1008.

- LANBECK, P., TENNVALL, G., RESMAN, F. (2016). A cost analysis of introducing an infectious disease specialist-guided antimicrobial stewardship in an area with relatively low prevalence of antimicrobial resistance. *BMC Health Services Research*, 16, p. 311.
- LANDELLE, C., ITEN, A., UCKAY, I. *et al.* (2014). Does colonization with methicillin-susceptible *Staphylococcus aureus* protect against nosocomial acquisition of methicillin-resistant *S. aureus*? *Infect Control Hosp Epidemiol*, 35(5), pp. 527-533.
- LAXMINARAYAN, R., DUSE, A., WATTAL, C. *et al.* (2013). Antibiotic resistance-the need of global solutions. *Lancet Infect Dis*, 13(12), pp. 1057-1098.
- LIU, C., BAYER, A., COSGROVE, E. *et al.* (2011). Clinical practice guidelines by the Infectious Diseases Society of America for the treatment of methicillin-resistant *Staphylococcus aureus* infections in adults and children. *Clin Infect Dis*, 52(3), pp. e18-e55.
- LOPEZ-CORTES, L. E., DEL TORO, M. D., GÁLVEZ-ACEBAL, J. *et al.* (2013). Impact of an evidence-based bundle intervention in the quality-of-care management and out-come of *Staphylococcus aureus* bacteriemia. *Clin Infect Dis*, 57 (9), pp. 1225-1233.
- MCGREGOR, J. y FURUNO, J. (2014). Optimizing Research Methods Used for the Evaluation of Antimicrobial Stewardship Program. *Clin Infect Dis*, 59(Suppl 3), pp. S185-S192.
- MCQUILLEN, D. P., PETRAK, R. M., WASSERMAN, R. B. *et al.* (2008). The value of infectious diseases specialists: non-patient care activities. *Clin Infect Dis*, 47(8), pp. 1051-1063.
- MOLINA, J., VALCERLOS, E., WON, R. *et al.* (2017). Long-Term Impact of an Educational Antimicrobial Stewardship Program on Hospital-Acquired Candidemia and Multidrug-Resistant Bloodstream Infections: A Quasi-Experimental Study of Interrupted Time-Series Analysis. *Clin Infect Dis*, 65(12), pp. 1992-1999.
- MORALES, E., COTS, F., SALA, M. *et al.* (2012). Hospital costs of nosocomial multi-drug resistant *Pseudomonas aeruginosa* acquisition. *BMC Health Serv Res*, 12, p. 122.
- NILHOLM, H., HOLMSTRAND, L., AHL, J. *et al.* (2015). An audit-based, infectious disease specialist-guided antimicrobial stewardship program profoundly reduced antibiotic use without negatively affecting patient outcomes. *Open Forum Infectious Diseases*, 2 (2), p. ofv042.
- NOWAK, M. A., NELSON, R. E., BREIDENBACH, J. L. *et al.* (2012). Clinical and economic outcomes of a prospective antimicrobial stewardship program. *Am J Health Syst. Pharm*, 69(17), pp. 1500-1508.
- PROVONOST, P., NEEDHAM, D., BERENHOLTZ, S. *et al.* (2006). An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med*, 355(26), pp. 2725-2732.
- ROBERTS, R. R., SCOTT, R. D., CORDELL, R. *et al.* (2003). The use of economic modeling to determinate the hospital costs associated with nosocomial infections. *Clin Infect Dis*, 36(11), pp. 1424-1432.
- RODRÍGUEZ-BAÑO, J., PAÑO-PARDO, J. R., ALVAREZ-ROCHE, L. *et al.* (2012). Programas de optimización de uso de antimicrobianos (PROA) en hospitales españoles: documento de consenso GEIH-SEIMC, SEFH y SEMSPH. *Enferm Infecc Microbiol Clin*, 30(1), pp. 22.e1-22.e23.
- SCHARAMM, G. E., JOHNSON, J. A., ROHERTY, J. A. *et al.* (2006). Methicillin-resistant *Staphylococcus aureus* sterile-site infection: the importance of appropriate initial antimicrobial treatment. *Crit Care Med*, 34(8), pp. 2069-2074.
- SHORR, A. F., MICEK, S. T. y KOLLEF, M. H. (2008). Inappropriate therapy for methicillin-resistant 23. *Staphylococcus aureus*: resource utilization and cost implications. *Crit Care Med*, 36(8), pp. 2335-2340.
- TURNER, R., VALCERLOS, E., WON, R. *et al.* (2016). Impact of infectious diseases consultation on clinical outcomes of patient with *Staphylococcus aureus* bacteriemia in a community health system. *Antimicrob Agents Chemoter*, 60(10), pp. 5682-5687.

WHO. (2015). *Antimicrobial resistance*, Fact sheet, No. 194. [update May 2015; cited 2016 January 4th. Fact sheet]. Disponible en: <http://www.who.int/mediacentre/factsheets/fs194/en/>



9

RESULTADOS EN SALUD EN LA INFECCIÓN POR VIH

ALEJANDRO SANTIAGO PÉREZ, JOSÉ MANUEL MARTÍNEZ SESMERO,
MARÍA JESÚS VIVANCOS GALLEGO

■ 9.1. INTRODUCCIÓN

La incidencia de SIDA ha disminuido progresivamente desde 1996. No obstante, España sigue siendo uno de los países con mayor incidencia de SIDA en Europa Occidental, con un total de 595 casos de sida notificados en 2015, lo que supone una tasa de 1,8/100.000 habitantes. Paralelamente la mortalidad por VIH y SIDA se ha ralentizado en los últimos años. Si tenemos en cuenta los últimos datos de incidencia y mortalidad en España, el incremento anual del número de personas que viven con VIH en España rondaría los 2.500-3.000 casos. Como consecuencia de la expansión de los tratamientos antirretrovirales (TAR), se ha modificado la historia natural de la enfermedad. El perfil de las personas que viven con VIH ha cambiado, y actualmente, la mitad de las personas que viven con VIH son mayores de 50 años. Debido al aumento de la esperanza de vida y a los nuevos casos de personas con VIH en tratamiento, se ha producido un incremento en el gasto en TAR que podría continuar creciendo si no se reducen las nuevas infecciones (Área de Vigilancia de VIH y Comportamientos de Riesgo, 2016; European Centre for Disease Prevention and Control [ECDC], 2016; Área de vigilancia de VIH y conductas de riesgo, 2016).

En el ámbito internacional hay un acuerdo general para lograr los resultados explicitados por ONUSIDA para el año 2020. Se trata del Objetivo 90/90/90: el 90 % de las personas que viven con VIH están diagnosticadas, el 90 % de las personas diagnosticadas con VIH reciben tratamiento antirretroviral y el 90 % de las personas en tratamiento logran la supresión viral. En 2016, el 18 % de las personas con VIH en España no conocía su estado serológico, y de los pacientes diagnosticados, un 8 % no accedía al tratamiento antirretroviral. De los pacientes en tratamiento, un 12 % no alcanzaban la supresión viral, por lo que en estos tres pasos se perdería al 34 % de las personas con VIH en España (ECDC, 2017; Gourlay *et al.*, 2017).

■ 9.2. RESULTADOS COMUNICADOS POR LOS CLÍNICOS (CLINICIAN-REPORTED OUTCOMES)

Eficacia en ensayos clínicos

Para los pacientes *naïve* que reciben tratamiento por primera vez, existen múltiples opciones de tratamiento efectivas, bien toleradas y apropiadas para su uso. A pesar de que un tratamiento activo y bien tolerado puede detectarse en 24 sema-

nas o menos, algunas diferencias en eficacia virológica, la aparición de resistencias o problemas de tolerabilidad pueden detectarse cuando se sigue a los pacientes hasta las 48 semanas o más. Los regímenes iniciales suelen ser el mejor fármaco disponible ya que la falta de eficacia con este podría afectar al uso de los siguientes tratamientos por desarrollo de resistencias. Los tratamientos en pacientes *naive*, por tanto, son evaluados de forma estricta y son comparados con otro fármaco con experiencia y eficaz como control (US Department of Health and Human Services, & Food and Drug Administration, 2015).

Los objetivos de eficacia en los ensayos clínicos en fase 2 o 3 son:

Grupo 1. Pacientes naive, sensibles a cualquier tratamiento o pretratados sin fracaso virológico previo. Proporción de pacientes con ARN-VIH menor del límite de cuantificación a las 48 semanas usando una prueba aprobada por la FDA.

Grupo 2. Pacientes con resistencia a varios fármacos o varias clases de fármacos sin posibilidad de construir un régimen supresor. Proporción de pacientes con descenso inicial de ARN-VIH respecto a su basal de 0,5 log₁₀ (por ejemplo a las dos semanas). También se puede usar el descenso de 1 log desde su basal.

Grupo 3. Pacientes con resistencias pero con posibilidad de construir un régimen supresivo. Proporción de pacientes con ARN-VIH menor del límite de cuantificación a las 48 semanas usando un test aprobado por la FDA. Se puede usar la semana 24 para comparaciones de superioridad cuando el tratamiento nuevo aporta una ventaja respecto al tratamiento previo.

Ensayos clínicos de cambio. Proporción de pacientes con ARN-VIH mayor o igual del límite de cuantificación a las 48 semanas usando un test aprobado por la FDA. Se diferencia de los anteriores (grupo 1 y 3) en que el objetivo se centra en el fracaso virológico y no en el éxito. Esto se debe a que en los ensayos clínicos de cambio los pacientes inician con ARN-VIH por debajo del límite de cuantificación. El objetivo de interés, en otras palabras, es la proporción de individuos con carga viral suprimida basalmente que pierden el control virológico tras cambiar a un nuevo tratamiento.

Los objetivos secundarios son:

- Cambios medios en la carga viral respecto a su basal en pacientes pretratados.
- Cambios en el recuento de células CD4/ μ L, con respecto a la visita basal.

Efectividad clínica

Se define efectividad a la determinación de una carga viral indetectable (< 50 copias/ml) en la semana 48 de tratamiento.

El recuento de linfocitos CD4 y la carga viral plasmática (CVP) son los paráme-

tros habituales que se emplean para monitorizar la eficacia del TAR. El objetivo del TAR es suprimir la replicación viral. La CVP desciende rápidamente al inicio del TAR, y aunque los pacientes con CVP muy elevadas pueden tardar más de 24 semanas en conseguir valores indetectables, la mayoría de las combinaciones de TAR logran CVP indetectables (< 50 copias/ml) en la semana 24 o 48 de tratamiento.

Fórmula:

Número de pacientes que inician TAR y logran una CVP < 50 copias/ml en la semana 48 de tratamiento x 100 / número de pacientes que inician TAR.

Se incluyen todos los pacientes infectados por VIH que inician TAR en el período evaluado, permitiéndose cambios de pauta durante el primer año. Se excluyen a los pacientes fallecidos y los que cambian de lugar de residencia. No se excluyen las pérdidas de seguimiento por causas distintas (Von Wichmann *et al.*, 2010; Grupo de Estudio del SIDA-SEIMC (GeSIDA). (Actualización enero 2017)).

Fracaso virológico

Se define fracaso virológico (FV) a dos determinaciones consecutivas de CVP >50 copias/mL, transcurridas 24 semanas desde el inicio del TAR. El FV puede ocurrir con o sin selección de mutaciones de resistencia (MR). Si la CVP basal es muy elevada, pueden ser necesarias más de 24 semanas de TAR para alcanzar la CVP <50 copias/mL, particularmente en regímenes con IP/r.

Los factores que influyen en el fracaso terapéutico son a) dependientes del paciente: adherencia al tratamiento, dificultad de acceso a la atención sanitaria o de seguimiento de los controles médicos; b) dependientes del fármaco: errores de dosificación, potencia del TAR, barrera genética, interacciones farmacocinéticas y c) dependientes del virus: preexistencia de mutaciones de resistencia (MR) transmitidas o adquiridas.

El fracaso virológico a ≥ 2 pautas de TAR supone una menor posibilidad de éxito en los TAR sucesivos y una mayor incidencia de SIDA y muerte. No obstante, la disponibilidad de nuevos fármacos permite plantear en esta situación el objetivo de lograr la máxima supresión viral (< 50 copias/ml).

En el fracaso virológico avanzado la mayoría de los pacientes han fracasado a varias líneas de TAR y presentan MR a dos o más familias de FAR (Von Wichmann *et al.*, 2010; Grupo de Estudio del SIDA-SEIMC (GeSIDA). (Actualización enero 2017)).

Fórmula:

Número de pacientes con multfracaso y CVP < 50 copias/ml en la semana 48 x 100 / número de pacientes con multfracaso.

Seguridad, tolerabilidad y monitorización de efectos adversos

En las primeras etapas de la enfermedad, cuando la mortalidad era muy elevada, se priorizó la eficacia por encima de la seguridad y la tolerancia. Los fármacos controlaban la infección, pero con muchos efectos secundarios a corto plazo, y la causa fundamental del abandono del tratamiento era la intolerancia, sobre todo digestiva. Desde mediados de 1997 empezaron a aparecer efectos adversos a más largo plazo que producían problemas psicológicos, de autopercepción y un aumento de la estigmatización. En la actualidad, disponemos de fármacos eficaces y seguros, con posología adecuada, y el inicio del tratamiento se realiza de forma precoz.

Cambios en la función renal

Los pacientes con infección por el VIH pueden desarrollar diversas nefropatías glomerulares, vasculares, tubulointersticiales y obstructivas relacionadas en algunos casos con el propio virus, con los fármacos administrados o con las coinfecciones que presentan.

La incorporación de la evaluación periódica de la función renal a la atención programada de los pacientes infectados por el VIH facilita el diagnóstico precoz de la enfermedad renal crónica (ERC) oculta y permitir la aplicación de medidas orientadas a evitar su progresión.

El estudio de la función renal se realiza en el diagnóstico de la infección por el VIH, inmediatamente antes de iniciar el TAR y durante el seguimiento posterior. La periodicidad con la que debe repetirse durante el seguimiento puede depender de la existencia o no de factores de riesgo para el desarrollo de nefropatía. En general, se recomienda que se realice en todas las revisiones programadas de los pacientes que reciben TAR, y muy especialmente en los que toman Tenofovir-Disoproxil-Fumarato (TDF) (Górriz *et al.*, 2010).

Cambios en el perfil óseo

La osteoporosis es más frecuente en pacientes con infección por el VIH-1, especialmente si reciben TAR. El número de fracturas también es superior en pacientes con infección VIH-1 que en sujetos de similares características sin esta infección. La base patogénica que justifica esta reducción en la densidad mineral ósea (DMO) es un incremento del remodelamiento óseo, que condiciona una pérdida progresiva de la DMO. Los mecanismos que justifican esta reducción de la DMO son similares a los descritos en población general, entre los que se incluyen la edad, el tabaquismo, el bajo peso corporal, la insuficiencia renal, la diabetes, la hepatitis C, el uso de opiáceos, el consumo excesivo de alcohol o la menopausia.

Se recomienda realizar una DEXA en pacientes con infección por VIH:

- Presencia de factores mayores de riesgo de presentar fracturas (uso prolongado de glucocorticoides, historia de fracturas óseas por fragilidad y elevado riesgo de caídas).
- Mujeres posmenopáusicas o varones con hipogonadismo confirmado.
- Varones de edad ≥ 50 años.
- Si en la evaluación mediante el algoritmo FRAX presenta un riesgo a diez años de fractura de cadera $>3\%$ y/o $>10\%$ de fractura osteoporótica mayor.

Algunos aspectos propios de la infección por VIH-1 como valores bajos de linfocitos CD4+ o el efecto de algunos fármacos pueden influir en la mayor incidencia de fracturas. Los regímenes basados en TDF reducen la DMO en mayor medida a los que contienen ABC o que los regímenes sin ITIAN (Negredo *et al.*, 2018).

Cambios en el perfil metabólico

Los pacientes con infección por el VIH tienen mayor riesgo de enfermedad cardiovascular (ECV) que la población general. Aunque el mecanismo no se conoce con certeza, intervienen diversos factores, que incluyen factores genéticos, los factores de riesgo cardiovascular (FRCV) tradicionales, el TAR y parámetros relacionados con el propio VIH, por cambios inflamatorios e inmunológicos.

La dislipemia se caracteriza en un aumento de los niveles plasmáticos de colesterol total, colesterol-LDL y, sobre todo, triglicéridos, es la anomalía metabólica que con mayor frecuencia se asocia al TAR. Aunque tradicionalmente se ha relacionado esta dislipemia con los IP, no todos ellos tienen el mismo impacto sobre los lípidos.

Si queremos conocer la prevalencia de los distintos FRCV en los pacientes con VIH, podemos fijarnos en los datos de la cohorte D:A:D. El factor de riesgo más frecuente es el hábito tabáquico, seguido de la hipercolesterolemia y la hipertrigliceridemia.

En un paciente con infección por VIH se debe realizar inicialmente una estimación del riesgo cardiovascular utilizando la ecuación de Framingham, y según los resultados obtenidos intervenir, ya sea recomendando únicamente medidas higiénico-dietéticas, pautando además fármacos hipolipemiantes (estatinas y/o fibratos), ó modificando el tratamiento antirretroviral en función de la alteración metabólica observada (Grupo de Expertos del Grupo de Estudio sobre Alteraciones Metabólicas [GEAM] de la Secretaría del Plan Nacional sobre el Sida (SPNS) y del Grupo de Estudio de Sida [GeSIDA], actualización febrero 2017).

Alteraciones morfológicas

La lipodistrofia es la redistribución grasa puede consistir en una pérdida de grasa periférica (miembros superiores, inferiores, nalgas y/o de bola adiposa de Bichat);

acúmulo de grasa en una o varias localizaciones o una lipomatosis múltiple; y lo más habitual, una combinación de pérdida y acúmulo de grasa.

La importancia de la lipodistrofia (LD) radica sobre todo en los problemas a largo plazo que produce: estigmatización, cuadros depresivos, suspensión de tratamiento, o bien retraso en el inicio o en la continuidad de este por este motivo.

Calidad de vida relacionada con la salud (CVRS)

La valoración de la calidad de vida relacionada con la salud en el (CVRS) proporciona una información adicional a la información clínica tradicional sobre los resultados de las intervenciones sanitarias dando una visión más integral del estado de salud de un individuo. Con la finalidad de medir la CVRS en el paciente VIH se han desarrollado diversos cuestionarios, válidos y fiables, específicos para medir la CVRS en pacientes VIH

El “HIV Symptom Index” es un cuestionario con *ítems* que se usa para valorar la pérdida de la calidad de vida en pacientes VIH y detectar en qué ámbito de la salud habría que realizar una intervención.

En España los dos cuestionarios específicos para VIH que han sido traducidos y validados para medir la CVRS en VIH son el MOS-HIV+ y el MQOL-HIV (Llach y Suriñach, 2005).

- Cuestionario MOS-HIV. Contiene 35 preguntas que incluyen 11 dimensiones de salud: Percepción General de la Salud, Dolor, Función Física, Función de Rol, Función Social, Salud Mental, Energía/ Fatiga, Preocupación por la Propia salud, Función Cognitiva, Calidad de Vida y Transición de Salud. El MOS-HIV es el cuestionario específico más usado en la evaluación del impacto de las intervenciones y de los tratamientos en pacientes VIH.
- Cuestionario MQOL-HIV25. Consiste en 40 ítems que determinan 10 dimensiones de salud incluyendo Salud Mental, Salud Física, Función Física, Función Social, Soporte Social, Función cognitiva, Estado Financiero, Relación de Pareja, Función Sexual y Atención Médica. Todos los ítems son respondidos sobre una escala de siete u ocho opciones de respuesta y hacen referencia a las dos últimas semanas. No ha sido tan utilizado como el MOS-HIV.

Cumplimiento terapéutico y adherencia

El cumplimiento terapéutico es una de las variables que mejor predice la respuesta virológica. La falta de adherencia aumenta el riesgo de resistencias y de morbilidad asociada a la infección por el VIH. La adherencia terapéutica o cumplimiento ha sido definida por la OMS (Sabate, 2003) como “el grado en que la conducta de un paciente, en relación a la toma de medicación, el seguimiento de una dieta o la modificación en los hábitos de vida, se ajusta a las recomendaciones acordadas con el profesional sanitario”. Más concretamente, la International Society for Health Economics and Outcomes Research (ISPOR), Medication Compliance

and Persistence Workgroup (Cramer *et al.*, 2008), la define como “la medida en que el paciente toma la medicación tal y como le ha sido prescrita (en dosis e intervalo posológico)”.

La valoración de la adherencia supone estimar la capacidad del paciente de garantizar un cumplimiento riguroso. La falta de la adherencia en los tratamientos crónicos es un problema universal y real que compromete la efectividad de estos y se traduce en unos peores resultados en salud con un impacto negativo en el gasto sanitario.

Los métodos para la valoración de la adherencia se han clasificado tradicionalmente como directos e indirectos.

Valoración de la adherencia

Se puede llevar a cabo por cualquiera de estos métodos:

- Recuento de medicación o valoración directa y subjetiva por parte del profesional sanitario.
- Indirectos, como los sistemas de control electrónico o los cuestionarios estructurados validados –SMAQ, SERAD–.
- Registro farmacéutico, en el que consten: a) las no asistencias y b) el número de dosis devueltas.

Entre los métodos directos está la determinación plasmática de fármacos antirretrovirales. Aunque se considera el método más objetivo, presenta muchas e importantes limitaciones, como son la dificultad para establecer el rango óptimo, la escasez de estudios farmacocinéticos poblacionales, la variabilidad intra e interindividual y sobretodo, que no está disponible en la mayoría de los hospitales. Sin embargo, puede ser de gran utilidad en determinadas situaciones, como en el caso de toxicidad o interacciones. Otro método directo sería la evolución clínica y los datos analíticos, a través de la monitorización de la carga viral plasmática (CVP) y los CD4. No debería considerarse un método de estimación de la adherencia, sino más bien, se debe analizar sistemáticamente junto a los métodos de medida de la adherencia de forma prospectiva.

Entre los métodos indirectos está la valoración del profesional sanitario, que tiende a sobreestimar la adherencia y es la utilizada cuando no se dispone de otro método. Lo hacemos todos intuitivamente, pero resulta una valoración muy subjetiva y no se debería utilizar.

Los métodos que se utilizan habitualmente para la medida de la adherencia son: los dispositivos electrónicos, los registros de dispensación de Farmacia (RDF), el recuento de medicación sobrante, los cuestionarios y diarios de los pacientes (Knobel *et al.*, 2002).

Dispositivos electrónicos

Los sistemas de control electrónico de apertura de los envases (tipo MEMS o EDEM) son dispositivos que llevan una tapa que contienen un microprocesador que registra la hora y el día en que se ha abierto el envase. La apertura del envase se relaciona con la toma de medicación y esos datos son procesados posteriormente en un programa informático.

Presenta algunas limitaciones, como son, el elevado coste económico y que sólo son aplicables a determinados envases y muchas veces obliga a cambiar el envase original por otro que se adapte al dispositivo, además habitualmente solo se mide la adherencia de un fármaco antirretroviral y no de todo el tratamiento. Otras características son que exige la colaboración del paciente; pueden ser manipulados y la apertura del envase no implica necesariamente la toma de la medicación, ni puede excluirse la toma a pesar de no quedar registrada en el dispositivo.

Recuento de medicación sobrante

Supone contar la medicación sobrante que trae el paciente y por lo tanto depende de su colaboración. Es un método barato y fácil, permite una medida cuantitativa y es objetivo, pero requiere tiempo y personal, y puede ser manipulable por el paciente, por lo que aun así sobreestima la adherencia muchas veces se ha utilizado como patrón.

Registros de dispensación de Farmacia (RDF)

Es un método indirecto que supone que un paciente no toma la medicación que no le es dispensada y que toma de forma adecuada aquella que se le dispensa. Se ha visto una buena correlación con los resultados virológicos y aceptable especificidad y sensibilidad. Exige que la dispensación se realice de forma centralizada, por un único dispensador. Sus principales limitaciones son que la dispensación de la medicación no es sinónimo de cumplimiento correcto.

El RDF es un método factible en países donde los fármacos antirretrovirales son dispensados en las farmacias de los hospitales. Es relativamente económico y permite establecer registros de forma rutinaria e informatizada con un seguimiento longitudinal de los pacientes.

Cuestionarios

La adherencia se valora según la respuesta del paciente a unas preguntas previamente definidas, que varían desde una única pregunta abierta, hasta un conjunto de preguntas estandarizadas abiertas y cerradas, que son más complejas de aplicar. Se pueden realizar vía telefónica, en entrevistas personales o incluso a través de cuestionarios escritos autoadministrados. Son métodos sencillos, que requieren pocos recursos, asequible y adaptable a las características de cada centro. Además de preguntas relacionadas con el número de tomas olvidadas, que intentan cuantifi-

car el grado de adherencia, existen otras que intentan valorar la calidad de la toma (horarios, relación con comidas, motivos de no adherencia, etc.). Las principales limitaciones son la subjetividad, su baja sensibilidad, y que solo mide la adherencia en un período de tiempo corto (4 últimos días, 7 últimos días, 30 días), diferente en función del cuestionario. Además, depende de la memoria humana y la capacidad para recordar disminuye con el tiempo.

Entre los cuestionarios validados en población española, cabe destacar el cuestionario SMAQ (*Simplified medication adherence questionnaire*), validado en pacientes con IP no potenciado y el SERAD, que permite una valoración cuantitativa y cualitativa de la adherencia (Muñoz-Moreno *et al.*, 2007). Otro cuestionario validado en pacientes VIH es el ACTG (*AIDS Clinical Trials Group*), ampliamente utilizado en EE. UU.; incluso la escala analógica visual se ha utilizado con este fin.

Selección del método y combinaciones de métodos

En líneas generales, los cuestionarios, los recuentos de medicación y los registros de dispensación de Farmacia, proporcionan valores sobrestimados de adherencia; mientras que los sistemas tipo MEMS proporcionan valores infraestimados. El problema derivado de la sobreestimación es que no se identifica a muchos pacientes no adherentes, perdiéndose así la oportunidad de intervenir sobre su conducta.

Debido a las limitaciones que presentan estos métodos se recomienda combinar varios de ellos para obtener información con la mayor exactitud posible.

Grado de persistencia

La persistencia se define como el número de días de utilización continua de la medicación durante un periodo específico. La tasa de persistencia se calcula dividiendo el número de días que el paciente ha sido persistente entre el número de días del periodo de seguimiento.

Complejidad del tratamiento y resultados en salud

La complejidad del tratamiento conlleva a una disminución de la adherencia y como consecuencia a una disminución de la eficacia. El *Medication Régimen Complexity Index* (MRCI) es un instrumento de 65 ítems para valorar la complejidad de los tratamientos e incluye una evaluación de cada medicación. Esta herramienta permite detectar a los pacientes en riesgo de disminuir su adherencia debido a la dificultad del tratamiento (Hirsch *et al.*, 2014).

Simplificación del tratamiento

Entendemos por simplificación del TAR el cambio de un régimen con el que se ha conseguido una adecuada respuesta virológica e inmunológica por otro que mantenga esta eficacia y permita reducir la complejidad del tratamiento, proporcionando así un beneficio en la calidad de vida del paciente que puede traducirse en una mayor adherencia terapéutica. Al plantearse la posibilidad de una simplifi-

Figura 1.

THE ANTIRETROVIRAL REGIMEN COMPLEXITY INDEX

	A	B	C	D	E	F	G	H	I	J	K	
SCHEDULE												
1 ARV medications	Enter total # of meds in column 1 →										X 0.5	
2 Doses per day	Enter the highest number of doses per day in column K →										→	
3 Dosing consistency	If \geq medications are given 2X/day and ≤ 1 are given ≥ 3 X/day, enter "2" in column K →										→	
4 Pills per DOSE												
5 Pills per DAY (row 2 x row 4)	+	+	+	+	+	+	+	Add columns B-G and enter the total number of pills per day in column I →			X 0.25	
6 Liquid	If any boxes are checked in this row, enter "1" in column K →										→	
7 Injection	If any boxes are checked in this row, enter "2" in column K →										→	
8 G-tube	If all medications (except injections) are given via g-tube, enter "1" in column K →										→	
ADMINISTRATION												
9 With food	If any boxes are checked in this row, enter "1" in column K →										→	
10 Empty stomach	If any boxes are checked in this row, enter "1,5" in column K →										→	
11 Dietary content (specify guideline)	Enter the total number of different diet content rules in column I →										X 2	
12 If 2 <i>different</i> instructions are checked in rows 9-11, columns B-G, enter "2" in column K →											→	
12 If 3 <i>different</i> instructions are checked in rows 9-11, columns B-G, enter "3" in column K →											→	
13 Refrigeration	If any boxes are checked in this row, enter "1" in column K →										→	
14 Reconstitution (daily)	Enter the total number of boxes checked in this row in column I →										X 2	
15 Reconstitution (daily)	Enter the total number of boxes checked in this row in column I →										X 0.5	
PREPARATIONS												
ARC Index								ARC INDEX =		(add rows 1-15)		

cación del TAR es indispensable no renunciar a la eficacia del tratamiento previo. Debe haberse demostrado que la nueva pauta terapéutica conseguirá mantener indetectable el ARN viral en plasma y la cifra de linfocitos CD4 como mínimo en la misma proporción de pacientes y como mínimo con la misma duración que el tratamiento que se desea sustituir. Los objetivos de la simplificación del TAR son mejorar la calidad de vida del paciente, la adherencia al tratamiento antirretroviral y reducir la complejidad (Ribera *et al.*, 2002). Podremos simplificar el TAR incidiendo en cinco aspectos:

- a) Número de fármacos.
- b) Número de formas farmacéuticas.
- c) Número de tomas al día.
- d) Restricciones alimentarias.
- e) Características organolépticas.

En varios estudios se ha demostrado que la adherencia al TAR se relaciona claramente con el número de comprimidos y con el número de tomas al día. En algunos pacientes, las restricciones alimentarias y las características organolépticas de los comprimidos también pueden dificultar la adherencia al TAR. La escala *Antiretroviral Régimen Complexity Index* (Figura 1) se usa para estimar la complejidad del TAR y valorar la posibilidad de simplificación (Martin *et al.*, 2017).

■ 9.3. RESULTADOS PERCIBIDOS POR LOS PACIENTES (PATIENT-REPORTED OUTCOMES)

El uso de resultados percibidos por los pacientes (PRO) aporta una información adicional a los proveedores de salud, destinados a mejorar los resultados de la atención prestada y de los propios resultados en salud de los pacientes. Se pueden definir como medidas basadas directamente en información generada por el paciente (Engler, Lessard y Lebouche, 2017). En la infección por VIH se usan habitualmente incluidos como parte de la propia investigación en los ensayos clínicos (Gathe *et al.*, 2015). También serían resultados candidatos a tener en cuenta en la práctica clínica diaria (Høgh Kølback Kjæra *et al.*, 2018) para guiar las decisiones en la atención de los pacientes. Son muchas las dimensiones que se han estudiado, ya que cualquier dato o enfoque que venga del paciente se le puede llamar PRO. Resumimos aquí las más relevantes.

Calidad de vida relacionada con la salud

El término calidad de vida relacionada con la salud (CVRS) surge como un concepto que hace referencia a valoraciones de la percepción de la salud por parte del individuo, recogiendo tanto aspectos objetivos como subjetivos. Se han desarrollado diversos instrumentos de medición de CVRS en diferentes patologías, entre

ellas la infección por VIH. Los diversos instrumentos tienen en cuenta distintas dimensiones de la CVRS y suelen clasificarse en cuestionarios genéricos y específicos. Estos instrumentos deben cumplir una serie de características para ser considerados útiles, permitiendo la posibilidad de registrar las percepciones del paciente de una forma cuantitativa o semicuantitativa, que pueda comunicarse y utilizarse para describir, evaluar o comparar.

Las dimensiones que son importantes para la medición de la CVRS son: el funcionamiento social, físico, y cognitivo; la movilidad y el cuidado personal; y el bienestar emocional (Herdman, 2000). Por otro lado, la CVRS constituye una importante variable de medida subjetiva del impacto que la enfermedad y su tratamiento producen en la vida del sujeto. Su valoración nos permite detectar alteraciones e intervenir precozmente, así como establecer comparaciones entre las distintas opciones terapéuticas. La evaluación de la CVRS en las ciencias biomédicas nos aporta un resultado final de salud que se centra en la persona, no en la enfermedad; en cómo se siente el paciente, independientemente de los datos clínicos.

Teniendo en consideración que la CVRS es un concepto multidimensional, desde la percepción del paciente y que cada dimensión de la CVRS cambia con el tiempo, se han diseñado diversos instrumentos para establecer una aproximación a la medición de la CVRS. Estos instrumentos se pueden dividir en genéricos y específicos. Los instrumentos genéricos son independientes del diagnóstico, por lo que ofrecen la oportunidad de ser aplicables a cualquier tipo de población o afección (Soto y Failde, 2004).

De entre los disponibles se han utilizado para la infección por VIH o SIDA: *Medical Outcomes Study Short Form-36* y sus variantes, *Multidimensional Quality of Life Questionnaire*, *EuroQol*, *Fanning Quality of Life scale*, *Index of multiple deprivation (IMD)* (Degroote, Vogelaers y Vandijck, 2014; Simpson *et al.*, 2013).

Los específicos se centran en aspectos de la calidad de vida propios de una enfermedad o síndrome concreto. No tienen, por tanto, la amplitud de los instrumentos genéricos, pero sí pueden ser más sensibles a aspectos de la calidad de vida determinados por efectos de una enfermedad concreta (Høgh Kølbæk Kjæra *et al.*, 2018). En el caso de la infección por VIH se han relacionado con la CVRS parámetros como: TAR, adherencia al TAR, estigma, tiempo desde el diagnóstico, y hábitos más frecuentes que en la población general como uso de drogas recreativas o consumo de alcohol.

Sobre estos cuestionarios también existe una amplia variedad de posibilidades como son: *Medical Outcomes Study-HIV*, *HIV targeted quality of life instrument*, *Functional Assessment of Human Immunodeficiency Virus Infection*, *HIV Overview of Problems Evaluation System*, *Health Index and PROQOL-HIV* y *World Health Organization Quality of Life HIV BREF* (WHOQOL-HIV BREF) (Herdman, 2000).

Satisfacción con el tratamiento

Existen muchas medidas relacionadas con el TAR. Su importancia radica en la relación que tiene con la adherencia al tratamiento, y por lo tanto con variables de salud relevantes en esta población.

Unas aluden más a la carga del tratamiento en cuanto a número de formas farmacéuticas o la complejidad del régimen (Grupo de Estudio del SIDA-SEIMC (GeSIDA). (Actualización enero 2017)). Algunas herramientas se han centrado en las creencias acerca de los pacientes respecto a su tratamiento (Górriz *et al.*, 2010), y otras se objetivan más hacia una satisfacción global con el tratamiento, como el *HIV Treatment Satisfaction Questionnaire* (HIVTSQ) (Negredo *et al.*, 2018).

Existen muchas medidas relacionadas con el TAR. Su importancia radica en la relación que tiene con la adherencia al tratamiento, y por lo tanto con variables de salud relevantes en esta población.

Unas aluden más a la carga del tratamiento en cuanto a número de formas farmacéuticas o la complejidad del régimen (Hirsch *et al.*, 2014). Algunas herramientas se han centrado en las creencias acerca de los pacientes respecto a su tratamiento (Smith *et al.*, 1997), y otras se objetivan más hacia una satisfacción global con el tratamiento, como el *HIV Treatment Satisfaction Questionnaire* (HIVTSQ) ([Woodcock y Bradley, 2001).

Preferencias de los pacientes

De nuevo nos encontramos con que gran parte de la literatura entorno a las preferencias de los pacientes se muestran en relación con el TAR, como es en el caso del uso de los *Beliefs about Medicines Questionnaire-ART version* (BMQ-ART) o la *HAART Intrusiveness Scale*. Otras, en cambio se han centrado en la atención prestada por los servicios sanitarios (Engler, Lessard y Lebouche, 2017; Marrone *et al.*, 2016). Existen incluso estudios que han tomado en consideración las preferencias globales de colectivos concretos de población, como en el caso de las mujeres durante el embarazo (Lytvyn *et al.*, 2017).

Discapacidad funcional/Dependencia

Algunas de las herramientas para conocer y medir los PROs en este sentido se han desarrollado como escalas para conocer la CVRS, como es el caso del *Functional Assessment of Human Immunodeficiency Virus Infection* (FAHI) o la *HIV Overview of Problems Evaluation System* (HOPES) (Engler, Lessard y Lebouche, 2017).

Existe una escala específica, la *HIV Disability Questionnaire Study* (HDQ), validada por O'Brien y colaboradores en 2013 (Whalen *et al.*, 1994).

Medición de la carga de la enfermedad

Existen diversas herramientas para la valoración de síntomas que se han empleado en pacientes con infección por HIV, siendo la de Whalen (Marrone *et al.*,

2016) y la *HIV Costs and Services Utilization Study* (HCSUS) (Mathews *et al.*, 2000) las más empleadas. La de Whalen y colaboradores consta de 12 *ítems* que valoran características psicométricas. La HCSUS se basa en la molestia que causan 14 síntomas habituales en portadores de la enfermedad. Ambas se encuentran validadas, pero son de poca aplicabilidad hoy en día, dado que son anteriores al tratamiento de alta eficacia, y se encuentran en desuso.

Grado de comorbilidad/fragilidad de los pacientes

No se puede plantear la edad cronológica como un marcador absoluto de envejecimiento ya que el envejecimiento no es un hecho homogéneo. Algunos individuos llegan a edades avanzadas de la vida en plenas facultades. Otros en cambio, padecen enfermedades y déficits que erosionan gravemente su funcionalidad y su calidad de vida o lo que es lo mismo, sujetos de la misma edad cronológica pueden diferir en el grado de preservación de sus funciones biológicas.

En el caso de los pacientes con infección por VIH supone un reto, dado que la evidencia del envejecimiento precoz del sistema inmune asociado a la infección por el VIH avala el aceptar como edad de corte, para definir edad avanzada en este grupo poblacional, la edad de 50 años. Esta marca el comienzo del declinar en la respuesta inmunológica al tratamiento. La población con infección por el VIH con edad mayor de 65 años es población muy envejecida (Polo, Blanco y Verdejo, 2015).

Entre los más conocidos está el Índice VACS, el *Electronic Frailty Index* (eFI), el *Frailty Phenotype* y la *Edmonton Frail Scale* (Clegg *et al.*, 2016)

■ 9.4. RESULTADOS COMUNICADOS POR OBSERVADORES (OBSERVER-REPORTED OUTCOMES)

Los resultados reportados por observadores (ObsRO) van adquiriendo mayor peso a la hora de poder contextualizar los resultados en salud en el marco de las enfermedades infecciosas (Powers *et al.*, 2016). No necesitan personal especializado ni entrenado para ello, y pueden salvar las situaciones en las que el paciente no puede expresarse para ello. Así como en otras patologías, los ObsRO se han empleado en estudios sobre todo en población pediátrica. En determinados estudios no se han dispuesto de herramientas específicas (Cruz *et al.*, 2014), sino que se emplean cuestionarios de adherencia, de calidad de vida, o salud mental idénticos a los empleados por los pacientes en los PROs.

■ 9.5. RESULTADOS EN SALUD ECONÓMICOS (ECONOMIC OUTCOMES ASSESSMENT)

La evaluación económica, comúnmente denominada análisis coste-efectividad (ACE), compara en términos de costes y resultados en salud tecnologías sanitarias

alternativas. En España, el Real Decreto Ley 16/2012 estipula que las decisiones de financiación sobre tecnología sanitaria deben estar presididas por los criterios de evidencia científica de coste-efectividad y por la evaluación económica.

En la situación actual, de fuertes restricciones presupuestarias y de mayor cultura económica, los análisis han de ir más allá de la simple comparación aislada de los costes y deben considerar estos últimos en función de los resultados conseguidos por las distintas opciones existentes (eficiencia). Una intervención sanitaria que no sea eficaz, que se utilice incorrectamente, o que pueda ser sustituida por otra con iguales resultados y menor coste tiene unas consecuencias económicas que se valoran en términos de renuncia, definiéndose coste de oportunidad como el valor de la mejor alternativa a la que se renuncia al tomar una decisión. En un contexto de recursos limitados, como es el sanitario, el objetivo de eficiencia supone minimizar el coste de oportunidad.

Introducir el criterio de eficiencia en relación con la terapia antirretroviral es factible como ayuda a la toma de decisiones puesto que disponemos de un escenario formado por múltiples alternativas de tratamiento y consumo de recursos posiblemente diferentes. Se dispone además de información epidemiológica de calidad que permite converger desde variables intermedias (viroológicas e inmunológicas) a variables finales de evolución a SIDA y mortalidad.

Cambio en el coste de la enfermedad

El tratamiento antirretroviral de alta efectividad cambia el curso clínico de la infección del virus de la inmunodeficiencia humana (VIH). La utilización de los recursos sanitarios ha cambiado drásticamente desde la introducción de los inhibidores de la proteasa (IP) en la terapéutica. Como consecuencia de la mejora en los parámetros clínicos: descenso en la Carga Viral (CV) y aumento de los recuentos de linfocitos T CD4+ (CD4), existe de forma paralela un descenso en las condiciones, tanto VIH-específicas, como las VIH-no específicas: Según un estudio en nuestro medio, las específicas de un 32,9 hasta un 9,7 % y de las no específicas de 67,7 a 25,2 % ($p < 0,001$ en ambas) (Velasco *et al.*, 2000).

En general, los ingresos hospitalarios por paciente y año se reducen de 0,64 a 0,27 ($p = 0,007$) por paciente y año, y las estancias hospitalarias de 18,3 a 9,3 días ($p < 0,001$). Las visitas al servicio de Urgencias disminuyen de 1,28 a 0,67 por paciente y año ($p < 0,001$), y las visitas a la consulta específica de HIV aumentan de 4,65 a 6,70 ($p < 0,001$) según un estudio de nuestro entorno (Velasco *et al.*, 2000).

De forma paralela los costes de la terapia antirretroviral aumentan debido al uso de IP. El resultado global del coste sanitario aumenta en torno a un 20 % para los pacientes con VIH. No obstante, en el mismo periodo de tiempo, el aumento en los pacientes sin VIH fue del 40 %. Estudios basados en la población de la Cohorte Suiza (Sendi *et al.*, 1999) indicarían que si bien, el aumento en la ratio coste-efectividad es elevado: 22.110\$, el aumento en la productividad de los pacientes compensaría este aumento en el gasto sanitario.

En una línea similar, otros estudios de la misma época revelan un similar patrón en el cambio de la asistencia sanitaria a los pacientes con infección por VIH: Descenso en el número de hospitalizaciones, descenso en la estancia hospitalaria, aumento del seguimiento ambulatorio de los pacientes y, sobre todo, aumento en el coste del tratamiento ambulatorio (Fleishman *et al.*, 2005). El resultado de este balance no siempre se ha asociado con un aumento en el global del gasto sanitario, ya que el descenso del coste de las infecciones oportunistas por la mejora del estado inmunológico de los pacientes en tratamiento con IP lleva también una reducción del gasto sanitario que compensa el mayor gasto en coste del tratamiento antirretroviral (Gebo *et al.*, 1999).

Años de Vida Ajustados por Calidad

Hemos hablado ya sobre la posibilidad y las herramientas de medida de la CVRS en los pacientes con infección por VIH. Junto con el incremento de años de vida debidos al tratamiento, tenemos los años de vida ajustados por calidad de vida. Se trata de una manera alternativa de medir el beneficio de un programa sanitario que evita algunos problemas de las unidades naturales, sin recurrir a las unidades monetarias. Estas escalas suelen estar basadas en la psicometría. Se trata de utilizar conceptos teóricos procedentes de la psicología para medir una variable subjetiva tal como la CVRS. La calidad de vida es un concepto muy amplio que incluye multitud de aspectos decisivos para el bienestar de los seres humanos como la libertad, la renta y la riqueza, la calidad ambiental, etcétera. Una de las dimensiones más importantes de la calidad de vida es la salud, y la CVRS hace referencia justamente a aquellos aspectos de la calidad de vida que están relacionados directamente con la salud de los individuos.

El TAR ha demostrado ser efectivo para disminuir la CV, aumentar o mantener el recuento de linfocitos CD4 y prevenir los eventos clínicos como las infecciones oportunistas. El uso de estas potentes terapias antirretrovirales ha producido una marcada disminución en las tasas de hospitalización, morbilidad y mortalidad, convirtiendo a la infección por VIH de una enfermedad aguda y mortal a corto plazo en una enfermedad crónica y mortal a medio-largo plazo. En esta nueva situación es necesario considerar el balance entre los beneficios del enlentecimiento de la progresión de la enfermedad y los posibles efectos secundarios de los antirretrovirales a medio y largo plazo. La medida de la CVRS del paciente puede ser un elemento clave.

Además, existe la necesidad de comparar pautas de tratamientos antirretrovirales de parecida eficacia, pero que pueden diferir en frecuencia o intensidad de efectos adversos (con el consecuente impacto sobre la adherencia) y el estudio de la efectividad de nuevos tratamientos antirretrovirales, son factores que han contribuido decisivamente a que aumente el número de instrumentos dedicados a medir la CVRS. Por esta razón la valoración de la CVRS ha sido incorporada de un modo creciente como medida del resultado en los ensayos clínicos de tratamientos para pacientes VIH.

Cociente coste-efectividad incremental

Teniendo los costes y los años de vida ganados con el TAR ajustados a la CVRS (QALYs), podemos calcular las diferencias en el cociente coste-efectividad (ICER) entre los distintos tratamientos disponibles.

En la actualidad existen varios fármacos de diferentes familias autorizados para el tratamiento del VIH. El Tratamiento Antirretroviral (TAR) se compone habitualmente de dos medicamentos Inhibidores de la Transcriptasa Inversa Análogos de Nucleósidos (ITIAN), y un tercer agente, tanto Inhibidores de la Transcriptasa Inversa No Análogos de Nucleósido (ITINAN), IP, Antagonistas del receptor CCR5 o un Inhibidor de la Integrasa (IINT) (Restelli *et al.*, 2017). Dentro de estas clases de fármacos usados como tercer agente dolutegravir (DTG) y raltegravir (RAL) son IINT de referencia, así como darunavir potenciado con ritonavir (DRV/r) dentro de los IP, y dentro de los ITINAN el efavirenz (EFV) ha sido un fármaco de inicio referente durante muchos años.

Se han propuesto modelos (ECDC, 2017) que indican respecto a DRV/r, DTG incrementa costes y QALYs con un ICER de 38.586 euros/QALY en pacientes naïve, y de 6.170 euros/QALY en paciente pretratado con alta CV (mayor de 100.000 cop/mL). El escenario con DTG respecto a EFV tendría un ICER de 33.664 euros/QALY en pretratados. El uso de DTG frente a RAL como IINT tendría un ICER de 12.074 euros/QALY.

Como se ha mencionado anteriormente muchos estudios en el plano de la farmacoeconomía no han llegado a valorar costes indirectos, que invertirían la balanza de gastos hacia el beneficio económico, gracias a la valoración de los ingresos derivados del aumento de la productividad de los pacientes, más allá del puro beneficio social sobre el aumento de la salud (Fleishman *et al.*, 2005).

Para ello, algunos autores han realizado cuestionarios a los pacientes ambulatorios con VIH en tratamiento antirretroviral (Gebo *et al.*, 1999). El promedio de ingresos por paciente se sitúa en 1.225 € (datos de nuestro entorno) en el momento del diagnóstico, y pasa a ser de 786€. A pesar de que existe un descenso en los ingresos, los pacientes se siguen manteniendo como productivos. A esto se suma una mejora en la CVRS con el tratamiento combinado con IP respecto a la anterior situación, medida en este caso con la herramienta EuroQOL-5D.

Impacto presupuestario para el Sistema Nacional de Salud (SNS)

La cronificación de los pacientes con infección por VIH hace que el número de pacientes se incremente de manera progresiva (Oliva, Roa y del Llano, 2003) y, por tanto, el coste global de su tratamiento es asumido en nuestro país por el SNS. El coste de los tratamientos, aunque de forma individual se abarata con el tiempo, supone una importante partida dentro del gasto de farmacia en hospital (Rank *et al.*, 1998).

El impacto presupuestario dependerá de la proporción de utilización de los distintos regímenes, el coste de estos, el número de los pacientes tratados, su grado de adherencia y de la persistencia al tratamiento. De hecho, existen trabajos que han intentado aproximarse a este cálculo tan complejo (Ventura-Cerda *et al.*, 2010).

■ 9.6. RESULTADOS EN SALUD EN EL MUNDO REAL (RWD-RWE-DVR)

Los datos de vida real (DVR) son aquellos datos utilizados para la toma de decisiones no obtenidos partir de ensayos clínicos aleatorizados, sino de la práctica clínica. Los DVR pueden generar conocimiento complementario al que proporcionan los ensayos clínicos aleatorizados, permitiendo una más fácil generalización de la información, aunque con riesgos de sesgo que hay que conocer y valorar.

Registro de pacientes (Cohortes)

Los estudios de cohortes han sido fundamentales para describir la historia natural de la infección por VIH y evaluar la efectividad de los tratamientos (Jarrin, Bolumar y del Amo, 2010). El diseño más adecuado para la investigación de efectos causales en biomedicina es el ensayo clínico, debido a que hace que la asignación aleatoria a recibir la exposición de interés valide que la diferencia entre los sujetos es sólo debida a esa exposición. No obstante, suelen tener deficiencias en su validez externa. En el mundo real, los pacientes no están seleccionados, y pese a sus propios sesgos, los estudios de cohortes son los de diseño observacional más adecuados para establecer la inferencia causal.

Existen diversas cohortes y se pueden clasificar según: tipos de exposiciones, tipos de episodios, como abiertas o cerradas, unicéntricas o multicéntricas, de seroconvertidores o seroprevalentes, etc. En nuestro medio, la cohorte VACH aúna a 23 centros hospitalarios que abarcan al 15 % de la población atendida y fue pionera en describir la epidemiología de la infección por VIH (Suárez-Lozano *et al.*, 2002).

Otra de las cohortes relevantes de nuestro medio es la cohorte de la Red Española de Investigación en sida (CoRIS). En el marco de Red Española de Investigación en sida (RIS), se creó una cohorte prospectiva multicéntrica de pacientes VIH-positivos sin tratamiento antirretroviral previo, asociada a un repositorio de muestras biológicas (Biobanco). CoRIS permitirá describir el panorama de la infección por el VIH en los próximos años en España y servir de plataforma para contestar nuevas preguntas de investigación de grupos de investigación epidemiológica, clínica y básica (Caro-Murillo *et al.*, 2007).

Resumimos a continuación las bases de datos y registros en idioma castellano más relevantes con recursos para profesionales y pacientes:

- Organización Mundial de la Salud. Dispone de un apartado sobre VIH y SIDA con información general y técnica. Dispone de acceso a sus campañas y publicaciones relacionadas¹.
- ONUSIDA. Mucha información sobre campañas propias, además de recursos como publicaciones propias y estadísticas sobre el estado de la epidemia².
- Ministerio de Sanidad Servicios Sociales e Igualdad. En su web³ podemos encontrar la información sobre el Plan Nacional sobre el SIDA en nuestro país. Diferencia los recursos para ciudadanos y profesionales y dispone de una biblioteca con publicaciones propias, además de un portal estadístico para conocer datos de la infección en nuestro país.
- *Centre for Disease Control* en español. En su web⁴ podremos encontrar, sobretodo, recursos para pacientes. Dispone de hojas informativas sobre temas básicos de VIH y SIDA.
- InfoSIDA. Promovida por el Departamento de Salud y Servicios Humanos de EE. UU. Dispone de recursos como aplicaciones móviles gratuitas, información sobre guías clínicas, ensayos clínicos, medicamentos y la propia enfermedad. Tiene hojas informativas con infografías para pacientes y familiares o amigos⁵.

■ 9.7. CONCLUSIONES

- En la infección por VIH, los resultados de eficacia y efectividad comunicados se basan en la determinación de la carga viral indetectable (<50 copias/ml) en la semana 48 del tratamiento. La adherencia predice la respuesta al mismo, y su ausencia puede llevar al fracaso terapéutico y a la aparición de mutaciones de resistencia. Existen múltiples métodos para valorar el cumplimiento del tratamiento. La seguridad, la tolerancia y los efectos adversos del tratamiento son variables que pueden predecir el control de la infección ya que pueden comprometer el cumplimiento de este.
- Los pacientes valoran su salud en términos de calidad de vida. Las diferentes herramientas usadas tienen en cuenta las preferencias del paciente, la satisfacción del tratamiento, las características cognitivas, físicas, sociales, el grado de comorbilidad y la dependencia, entre otros.
- En situaciones en las que el paciente no puede expresarse (por ejemplo, en población pediátrica), los resultados de salud pueden ser comunicados

¹ Disponible en <http://www.who.int/hiv/es/>

² Disponible en <http://www.unaids.org/es>

³ <http://www.msssi.gob.es/ciudadanos/enfLesiones/enfTransmisibles/sida/home.htm>

⁴ <https://www.cdc.gov/hiv/spanish/index.html>

⁵ Disponible en infosida.nih.gov

por otros observadores. Habitualmente se usan cuestionarios similares a los empleados para los pacientes.

- La infección por VIH actualmente se considera una condición crónica con un elevado impacto económico. Los años de vida se deben ajustar por calidad de vida y por la existencia o no de efectos adversos. Para comparar los diferentes regímenes se pueden usar cálculos de coste-efectividad. Otros costes que se deben tener en cuenta tienen que ver con la mejora de la productividad al evitar ingresos hospitalarios del paciente.
- Los registros y las cohortes de vida real de individuos con infección por VIH recogen los datos de la práctica clínica habitual y son muy útiles. “CoRIS” es la cohorte prospectiva multicéntrica que recoge los datos y muestras biológicas de los individuos VIH positivos en España antes de iniciar tratamiento y de su evolución posterior.

■ REFERENCIAS

- ÁREA DE VIGILANCIA DE VIH Y COMPORTAMIENTOS DE RIESGO. (2016). *Vigilancia Epidemiológica del VIH y Sida en España: Sistema de Información sobre Nuevos Diagnósticos de VIH y Registro Nacional de Casos de Sida*. Madrid: Plan Nacional sobre el Sida – S.G. de Promoción de la Salud y Epidemiología/Centro Nacional de Epidemiología, ISCIII. Disponible en: https://www.msssi.gob.es/ciudadanos/enfLesiones/enfTransmisibles/sida/vigilancia/InformeVIH_SIDA_2016.pdf (acceso mayo 2018).
- ÁREA DE VIGILANCIA DE VIH Y CONDUCTAS DE RIESGO. (2016). *Mortalidad por VIH y Sida en España, año 2014. Evolución 1981-2014*. Madrid: Centro Nacional de Epidemiología/Subdirección General de Promoción de la salud y Epidemiología – Plan Nacional sobre el Sida. Disponible en: <https://www.msssi.gob.es/ciudadanos/enfLesiones/enfTransmisibles/sida/vigilancia/MortalidadXVIH2014.pdf> (acceso mayo 2018).
- CARO-MURILLO, A. M., CASTILLA, J., PÉREZ-HOYOS, S. *et al.* Spanish cohort of naïve HIV-infected patients (CoRIS): rationale, organization and initial results. *Enferm Infecc Microbiol Clin*, 25(1), pp. 23-31.
- CLEGG, A., BATES, C., YOUNG, J. *et al.* (2016). Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age Ageing*, 45(3), pp. 353–360.
- CRAMER, J. A., ROY, A., BURRELL, A. *et al.* (2008). Medication Compliance and Persistence: Terminology and Definitions. *Value Health*, 11(1), pp. 44-47.
- CRUZ, M. L. S., CARDOSO, C. A. A., DARMONT, M. Q. *et al.* (2014). Viral suppression and adherence among HIV-infected children and adolescents on antiretroviral therapy: results of a multicenter study. *J Pediatr (Rio J)*, 90(6), pp. 563-571.
- DEGROOTE, S., VOGELAERS, D. y VANDIJCK, D. M. (2014). What determines health-related quality of life among people living with HIV: an updated review of the literature. *Arch Public Health*, 72(1), p. 40.
- ENGLER, K., LESSARD, D., LÉBOUCHE, B. A. (2017). Review of HIV-Specific Patient-Reported Outcome Measures. *Patient*, 10(2), pp. 187–202.
- EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL (ECDC). (2016). *HIV/AIDS surveillance in Europe 2015*. Disponible en: <https://ecdc.europa.eu/en/publications-data/hiv/aids-surveillance-europe-2015> (acceso enero 2018).

- . (2017). *Continuum of HIV care Monitoring implementation of the Dublin Declaration on Partnership to Fight HIV/AIDS in Europe and Central Asia: 2017 progress report*. Disponible en: <https://eccd.europa.eu/sites/portal/files/documents/Continuum-of-HIV-care-2017.pdf> (acceso enero 2018).
- FLEISHMAN, J. A., GEBU, K. A., REILLY, E. D. *et al.* (2005). Hospital and Outpatient Health Services Utilization Among HIV-Infected Adults in Care 2000–2002. *Med Care*, 43(9 Suppl), pp. III-40 –III-52.
- GATHE, J., ARRIBAS, J. R., VAN LUNZEN, J. *et al.* (2015). Patient-Reported Symptoms over 48 Weeks in a Randomized, Open-Label, Phase 3b Non-inferiority Trial of Adults with HIV Switching to Coformulated Elvitegravir, Cobicistat, Emtricitabine, and Tenofovir DF Versus Continuation of Ritonavir-Boosted Protease Inhibitor with Emtricitabine and Tenofovir DF. *Patient*, 8(5), pp. 445–454.
- GEBU, K. A., CHAISSON, R. E., FOLKEMERA, J. G. *et al.* (1999). Costs of HIV medical care in the era of highly active antiretroviral therapy. *AIDS*, 13(8), pp. 963–969.
- GÓRRIZ, J. L., GUTIÉRREZ, F., TRULLAS, J. C. *et al.* (2010). Executive summary of the consensus document on the management of renal disease in HIV-infected patients. *Nefrología*, 34(6), pp. 768-788.
- GOURLAY, A., NOORI, T., PHARRIS, A., *et al.* (2017) The Human Immunodeficiency Virus Continuum of Care in European Union Countries in 2013: Data and Challenges. *Clin Infect Dis*, 64(12), pp. 1644–1656.
- GRUPO DE ESTUDIO DEL SIDA-SEIMC (GeSIDA). (Actualización enero 2017). *Documento de consenso de GeSIDA/Plan Nacional sobre el Sida respecto al tratamiento antirretroviral en adultos infectados por el virus de la inmunodeficiencia humana*.
- GRUPO DE EXPERTOS DEL GRUPO DE ESTUDIO SOBRE ALTERACIONES METABÓLICAS (GEAM), DE LA SECRETARÍA DEL PLAN NACIONAL SOBRE EL SIDA (SPNS) Y DEL GRUPO DE ESTUDIO DE SIDA (GeSIDA). (Actualización febrero 2017). *Documento de consenso sobre alteraciones metabólicas y riesgo cardiovascular en pacientes con infección por el VIH*. Disponible en: http://gesida-seimc.org/wpcontent/uploads/2017/02/gesida-guiasclinicas-2017-alteraciones_Metabolicas_y_RCV_VIH.pdf
- HERDMAN, M. (2000). The measurement of health-related quality of life. *Med Clin (Barc)*, 114(Suppl 3), pp. 22-25.
- HIRSCH, J. D., METZ, K. R., HOSOKAWA, P. W. *et al.* (2014). Validation of a Patient-Level Medication Regimen Complexity Index as a Possible Tool to Identify Patients for Medication Therapy Management Intervention. *Pharmacotherapy*, 34(8), pp. 826-835.
- HØGH KØLBÆK KJÆRA, A. S., RASMUSSENA, T. A., HJOLLUND, N. H. *et al.* (2018). Patient-reported outcomes in daily clinical practice in HIV outpatient care. *Int J Infect Dis*, 69, pp. 108-114.
- JARRIN, I., BOLUMAR, F. y DEL AMO, J. (2010). Los estudios de cohortes y su contribución al conocimiento de la progresión de la infección por el virus de la inmunodeficiencia humana (VIH): Principales características y limitaciones. *Enferm Infecc Microbiol Clin*, 28(5), pp. 304–309.
- KNOBEL, H., CODINA, C., MIRO, J. M. *et al.* (2002). Cómo y por qué debe monitorizarse la adherencia al tratamiento antirretroviral en la actualidad. *Enferm Infecc Microbiol Clin*, 20, pp. 481-483.
- LLACH, X. B. y SURIÑACH, N. L. (2005). *Calidad de vida relacionada con la salud en el paciente VIH* (p. 79). Imprime: Gráficas Enar, SA.
- LYTVYN, L., SIEMIENIUK, R. A., DILMITIS, S. *et al.* (2017). Values and preferences of women living with HIV who are pregnant, postpartum, or considering pregnancy on choice of antiretroviral therapy during pregnancy. *BMJ Open*, 7, p. e019023.
- MARRONE, G., MELLGREN, S., LARS, E. *et al.* (2016). High Concordance between Self-Reported Adherence, Treatment Outcome and Satisfaction with Care Using a Nine-Item Health Questionnaire in Inf Care HIV. *PLoS ONE*, 11(6), p. e0156916.
- MARTIN, S., WOLTERS, P. L., CALABRESE, S. K. *et al.* (2007). The Antiretroviral Regimen Complexity Index: a novel method of quantifying regimen complexity. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 45(5), pp. 535-544.

- MATHEWS, W. C., McCUTCHAN, J. A., ASCH, S. *et al.* (2000). National estimates of HIV-related symptom prevalence from the HIV Cost and Services Utilization Study. *Med Care*, 38(7), p. 750-762.
- MUÑOZ-MORENO, J. A., FUMAZ, C. R., FERRER, M. J. *et al.* (2007). Assessing self-reported adherence to HIV therapy by questionnaire: The SERAD (self-reported adherence) study. *AIDS Res Human Retroviruses*, 23(10), pp. 1166-1175.
- NEGREDO, E., DOMINGO, P., JOSÉC, M. *et al.* Executive summary of the consensus document on osteoporosis in HIV-infected individuals. *Enferm Infecc Microbiol Clin*, 36(5), pp. 312-314.
- OLIVA, J., ROA, C., DEL LLANO, J. (2003). Indirect costs in ambulatory patients with HIV/AIDS in Spain: a pilot study. *Pharmacoeconomics*, 21(15), pp.1113-1121.
- POLO, R., BLANCO, J. R. y VERDEJO, C. (2015). *Documento de consenso sobre edad avanzada e infección por el virus de la inmunodeficiencia humana*. Grupo de expertos de la Secretaría del Plan Nacional sobre el sida (SPNS), Sociedad Española de Geriatría y Gerontología (SEGG). Edita MSSSI.
- POWERS, J. H., HOWARD, K., SARETSKY, T. *et al.* (2016). Patient-Reported Outcome Assessments as Endpoints in Studies in Infectious Diseases. *Clin Infect Dis*, 63(Suppl 2), pp. S52-S56.
- RANK, J., PALELLA, JR., DELANEY, K. *et al.* (1998). Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *N Engl J Med*, 338, pp. 853-860.
- RESTELLI, U., RIZZARDINI, G., ANTINORI, A. *et al.* (2017). Cost-effectiveness analysis of dolutegravir plus backbone compared with raltegravir plus backbone, darunavir+ritonavir plus backbone and efavirenz/tenofovir/emtricitabine in treatment naïve and experienced HI V-positive patients. *The Clin Risk Manag*, 13, pp. 787-797.
- RIBERA, E., AGUIRREBENGOA, K., MIRALLES, C. *et al.* (2002). Simplificación del tratamiento antirretroviral. *Enferm Infecc Microbiol Clin*, 20(Supl 2), pp. 48-57.
- SABATE, E. (Ed.). (2003). *Adherence to Long Term Therapies: Evidence for Action*. Switzerland: World Health Organization (citado 15-12-2014). Disponible en: http://www.who.int/chp/knowledge/publications/adherence_report/en/
- SENDI, P. P., BUCHER, H. C., HARR, T. *et al.* (1999). Cost effectiveness of highly active antiretroviral therapy in HIV-infected patients. Swiss HIV Cohort Study. *AIDS*, 13(9), pp. 1115-1122.
- SIMPSON, K. N., HANSON, K. A., HARDING, G. *et al.* (2013). Patient reported outcome instruments used in clinical trials of HIV-infected adults on NNRTI based therapy: a 10-year review. *Health and Quality of Life Outcomes*, 11, p.164.
- SMITH, M. Y., RAPKIN, B. D., MORRISON, A. *et al.* (1997). Zidovudine adherence in persons with AIDS. The relation of patient beliefs about medication to self-termination of therapy. *J Gen Intern Med*, 12(4), pp. 216-23.
- SOTO, M. y FAILDE, I. (2004). La calidad de vida relacionada con la salud como medida de resultados en pacientes con cardiopatía isquémica. *Rev. Soc. Esp. Dolor*, 11, pp. 505-514.
- SUÁREZ-LOZANO, I., FAJARDO, J. M., GARRIDO, M. *et al.* (2002). Epidemiological trends of HIV infection in Spain: preventative plans have to be oriented to new target populations (Spanish VACH Cohort). *AIDS*, 16(18), pp. 2496-2499.
- US DEPARTMENT OF HEALTH AND HUMAN SERVICES, & FOOD AND DRUG ADMINISTRATION. (2015). *Human immunodeficiency virus-1 infection: developing antiretroviral drugs for treatment guidance for industry*. Disponible en: <http://www.fda.gov/downloads/drugs/guidancecom-pliance/regulatoryinformation/guidances/ucm355128.pdf> (acceso 08-04-2016).
- VELASCO, M., GÓMEZ, A., FERNÁNDEZ, C. *et al.* (2000). Economic impact of HIV protease inhibitor therapy in the global use of health-care resources. *HIV Med*, 1(4), pp. 246-251.
- VENTURA-CERDA, J. M., AYAGO-FLORES, D., VICENTE-ESCRIG, E. *et al.* (2010). Costes y adherencia del tratamiento antirretroviral. *Farm Hosp*, 34(6), pp. 284-292.

- VON WICHMANN, M. Á., LOCUTURA, J., BLANCO, J. R. *et al.* & GRUPO DE ESTUDIO DEL SIDA. (2010). Indicadores de calidad asistencial de GESIDA para la atención de personas infectadas por el VIH/sida. *Enferm Infecc Microbiol Clín*, 28, pp. 6-88.
- WHALEN, C. C., ANTANI, M., CAREY, J. *et al.* (1994). An index of symptoms for infection with human immunodeficiency virus: reliability and validity. *J Clin Epidemiol*, 47(5), pp. 537-546.
- WOODCOCK, A. y BRADLEY, C. (2011). Validation of the HIV treatment satisfaction questionnaire (HIVTSQ). *Qual Life Res*, 10(6), pp. 517-531.



10

RESULTADOS EN SALUD EN PACIENTES CON HEPATITIS C CRÓNICA

MARÍA BUTI FERRET, MARÍA QUERALT GORGAS,
MIGUEL A. RODRÍGUEZ SAGRADO

10.1. INTRODUCCIÓN

La hepatitis C es un problema de salud global que afecta a más de 170 millones de personas en todo el mundo. La prevalencia actual en Europa varía según el país, y oscila entre el 2-3 % de la población total (Myers *et al.*, 2015; Ansaldi *et al.*, 2014; Negro, 2014).

La prevalencia real en España no es bien conocida. En el año 2013 se estimaba que 473.000 personas tenían hepatitis C, definida por la presencia de RNA-VHC (Omata *et al.*, 2016).

La evolución natural de la enfermedad hace que el 60-90 % de los pacientes infectados desarrollarán una hepatitis C crónica y, de estos, el 20 % desarrollarán una cirrosis (Myers *et al.*, 2015; Ansaldi *et al.*, 2014 y Omata *et al.*, 2016). La descompensación de la cirrosis puede causar insuficiencia hepática o carcinoma hepatocelular (3-5 % a los 5 años) (Razavi *et al.*, 2014).

Por tanto, la detección temprana y el tratamiento son muy importantes para la prevención de estos eventos (Ansaldi *et al.*, 2014; Omata *et al.*, 2016; Reid, Price y Tien, 2017; Fernández Rodríguez y Gutierrez Garcia, 2014; Benítez-Gutiérrez *et al.*, 2016). Algunos autores estiman que la hepatitis C crónica es responsable de aproximadamente un millón de muertes por año en todo el mundo (Ansaldi *et al.*, 2014).

La carga social y económica de la infección por el virus de la hepatitis C (VHC) es considerable y ha llevado a la Organización Mundial de la Salud (OMS) a recomendar estudios locales para definir el alcance del problema y dar prioridad a las medidas preventivas (Baylis, Heath y Collaborative Study Group, 2011; Lauffenburger *et al.*, 2014), entre la que se destaca el tratamiento de los pacientes infectados para controlar la transmisión de la enfermedad.

El arsenal terapéutico cuenta con los nuevos fármacos antivirales de acción directa (AAD) que son mucho más seguros, de administración oral y con una altísima eficacia, que permiten la curación de más del 95 % de los pacientes (Ministerio de Sanidad, Consumo y Bienestar Social, [2018]). Sin embargo, estos nuevos regímenes de AAD implican un incremento en el presupuesto de los sistemas de salud a corto plazo, debido al elevado coste de dichos fármacos y al aumento significativo de pacientes candidatos a recibir tratamiento por sus elevadas tasas de eficacia y excelente perfil de seguridad. Sin embargo, estos tratamientos implican una dismi-

nución en las necesidades de atención sanitaria (menos visitas médicas, de enfermería, de farmacia, analíticas, etc.).

Por este motivo en España y en otros países (Alves *et al.*, 2016; Association Francaise pour l'Etude du Foie (AFEF), 2015; Australian Government Department of Health, 2018) se han establecido políticas sanitarias que, basadas en la limitación de recursos económicos disponibles, permitan el tratamiento prioritario de los pacientes con mayor grado de fibrosis (y por tanto con mayor riesgo de evolución de la enfermedad) para conseguir beneficios a corto plazo en forma de curación de la enfermedad y disminución de las complicaciones clínicas derivadas de la enfermedad hepática o de las enfermedades extrahepáticas de la infección, y para conseguir beneficios a largo plazo, por los costes evitados del tratamiento de posibles complicaciones futuras derivadas de la evolución de la enfermedad. En fases posteriores, se ha generalizado el uso de los AAD a todos los pacientes infectados.

En España, el Plan Estratégico para el Abordaje de la hepatitis C (PEACH) ha permitido que numerosos pacientes hayan tenido acceso a los tratamientos de la hepatitis C (Ministerio de Sanidad, Consumo y Bienestar Social, 2020).

■ 10.1.1. Resultados en salud clínicos (*Clinical outcomes assessments*)

a. Resultados comunicados por los clínicos (*Clinician-reported outcomes*)

Los AAD han revolucionado el tratamiento de la hepatitis crónica C. Las tasas de curación de la infección crónica por VHC definidas por la respuesta virológica persistente (RVP) en los sujetos incluidos en los ensayos clínicos de los fármacos más recientemente aprobados como sofosbuvir/velpatasvir (SOF/VEL) o glecaprevir/pibrentasvir (GLE/PIB) o elbasvir /grazoprevir (EBV/GZR) son muy elevadas con duraciones de tratamiento entre 8 y 12 semanas y un perfil de seguridad y tolerancia excelente (European Association for the Study of the Liver, 2018). Incluso en pacientes con cirrosis descompensada en los que el interferón estaba contraindicado, los nuevos antivirales orales son capaces de eliminar la infección por VHC de una manera muy eficaz y segura (Mangia *et al.*, 2018). Sin embargo, las tasas de RVP con los antivirales orales en pacientes descompensados son más bajas que en los pacientes con cirrosis, especialmente en los pacientes infectados por genotipo 3 (European Association for the Study of the Liver, 2018). Pocas incertidumbres quedan en el tratamiento de los pacientes con hepatitis C como la necesidad de añadir ribavirina (RBV) al régimen antiviral en los casos de cirrosis descompensada y el retratamiento de los pacientes que han fracasado a los antivirales más potentes como SOF/VEL o GLE/PIB.

La erradicación viral en pacientes con hepatitis crónica C se asocia a la mejoría en la evolución clínica a medio y largo plazo tanto en morbilidad como en mortalidad, así como en resultados informados por el paciente (PROs) (European Association for the Study of the Liver, 2018). Sin embargo, en pacientes con

cirrosis descompensada, no está claro si el tratamiento y el aclaramiento del VHC es a largo plazo beneficioso, especialmente en aquellos sujetos con enfermedad hepática descompensada en los que la enfermedad hepática puede estabilizarse y no progresar a una fase en la que esté indicado el trasplante hepático pero el sujeto puede quedar con una calidad de vida pobre, lo que se conoce como “el Purgatorio MELD” (European Association for the Study of the Liver, 2018).

La experiencia clínica con los últimos tratamientos aprobados en práctica clínica empieza a ser importante. El régimen de SOF/VEL, combina un análogo de los nucleótidos con un inhibidor de NS5A con actividad pangenotípica, y se administra en forma de un comprimido al día durante 12 semanas en todos los pacientes independientemente del genotipo viral y del grado de fibrosis. Mangia *et al.*, han comunicado los resultados en 1.099 pacientes tratados con esta combinación durante 12 semanas (Mangia *et al.*, 2018). La mayoría eran hombres (55 %), edad media 63,2 con una elevada proporción de casos con edad superior a 70 años (39 %) y 223 (43,5 %) tenían cirrosis o fibrosis avanzada. El genotipo 1 se detectó en 36,2 % casos, el genotipo 2 en 46,8 %, el genotipo 3 en 17,2 % y genotipo 4 en 9,8 %. La tasa de RVP por intención de tratar fue del 98,6 % y en los pacientes que terminaron el seguimiento del 99 %. No se observaron diferencias en las tasas de curación en relación con genotipo. Solo cuatro pacientes interrumpieron prematuramente el tratamiento y no se reportaron efectos adversos graves (Mangia *et al.*, 2018).

En pacientes infectados por genotipo 3, la combinación SOF/VEL durante 12 semanas fue evaluada en un estudio de fase 3 (Astral 3) que mostro tasas de RVP muy elevadas, del 95 %, si bien estas tasas fueron numéricamente inferiores (91 %) en los pacientes con cirrosis comparadas con las observadas en los pacientes sin cirrosis (97 %). Esto motivó la realización de un estudio multicéntrico español que evaluó la eficacia y seguridad de SOF/VEL ± RBV durante 12 semanas en pacientes infectados por genotipo 3 con cirrosis hepática. Un total de 204 sujetos con cirrosis y genotipo 3 fueron incluidos. Las tasas de RVP fueron del 91 % para el brazo SOF/VEL y del 96 % para el brazo SOF/VEL+RBV. Las tasas de respuestas fueron idénticas para los pacientes previamente tratados con SOF/VEL o SOF/VEL+RBV. Entre los pacientes que recibieron SOF/VEL se observaron 5 recaídas mientras que en el brazo SOF/VEL+RBV se observaron 2 recaídas. Un paciente en cada grupo interrumpió el tratamiento, en el brazo SOF/VEL debido a ansiedad y en el brazo SOF/VEL+RBV debido a aumento en los niveles de bilirrubina. Los efectos adversos más comunes fueron astenia (27 %), cefalea (24 %) e insomnio (12 %). No se observaron efectos adversos graves (Esteban *et al.*, 2018).

La combinación glecaprevir (un inhibidor de NS3/4A) y pibrentasvir (un inhibidor de NS5) es un régimen pangenotípico con tasas de RVP en los estudios de registro del 98 %. Esta combinación se administra durante 8 semanas en los pacientes sin cirrosis y durante 12 semanas en los pacientes con cirrosis. Los primeros datos de práctica clínica fueron presentados por Berg *et al.*, en un estudio realizado en Alemania que incluyó 414 pacientes, la mayoría no tratados previamente y sin cirrosis hepática, por lo que recibieron 8 semanas de tratamiento. Los datos preli-

minares en 96 pacientes muestran una tasa de RVP del 97 %. Tres pacientes interrumpieron el tratamiento por efectos adversos no relacionados con la medicación siendo la tasa de curación del 100 % (Berg *et al.*, 2018). Un estudio similar con la misma combinación realizado en Italia incluyó 723 pacientes, con edades entre 21 y 89 años, 49 % varones, la mayoría no cirróticos (93 %) y la mitad infectados por el genotipo 1. La tasa de RVP en los primeros 44 pacientes que han completado las 8 semanas de tratamiento fue del 97,7 %. No se reportaron efectos adversos relevantes (D'Ambrosio *et al.*, 2018).

Finalmente, la combinación elbasvir/grazoprevir, fármacos coformulados en 1 comprimido que se administra una vez al día durante 12 semanas ha sido evaluada en práctica clínica en pacientes infectados por genotipo 1 y 4. Esta combinación ha sido aprobada para utilizarse en sujetos infectados por genotipo 1 y/o 4. Los resultados de práctica clínica recogidos en la base de datos de la Asociación Española para el Estudio del Hígado (AEEH) incluyeron 804 pacientes, la mayoría (72 %) genotipo 1b, sin cirrosis hepática (86 %) que recibieron EBV/GZV durante 12 semanas. Las tasas de RVP en los pacientes que han finalizado el seguimiento fue del 96,9 % (570/588), sin diferencias entre genotipo 1 y 4. Tres pacientes interrumpieron el tratamiento y los efectos más comunes fueron astenia o síntomas gastrointestinales de intensidad leve o moderada (Hernández-Conde *et al.*, 2018).

En resumen, estos estudios demuestran que, en práctica clínica, en una población de sujetos con hepatitis crónica C más heterogénea y con comorbilidades, los resultados reportados por los clínicos son muy similares a los observados en los estudios de registro en términos de eficacia clínica, tolerabilidad y seguridad. Ello hace que el cumplimiento del tratamiento sea muy bueno.

b. Resultados percibidos por los pacientes (*Patient-reported outcomes*)

Los PRO incluyen cualquier resultado derivado de los informes de los pacientes. Más específicamente, los PRO se definen como cualquier informe del estado de salud del paciente que comunica directamente el mismo, sin que su respuesta sea interpretada por un médico o cualquier otra persona (European Association for the Study of the Liver, 2018). Los PRO incluyen conceptos tales como calidad de vida, adherencia al tratamiento y toxicidad percibida reportada como síntomas auto informados, que son medidas de la experiencia de los pacientes con el tratamiento y un complemento clave de los resultados clínicos (U.S. Department of Health and Human Services FDA *et al.*, 2006). También incluyen otras mediciones de la experiencia subjetiva de los pacientes con enfermedades y tratamientos que, al igual que los conceptos mencionados anteriormente, no pueden ser capturados directamente por marcadores clínicos objetivos, pero que son útiles desde una perspectiva clínica y de salud pública, como el estigma percibido de la enfermedad, el acceso y la satisfacción con la atención (U.S. Department of Health and Human Services FDA *et al.*, 2006) o la productividad laboral (Younossi *et al.*, 2016a). Estos últimos tienen consecuencias inmediatas en la vida cotidiana y la condición social de las

personas, y al igual que los anteriores tienen repercusiones económicas a nivel de toda la sociedad.

Es conocido que los pacientes con hepatitis crónica C tienen los PRO deteriorados en comparación con las personas no infectadas y reportan una peor calidad de vida relacionada con la salud (CVRS), especialmente en la dimensión vitalidad. La CVRS deteriorada en pacientes con hepatitis crónica C se observa incluso en aquellos sujetos con enfermedad hepática leve, y se debe principalmente a una fatiga persistente y otras manifestaciones extrahepáticas, incluidas el deterioro del funcionamiento cognitivo debido a la acción de virus de la hepatitis C en el sistema nervioso central, así como comorbilidades asociadas frecuentes como depresión. La enfermedad hepática avanzada, y especialmente la cirrosis, empeoran la CVRS, principalmente en algunas dimensiones de CVRS como la actividad, energía, vitalidad y fatiga. La evaluación de los PRO ayuda a conocer el impacto del tratamiento de la hepatitis C. Esto fue específicamente importante en la era de terapias basadas en interferón, ya que estas terapias se asociaron con múltiples efectos secundarios.

Los primeros análisis de los PRO en pacientes que recibieron AAD se centraron en la experiencia directa de los pacientes en los síntomas auto reportados de calidad de vida o fatiga.

Diferentes instrumentos psicométricos se utilizan para evaluar PRO en pacientes con hepatitis C siendo los más utilizados la encuesta de 36 *ítems* de forma abreviada del estudio de resultados médicos (SF-36), el Perfil de Impacto de Enfermedad (SIP), y el Cuestionario de Productividad del Trabajo y Deterioro de la Actividad Problema de salud específico (WPAI: SHP). Además, se han utilizado instrumentos específicos de enfermedad hepática como *Chronic Liver Disease Questionnaire-HCV* (CLDQ-HCV), y la calidad de vida de la enfermedad hepática ([LDQOL] –forma corta) e instrumentos específicos de fatiga como el formulario de inventario de síntomas de fatiga multidimensional (MFSI-SF) y la evaluación funcional de la terapia de enfermedades crónicas: fatiga (FACIT F).

La mayoría de los datos PRO de la literatura publicada sobre AAD se refieren a ensayos clínicos que evaluaron regímenes basados en sofosbuvir. Los hallazgos muestran mejoras tempranas y sustanciales de PRO con regímenes libres de interferón y RBV (European Association for the Study of the Liver, 2018). Los ensayos ASTRAL-1, -2, -3 y -4 incluyeron pacientes con hepatitis C infectados con diferentes VHC genotipos, tratados con SOF/VEL con o sin RBV. En un análisis *post-hoc*, Younossi *et al.*, compararon los cambios entre los PRO basales y después del tratamiento en pacientes sin cirrosis, con cirrosis compensada, y con cirrosis descompensada (Younossi *et al.*, 2016b). Los resultados mostraron una mejora significativa en las puntuaciones de los PRO (CVRS, fatiga, trabajo y productividad) durante el tratamiento y después de alcanzar la RVP especialmente en pacientes con cirrosis descompensada.

El tratamiento con elbasvir/grazoprevir en pacientes naive con hepatitis crónica C impactó de forma positiva en la CVRS, los niveles de fatiga, la productividad laboral comparado con placebo (Younossi *et al.*, 2016b). Estos hallazgos fueron confirmados en pacientes sin tratamiento previo con coinfección por el virus de la inmunodeficiencia humana (VIH) (ensayo C-EDGE CO-INFXN). Además, en personas con hepatitis crónica C que se inyectan drogas y reciben tratamiento sustitutivo con metadona, la administración de EBV/GZR se asoció con cambios más favorables en CVRS que el uso de placebo, para las subescalas de Salud general, Vitalidad, Salud mental y en las puntuaciones de Componente Mental de la escala MOS SF-36 (Scheiner *et al.*, 2016). En todos los estudios, los pacientes que reciben tratamiento con AAD y ribavirina informan de peores PRO que aquellos que reciben solo AAD, debido a los efectos adversos de la ribavirina, como la anemia y la fatiga asociada.

No existen datos del mundo real sobre PRO en pacientes con hepatitis crónica C tratados con AAD, especialmente en poblaciones vulnerables con menor acceso a la atención, tales como usuarios de drogas por vía parenteral. Están en curso varios estudios de cohortes de pacientes con hepatitis C algunos de ellos incluyendo coinfectados con el VIH que están recolectando datos de PRO. Por ejemplo, en la cohorte francesa HEPATHER, HRQL se recogen los PRO de pacientes con hepatitis C que reciben diferentes regímenes antivirales con el fin de estimar la CVRS para realizar posteriormente un análisis de costo-efectividad.

En el estudio francés del ANRS CO13, la cohorte HEPAVIH, también colecta datos de CVRS, fatiga, síntomas autoinformados y conductas adictivas antes y después del tratamiento con AAD en pacientes con hepatitis C coinfectados con VIH y en seguimiento en clínicas ambulatorias (Arduino *et al.*, 2016). Los PRO relacionados con conductas adictivas, como el consumo de drogas, tabaco o alcohol, también son muy útiles no solo antes y durante el tratamiento contra el VHC, sino también después de lograr la eliminación del VHC, como estos comportamientos pueden agravar las comorbilidades relacionadas con el VHC.

La investigación de PRO en los pacientes con hepatitis C tratados con AAD debe continuarse especialmente en sujetos que pertenecen a grupos vulnerables y en el contexto de reinfección por VHC. Los datos actualmente disponibles principalmente provienen de ensayos clínicos, que tienen varias limitaciones. Primero, el período de observación es limitado a la duración de los ensayos, siendo lo más común de entre 6 meses y tres años. Segundo, la evaluación de PRO en ensayos clínicos generalmente se programa coincidiendo con los resultados de eficacia, pero rara vez después. Por lo tanto, se necesitan estudios con un seguimiento más largo para documentar los cambios en PRO tras finalizar el tratamiento. Los estudios de seguimiento más largos son también necesarios para documentar las reinfecciones múltiples del VHC y para implementar programas de reducción de daños en los subgrupos de personas con mayor riesgo de reinfección, incluido los hombres que tienen sexo con hombres y los usuarios de drogas por vía parenteral. Otra limitación de los ensayos clínicos es que se han llevado a cabo entre poblaciones seleccionadas, cumpliendo criterios de inclusión restrictivos, lo que limita la

generalización de los resultados obtenidos para toda la población de pacientes con hepatitis C. Todavía faltan datos sobre los PRO con AAD en subpoblaciones claves de pacientes con hepatitis C como pacientes con trastornos psiquiátricos y/o trastornos adictivos. Estos datos serán importantes para diseñar e implementar modelos efectivos de atención para todos los pacientes con hepatitis C incluyendo aquellos con dificultades sociales. Además, ayudarán a describir mejor la vinculación con la atención médica documentando el acceso al diagnóstico y al tratamiento del VHC, así como a conocer las tasas de pacientes que logran la eliminación del VHC en subpoblaciones vulnerables. Finalmente, los estudios de cohorte observacionales jugarán un papel clave en la documentación de los problemas restantes sobre los resultados informados por los pacientes con AAD para el tratamiento de hepatitis C y serán esenciales para ayudar a implementar estrategias efectivas de reducción de riesgos para limitar la reinfección del VHC y promover el acceso al tratamiento universal.

Finalmente, el cumplimiento o adherencia al tratamiento es básica para que los AAD sean eficaces. Para que la adherencia a un tratamiento pueda considerarse buena se deben tomar como mínimo más del 95 % de las dosis tal como fueron pautadas, a su hora y cumpliendo con las pautas de ingestión que se han marcado por el profesional sanitario que supervisa el tratamiento.

Es bastante difícil establecer en qué medida el nivel de adherencia condiciona la respuesta al tratamiento y las probabilidades de curación pero de hecho es casi lo único en lo que el paciente, con su cumplimiento, puede influir de cara al éxito de su tratamiento, en la calidad del nivel de cumplimiento.

■ 10.1.2. Resultados en salud económicos (*Economic Outcomes assessment*)

Los avances en farmacoterapia condicionan la evolución de la financiación de los tratamientos en los sistemas sanitarios. Por ello, las políticas sanitarias actuales tienden hacia la racionalización del uso de los medicamentos, en términos de seguridad, eficacia, efectividad, eficiencia, equidad, beneficio social, calidad de vida de los pacientes y adecuación de costes. La necesidad de conocer los resultados en salud de las alternativas terapéuticas surge de la condición de administrar los recursos, que son escasos, frente a las necesidades, que son ilimitadas o por lo menos superiores a los recursos. Siempre y cuando los resultados de evaluar una nueva alternativa terapéutica en términos de seguridad, eficacia, efectividad y calidad de vida sean positivos, los resultados en salud económicos son un pilar fundamental de apoyo a la posterior toma de decisiones.

En este sentido, (Turnes, Domínguez-Hernandez y Casado, 2017a y 2017b) realizaron en 66.777 pacientes un análisis de coste utilidad del acceso a los nuevos fármacos tras dos años de implantación del PEACH en España (años 2015-2016)

mediante un árbol de decisión y un modelo de Markov basados en un análisis previo, comparando los tratamientos nuevos (pos-AAD: a partir de la comercialización de simeprevir) con los anteriores (pre-AAD: regímenes con interferón y ribavirina en doble o triple terapia con telaprevir y boceprevir). La respuesta viral sostenida en semana 12 postratamiento se consiguió en el 95-98 % en la estrategia pos-AAD y en el 48-78 % en la estrategia pre-AAD (estas últimas extraídas de los resultados de los estudios más relevantes con dicha estrategia pre-AAD). Los costes totales invertidos en la estrategia pos-AAD fueron de 1.602 millones de euros en tratamiento farmacológico, 83 millones en la monitorización y 248 millones en el manejo de la enfermedad (1.932 millones de euros en total de la estrategia). Por otro lado, los costes totales invertidos en la estrategia pre-AAD fueron de 1.532 millones de euros en tratamiento farmacológico, 46 millones en la monitorización y 1.197 millones en el manejo de la enfermedad (1.532 millones de euros en total de la estrategia). Los años de vida ajustados por calidad (AVAC) ganados en cada estrategia fueron de 1.196.483 en la estrategia pos-AAD y 962.556 en la pre-AAD. Como es lógico, los costes farmacológicos utilizados en la estrategia pos-AAD fueron mayores que en la pre-AAD, y los costes del tratamiento de las complicaciones fueron mayores en la estrategia pre-AAD. La estrategia pos-AAD comparada con la estrategia pre-AAD, se asoció con un aumento en los AVAC de 17,9 vs 14,4, asociándose a mayores costes (28.935 euros vs 22.947 euros). La estrategia pos-AAD mostró una relación coste-utilidad incremental de 1.709 euros por AVAC por paciente. También estiman que durante la vida de los pacientes la estrategia pos-AAD evitaría 13.791 casos de cirrosis descompensada (lo que supondría una reducción del 87 %), 8.573 casos de carcinoma hepatocelular (reducción del 74 %) y 1.761 trasplantes hepáticos (reducción del 84 %). De esta forma, se reduciría un 82 % las muertes por causa hepática, evitándose 14.659 fallecimientos. El valor monetario del incremento en los AVAC conseguidos con la estrategia pos-AAD frente a la pre-AAD en toda la cohorte de pacientes, a partir del número de los AVAC incrementales (233.932) considerando un umbral de eficiencia entre 20.000 y 30.000 euros, osciló entre 4.679 y 7.018 millones de euros evitados.

Un estudio similar (Marco, Domínguez-Hernández y Casado, 2018) realizado en 4.408 pacientes reclusos dependientes del Instituto Catalán de Salud, da resultados comparables, evitándose 951 casos de cirrosis descompensada (lo que supondría una reducción del 92 %), 650 casos de carcinoma hepatocelular (reducción del 83 %) y 132 trasplantes hepáticos (reducción del 90 %), reduciéndose finalmente en un 88 % los casos de muerte por causa hepática, evitándose 1.100 muertes al utilizarse la estrategia pos-AAD. Igualmente, Chhatwall *et al.* (2018), concluyen que la generalización del tratamiento en la población reclusa de todas las cárceles españolas sería costo efectivo, al evitarse en los próximos 25 años en España 70 nuevos casos incidentes de VHC, 270 de cirrosis descompensada, 450 de carcinoma hepatocelular y 670 muertes relacionadas con el VHC.

Los distintos acuerdos con los laboratorios que comercializan estos fármacos (a veces muy complicados con novedosas fórmulas de pago) han permitido la uni-

versalización de los tratamientos, pero han introducido un sesgo en todos los estudios, ya que se desconocen los precios realmente pagados por el Sistema Sanitario por estos fármacos. Todos los estudios utilizan el precio de los fármacos basados en los datos publicados por el Ministerio de Sanidad. Otras limitaciones de los estudios son las derivadas de la modelización.

Además, no han tenido en cuenta el impacto social de la enfermedad, de forma que no se ha considerado la pérdida de productividad en pacientes con hepatitis C y los costes futuros no relacionados con la enfermedad. Estos pacientes, debido al manejo clínico de la enfermedad, se ausentan numerosas horas de sus puestos de trabajo. Distintos estudios indican que el absentismo y el presentismo laboral de un paciente con hepatitis C en España supondría un coste anual total de 2.395 euros (Younossi *et al.*, 2016a). Por tanto, la elevada eficacia y seguridad de estos fármacos implica una mayor productividad laboral de estos pacientes, reduciendo con ello los costes relacionados con la pérdida de productividad desde el punto de vista de la sociedad española (Younossi *et al.*, 2016a; Oliva-Moreno *et al.*, 2015).

Buti *et al.* (2017) estudiaron el coste-efectividad de iniciar tratamiento en pacientes con genotipo 1 con ledipasvir/sofosbuvir en fases tempranas de la enfermedad cuando el grado de fibrosis es bajo (F0-F1) versus el inicio en fases más evolucionadas de la enfermedad (F2-F4). Los resultados demuestran que empezar tratamiento en fases iniciales es una estrategia dominante [19,80 años de vida ganados (AVG) y 19,85 años de vida ajustados por calidad (AVAC)] frente a comenzar tratamiento en fases más avanzadas de la enfermedad (18,63 AVG y 16,25 AVAC), lo que se tradujo en un ahorro de 9.228 euros por paciente. En una cohorte de 1.000 pacientes F0-F1 genotipo 1 en tratamiento con ledipasvir/sofosbuvir se disminuyó el número de casos de cirrosis descompensada en un 93 %, de carcinoma hepatocelular en un 97 %, lo que reduciría un 95 % las muertes por causa hepática y prevendría 6 trasplantes de hígado en comparación con el inicio de tratamiento en fases tardías de la enfermedad (F2-F4). En estos pacientes con genotipo 1 con hepatocarcinoma el inicio de tratamiento con ledipasvir/sofosbuvir en fases tempranas comparado con el inicio en fases avanzadas de la enfermedad se tradujo en un aumento de la efectividad de 1,22 AVG y 3,55 de los AVAC.

En cuanto a la comparación entre distintos regímenes de tratamientos con AAD, hay poco publicado, siendo fundamentalmente los análisis de minimización de costes. En ellos se presuponen que todos los tratamientos tienen una eficacia similar y una tolerancia excelente (y por tanto que no producen costes asociados a las comorbilidades derivadas del tratamiento con AAD) y solo tienen en cuenta el coste directo de adquisición del medicamento. Margusino-Framiñán *et al.* (2018) concluyen que para el tratamiento en pacientes genotipo 1 sin cirrosis, el tratamiento durante 8 semanas con pibrentasvir/glecaprevir es la opción más costo-efectiva seguida del tratamiento durante 12 semanas con veltapasvir/sofosbuvir. Al igual que otras comparaciones indirectas tiene la limitación de que no hay ensayos clínicos que comparen directamente las alternativas terapéuticas incluidas en las recomendaciones de tratamiento de las sociedades científicas, lo que hace que

las características basales de los pacientes incluidos en los distintos estudios puedan incluir sesgos.

La disponibilidad de AAD en España para el tratamiento de la infección por VHC desde abril de 2015 puede considerarse una innovación disruptiva. De acuerdo con este concepto, formulado por primera vez por el profesor Clayton Christensen en 1995 (Christensen y Rosenbloom, 1995), la innovación radical o disruptiva eventualmente hace cambios profundos en la sociedad al introducir simplicidad, conveniencia, accesibilidad y asequibilidad en sectores donde las condiciones previamente existentes están asociadas a barreras, altos costos y/o algunos tipos de restricciones. Los AAD han permitido curar a pacientes con estadios iniciales de la enfermedad, así como han curado a pacientes tradicionalmente considerados difíciles de tratar y sin alternativas terapéuticas disponibles, avocados al trasplante o a la muerte, con los consiguientes beneficios personales y para la sociedad.

■ 10.1.3. Medida de resultados en salud en hepatitis

La eliminación de la hepatitis C es un objetivo marcado por la OMS para 2030. Esto implica una estrategia global de prevención, tratamiento, financiación, seguimiento y control para conseguir los resultados terapéuticos deseados (WHO, 2016).

La información sobre resultados en salud y consumo de recursos en la práctica clínica es una herramienta esencial para la toma de decisiones. Los resultados en vida real pueden ser diferentes de los obtenidos en los ensayos clínicos ya que algunas poblaciones pueden estar menos representadas, los pacientes presentan menos comorbilidades o la adherencia terapéutica en vida real es menor. Estos estudios permiten obtener datos de efectividad y seguridad en poblaciones que no fueron reclutadas en ensayos clínicos, determinar el beneficio del tratamiento en comorbilidades asociadas a la infección por VHC o identificar los factores predictivos de respuesta, para así ayudar a la toma de decisiones futuras sobre el tratamiento (Backus *et al.*, 2016).

Para la monitorización de la efectividad terapéutica de los tratamientos pueden emplearse diferentes diseños, desde estudios observacionales retrospectivos (por ejemplo, datos recogidos a partir de las historias clínicas, registros de pruebas de laboratorio, consumo de medicamentos, etc...), o estudios prospectivos (registros de pacientes y tratamientos, encuestas de salud, etc...) hasta ensayos clínicos pragmáticos (incluyendo pacientes procedentes de la práctica clínica habitual) (The Association of British Pharmaceutical Industry, 2011).

En la hepatitis, la mayoría de los estudios basados en *Real World Evidence* (RWE) son estudios observacionales con cohortes prospectivas de pacientes que son seguidos durante un periodo de tiempo para ver la efectividad, la seguridad y/o la eficiencia del tratamiento. Estas cohortes pueden incluir diferentes tipos de pacientes: mono infectados, coinfectados, con diferente grado de fibrosis, diferen-

tes genotipos o con enfermedad hepática avanzada o distintas comorbilidades (Hézode, 2018).

Los resultados de efectividad y seguridad de diferentes estudios de RWE con diversas combinaciones de fármacos han proporcionado tasas globales de RVP de hasta el 95-100 % y con buena tolerancia al tratamiento. Las combinaciones recomendadas para pacientes coinfectados VIH/VHC son las mismas que para los pacientes mono infectados y muestran tasas de respuesta similares. Estos datos confirman los resultados de los ensayos clínicos publicados (Hézode, 2018; Afdhal y Serfaty, 2016; Berenguer *et al.*, 2018).

Además, diversos estudios publicados valoran la eficiencia del tratamiento de la hepatitis C desde diferentes perspectivas, resultando ser más coste-efectivo cuanto antes se inicie el tratamiento (Turnes, Domínguez-Hernández y Casado, 2017a; Pinsky *et al.*, 2018; Deuffic-Burban *et al.*, 2018; Ruggeri *et al.*, 2018; Cuadrado *et al.*, 2018; Buti *et al.*, 2018).

En España, la irrupción de los nuevos tratamientos de la hepatitis C y el impacto social y político que generó en su día el acceso a estos medicamentos, propició una rápida actuación del sistema público de salud para facilitar el acceso de los pacientes a estos fármacos. Entre las actuaciones regulatorias realizadas hay que señalar la creación del PEAHC en el Sistema Nacional de Salud (SNS) en marzo de 2015 y un registro de pacientes de hepatitis C tratados con los nuevos fármacos, como requisito necesario para la monitorización de la efectividad terapéutica y el análisis de los resultados en salud. Este registro se denominó “Sistema de Información de monitorización Terapéutica de pacientes con Hepatitis C crónica” (SITHePaC) y supone una recogida protocolizada de datos de seguimiento terapéutico de pacientes tratados con los nuevos antivirales de acción directa (Ministerio de Sanidad, Consumo y Bienestar Social, [2018]).

A fecha 30 de junio de 2018 el número de pacientes tratados en España es de 113.147, de los cuales 65.689 pacientes están registrados en SITHePaC, y de estos se dispone de datos de efectividad del 44 % de los registros, siendo la RVP a las 12 semanas de la finalización del tratamiento del 95,52 % (Ministerio de Sanidad, Consumo y Bienestar Social, [2018]).

La evaluación preliminar de los resultados de los dos primeros años (2015-2016) de aplicación del PEAHC, desde la doble perspectiva de la eficiencia en términos de coste/utilidad del plan y de reducción de eventos clínicos futuros, estima que se evitarán entre el 74-87 % de complicaciones clínicas secundarias a la enfermedad hepática a largo plazo (88 % de casos de cirrosis descompensada, 74 % de hepatocarcinoma, 84 % trasplantes hepáticos y 82 % muertes por causa hepática) con un aumento de 3,51 AVAC (años de vida ajustados por calidad) y una buena relación de coste-utilidad incremental de 1.451€/AVAC. Por ello, la implantación del PEAHC en España puede considerarse una opción eficiente para el tratamiento de la hepatitis C, obteniendo un coste por AVAC de 28.027€, inferior al umbral de

eficiencia utilizado en España, y con una reducción significativa de eventos clínicos asociados a la enfermedad (Turnes, Domínguez-Hernández y Casado, 2017a).

Aunque el PEAHC ha representado un éxito de nuestro sistema sanitario en términos de accesibilidad a tratamientos innovadores y de resultados en salud de su aplicación, extender estos tratamientos al resto de la población infectada en nuestro país mediante un programa de cribado poblacional es, no solo prioritario, sino también una medida eficiente en términos de inversión en salud.

Varios estudios publicados muestran los resultados positivos del inicio precoz del tratamiento de la hepatitis C. Un análisis coste-efectividad realizado en Escocia concluye que el inicio del tratamiento en las fases tempranas de la enfermedad es más coste-efectivo que cuando existe un grado de fibrosis avanzada o cirrosis. Si el tratamiento se inicia cuando el grado de fibrosis es F0 o F1 el riesgo de por vida de desarrollar una cirrosis descompensada es del 4 %, dicho riesgo aumenta al 11,6 % cuando el grado de fibrosis es F4. Incluso diferir el tratamiento hasta el estadio F2 supone un incremento de riesgo de hasta un 8,9 %. Los riesgos acumulados de por vida de sufrir carcinoma hepatocelular o de fallecer por causas hepáticas aumentan de forma espectacular, en un 35,2 % y un 41,1 % respectivamente, en el caso de las personas que inician el tratamiento durante el estadio F4 de fibrosis (Pinsky *et al.*, 2018).

En un estudio realizado en Francia, en el que se comparaban diferentes estrategias de cribado en función del riesgo o de la edad del paciente, muestra que el cribado universal es la estrategia más eficaz y que es coste-efectiva cuando se inicia el tratamiento independientemente del estadio de fibrosis del paciente (Deuffic-Burban *et al.*, 2018). Otro estudio realizado en una cohorte italiana concluye que el acceso temprano al tratamiento con independencia del grado de fibrosis, en comparación al tratamiento de estadios más avanzados de fibrosis F3-F4 también resulto ser coste-efectivo (Ruggeri *et al.*, 2018).

En España dos estudios recientes sobre cribado de la hepatitis C acreditan el coste-efectividad del cribado universal en adultos. Un estudio de la primera estrategia de cribado propuesta para una cohorte real (cohorte Ethon, 12.515 personas) concluye que el tratamiento es coste-efectivo en cualquier grupo de edad, y especialmente entre los 45 y 49 años. (Cuadrado *et al.*, 2018). Otro estudio, basado en un modelo matemático de decisión, concluye que el cribado y el tratamiento de la población general de entre 20 y 79 años es más coste-efectivo que el de la misma población con factores de riesgo. En concreto, la relación coste-utilidad incremental de la estrategia de cribado de la población general frente a la población con factores de riesgo fue de 8.914 euros, muy por debajo del umbral de eficiencia aceptado en España (Buti *et al.*, 2018).

Lograr la curación de la hepatitis C no solamente tiene ventajas individuales, como evitar la progresión de la enfermedad hepática, reducir el número de casos de hepatocarcinoma, mejorar la calidad de vida y disminuir las manifestaciones extra-

hepáticas de la infección, sino también logra beneficios sociales, como aumentar la productividad laboral o reducir la infectividad del virus (Deuffic-Burban *et al.*, 2018).

De este modo, las estrategias de cribado universal junto con una correcta política de utilización, si bien no permitirán erradicar el virus de la hepatitis C por no existir vacuna, si reducir su incidencia y prevalencia a valores mínimos.

■ 10.1.4. Papel de la evaluación de resultados en salud en la gestión sanitaria

Uno de los retos fundamentales de todos los sistemas sanitarios es conseguir los mejores resultados en salud garantizando la sostenibilidad, en un entorno de recursos limitados y con un aumento progresivo de los costes asociados a la introducción de nuevas tecnologías.

Los resultados en salud pueden usarse para comparar múltiples estrategias terapéuticas, identificar los mejores resultados clínicos, valorar el uso de recursos y evaluar los servicios de salud. Los estudios de RWE ayudan a contrastar los resultados clínicos de los tratamientos en vida real con los derivados de los ensayos clínicos. Por ello, cada vez más las agencias evaluadoras y la administración sanitaria demandan la verificación de los datos provenientes de ensayos aleatorizados para reducir la incertidumbre. De estos principios derivan los modelos de pago por resultados, fundamentados principalmente en la práctica clínica real (Garrison *et al.*, 2007).

Los estudios de RWE son un campo prometedor debido a sus cualidades inherentes y a la progresiva mejora de las metodologías de extracción y análisis de datos. Los diseños de los estudios observacionales cada vez ofrecen resultados de mayor calidad y validez, debido a un mejor control de los sesgos y de los factores de confusión y a que las bases de datos son más extensas y fiables (Garrison *et al.*, 2007).

La medición de resultados en salud de las innovaciones terapéuticas permite compatibilizar el acceso a la innovación con la sostenibilidad de los sistemas sanitarios, mediante fórmulas flexibles de financiación como los acuerdos de riesgo compartido. Estos acuerdos vinculan el precio y la financiación de ciertas innovaciones sanitarias, —especialmente medicamentos de alto coste—, a sus resultados financieros (acuerdos financieros, techos de gasto, etc) y/o a resultados en términos de salud (acuerdos basados en resultados de salud) y consiguen que el sistema sanitario sea más eficiente.

En España, con la introducción de los nuevos antivirales de acción directa en el año 2015, se realizaron diferentes acuerdos de techo de gasto entre las compañías farmacéuticas y el Ministerio de Sanidad, que permitieron incrementar el acceso a

los tratamientos a un mayor número de pacientes. Posteriormente, se ha realizado un análisis del coste-efectividad incremental de estos tratamientos siendo positivo debido a su mayor ganancia en AVAC (Turnes, Domínguez-Hernandez y Casado, 2017a).

La valoración de las nuevas terapias implica no solo la valoración de la efectividad en términos de salud, sino también los ahorros que se generan en el medio y largo plazo al evitar otros costes sanitarios y mejorar la productividad por la disminución del absentismo laboral debido a una mejor recuperación o menores recaídas. Es preciso tener en cuenta todas estas variables ya que la incorporación de la innovación, cuando responde a criterios de eficiencia, no es un gasto, sino una inversión, y que, más allá de los resultados clínicos proporciona ahorros sanitarios directos e indirectos. Es probable que se requiera un cambio cultural, que haga que los decisores se conciencien sobre la importancia de los datos en vida real en múltiples aspectos de la política y la asistencia sanitaria en España, tales como la planificación sanitaria, la provisión de servicios sanitarios o la autorización y el reembolso de nuevas tecnologías.

■ 10.1.5. Transformación digital y *big data* en la obtención de datos de resultados en salud

Estamos ante un proceso de transformación digital que permite a su vez analizar y procesar ingentes cantidades de datos clínicos (*big data*) que van a transformar de forma radical los procesos de gestión y de decisión y, por tanto, la eficiencia del sistema sanitario en todo el mundo. La aplicación de herramientas que permitan procesar estas enormes cantidades de datos y aportar inteligencia y predicción permitirá extraer información aplicable a la práctica clínica diaria, a la toma de decisiones en gestión sanitaria y a la gestión de la salud poblacional.

Algunos de los beneficios del *big data* son la obtención de información acerca de las características de la utilización de una intervención sanitaria en la práctica clínica; la estimación de su efectividad y seguridad a largo plazo; la obtención de resultados clínicos en poblaciones heterogéneas, que reflejen la distribución de los pacientes observados en la práctica clínica; el estudio de nuevas variables relevantes para los pacientes, tales como la calidad de vida, la satisfacción con los tratamientos o la funcionalidad; la posibilidad de obtener datos utilizables para cálculo de los costes de los servicios de salud, la evaluación económica o el estado de salud de la población (Roski, Bo-Linn y Andrews, 2014).

No obstante, todavía hay temas a resolver tales como la propiedad de estos datos, y las particularidades de las metodologías en la generación de estos (registros estructurados, semiestructurados o de campo libre, lenguaje natural, etc.). Otra de las limitaciones más comunes del RWE son las metodológicas debido a los problemas de validez interna con posibilidad de sesgos, baja calidad de algunos datos

individuales o bien la falta de estandarización y la fragmentación de la información en diferentes fuentes, que pueden dificultar enormemente la interrelación de los datos.

Es obvio que es una prioridad mejorar la calidad y el volumen de datos obtenidos de la práctica clínica real, y para ello, debe planificarse y priorizarse la mejora de las infraestructuras tecnológicas, la generalización del uso de la historia clínica electrónica y su interoperabilidad entre diferentes comunidades autónomas. Asimismo, debe fomentarse la calidad y la exhaustividad de los datos, no solamente con mejoras tecnológicas, sino incorporando sistemas de control de calidad y proporcionando incentivos a los profesionales sanitarios. Algunas comunidades autónomas como Aragón, Madrid, Navarra y Cataluña han puesto en marcha registros de pacientes y tratamientos para el análisis de datos y una mejor planificación, evaluación y gestión del medicamento, en diferentes áreas terapéuticas, entre ellas la hepatitis C. Veremos si estas iniciativas permiten obtener datos de calidad que puedan servir en la toma de decisiones para la mejora de la salud de la población.

En resumen, los resultados en salud del tratamiento del VHC, tanto en los ensayos clínicos como en los datos en vida real muestran tasas de RVP cercanas al 100 %, con una buena tolerancia al tratamiento y con un coste-efectividad aceptable, siendo, por tanto, una buena estrategia de inversión en salud ya que permite mejorar la calidad de vida de los pacientes, evitar complicaciones asociadas a la infección y disminuir la transmisión.

■ REFERENCIAS

- AFDHAL, N.H. y SERFATY, L. (2016). Cohort studies and registries in treatment of HCV the impact of real life effectiveness. *Gastroenterology*. Disponible en: doi:10.1053/j.gastro.2016.07.024
- ALVES, E. C., FILIPE, H. M., MARTINS, J. C., RODRIGUES, J., ANDREOZZI, V., VANDEWALLE, B. *et al.* (2016). *Impact of the universal drug coverage program to eradicate the burden of hepatitis C in Portugal*. 25th Conference of the APASL. Tokyo, 22 de febrero de 2016.
- ANSALDI, F., ORSI, A., STICCHI, L., BRUZZONE, B. e ICARDI, G. (2014). Hepatitis C virus in the new era: Perspectives in epidemiology, prevention, diagnostics and predictors of response to therapy. *World J Gastroenterol*, 20, pp. 9633-9652.
- ARDUINO, J. M., SHIBOLET, O., LITWIN, A. H., GREBELEY, J., ALTICE, F., NWANKWO, C. *et al.* (2016). C-EDGE CO-STAR: Favourable Impact of Elbasvir and Grazoprevir on Health-Related Quality of Life in Treatment-Naïve HCV Infected Persons Who Inject Drugs Receiving Opioid Agonist Therapy. EASL 2016 (April 14-17). Barcelona.
- ASSOCIATION FRANCAISE POUR L'ETUDE DU FOIE (AFEF). (2015). *Recommandations AFEF sur la prise en charge des hépatites virales C*. Disponible en: <http://www.afef.asso.fr/> (Consultado agosto 2018).
- AUSTRALIAN GOVERNMENT DEPARTMENT OF HEALTH. (2018). *Fourth National Hepatitis C Strategy 2014-2017*. Disponible en: <https://consultations.health.gov.au/> (Consultado agosto 2018).
- BACKUS, L. I., BELPERIO, P. S., SHAHOUMIAN, T. A., LOOMIS, T. P. y MOLE, L. A. (2016). Real-world effectiveness and predictors of sustained virological response with all-oral therapy in 21,242 hepatitis C genotype-1 patients. *Antivir Ther*, 44, pp. 400-410. Disponible en: <https://doi.org/10.3851/IMP3117>

- BAYLIS, S. A. y HEATH, A. B. (2011). Collaborative Study Group. World Health Organization collaborative study to calibrate the 3rd International Standard for Hepatitis C virus RNA nucleic acid amplification technology (NAT)-based assays. *Vox Sang*, 100, pp. 409-417.
- BENÍTEZ-GUTIÉRREZ, L., BARREIRO, P., LABARGA, P., DE MENDOZA, C., FERNÁNDEZ-MONTERO, J. V., ARIAS, A. *et al.* (2016). Prevention and management of treatment failure to new oral hepatitis C drugs. *Expert Opin Pharmacother*, 17, pp. 1215-1223.
- BERENQUER, J., GIL-MARTÍN, Á., JARRÍN, I., MORENO, A., DOMÍNGUEZ, L., MONTES, M. *et al.* (2018). All-oral direct-acting antiviral therapy against hepatitis C virus (HCV) in human immunodeficiency virus/HCV-coinfected subjects in real-world practice: Madrid coinfection registry findings. *Hepatology*, 68(1), pp. 32-47. Disponible en: doi: 10.1002/hep.29814
- BERG, T., NAUMANN, U., STÖRER, A., SICK, C., TEUBER, G., SCHIFFELHOLZ, W. *et al.* (2018). First real-world data on safety and effectiveness of glecaprevir/pibrentasvir for the treatment of patients with chronic hepatitis C virus infection: data from the German Hepatitis C-Registry. EASL 2018.
- BUTI, M., DOMÍNGUEZ-HERNÁNDEZ, R., CASADO, M. A., SABATER, E. y ESTEBAN, R. (2018). El cribado y tratamiento del virus de la hepatitis C en población general española entre 20 y 79 años de edad es coste-efectivo. *Gastroenterol Hepatol*, 41(SE1), p. 96.
- BUTI, M., DOMÍNGUEZ-HERNÁNDEZ, R., OYAGÜEZ, I., CASADO, M. A. y ESTEBAN, R. (2017). Cost-effectiveness analysis of ledipasvir/sofosbuvir in patients with chronic hepatitis C: Treatment of patients with absence or mild fibrosis compared to patients with advanced fibrosis. *J Viral Hepat*, 24(9), pp. 750-758.
- CHHATWALL, J., SAMUR, S., LI, K., HE, T., LLERENA, S., COBO, C. *et al.* (2018). Improved health outcomes from hepatitis C treatment scale-up in Spain's prisons: A cost-effectiveness study. *Journal of Hepatology*, 68, pp. S105-S364.
- CHRISTENSEN, C. M. y ROSENBLUM, R. S. (1995). Explaining the attacker's advantage: Technological paradigms, organizational dynamics, and the value network. *Research Policy*, 24, pp. 233-257. Disponible en: DOI: 10.1016/0048-7333(93)00764-K
- CUADRADO, A., PERELLO, C., LLERENA, L., ESCUDERO, M. D., GÓMEZ, M., ESTÉBANEZ, A. *et al.* (2018). *Diseño y coste efectividad de una política de eliminación del VHC basada en un estudio epidemiológico actualizado (Cohorte Ethon)*. Madrid: Congreso EAAH 2018.
- D'AMBROSIO, R., PASULO, L., PUOTI, M., SHIIVINI, P., VIGANÒ, P., VINCI, M. *et al.* (2018). Real-life effectiveness and safety of Glecaprevir/Pibrentasvir among 723 Italian patients with chronic hepatitis c: the Navigator-II study. EASL 2018, GS-013.
- DEUFFIC-BURBAN, S., HUNEAU, A., VERLEENE, A., BROUARD, C., PILLONEL, J., LE STRAT, Y. *et al.* (2018). Assessing the cost-effectiveness of hepatitis C screening strategies in France. *Journal of Hepatology*. Disponible en: <http://dx.doi.org/10.1016/j.jhep.2018.05.027> (acceso 1 julio 2018).
- ESTEBAN, R., PINEDA, J. A., CALLEJA, J. L., CASADO, M., RODRÍGUEZ, M., TURNES, J. *et al.* (2018). Efficacy of Sofosbuvir and Velpatasvir, With and Without Ribavirin, in Patients With HCV genotype 3 Infection and Cirrhosis. *Gastroenterology*, 26, pii: S0016-5085(18)34684-5.
- EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER (EASL). (2018). Recommendations on Treatment of Hepatitis C 2018. *J of Hepatol* 2018 (in press). Disponible en; <https://doi.org/10.1016/j.jhep.2018.03.026>
- FERNÁNDEZ RODRIGUEZ, C. M. y GUTIERREZ GARCIA, M. L. (2014). Impact of anti-viral therapy on the natural history of hepatitis C virus. *Gastroenterol Hepatol*, 37, pp. 583-592.
- GARRISON, L. P., NEUMANN, P. J., ERICKSON, P., MARSHALL, D. y MULLINS, C. D. (2007). Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force Report. *Value Health*, 10, pp. 326-335.
- HERNÁNDEZ-CONDE, M., FERNÁNDEZ, I., PERELLO, C. *et al.* (2018). *Eficacia y seguridad de Elbasvir/grazoprevir en práctica clínica real: Resultados de la cohorte Hepa-C*. Congreso de la Asociación Española para el Estudio del Hígado 2018, poster: 102.

- HÉZODE, C. (2018). Treatment of hepatitis C: Results in real life. *Liver International*, 38(Suppl. 1), pp. 21–27.
- LAUFFENBURGER, J. C., MAYER, C. L., HAWKE, R. L., BROUWER, K. L., FRIED, M. W. y FARLEY, J. F. (2014). Medication use and medical comorbidity in patients with chronic hepatitis C from a US commercial claims data-base: High utilization of drugs with interaction potential. *Eur J Gastroenterol Hepatol*, 26, pp. 1073-1082.
- MANGIA, A., PIAZZOLLA, V., LOSAPPIO, R., MAZZOLA, M., VISSAGI, E., BACCA, D. *et al.* (2018). High SVR rates in patients with and without cirrhosis treated in real life with Sofosbuvir/Velpatasvir (SOF/VEL) combination for 12 weeks without Ribavirin (RBV). EASL 2018; THU-323. Disponible en: DOI: [https://doi.org/10.1016/S0168-8278\(18\)30763-3](https://doi.org/10.1016/S0168-8278(18)30763-3)
- MARCO, A., DOMÍNGUEZ-HERNÁNDEZ, R. y CASADO, M. A. (2018). *Análisis coste-efectividad del tratamiento de la hepatitis C crónica en población reclusa*. Poster EI15. XII Congreso Nacional y XX Jornadas de la Sociedad Española de Sanidad Penitenciaria. Cartagena, mayo 2018.
- MARGUSINO-FRAMIÑÁN, L., CASTRO-IGLESIAS, A., MENA-DE-CEA, A., CID-SILVA, P., RODRÍGUEZ-OSORIO, I., PERNAS-SOUTO, B. *et al.* (2018). Cost/Effectiveness Analysis of Direct-Acting Antivirals for Hepatitis C Virus in Treatment-Naive Genotype 1 without Cirrhosis Patients in Spanish Population. *Value in Health*, 21(1), p. S85.
- MINISTERIO DE SANIDAD, CONSUMO Y BIENESTAR SOCIAL. (2018) *Plan Estratégico para el Abordaje de la Hepatitis C en el SNS*. Madrid. Disponible en: <https://www.msssi.gob.es/en/ciudadanos/enfLesiones/enfTransmisibles/hepatitisC/home.htm> (consultado agosto 2018).
- MYERS, R. P., SHAH, H., BURAK, K. W., COOPER, C., y FELD, J. J. (2015). An update on the management of chronic hepatitis C: 2015 consensus guidelines from the Canadian Association for the Study of the Liver. *Can J Gastroenterol Hepatol*, 29, pp. 19-34.
- NEGRO, F. (2014). Epidemiology of hepatitis C in Europe. *Dig Liver Dis*, 46(Suppl 5), pp. S158-1564.
- OLIVA-MORENO, J., PEÑA-LONGBARDO, L. M., ALONSO, S., FERNÁNDEZ-BOLAÑOS, A., GUTIÉRREZ, M. L., HIDALGO-VEGA, Á. *et al.* (2015). Labour productivity losses caused by premature death associated with hepatitis C in Spain. *Eur J Gastroenterol Hepatol*, 27, pp. 631-710.
- OMATA, M., KANDA, T., WEI, L., YU, M. L., CHUANG, W. L., IBRAHIM, A. *et al.* (2016). APASL consensus statements and recommendations for hepatitis C prevention, epidemiology, and laboratory testing. *Hepatol Int*, 10, pp. 681-701.
- PINSKY, B., MITCHELL, D., SANCHEZ-GONZALEZ, Y., JOHNSON, S. y SAAB, S. (2018). Early versus delayed hepatitis C treatment provides increased health benefits at lower costs: a pan-genotypic cost-effectiveness analysis set in Scotland. The International Liver Congress, abstract PS-058. *Journal of Hepatology*, 68, p. S32. París.
- RAZAVI, H., WAKED, I., SARRAZIN, C., MYERS, R. P., IDILMAN, R., CALINAS, F. *et al.* (2014). The present and future disease burden of hepatitis virus (HCV) infection with today's treatment paradigm. *J Viral Hepat*, 21(Suppl 1), pp. 34-59.
- REID, M., PRICE, J. C. y TIEN, P. C. (2017). Hepatitis C virus infection in the older patient. *Infect Dis Clin North Am*, 31, pp. 827-838.
- ROSKI, J., BO-LINN, G. W. y ANDREWS, T. A. (2014). Creating value in health care through big data: opportunities and policy implications. *Health Affairs*, 33, pp. 1115-1122.
- RUGGERI, M., CORETTI, S., ROMANO, F., KONDILI, L. A., VELLA, S. y CICCETTI, A. (2018). Economic Evaluation of the Hepatitis C Virus Treatment Extension to Early-Stage Fibrosis Patients: Evidence from the PITER Real-World Cohort. *Value Health*, 21(7), pp. 783-791. Disponible en: doi: 10.1016/j.jval.2017.10.021
- SCHEINER, B., SCHWABL, P., STEINER, S., BUCSICS, T., CHROMY, D., AICHELBERG, M. C. *et al.* (2016). Interferon-free regimens improve health-related quality of life and fatigue in HIV/HCV-coinfected patients with advanced liver disease. *Medicine (Baltimore)*, 95(27), p. e4061.

- THE ASSOCIATION OF BRITISH PHARMACEUTICAL INDUSTRY. (2011). *The vision for real world data. Harnessing the opportunities in the UK. Demonstrating value with real world data*. White paper. London: ABPI.
- TURNES, J., DOMÍNGUEZ-HERNANDEZ, R. y CASADO, M. A. (2017a). Cost-effectiveness analysis of two treatment strategies for chronic hepatitis C before and after access to direct-acting antivirals in Spain. *Gastroenterol Hepatol*, 40(7), pp. 433-446.
- (2017b). *Evaluación del impacto de los agentes antivirales de acción directa tras dos años del plan estratégico para el abordaje de la hepatitis C en España: Resultados preliminares del análisis coste-utilidad*. XXXVII Jornadas de economía de la salud. Poster nº 48. Barcelona, septiembre de 2017.
- U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES FDA CENTER FOR DRUG EVALUATION AND RESEARCH, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES FDA CENTER FOR BIOLOGICS EVALUATION AND RESEARCH, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES FDA CENTER FOR DEVICES AND RADIOLOGICAL HEALTH. (2006). Guidance for industry: patient-reported outcome measures: use in medical product development to support labelling claims: draft guidance. *Health Qual. Life Outcomes*, 4, 79.
- WHO. (s. f.). *Global Health Sector Strategy on viral hepatitis 2016*. Disponible en: http://apps.who.int/gb/ebwha/pdf_files/WHA69/A69_32-en.pdf (Consultado agosto 2018).
- YOUNOSI, Z., BROWN, A., BUTI, M., FAGIOLI, S., MAUSS, S., ROSENBERG, W. *et al.* (2016a). Impact of eradicating hepatitis C virus on the work productivity of chronic hepatitis C (CH-C) patients: an economic model from five European countries. *J. Viral Hepat*, 23(3), pp. 217–226.
- YOUNOSI, Z. M., STEPANOVA, M., SULKOWSKI, M., FOSTER, G. R., REAU, N., MANGIA, A. *et al.* (2016b). Ribavirin-Free Regimen With Sofosbuvir and Velpatasvir Is Associated With High Efficacy and Improvement of Patient-Reported Outcomes in Patients With Genotypes 2 and 3 Chronic Hepatitis C: Results From Astral-2 and -3 Clinical Trials. *Clin. Infect. Dis*, 63(8), pp. 1042-1048.



11

RESULTADOS EN SALUD EN REUMATOLOGÍA

EMILIO MONTE BOQUET, KARLA ARÉVALO RUALES,
JORGE JUAN FRAGÍO GIL, ROXANA GONZÁLEZ MAZARÍO,
MARTA DE LA RUBIA NAVARRO, CRISTÓBAL PAVEZ PERALES

■ 11.1. RESULTADOS COMUNICADOS POR LOS CLÍNICOS (CLINICIAN REPORTED OUTCOMES)

El conocimiento cada vez más amplio de la fisiopatología de las distintas enfermedades reumáticas, como la artritis reumatoide (AR), la espondilitis anquilosante (EA) o la artritis psoriásica (APs), ha permitido el desarrollo de nuevas opciones terapéuticas, que permiten mejorar la calidad de vida y la supervivencia de los pacientes, así como reducir la morbilidad (Ugarte-Gil *et al.*, 2013).

El conjunto de estos tratamientos se conoce como “terapia biológica” y se dirige a dianas celulares concretas, como pueden ser: marcadores solubles (citoquinas, quimioquinas, complemento), líneas celulares (célula T o B), moléculas que permiten la supervivencia de dichas células (BLySS, BAFF), receptores de superficie (CTLA4, BAFF-R) y vías de señalización intracelular (Tirosinkinasa de Bruton) (Catay y Soriano, 2014).

El primer fármaco biológico aprobado fue el etanercept (proteína de fusión que impide la acción del TNF-alfa), en el año 1998. Desde ese momento, no ha cesado la investigación y el desarrollo en torno a estas enfermedades y sus potenciales dianas terapéuticas. En la actualidad existe un amplio abanico de fármacos disponibles con diferentes estructuras y mecanismos de acción (Catay y Soriano, 2014).

Este capítulo abordará datos de efectividad clínica, tolerabilidad, seguridad, cumplimiento y grado de persistencia de los fármacos biológicos utilizados en pacientes diagnosticados de AR, EsA y APs.

■ 11.1.1. Efectividad clínica

■ 11.1.1.1. Artritis reumatoide

La AR es una enfermedad crónica, inflamatoria, inmunomediada, cuyo órgano diana es la membrana sinovial que envuelve las articulaciones sinoviales. Se postula que en su patogenia está implicada una disregulación del sistema inmunitario, que da lugar a inflamación crónica y persistente, causante de la destrucción

articular. Sin embargo, la articulación no es el único órgano afecto, pudiendo afectarse también otros niveles (manifestaciones extraarticulares) como piel, pulmón, riñón, hueso, corazón, ojos, etc.

El tratamiento tiene como objetivo principal la remisión de la enfermedad o, en su defecto, alcanzar la mínima actividad posible, en un periodo de seis meses. El tratamiento de esta enfermedad está basado en la utilización de los FAME (fármacos antirreumáticos modificadores de la enfermedad) sintéticos o tradicionales, destacando el metotrexato (MTX), asociado a corticoterapia. En caso de fracaso terapéutico, se puede asociar a otros FAME, como, por ejemplo, leflunomida, hidroxiquina, sulfasalazina o bien iniciar terapia biológica, asociada o no a MTX (Sociedad Española de Reumatología, 2018; Aletaha y Smolen, 2018).

En pacientes *naive* a MTX con elevada actividad de la enfermedad, las respuestas ACR70 para el tratamiento con biológicos más MTX se encuentran entre 35-40 %, considerándose ACR70: mejoría igual o superior al 70 % en el recuento de articulaciones dolorosas y en el recuento de articulaciones tumefactas, así como en al menos tres de los siguientes parámetros: VSG o PCR, valoración global del médico, valoración global del paciente, valoración del paciente del grado de dolor, discapacidad física). En pacientes con insuficiente respuesta a MTX la respuesta está en torno al 20 % y si además no responden adecuadamente a anti-TNF, el porcentaje se sitúa en torno al 10-15 %. Con esta información, se puede afirmar inicialmente que los fármacos biológicos combinados con MTX parecen los más efectivos en pacientes en los que no se ha utilizado MTX previamente, pero no se deben utilizar antes que este último, ya que los datos mencionados anteriormente incluyen pacientes que podían haber respondido a MTX en monoterapia. Este es el motivo por el cual el tratamiento biológico no se recomienda como opción inicial. Por otra parte, en caso de no alcanzar la remisión o la mínima actividad de la enfermedad, es posible cambiarlo a otro que sí tenga beneficio (Aletaha y Smolen, 2018).

En el momento actual, hay múltiples fármacos biológicos aprobados para el tratamiento de la AR. Todos ellos tienen una eficacia similar de forma indirecta y ésta depende realmente de las características de la enfermedad, más que del fármaco empleado. Hoy, no es posible recomendar un fármaco por encima de otro en ningún caso, dada la falta de estudios comparativos entre los mismos. La elección por tanto está basada en la situación del enfermo, la forma de administración del fármaco, la experiencia clínica, el fallo o no a otros biológicos, enfermedades concomitantes que presente el enfermo, reacciones adversas y preferencias del propio paciente (Sociedad Española de Reumatología, 2018).

Estas terapias previenen el daño articular irreversible hasta en el 90 % de los pacientes, especialmente si el inicio del tratamiento es precoz. La historia natural de la enfermedad se caracteriza por la destrucción articular, que está directamente asociada a la progresión de la enfermedad, asociación que desaparece con el trata-

miento biológico. Esto conlleva a una prevención de la discapacidad funcional y una mejora en la calidad de vida de los enfermos. La estrategia terapéutica dirigida a la diana reduce la actividad de la enfermedad hasta en un 50 % en los tres primeros meses, permitiendo alcanzar la remisión dentro de los 6 meses, con la consecuente mejoría en la calidad de vida y la prevención de la discapacidad funcional (Aletaha y Smolen, 2018).

■ 11.1.1.2. *Espondilitis anquilosante*

La EA es una enfermedad inflamatoria articular, crónica, perteneciente al grupo de las espondiloartritis, cuya característica principal es la afectación de las articulaciones sacroilíacas y del esqueleto axial, aunque puede afectar también a otras articulaciones periféricas, fundamentalmente de miembros inferiores. Además, puede acompañarse de otras manifestaciones fuera del sistema musculoesquelético, como uveítis, fibrosis pulmonar, psoriasis o enfermedad inflamatoria intestinal (Sociedad Española de Reumatología, 2018).

La etiología de la enfermedad no está del todo esclarecida, pero actualmente la investigación se sitúa especialmente en torno al eje IL-23/IL-17.

En este contexto, el tratamiento inicial de los pacientes con EA, son los antiinflamatorios no esteroideos (AINE), incluso si fuera necesario a dosis elevadas, de forma mantenida. En caso de que se objetive buena respuesta y control de la sintomatología, podrán ser utilizados a demanda, para minimizar los efectos secundarios (Sociedad Española de Reumatología, 2018). El objetivo principal es, como en el caso de la AR, la optimización de la calidad de vida, reducción del daño estructural y de la discapacidad funcional.

La terapia con AINE fracasa en un 20-30 % de los pacientes o bien debe ser discontinuada por efectos secundarios (gastrointestinales, cardiovasculares, renales). En este punto deben plantearse otras opciones terapéuticas como la terapia biológica, ya que los FAME convencionales como el MTX o la sulfasalazina no se han mostrado eficaces para el tratamiento de la afectación axial de la EA (De La Forest Divonne, Gottenberg y Salliot, 2017).

Desde la publicación de un metaanálisis por parte de la Cochrane en el año 2015, que no encontró diferencias en cuanto a eficacia entre los distintos anti-TNF (infliximab, adalimumab, golimumab y etanercept), han surgido diferentes estudios con otros anti-TNF (certolizumab) y con otros fármacos no pertenecientes a la familia anti-TNF como secukinumab (anti-IL17), ustekinumab (anti-IL23), apremilast (inhibidor de la fosfodiesterasa 4) y tofacitinib (inhibidor de la JAK kinasa).

Hasta el momento, únicamente existen comparaciones indirectas entre los distintos biológicos, por lo que ante fallo de AINE, cualquiera de ellos puede ser utilizado, atendiendo a su perfil de seguridad y las características de cada enfermo (De

La Forest Divonne, Gottenberg y Salliot, 2017). De hecho, en comparación con placebo se ha visto que tanto los anti-TNF como el secukinumab alcanzan la respuesta ASAS20 (definida como la mejoría de al menos el 20 % y mejoría absoluta de al menos 10 unidades en una escala del 0-100 en tres de los siguientes: percepción global del paciente, dolor, función [BASFI] e inflamación [dos últimas preguntas del BASDAI]) con un *Ods* ratio de 4,31 (95 % CI 3,57-5,20) y 3,47 (95 % CI 2,23-5,35) respectivamente.

■ 11.1.1.3. Artritis psoriásica

La APs es una enfermedad inflamatoria articular, crónica, mediada por el sistema inmune y también perteneciente al grupo de las espondiloartritis (Sociedad Española de Reumatología, 2018). En esta entidad se afectan fundamentalmente el sistema musculoesquelético y la piel, pudiendo producir artritis erosiva, dactilitis, entesitis, lesiones cutáneas eritemato-descamativas e hiperqueratósicas, que afectan sobre todo a las áreas extensoras como codos o rodillas. Los enfermos pueden desarrollar múltiples comorbilidades, fundamentalmente de tipo cardiovascular. Además, la enfermedad no diagnosticada y no tratada adecuadamente, puede producir importante destrucción articular, lo que conlleva discapacidad funcional, deterioro de la calidad de vida y aumento de los costes sociosanitarios (Sociedad Española de Reumatología, 2018).

La fisiopatología de la APs comparte muchas similitudes con la EA, por lo que las moléculas que constituyen dianas terapéuticas son, en términos generales, las mismas. Junto al anti-TNF ha cobrado gran importancia el eje IL-23/IL-17, al igual que en la EA, por lo que actualmente existe una gran variedad de fármacos disponibles y también nuevas moléculas en desarrollo, que aumentan el abanico de opciones terapéuticas (Caso *et al.*, 2015).

Al igual que en la espondilitis, hoy en día no es posible afirmar categóricamente que un fármaco sea más efectivo o eficaz que otro, ya que no existe evidencia de que esto sea así. El uso de los fármacos biológicos, al igual que en la AR o en la EA, ha abierto un nuevo horizonte a los pacientes, reduciendo la comorbilidad, discapacidad y mejorando la calidad de vida (Caso *et al.*, 2015).

■ 11.1.2. Tolerabilidad y seguridad: efectos adversos producidos

Como ya se ha comentado, la reumatología ha experimentado un gran cambio en cuanto a opciones de tratamiento en las últimas décadas. Actualmente se dispone de corticoides, FAME sintéticos, fármacos biológicos y recientemente pequeñas moléculas de terapia dirigida. Este cambio ha generado un gran impacto en la calidad de vida de los pacientes. Los corticoides tienen todavía un rol importante en el tratamiento de las enfermedades reumáticas, aunque posiblemente en menor dosis que en el pasado. En el tratamiento de AR, APs y en EA las dosis de corti-

coides suelen ser medias o bajas (van der Goes, Jacobs y Bijlsma, 2014). Existen varios efectos secundarios asociados a su uso, entre los más importantes se destacan la osteoporosis inducida por esteroides, osteonecrosis, aumento de peso, supresión del eje hipotálamo-hipofisis-adrenal, aumento de la tasa de infecciones, aparición de cataratas, efectos neuropsiquiátricos, entre otros. A pesar de sus efectos secundarios, la tolerabilidad de estos fármacos es buena y en varios estudios se ha mostrado que sus efectos secundarios pueden reducirse con dosis menores a 10 mg/día (Da Silva *et al.*, 2006).

En cuanto a los FAME sintéticos, sus efectos adversos más comunes son la intolerancia gastrointestinal y reacciones cutáneas (Cummins *et al.*, 2015). De este grupo se puede destacar el MTX, que es conocido además de sus efectos adversos gastrointestinales, la elevación de enzimas hepáticas, por lo que requiere monitorización de forma periódica. Asimismo, puede producir toxicidad medular y afectación pulmonar.

En relación a los anti-TNF, infliximab puede causar reacciones infusionales, y los de administración subcutánea pueden causar reacciones locales. Es conocido que estos fármacos puedan aumentar el riesgo de infecciones o reactivar infecciones latentes como tuberculosis o hepatitis, por lo que se recomienda tomar medidas profilácticas en los pacientes en riesgo.

Existen otros fármacos biológicos no anti-TNF que tienen indicación en ciertas patologías. Por ejemplo, en AR, abatacept es bastante bien tolerado, siendo sus efectos secundarios más comunes la cefalea y las náuseas, aunque se ha visto también un cierto incremento de las infecciones de tracto respiratorio superior (Blair y Deeks, 2017). Otro fármaco usado en AR es el rituximab, que también puede presentar reacciones infusionales y mayor riesgo de infecciones. También tiene indicación tocilizumab que suele presentar efectos adversos a nivel analítico con aumento de colesterol y triglicéridos y de enzimas hepáticas. También son comunes la neutropenia y leucopenia con su uso (Mora Arias, Ortega y Bourlon, 2010).

Por otra parte, para la APs y la EA también tiene indicación secukinumab, cuyos efectos secundarios más importantes son el aumento en la tasa de infecciones de tracto respiratorio superior y la cefalea. También se han reportado casos de infección fúngica candidiásica esofágica, oral y genital (Pavelka *et al.*, 2018). Ustekinumab tiene indicación para APs, su tolerabilidad es buena y los efectos secundarios son similares a los del resto de biológicos (aumento del riesgo de infecciones, cefalea, artralgias). Otros efectos asociados a su uso que han sido reportados han sido eventos cardiovasculares y malignidad (Krulig E y Gordon, 2010).

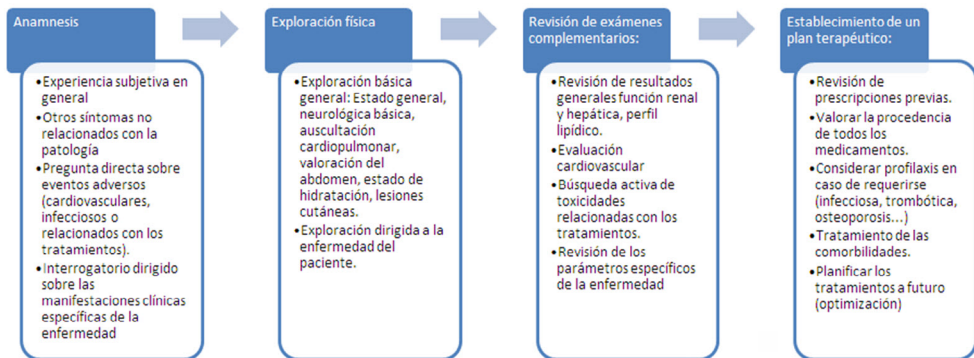
■ 11.1.3. Síntomas y signos clínicos

Desde el punto de vista del clínico, en cada acto médico (ya sea en consultas, hospital de día u hospitalización) se debe hacer siempre una valoración general

que comience por la experiencia subjetiva del paciente, dirigiendo progresivamente la anamnesis hacia las manifestaciones clínicas específicas del proceso que adolezca. Del mismo modo, se seguiría el mismo esquema de actuación a la hora de realizar una exploración física, la revisión de exámenes complementarios y el establecimiento de un plan terapéutico (Figura 1). El manejo desde lo general a lo particular es siempre recomendable para realizar una valoración multidimensionalidad de la enfermedad.

Figura 1.

VALORACIÓN MULTIDIMENSIONAL DE LA ENFERMEDAD Y ABORDAJE DEL PACIENTE



Fuente: Elaboración propia.

En muchas ocasiones, la valoración de una enfermedad inflamatoria crónica incluye el empleo de índices de actividad clínicos, analíticos, de imagen o más frecuentemente y extendido, combinados. El estandarizar la medida de actividad de la enfermedad es extraordinariamente útil y necesario por diferentes motivos.

- A nivel individual: permite un adecuado seguimiento del paciente, pudiendo medir el efecto que están teniendo las diferentes intervenciones sobre la enfermedad.
- A nivel del equipo asistencial: permite la comunicación en un “mismo lenguaje” que prima lo objetivo sobre lo subjetivo, sabiendo siempre que los índices van a tener un componente subjetivo.
- A nivel gestor e investigador: permite valorar a nivel colectivo el trabajo del equipo. Además, facilita la explotación de la información.

Más allá de la medición de la actividad de la enfermedad, en la actualidad existen numerosos índices que permiten evaluar múltiples facetas de cada entidad, a continuación, se exponen diferentes medidas extendidas en AR, APs, espondilitis anquilosante (EA) o psoriasis cutánea:

- Artritis reumatoide: DAS28 (*Disease Activity Score* que evalúa, entre otros, dolor y tumefacción en 28 articulaciones), CDAI (*Clinical Disease Activity Index*), SDAI (*Simplified Disease Activity Index*), para medir actividad clínica, HAQ (*Health Assessment Questionnaire*) como medida de discapacidad.
- Artropatía psoriásica: DAPSA (*Disease Activity for Psoriatic Arthritis*), MDA (*Minimal Disease Activity*).
- Espondilitis anquilosante: BASDAI (*Bath Ankylosing Spondylitis Disease Activity Index*), BASFI (*Bath Ankylosing Spondylitis Functional Index*), ASDAS (*Ankylosing Spondylitis Disease Activity Score*).
- Psoriasis cutánea: PASI (*Psoriasis Area and Severity Index*).

Además, también se dispone de herramientas para evaluar otros aspectos específicos como el cansancio o el riesgo cardiovascular. Estos y otros índices se comentarán más adelante en el apartado de resultados reportados por el paciente.

■ 11.1.4. Adherencia terapéutica

En la actualidad la literatura presenta diversos términos que se refieren al mismo concepto: observancia, adherencia o cumplimiento terapéutico, son algunas de ellos. También existen múltiples definiciones. De manera simple podemos definirlo como la concordancia existente entre las instrucciones que el facultativo entrega y las conductas que sigue el paciente. Una de las definiciones más aceptadas es la que estableció la Organización Mundial de la Salud (OMS) en 2003, a partir de la definición previa de Haynes y Sackett, en la cual la adherencia o cumplimiento terapéutico se refiere a “el grado en el que la conducta del paciente, en relación con la toma de medicación, el seguimiento de una dieta o la modificación de los hábitos se corresponde con las recomendaciones acordadas con el profesional sanitario” (Dillaa *et al.*, 2009).

En la mayoría de los estudios el buen cumplimiento terapéutico se define como tomar o aplicarse el 80 % o más de la medicación durante la duración del estudio (Lavielle *et al.*, 2018).

En cuanto a los fármacos biológicos, como ya se ha comentado anteriormente, son ampliamente utilizados como tratamiento de la AR, PsA y EA, con buenos resultados. No obstante, los beneficios demostrados en los estudios clínicos dependen de manera crucial de que los pacientes adhieran al tratamiento en la práctica clínica real (Bhoi *et al.*, 2017). En efecto, la falta de cumplimiento terapéutico en estas

enfermedades impacta negativamente en los resultados en salud (progresión de la enfermedad, necesidad de tratamientos más agresivos, etc.) y se asocia al aumento del coste sanitario (López-González *et al.*, 2015).

En el caso de los fármacos biológicos, en general, en los estudios existentes, la adherencia se ha medido a través de métodos indirectos, siendo los principales la tasa de posesión de la medicación (MPR) y el porcentaje de días cubiertos por receta (PDC) (Murage *et al.*, 2018). Sin embargo, en la práctica clínica diaria se pueden usar herramientas más simples como, por ejemplo, la evaluación por el profesional sanitario (juicio clínico, evolución de la enfermedad, etc.) y el autoreporte de los pacientes (por ejemplo, mediante el cuestionario de Morisky-Green) (Áreas Del Águila y Monte Boquet, 2017).

En cuanto a la AR, la adherencia varía sustancialmente entre estudios, estimándose que la adherencia a biológicos está entre el 11 % y el 88 % (López-González *et al.*, 2015). En lo que refiere a la PsA y EA, la literatura es escasa. Para EA el estudio de Hromadkova *et al.*, mostró una adherencia del 38.3 % para fármacos biológicos en general. En el caso de PsA el estudio de Ellis *et al.*, mostró una adherencia del 76 % para golimumab (Hromadkova, Soukup y Vlcek, 2015; Ellis, Bolge y Rice, 2014).

Se ha estudiado cuáles son los factores predisponentes al incumplimiento terapéutico. Dentro de estos existen los no modificables, que sirven para identificar a aquellos pacientes con más riesgo de no adherir al tratamiento, y así poder supervisarlos. Por otro lado, están los factores potencialmente modificables, que además de identificar pacientes, puede intentar controlarse para evitar el incumplimiento. Dentro de estos últimos se encuentran: la frecuencia de dosificación (algunos estudios han demostrar mejor adherencia a golimumab versus otros anti-TNF por su posología menos frecuente), la vía de administración (probablemente más adherencia de la vía subcutánea versus la intravenosa), el conocimiento del paciente sobre la enfermedad, la ansiedad y depresión, las creencias sobre la medicación y el pertenecer a asociaciones de pacientes (Áreas Del Águila y Monte Boquet, 2017).

Por otro lado, la no adherencia se puede clasificar en no intencionada e intencionada. La primera tiene que ver con olvidos, poca destreza en el manejo del fármaco o no disponibilidad de éste. La segunda es una decisión propia del paciente y se basa en las creencias del paciente sobre la enfermedad y el tratamiento; el paciente hace un análisis riesgo-beneficio, basado en las preocupaciones en relación con el tratamiento (dependencia y efectos adversos del fármaco) y los beneficios percibidos (alivio sintomático, preservación de la funcionalidad y cura de la enfermedad) (Van den Bemt, Zwikker y Van den Ende, 2012).

Los factores psicosociales, como las preocupaciones acerca del tratamiento, la falta de percepción de mejoría, la poca creencia en la necesidad del tratamiento, la comorbilidad depresiva o una deficiente relación entre el profesional sanitario y el paciente, son los factores para los que se ha encontrado mayor evidencia de impacto negativo sobre el cumplimiento terapéutico. La evidencia es menor sobre la

asociación entre los factores relacionados con el tratamiento y la no adherencia. No se ha evidenciado una relación con los factores demográficos y clínicos (Áreas Del Águila y Monte Boquet, 2017).

Diversos estudios han propuesto intervenciones para mejorar el cumplimiento terapéutico, las cuales se pueden clasificar en: cognitivo-conductuales, conductuales, educacionales y multifactoriales (Lavielle *et al.*, 2018):

- **Cognitivo-conductuales:** intentan producir en el paciente patrones de pensamiento que optimicen la adherencia. Proporcionan al paciente herramientas para identificar los factores personales de no adherencia y hallar las soluciones que puedan sobrepasar aquellos factores y, por lo tanto, mejorar el cumplimiento terapéutico.
- **Conductuales:** buscan mejorar la adherencia mediante acciones concretas sobre el paciente. Por ejemplo, simplificar la posología, sistemas de recordatorio o sistemas de recompensa.
- **Educacionales:** se basan en la entrega de información escrita (folletos, revista, etc.) u oral (entrevistas, sesiones grupales, etc) al paciente, con objeto de mejorar su conocimiento sobre la enfermedad y su tratamiento.
- **Multifactoriales:** son combinaciones de algunas de las anteriores.

Menos de la mitad de los estudios que han evaluado estrategias para mejorar adherencia en enfermedades reumáticas han mostrado algún impacto. Las intervenciones que incluyen educación han evidenciado mejorar la adherencia. Las intervenciones educacionales deben basarse en la literatura médica con evidencia y recomendando al paciente las fuentes de las que puede informarse sobre su patología. Se debe educar al paciente sobre los beneficios del tratamiento farmacológico, ya que los pacientes que consideran que el beneficio del tratamiento es alto, están mejor preparados para afrontar los riesgos de este. Estas intervenciones tienen como requisito mínimo para su funcionamiento la implicación activa y constante del profesional de salud y del paciente (Áreas Del Águila y Monte Boquet, 2017). Además, implicar al paciente en las decisiones que se toman con relación a su tratamiento es beneficioso, pues se ha demostrado que aumenta el nivel de cumplimiento terapéutico en AR y PsA (Lofland *et al.*, 2017).

Por último, un metaanálisis de Demonceau, J. *et al.*, ha reportado beneficio de retroalimentar a los pacientes sobre su grado de adherencia al tratamiento (medido a través de dispositivos electrónicos), asociando posteriormente intervenciones cognitivo-conductuales (Demonceau *et al.*, 2013).

■ 11.1.5. Grado de persistencia

Se define el grado de persistencia o supervivencia de una molécula como el tiempo que pasa desde la prescripción de un producto hasta su discontinuación. Las

causas que pueden motivar dicho evento pueden ser diversas, pero principalmente van a estar relacionadas con la seguridad y la durabilidad de la eficacia de una terapia a lo largo del tiempo.

Este concepto tiene especial relevancia en el ámbito del empleo de terapias de alto impacto económico (anticuerpos monoclonales, proteínas de fusión u otras terapias dirigidas de síntesis química). Este tipo de fármacos están supeditados a una vigilancia especial desde todos los puntos de vista y en esta línea, las sociedades científicas han elaborado registros multicéntricos a fin de conocer datos “de vida real” del empleo de estos fármacos. Las principales variables que se ha visto que motivan la suspensión de un tratamiento son:

- Falta de eficacia: en el supuesto de que el tratamiento nunca llegue a generar una respuesta clínica (conocido como fallo primario).
- Pérdida de eficacia: tras una buena respuesta clínica inicial, ésta se pierde con el tiempo (conocido como fallo secundario).
- Efectos adversos: infecciones, reacciones de hipersensibilidad, toxicidades.
- Mala adherencia: olvidos u omisiones de dosis.
- Otros factores.

El empleo racional de este tipo de fármacos debe incluirse en la medición de diferentes variables en el paciente. Desde un punto de vista filosófico, debemos entender la enfermedad como un problema a resolver mediante una intervención (el tratamiento) y para demostrar que dicha intervención tiene efecto, debemos medir variables de respuesta que sean lo más objetivas posibles y, sobre todo, estandarizadas.

Las sociedades científicas, como por ejemplo en el ámbito de la reumatología la American College of Rheumatology (ACR), la European League Against Rheumatology (EULAR) o la Sociedad Española de Reumatología (SER), suelen emitir documentos de consenso o de recomendación sobre el empleo de este tipo de valoraciones estandarizadas que permiten realizar un seguimiento multidimensional de cada enfermedad. Algunos ejemplos son los siguientes:

- Artritis reumatoide: principalmente se emplean el DAS28 como medida de actividad clínica y el HAQ como medida de discapacidad.
- Espondiloartropatías (EA, APs, etc): BASDAI, BASFI, ASDAS, DAPSA.
- Lupus Eritematoso Sistémico: SLEDAI (*Systemic Lupus Erythematosus Disease Activity Index*).
- Síndrome de Sjögren: clinESSDAI (*Clinical European League Against Rheumatism Sjögren's Syndrome Diseases Activity Index*), ESSPRI (*Sjögren Syndrome Patient Reported Index*).

Cada enfermedad tiene una naturaleza patogénica única y por tanto la respuesta se mide y se evalúa de manera diferente en cada una de ellas. Derivado de ello también el concepto de respuesta terapéutica difiere entre entidades, pero en general hay cierta homogeneidad con los conceptos de fallo primario y secundario, que ya se han mencionado previamente como causas de discontinuación de tratamientos. Este tipo de definiciones son objeto de controversia científica continua y en ocasiones es difícil saber a qué definiciones acogerse, aunque en el campo de la artritis reumatoide se acepta en general las siguientes definiciones:

- Respuesta clínica: descenso ≥ 1.2 puntos en el índice DAS28.
- Fallo primario: ausencia de respuesta clínica en las primeras 16 semanas de tratamiento.
- Fallo secundario: Incremento del índice DAS28 ≥ 0.6 puntos en los últimos seis meses.

Los factores que más se han vinculado al fallo primario son la obesidad, el tabaquismo, la elevada actividad clínica al diagnóstico, la larga evolución de la enfermedad previamente al inicio del tratamiento y factores genéticos, entre otras variables. Respecto al fallo secundario, sobre todo destacan la obesidad y el tabaquismo (Emery, 2012).

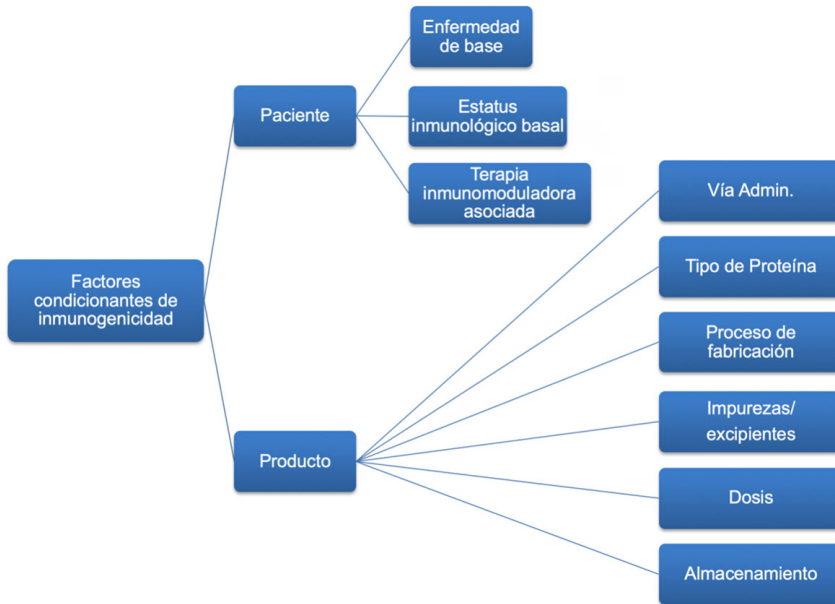
Parte de las causas de discontinuación que se han ido introduciendo se enmarcan en un concepto que se aplica a las terapias biológicas, a saber, la inmunogenicidad, definida como la capacidad de una sustancia para desencadenar una respuesta inmunitaria de tipo celular y humoral, con el desarrollo de memoria inmunológica. Esta propiedad es especialmente provechosa cuando se habla de inmunizaciones, pues gracias a la inmunogenicidad de dichos productos, se prepararía a un individuo frente a posteriores exposiciones a un agente nocivo (virus, bacterias encapsuladas) tras haber sido vacunado. Paradójicamente este atributo supone un problema cuando se emplean moléculas de síntesis biotecnológica basadas en proteínas, puesto que podría ser causa no sólo de una pérdida de respuesta, sino también una fuente de eventos adversos (Porter, 2001).

La hipótesis de que la inmunogenicidad pudiera ser un problema surge a principios de este siglo con el comienzo de la expansión del empleo de proteínas recombinantes (eritropoyetinas, insulinas, factores de crecimiento), cuando ya se comienza a demostrar que dichos productos generan respuestas inmunes en el individuo en forma de anticuerpos antifármaco (AAF) que podían comportarse como neutralizantes o no neutralizantes en función del determinante antigénico reconocido (Koren, Zuckerman y Mire-Sluis, 2002).

Pronto surgió la pregunta de qué factores influían en tal inmunogenicidad, identificándose tanto variables dependientes del paciente (peso, enfermedad, otros tratamientos asociados) como del producto (naturaleza, quimerismo, vía de admi-

Figura 2.

FACTORES CONDICIONANTES DE LA INMUNOGENICIDAD



Fuente: Elaboración propia.

nistración...). Algunos de estos factores se muestran en la Figura 2 (Sharma, 2007; Schellekens, 2005).

La mayor fuente de evidencia en el campo de la reumatología proviene del empleo de los fármacos inhibidores de la acción del factor de necrosis tumoral (TNF-alfa), especialmente con infliximab (anticuerpo monoclonal quimérico murino-humano), etanercept (proteína de fusión) y adalimumab (anticuerpo monoclonal humano). Los AAF neutralizantes disminuyen la acción farmacológica del tratamiento debido a que se dirigen frente al idiotipo y por tanto bloquearían la unión del anticuerpo monoclonal con su diana. El efecto de los anticuerpos no neutralizantes, más frecuentes, es heterogéneo y principalmente alterarían las propiedades farmacocinéticas del producto, por ejemplo, aumentando su aclaramiento. En general la incidencia de AAF es superior en moléculas quiméricas que, en tratamientos humanizados o humanos y especialmente baja en el caso de las proteínas de fusión, por carecer éstas de idiotipo (Kalden y Schulze-Koops, 2017). Así, en los diferentes registros de terapias biológicas esto ha quedado patente en tanto en cuanto las moléculas quiméricas como infliximab han demostrado tener menor persistencia que adalimumab o etanercept (Maneiro, Salgado y Gomez-Reino, 2013).

La inmunogenicidad y por tanto la producción de AAF es inferior también en el caso del empleo de inmunosupresión concomitante, con gran evidencia respaldando el uso de MTX concomitante o, en su defecto, otros fármacos como leflunomida o sulfasalazina (Krieckaert, Nurmohamed y Wolbink, 2012; Burmester *et al.*, 2015; Vogelzang *et al.*, 2015; Jani *et al.*, 2014).

Cabe mencionar que la inmunogenicidad es específica de cada producto y por tanto en el caso de haberse desarrollado AAF frente a un fármaco innovador, también reconocerían a una molécula biosimilar (Ben-Horin *et al.*, 2016).

11.2. RESULTADOS PERCIBIDOS POR LOS PACIENTES (PATIENT REPORTED OUTCOMES) EN REUMATOLOGÍA

Históricamente la reumatología, tanto en su ámbito clínico como de investigación, ha estado en la vanguardia de la medición de resultados percibidos por

Tabla 1.

PRO UTILIZADOS EN ENFERMEDADES REUMÁTICAS

Enfermedad	PRO habitualmente utilizados			Otros PRO
	Dolor	Función física	VGP	
Artritis reumatoide	Sí	Sí	Sí	Fatiga
Espondilitis anquilosante	Sí	Sí	Sí	Rigidez espinal
Artritis psoriásica	Sí	Sí	Sí	CVRS
Gota crónica	Sí	Sí	Sí	CVRS Limitación de la actividad
Gota aguda	Sí	No	Sí	Limitación de la actividad
Osteoartritis	Sí	Sí	Sí	CVRS
Lupus	No	No	No	CVRS
Vasculitis	No	No	No	CVRS
Esclerosis sistémica	Sí	Sí	Sí	CVRS Transición de salud (*) Sensibilidad
Fibromialgia	Sí	Sí	Sí	Fatiga Trastornos del sueño Depresión Deterioro cognitivo
Osteoporosis	No	No	No	Ninguno
Artritis idiopática juvenil	No	Sí	Sí	Ninguno
Miositis	No	Sí	Sí	Ninguno

Notas: VGP: valoración global del paciente de la actividad de su enfermedad; CVRS: calidad de vida relacionada con la salud.

(*): definido como el cambio de VGP a lo largo del tiempo

Fuente: Adaptado de van Tuyl y Boers, 2015.

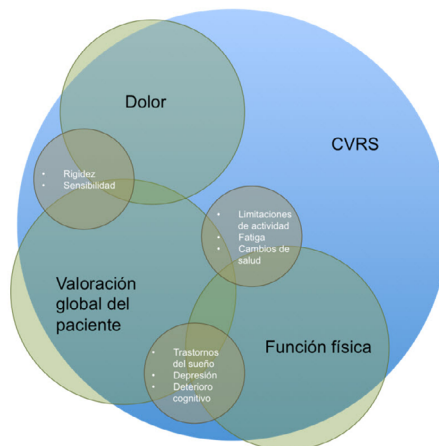
los pacientes (*patient reported outcomes* – PRO), promoviendo el cambio de paradigma de un estrecho modelo biomédico hacia un modelo biopsicosocial de la salud mucho más amplio. Si bien inicialmente los PRO se utilizaban fundamentalmente en investigación, en los últimos 15-20 años se han realizado importantes esfuerzos para poner en valor su utilidad en la práctica clínica y, aunque todavía no se ha alcanzado una situación óptima, son cada vez más aceptados y utilizados (Callahan, 2016).

El número de PRO utilizados en el campo de las enfermedades reumáticas en la actualidad es abundante y sigue creciendo, lo que implica que cada vez existe mayor dificultad para identificar cuáles son los más importantes y priorizar así su medición. En este sentido, se han propuesto algunos dominios que, como mínimo, deberían ser medidos (Tabla 1), identificándose tres PRO claves en reumatología, que son el dolor, la función física y la valoración global del paciente de la actividad de su enfermedad (van Tuyl y Boers, 2015). Estos tres PRO coinciden con los tres grandes dominios a tener en cuenta, los cuales, además, se solapan entre sí (tanto la función física como el dolor se reflejan en buena medida en la valoración global del paciente de la actividad de su enfermedad), así como con el concepto amplio de calidad de vida relacionada con la salud (CVRS) (Figura 3). Además, en algunas enfermedades reumáticas concretas, también es importante medir otros dominios específicos (van Tuyl y Boers, 2015).

Teniendo en cuenta lo expuesto, la cuestión que se plantea es qué es preferible medir, la valoración global del paciente de la actividad de su enfermedad, la

Figura 3.

RELACIONES ENTRE PRO EN ENFERMEDADES REUMÁTICAS



Fuente: Adaptado de van Tuyl y Boers, 2015.

CVRS o medidas específicas para cada dominio identificado. Y la respuesta a esta pregunta seguramente dependerá de la enfermedad de que se trate y el ámbito en el que se realice la medición (investigación o práctica clínica, evaluación de intervenciones dirigidas a dominios específicos, etc.) (van Tuyl y Boers, 2015).

Por otra parte, además de para medir el impacto de la enfermedad en el paciente y el efecto del tratamiento, los PRO se muestran como una herramienta de utilidad para medir la calidad de los cuidados prestados en el área de reumatología (Wahl y Yazdany, 2016).

■ 11.2.1. Artritis reumatoide

Las actuales guías para el manejo de la AR inciden en la necesidad de un diagnóstico y una intervención temprana, así como de instaurar un tratamiento con un objetivo concreto (estrategia *treat to target*) que debe ser la remisión clínica o, alternativamente, un estado de bajo nivel de actividad de la enfermedad. En los últimos años esta estrategia ha sido extremadamente efectiva en el control de la enfermedad, pero, paradójicamente, no lo ha sido tanto en la mejora del bienestar del paciente en relación con diferentes PRO como el dolor, la fatiga, la rigidez matutina o la percepción de la actividad de la enfermedad, probablemente debido a sus elevadas expectativas en relación con estos dominios. Una reciente revisión sobre las necesidades no cubiertas en los pacientes con AR demostró que, a pesar de estar con un tratamiento efectivo, muchos pacientes experimentan dolor y fatiga significativos, con un importante impacto negativo sobre su CVRS. Además, muchos pacientes presentan síntomas mentales (ansiedad, depresión) que no son tenidos en cuenta, así como problemas sexuales o de participación social o laboral que también impactan negativamente sobre su calidad de vida. En este sentido, otro estudio cifra en un 36 % el grado de discordancia entre pacientes y reumatólogos con relación a la percepción del grado de actividad de la enfermedad (Fautrel *et al.*, 2018).

Por ello, en la actualidad el reto es mejorar las percepciones individuales de cada paciente en relación a su enfermedad y su tratamiento y conseguir una asistencia más centrada en el paciente, mediante la toma de decisiones compartidas y la mejora de resultados considerados importantes por los propios pacientes. En este sentido, una buena utilización de PRO puede ser de gran utilidad para alinear los objetivos de los profesionales sanitarios (reducción de la inflamación, prevención del daño articular, mejoría de la función física y reducción de la mortalidad) con la perspectiva de los pacientes (control del dolor y fatiga y mantenimiento de la función física y de la CVRS). Adicionalmente, los PRO se pueden utilizar como señales de alarma de que el paciente no está controlando bien la situación y que es necesaria una visita con el reumatólogo (Fautrel *et al.*, 2018).

En la actualidad existe consenso internacional sobre la importancia de medir los tres dominios clave en el área de reumatología citados previamente: la función

física, el dolor y la valoración global del paciente de la actividad de la enfermedad (van Tuyl y Michaud, 2016). Una revisión sistemática publicada en 2016 reveló que estos tres dominios son medidos en la mayoría de los estudios (función física en el 68 %, valoración global del paciente en el 49 % y dolor en el 40 %), muy por encima de cualquier otro PRO (Kilic *et al.*, 2016). Sin embargo, otros muchos PRO también pueden ser utilizados (brotes, la remisión, rigidez, capacidad para trabajar, productividad laboral, automanejo del paciente, trastornos del sueño y el distrés emocional...). En la tabla 2 se recogen algunos ejemplos de mediciones de PRO en AR agrupados por dominios (van Tuyl y Michaud, 2016).

Tabla 2.

EJEMPLOS DE PRO UTILIZADOS EN ARTRITIS REUMATOIDE

	Dominio	Medidas de PRO
Medidas clave del ACR	Función física	HAQ, HAQ-II, MDHAQ, AIMS2, PROMIS PF-10, PF-20, and PF-CAT
	Dolor	VAS, NRS, RAPS, RPS, PROMIS pain intensity and pain interference
	Valoración global del paciente	PGA, NRS
Potenciales medidas clave	Fatiga	VAS, NRS, MAF, SF-36 vitality, FACITFS, PMS, BRAF NRS
	Rigidez	Morning stiffness duration, severity
	Brotes	Patient reported flare, worsening, and treatment change
Medidas multidimensionales de actividad de la AR		PAS, PAS-II, RAPID-3, RADAI, RAID
Otras	Sueño	VAS, NRS, AIS, MOS Sleep Measure, Pittsburgh Sleep Diary, WHI Insomnia Rating Scale
	Depresión	HADS, PHQ-8/9, BDI, MHI
	Trabajo	WLQ, WALs, SPS-6, EWPS, RA WIS, disability/work status

Notas: HAQ: Health Assessment Questionnaire; MDHAQ: modified HAQ; AIMS: Arthritis Impact Measurement Scales; PROMIS: Patient Reported Outcomes Measurement Information System; PF-10: physical function 10-item, PF-20: physical function 20-item; CAT: computerized adaptive test; VAS: visual analog scale; NRS: Numerical Rating Scale; RAPS: Rheumatoid Arthritis Pain Scale; RPS: Regional Pain Scale; PGA: Patient's Global Assessment of Disease Activity; MAF: Multidimensional Assessment of Fatigue scale; SF-36: Cuestionario de Calidad de Vida Relacionada con la Salud (Short-form 36); FACITFS: Functional Assessment of Chronic Illness Therapy Fatigue Scale; PMS: Profile of Mood States; BRAF: Bristol Rheumatoid Arthritis Fatigue; PAS: Patient Activity Scale; RAPID-3: Routine Assessment of Patient Index Data with 3 measures; RADAI: Rheumatoid Arthritis Disease Activity Index; RAID: Rheumatoid Arthritis Index of Disease Activity; AIS: Athens Insomnia Scale; MOS: Medical Outcome Study; WHI: Women's Health Initiative; HADS: Hospital Anxiety and Depression Scale; PHQ: Patient Health Questionnaire; BDI: Beck Depression Inventory; MHI: Mental Health Inventory; WLQ: Work Limitations Questionnaire; WALs: Workplace Activity Limitations Scale; SPS-6: 6-item Stanford Presenteeism Scale; EWPS: Endicott Work Productivity Scale; WIS: Work Instability Scale; RA: Rheumatoid Arthritis.

Fuente: Van Tuyl y Michaud, (2016).

Para la medición de la valoración global de la enfermedad por el paciente existen instrumentos como el PGA (*Patient's Global Assessment of Disease Activity*), el RAPID-3 (*Routine Assessment of Patient Index Data with 3 measures*) o el RAID (*Rheumatoid Arthritis Index of Disease Activity*). El PGA está fuertemente recomendado tanto por ACR como por EULAR y abarca dos conceptos, la salud y la actividad globales de la enfermedad. Se valora mediante una escala numérica (NRS), una escala numérica verbal o una escala visual analógica (VAS). El RAPID-3 incluye tres PRO de dolor (VAS), de actividad de la enfermedad (PGA VAS) y función física (HAQ: *Health Assessment Questionnaire*). El RAID contiene siete dominios (dolor, función física, fatiga, sueño, bienestar físico y emocional y afrontamiento de la enfermedad) y fue desarrollado por EULAR con el objetivo de combinar los PRO más importantes en una sola medida. El RAPID-3 muestra una buena correlación con el DAS28, el CDAI y el SDAI y puede utilizarse para predecir la progresión del daño estructural. Además, es más rápido de obtener que hacer un recuento de articulaciones, lo que puede ahorrar tiempo al reumatólogo, y puede ser utilizado por el propio paciente en su domicilio. Por su parte, el RAID muestra buena correlación con el PGA, el DAS28 y los componentes físicos y mentales del cuestionario de calidad de vida SF-36 (*Medical Outcomes Survey Short Form-36*). Además, muestra buena fiabilidad y sensibilidad al cambio, siendo capaz de discriminar entre enfermedad activa y no activa (Fautrel *et al.*, 2018).

La función física puede medirse mediante el HAQ o cualquiera de sus versiones. Aunque no es un cuestionario específico de AR, es muy sensible a los cambios, especialmente en caso de elevada actividad de la enfermedad, por lo que es un instrumento recomendado tanto por ACR como por EULAR. Es capaz de predecir presencia de comorbilidades, mortalidad, utilización de recursos sanitarios (incluyendo utilización de fármacos o la necesidad de cirugía), necesidad de medidas de soporte social, probabilidad de pérdida de empleo o productividad y costes (Fautrel *et al.*, 2018).

Para medir el dolor en pacientes con AR existen dos tipos de herramientas: simples, que miden la intensidad del dolor (VAS, NRS), y compuestas, que abarcan diferentes aspectos del dolor (severidad, frecuencia, duración y localización), tales como AIMS (*Arthritis Impact Measurement Scales*) o la subescala de dolor del SF-36. En la práctica, las herramientas simples suelen ser las más utilizadas y proporcionan una información global del dolor, reservándose las compuestas para objetivos más específicos (Fautrel *et al.*, 2018).

La CVRS puede verse seriamente comprometida en los pacientes con AR, por lo que cada vez es más recomendada su evaluación en la práctica clínica. Diferentes herramientas pueden ser de utilidad, tales como el SF-36, el RAQoL (*RA Quality of Life Scale*) el EuroQoL-5D o el HUI3 (*Health Utilities Index-Mark 3*). La medición de la CVRS proporciona información valiosa acerca del bienestar general, tanto físico como mental, del paciente. Además, algunas dimensiones del SF-36 (vitalidad, función emocional) pueden predecir la remisión, mientras que una disminución de la CVRS podría indicar una mayor probabilidad de abandono de la actividad

laboral o pérdida de productividad, lo que conlleva consecuencias importantes tanto para el propio paciente como para la sociedad (Fautrel *et al.*, 2018).

Otros PRO menos utilizados, pero que también pueden proporcionar información de valor en AR, son la fatiga, la productividad laboral, la sintomatología aceptable por el paciente o la presencia de señales de alerta.

■ 11.2.2. Artritis psoriásica

La APs es una enfermedad compleja asociada con la psoriasis en la que los pacientes pueden presentar una gran variedad de manifestaciones (artritis inflamatoria, espondilitis, entesitis, dactilitis, psoriasis, afectación ungueal...), con un impacto significativo sobre su función física, nivel de energía, relaciones sociales, carácter y calidad de vida. El tipo de manifestaciones que aparecen es muy variable y pueden hacerlo en distinto grado, de modo que la forma en que cada paciente experimenta la enfermedad y el modo en que esta afecta a su vida diaria es extremadamente variable. Por este motivo resulta difícil la evaluación de la enfermedad por parte del médico y, en este sentido, se considera que los PRO pueden tener una especial utilidad para complementar la información aportada por las medidas reportadas por el médico y por los datos de laboratorio (Orbai y Ogdie, 2016).

Son escasos los PRO que han sido validados en APs y todavía son menos los que han sido desarrollados específicamente para esta enfermedad. Por este motivo, la mayoría de PRO que se utilizan en APs han sido desarrollados para otras patologías o son genéricos, dirigidos a evaluar de forma general la salud de la población. Se considera que deben medirse PRO dirigidos a evaluar la valoración global del paciente de la actividad de su enfermedad, el dolor, la función física y la CVRS. Aparte de estos PRO, existen otros destinados a valorar la productividad laboral, la fatiga, aspectos psicológicos y otros síntomas. Algunos de los utilizados, agrupados por dominios, son los siguientes (Orbai y Ogdie, 2016):

- Dolor
 - Escala visual analógica
- Valoración global
 - Valoración global del paciente:
 - Piel
 - Articulaciones
 - Piel y articulaciones
- CVRS
 - Medical Outcomes Study Short Form-36

- Euro-Qol 5 Dimensions
- PsA Quality of Life Index
- Dermatologic Life Quality Index (DLQI)
- Ankylosing Spondylitis Quality of Life Index
- Impacto de la enfermedad
 - Arthritis Impact Measurement Scales
 - Psoriatic Arthritis Impact of Disease
- Actividad de la enfermedad
 - Routine Assessment of Patient Index Data (RAPID)
 - Rheumatoid Arthritis Disease Activity Index (RADAI)
 - Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)
- Discapacidad y función física
 - Health Assessment Questionnaire Disability Index
 - Bath Ankylosing Spondylitis Functional Index (BASFI)
 - Disabilities of Arm, Shoulder, and Hand Questionnaire
- Piel
 - Psoriasis Symptom Inventory
 - Worst Itch Numerical Rating Scale
- Fatiga
 - Functional Assessment Chronic Illness Therapy-Fatigue
 - Fatigue Visual Analog Scale/Numerical Rating Scale
- Productividad
 - Work Productivity Survey (arthritis specific)

■ 11.2.3. Espondiloartritis axial

Tanto la espondiloartritis axial (axEA) como su forma severa y prototípica, la EA, suponen un desafío desde el punto de vista de definición de PRO dada la ausencia de datos obtenidos de estudios epidemiológicos longitudinales y a la naturaleza inespecífica de los marcadores inflamatorios de laboratorio que se utilizan para monitorizar la actividad de la enfermedad. Las manifestaciones clínicas son muy variables y se asocian con dificultad para llevar a cabo actividades de la vida diaria, con la productividad laboral y con la CVRS. Por este motivo los PRO constituyen

una fuente de información eficiente para estimar la calidad de vida del paciente y la actividad de la enfermedad en el contexto de la práctica clínica (Nhan y Caplan, 2016; Kwan *et al.*, 2017; Png *et al.*, 2018).

Se ha identificado un elevado número de PRO, tanto generales como específicos, que pueden ser de utilidad en axEA. Sin embargo, en la práctica son muy pocos los utilizados, siendo el ASQoL (*Ankylosing Spondylitis Quality of Life*) y el BASFI (*Bath Ankylosing Spondylitis Functional Index*) los más utilizados en EA (Nhan y Caplan, 2016; Kwan *et al.*, 2017; Png *et al.*, 2018).

En la tabla 3 se recogen algunos ejemplos de PROs utilizados en axEA.

Tabla 3.

EJEMPLOS DE PRO UTILIZADOS EN ESPONDILOARTRITIS AXIAL

PRO	Componentes
ASDAS	Dolor de espalda, duración de la rigidez matutina, dolor/inflamación en articulaciones periféricas, bienestar general, PCR/VSG
ASQoL	Impacto de la enfermedad sobre sueño, estado de ánimo, motivación, actividades de la vida diaria, relaciones sociales
BASDAI	Dolor de espalda, fatiga, dolor/inflamación en articulaciones periféricas, sensibilidad, grado y duración de la rigidez matutina
BASFI	Anatomía funcional y capacidad para llevar a cabo actividades de la vida diaria
BAS-G	Bienestar en la última semana y en los 6 meses previos
BASMI	Rotación cervical, flecha nucal, flexión lumbar, flexión lumbar lateral, distancia intermaleolar
DFI	Actividades de la vida diaria
HAQ-S	Capacidad para vestirse, comer, caminar, desarrollar trabajos rutinarios, cuello y postura

Notas: ASDAS: Ankylosing Spondylitis Disease Activity Score; PCR: Proteína C Reactiva; VSG: Velocidad de Sedimentación Globular; ASQoL: Ankylosing Spondylitis Quality of Life; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; BAS-G: Bath Ankylosing Spondylitis Global score; BASMI: Bath Ankylosing Spondylitis Metrology Index; DFI: Dougados Functional Index; HAQ-S: Health Assessment Questionnaire for the Spondyloarthropathies.

Fuente: Nhan y Caplan, 2016.

REFERENCIAS

- ALETAHA, D., y SMOLEN, J. S. (2018). Diagnosis and Management of Rheumatoid Arthritis: A Review. *JAMA*, 320(13), pp. 1360-1372.
- ÁREAS DEL ÁGUILA, V. L. y MONTE BOQUET, E. (2017). Adherencia en enfermedades inflamatorias crónicas. En: Sociedad Española de Farmacia Hospitalaria (ed.). En: O. IBARRA BARRUETA y R. MORILLO VERDUGO (coords.), *Lo que debes sobre la adherencia al tratamiento*. Badalona: Euromedice Vivactis.
- VAN DEN BEMT, B. J., ZWIKKER, H. E. y VAN DEN ENDE, C. H. (2012). Medication adherence in patients with rheumatoid arthritis: a critical appraisal of the existing literature. *Expert Rev Clin Immunol*, 8(4), pp. 337-351.

- BEN-HORIN, S., YAVZORI, M., BENHAR, I. *et al.* (2016). Cross-immunogenicity: antibodies to infliximab in remicade-treated patients with IBD similarly recognise the biosimilar remsima. *Gut*, 65, pp.1132–1138.
- BHOI, P., BESSETTE, L., BELL, M. J. *et al.* (2017). Adherence and dosing interval of subcutaneous antitumour necrosis factor biologics among patients with inflammatory arthritis: analysis from a Canadian administrative database. *BMJ Open*, 7(9), p. e015872.
- BLAIR, H. A. y DEEKS, E. D. (2017). Abatacept: A Review in Rheumatoid Arthritis. *Drugs*, 77(11), pp. 1221-1233.
- BURMESTER, G. R., KIVITZ, A. J., KUPPER, H. *et al.* (2015). Efficacy and safety of ascending methotrexate dose in combination with adalimumab: the randomised CONCERTO trial. *Ann Rheum Dis*, 74(6), pp. 1037–1044.
- CALLAHAN, L. F. (2016). The History of Patient-Reported Outcomes in Rheumatology. *Rheum Dis Clin North Am*, 42(2), pp. 205-217.
- CASO, F., COSTA, L., DEL PUENTE, A. *et al.* (2015). Psoriatic arthritis and TNF inhibitors: advances on effectiveness and toxicity. *Expert Opin Biol Ther*, 15(1), pp. 1-2.
- CATAY, E. R. y SORIANO, E. (2014). Tratamiento biológico en reumatología. *Rev. Hosp. Ital*, 34(3), pp. 76-83.
- CUMMINS, L. y KATIKIREDDI, V. S., SHANKARANARAYANA, S. *et al.* (2015). Safety and retention of combination triple disease-modifying anti-rheumatic drugs in new-onset rheumatoid arthritis. *Intern Med J*, 45(12), pp. 1266-1273.
- DEMONCEAU, J., RUPPAR, T., KRISTANTO, P. *et al.* (2013). Identification and assessment of adherence-enhancing interventions in studies assessing medication adherence through electronically compiled drug dosing histories: a systematic literature review and meta-analysis. *Drugs*, 73(6), pp. 545-562.
- DILLAA, T., VALLADARES, A., LIZÁN, L. *et al.* (2009). Adherencia y persistencia terapéutica: causas, consecuencias y estrategias de mejora. *Aten Primaria*, 41(6), pp. 342–348.
- ELLIS, L., BOLGE, S. y RICE, P. (2014). Golimumab utilization patterns and refill adherence in patients with rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis. *Value Health*, 17(3): pp. A49–A50.
- EMERY, P. (2012). Optimizing outcomes in patients with rheumatoid arthritis and an inadequate response to anti-TNF treatment. *Rheumatology*, 51(Suppl 5), pp. v22-30.
- FAUTREL, B., ALTEN, R. y KIRKHAM, B. *et al.* (2018). Call for action: how to improve use of patient-reported outcomes to guide clinical decision making in rheumatoid arthritis. *Rheumatol Int*, 38(6), pp. 935-947.
- DE LA FOREST DIVONNE, M., GOTTENBERG, J. E. y SALLIOT, C. (2017). Safety of biologic DMARDs in RA patients in real life: A systematic literature review and meta-analyses of biologic registers. *Joint Bone Spine*, 84(2), pp. 133-140.
- VAN DER GOES, M. C., JACOBS, J. W. y BIJLSMA, J. W. (2014). The value of glucocorticoid co-therapy in different rheumatic diseases--positive and adverse effects. *Arthritis Res Ther*, 16(Suppl 2), p. S2.
- HROMADKOVA, L., SOUKUP, T. y VLCEK, J. (2015). Quality of life and drug compliance: their interrelationship in rheumatic patients. *J Eval Clin Pract*, 21(5), pp. 919–924.
- JANI, M., BARTON, A., WARREN, R. B. *et al.* (2014). The role of DMARDs in reducing the immunogenicity of TNF inhibitors in chronic inflammatory diseases. *Rheumatology*, 53(2), pp. 213-222.
- KALDEN, J. R. y SCHULZE-KOOPS, H. (2017). Immunogenicity and loss of response to TNF inhibitors: implications for rheumatoid arthritis treatment. *Nat Rev Rheumatol*, 13(12), pp. 707-718.
- KILIC, L., ERDEN, A., BINGHAM, C. O. *et al.* (2016). The Reporting of Patient-reported Outcomes in Studies of Patients with Rheumatoid Arthritis: A Systematic Review of 250 Articles. *J Rheumatol*, 43(7), pp. 1300-1305.

- KOREN, E., ZUCKERMAN, L. A. y MIRE-SLUIJ, A. R. (2002). Immune responses to therapeutic proteins in human-clinical significance, assessment and prediction. *Curr Pharm Biotechnol*, 3(4), pp. 349-360.
- KRIECKAERT, C. L., NURMOHAMED, M. T. y WOLBINK, G. J. (2012). Methotrexate reduces immunogenicity in adalimumab treated rheumatoid arthritis patients in a dose dependent manner. *Ann Rheum Dis*, 71(11), pp. 1914-1915.
- KRULIG, E. y GORDON, K. B. (2010). Ustekinumab: an evidence-based review of its effectiveness in the treatment of psoriasis. *Core Evid*, 5, pp. 11-22.
- KWAN, Y. H., FONG, W., TAN, V. I. C. *et al.* (2017). A systematic review of quality-of-life domains and items relevant to patients with spondyloarthritis. *Semin Arthritis Rheum*, 47(2), pp. 175-182.
- LAVIELLE, M., PUYRAIMOND-ZEMMOUR, D., ROMAND, X. *et al.* (2018). Methods to improve medication adherence in patients with chronic inflammatory rheumatic diseases: a systematic literature review. *RMD Open*, 4(2), p. e000684.
- LOFLAND, J. H., JOHNSON, P. T., INGHAM, M. P. *et al.* (2017). Shared decision-making for biologic treatment of autoimmune disease: influence on adherence, persistence, satisfaction, and health care costs. *Patient Prefer Adherence*, 11, pp. 947-958.
- LÓPEZ-GONZÁLEZ, R., LEÓN, L., LOZA, E. *et al.* (2015). Adherence to biologic therapies and associated factors in rheumatoid arthritis, spondyloarthritis and psoriatic arthritis: a systematic literature review. *Clin Exp Rheumatol*, 33(4), pp. 559-569.
- MANEIRO, J. R., SALGADO, E. y GÓMEZ-REINO, J. J. (2013). Immunogenicity of monoclonal antibodies against tumor necrosis factor used in chronic immune-mediated inflammatory conditions: systematic review and meta-analysis. *JAMA Intern Med*, 173(15), pp. 1416-1428.
- MORA ARIAS, T. T., ORTEGA, L. A. y BOURLON, R. A. (2010). Medicamentos biológicos en artritis reumatoide. *Med Int Mex*, 26(3), pp. 259-268.
- MURAGE, M. J., TONGBRAM, V., FELDMAN, S. R. *et al.* (2018). Medication adherence and persistence in patients with rheumatoid arthritis, psoriasis, and psoriatic arthritis: a systematic literature review. *Patient Prefer Adherence*, 12, pp. 1483-1503.
- NHAN, D. T. y CAPLAN, L. (2016). Patient-Reported Outcomes in Axial Spondyloarthritis. *Rheum Dis Clin North Am*, 42(2), pp. 285-299.
- ORBAI, A. M. y OGDIE, A. (2016). Patient Reported Outcomes in Psoriatic Arthritis. *Rheum Dis Clin North Am*, 42(2), pp. 265-283.
- PAVELKA, K., KIVITZ, A., DOKOUPILOVA, E. *et al.* (2018). Efficacy, safety, and tolerability of secukinumab in patients with active ankylosing spondylitis: a randomized, double-blind phase 3 study, MEASURE 3. *Arthritis Res Ther*, 19(1), p. 285.
- PNG, K., KWAN, Y. H., LEUNG, Y. Y. *et al.* (2018). Measurement properties of patient reported outcome measures for spondyloarthritis: A systematic review. *Semin Arthritis Rheum*, 48(2), pp. 274-282.
- PORTER, S. (2001). Human immune response to recombinant human proteins. *J Pharm Sci*, 90(1), pp. 1-11.
- SCHELLEKENS, H. (2005). Factors influencing the immunogenicity of therapeutic proteins. *Nephrol Dial Transplant*, 20(Suppl 6), p. vi3-9.
- SHARMA, B. (2007). Immunogenicity of therapeutic proteins. Part 1: Impact of product handling. *Biotechnol Adv*, 25(3), pp. 310-317.
- DA SILVA, J. A., JACOBS, J. W., KIRWAN, J. R. *et al.* (2006). Safety of low dose glucocorticoid treatment in rheumatoid arthritis: published evidence and prospective trial data. *Ann Rheum Dis*, 65(3), pp. 285-293.
- SOCIEDAD ESPAÑOLA DE REUMATOLOGÍA. (2018). *Tratado de Enfermedades Reumáticas de la SER*. Madrid, España: Editorial Panamericana.

- VAN TUYL, L. H. y BOERS, M. (2015). Patient-reported outcomes in core domain sets for rheumatic diseases. *Nat Rev Rheumatol*, 11(12), pp. 705-712.
- VAN TUYL, L. H. y MICHAUD, K. (2016). Patient-Reported Outcomes in Rheumatoid Arthritis. *Rheum Dis Clin North Am*, 42(2), pp. 219-237.
- UGARTE-GIL, M. F., ACEVEDO-VÁSQUEZ, E. *et al.* (2013). Terapia biológica en enfermedades reumatológicas. *Rev Med Hered*, 24(2), pp. 141-155.
- VOGELZANG, E., POWW, M. F., NURMOHAMED, M. H. *et al.* (2015). Adalimumab trough concentrations in patients with rheumatoid arthritis and psoriatic arthritis treated with concomitant disease-modifying antirheumatic drugs. *Ann Rheum Dis*, 74(2), pp. 474-475.
- WAHL E.R. y YAZDANY, J. (2016). Challenges and Opportunities in Using Patient-reported Outcomes in Quality Measurement in Rheumatology. *Rheum Dis Clin North Am*, 42(2), pp. 363-375.



12

RESULTADOS EN SALUD EN DERMATOLOGÍA

JOAQUÍN BORRÁS BLASCO, PABLO DE LA CUEVA DOBAO

■ 12.1. INTRODUCCIÓN

Los resultados en salud en dermatología están adquiriendo una relevancia creciente en los últimos años y no solo sirven de apoyo a las decisiones clínicas de los dermatólogos, sino que son de utilidad para los decisores sanitarios y farmacéuticos. Las patologías de la piel están asociadas con un elevado impacto sobre el paciente que afecta a su condición física, también social, emocional y psicológica. La dermatitis atópica, urticaria, psoriasis e hidrosadenitis, entre otras enfermedades dermatológicas, tienen una gran influencia en los individuos que las padecen, provocando un incremento en los trastornos del sueño y en la depresión que afectan el desempeño laboral y la vida cotidiana de los pacientes.

■ 12.2. PSORIASIS

La psoriasis es una enfermedad inflamatoria sistémica de la piel crónica, recidivante y con participación multiorgánica en la que se ve afectado el aparato locomotor. Tiene un efecto negativo significativo sobre el bienestar físico, emocional y psicológico de los pacientes. La psoriasis es una enfermedad crónica de distribución universal y recidivante del sistema inmunológico con diversas manifestaciones clínicas, aunque en general, se presenta en forma de placas eritematosas, descamativas, crónicas y simétricas (Daudén *et al.*, 2016).

La morfología de las lesiones y la región de la piel afectada determina el tipo de psoriasis del paciente: psoriasis en placas (psoriasis vulgaris), en gotas (guttata), invertida, pustulosa, eritrodérmica, del cuero cabelludo y ungueal (World Health Organization, 2016).

Existen numerosos instrumentos de medida de la intensidad y severidad de la psoriasis. Los parámetros tradicionalmente utilizados para valorar la psoriasis son: el PASI (*Psoriasis Area and Severity Index*), el BSA (superficie corporal afectada) y el PGA (Evaluación global del médico). Además de estos instrumentos de medida de la intensidad y severidad de la psoriasis, también se debe valorar la intensidad de las lesiones, su topografía (afectación facial, flexural, genital, etc.), la afectación psicossocial sobre el paciente y su efecto sobre su estado general (Puzenat *et al.*, 2010).

Psoriasis Area and Severity Index (PASI)

El índice PASI definido por Fredriksson y Pettersson en 1978, fue desarrollado como un instrumento de medida de la intensidad de la enfermedad en un trabajo que valoraba el tratamiento de la psoriasis grave con retinoides orales. Pese a sus limitaciones, en el estándar para la medida de la gravedad de la psoriasis, es el sistema más utilizado y permite su realización de forma bastante rápida. La Food and Drug Administration estableció la respuesta PASI 75 (porcentaje de pacientes que alcanzan una mejoría superior o igual al 75 % del PASI basal) como parámetro de eficacia en los ensayos clínicos de los nuevos fármacos biológicos. En la actualidad, el PASI 90 (que define el éxito del tratamiento según la European Medicines Agency) representa el objetivo terapéutico en las directrices españolas. (Chalmers, 2015).

El PASI es un índice de gravedad y área afectada de psoriasis que valora el eritema, la induración y la descamación de las lesiones por separado, en distintas partes de la superficie corporal: cabeza y cuello, extremidades superiores, tronco, extremidades inferiores incluyendo glúteo y las relaciona con el área afectada en cada una de ellas.

Este índice usa una fórmula que combina la valoración de cada lesión de psoriasis del 0 al 4 (0=ninguno, 1=leve, 2=moderado, 3=marcado, 4=muy marcado) en base a tres parámetros: eritema, infiltración y descamación, así como una evaluación ponderada en la que se asigna a cada área un porcentaje de la superficie corporal total: a la cabeza, un 10 %; a las extremidades superiores, un 20 %; al tronco, un 30 %, y a las extremidades inferiores, un 40 %. Además, La extensión de las lesiones en la zona determinada se valora de 0 a 6, siendo 0 la ausencia; 1 a <10 %, 2.10 a <30 %, 3.30 a <50 %, 4.50 a <70 %, 5.70 a <90 % y 6.90 a 100 %.

La fórmula con la que se calcula el PASI viene dada por:

$$\text{PASI} = 0,1 \times (\text{Ec} + \text{Ic} + \text{Dc}) + 0,2 \times (\text{Es} + \text{Is} + \text{Ds}) + 0,3 \times (\text{Et} + \text{It} + \text{Dt}) + 0,4 \times (\text{Ei} + \text{Ii} + \text{Di})$$

Existen herramientas y aplicaciones para facilitar el cálculo del PASI *online* y diferentes tutoriales para entender y practicar el uso de este índice: <http://aedv.es/calculadora-pasi/>.

El valor resultante es un número comprendido entre 0 y 72, a intervalos de 0,1 unidades. Una valoración de intensidad del PASI es considerar leve a los pacientes entre 0-5; moderada de 5-10 y grave en los PASI mayores de 10. Otra clasificación aceptada los divide en leve 0-7; moderada 7-12 y grave en PASI superiores a 12. No obstante, existe una gran variabilidad entre los distintos estudios, de ahí que se considere por muchos dermatólogos una psoriasis moderada/grave aquella con un PASI mayor a 10. Las tasas de respuesta PASI 50, 75, 90 y 100 corresponden a un 50, 75, 90 y 100 % respectivamente de los parámetros definidos.

Entre las ventajas del PASI podemos destacar:

- Comparabilidad histórica de las determinaciones
- Buena correlación con otras medidas
- Ha sido validado en múltiples ocasiones
- Es fácil de realizar, aunque no se emplea de manera habitual en la práctica clínica
- Buena correlación entre observadores y escasa variabilidad (2 %)

Los principales problemas del PASI son:

- La dependencia de la subjetividad del médico que puede dar lugar a distintos valores de PASI en un mismo paciente si la medición la realizan distintos dermatólogos
- No ha sido validado adecuadamente como indicador de la gravedad percibida por el paciente
- La falta de inclusión de los síntomas que produce la enfermedad,
- Ausencia de utilidad en formas como la pustulosa
- La sensibilidad de la escala es menor con valores bajos y no son proporcionales. (Jacobson y Kimball, 2004)

Body Surface area (BSA)

Se basa en el cálculo directo de la superficie corporal afectada que se mide por las lesiones de psoriasis en la palma de la mano del paciente (considera la palma de la mano como un 1 % de dicha superficie). Complementa el carácter cualitativo de las escalas de valoración global, por lo que a menudo se emplea conjuntamente. La psoriasis se considera leve cuando afecta menos del 3 % de la superficie cutánea, moderada cuando afecta entre el 3 y el 10 % y grave cuando afecta más del 10 %.

La principal limitación es la dificultad para calcular el porcentaje afectado en pacientes con áreas de piel indemne entre placas pequeñas o "guttata" y cuando hay zonas de discromía residual que no pueden considerarse afectadas por psoriasis.

Para la evaluación de un paciente, es muy útil usar los valores del BSA y del PASI. Así un paciente con PASI menor 10 y un BSA del 28 % nos indicaría una forma extensa de enfermedad con un eritema, descamación e infiltración discretos. Mientras que un paciente con un BSA bajo con un PASI elevado nos indicaría una enfermedad con placas eritematosas, gruesas y muy descamativas localizadas.

Evaluación global del médico (PGA, Physician`s Global Assessment)

Escala simple que proporciona una valoración subjetiva de la gravedad de la enfermedad sin tener en cuenta su extensión. Existen dos formas principales, la

forma dinámica, en la que se evalúa la mejoría global con respecto a la situación basal; su principal inconveniente es que se basa en la memoria del observador, por esta razón la valoración estática se ha convertido en el estándar.

Se emplea una puntuación comprendida entre 0-5 en base a la induración eritema y descamación. 0 Blanqueado, no hay signos de psoriasis, pero la decoloración postinflamatoria puede estar presente; 1 Casi blanqueado, elevación mínima de la placa, descamación y eritema; 2 Leve, elevación leve de la placa, descamación y eritema; 3 Moderada, elevación/infiltración moderada de la placa, descamación y eritema; 4 Grave, elevación de la placa muy marcada, descamación y eritema; 5 Muy grave. (Robinson, Kardos y Kimball, 2012).

Deben evaluarse por separado el eritema, la descamación y la infiltración de las lesiones en todo el cuerpo y calcular el promedio aproximado. Su ventaja es la rapidez de aplicación, y que presenta un carácter estático y no comparativo. (Pascoe *et al.*, 2015)

Tabla 1.

CLASIFICACIÓN DE LA PSORIASIS SEGÚN GRAVEDAD

	PASI	BSA	PGA	DLQI
<i>Leve.</i>				
Psoriasis estable, en placas, que afecta a menos del 10% de la superficie corporal	< 10	< 5 %	0 - 2	
<i>Moderada-severa.</i>				
Psoriasis que afecta a más del 10 % de la superficie corporal o si afecta a cara, manos y pies, genitales o pliegues.				
Afectación articular con limitación de la movilidad				
Alteración emocional que afecta vida diaria	>10	>5 %	3	>10
Psoriasis aguda, que incluye:				
Psoriasis con artritis psoriásica				
Psoriasis pustulosa generalizada				
Psoriasis en extensión rápida				
Eritrodermia psoriásica				

Self administered Psoriasis Area and Severity Index (SAPASI)

El PASI autoadministrado es una variante del PASI diseñado con la finalidad de que la puede cumplimentar el propio paciente y determinar la gravedad de sus placas de psoriasis.

Consiste en la silueta de un cuerpo humano donde el paciente sombrea las zonas afectadas y de tres escalas analógicas visuales que marcan el grado de eritema, descamación e infiltración en una escala del 0 a 12 mm. Su fórmula es $SAPASI = [(0,1 \cdot \text{área cabeza}) + (0,2 \cdot \text{área extremidades superiores}) + (0,3 \cdot \text{área de tronco}) + (0,4 \cdot \text{área de extremidades inferiores})] \cdot [0,0333x (\text{valoración del eritema} + \text{descamación} + \text{infiltración})]$. El SAPASI está bien validado, presenta una gran fiabilidad en pruebas repetidas, es adecuado para estudios epidemiológicos y existen estudios que correlacionan el PASI y el SAPASI con un buen índice de concordancia

Calidad de vida

Para el estudio de la calidad de vida en los pacientes con psoriasis se pueden usar instrumentos de medida genéricos, específicos de especialidad o específicos de psoriasis. Estos instrumentos genéricos incluyen las puntuaciones globales, los perfiles de salud y las medidas de utilidad, permitiendo comparar el impacto de las enfermedades sobre el estado de salud, hacer análisis de coste-utilidad, obtener valores poblacionales de referencia y evaluar programas para la asignación de recursos. (Bronsard *et al.*, 2010).

Índice de Calidad de Vida Dermatológica (DLQI)

En el caso de la psoriasis, el índice de calidad de vida dermatológica (DLQI) es un cuestionario autoevaluable que se emplea para medir la calidad de vida relacionada con la salud en adultos que viven con una enfermedad de la piel. Es la herramienta más comúnmente empleada para medir la calidad de vida de los pacientes con psoriasis y aborda aspectos como síntomas y sentimientos, actividades diarias, relaciones personales y efectos secundarios de los tratamientos. Fue el primer instrumento de medición de la calidad de vida específico en dermatología.

El DLQI es ampliamente utilizado tanto en investigación como en práctica clínica para valorar la calidad de vida del paciente psoriásico y se considera como uno de los índices (junto al PASI y el BSA) habitualmente utilizados para evaluar el estado del paciente, la gravedad de la enfermedad y el efecto del tratamiento.

El DLQI es una escala unidimensional para pacientes dermatológicos por encima de los 16 años, que consta de un cuestionario corto y simple que se compone de diez preguntas y que se responde en menos de 5 minutos. Las preguntas están relacionadas con la percepción del impacto que tiene la psoriasis en la calidad de vida en la última semana.

Las preguntas abarcan; síntomas y sentimientos (preguntas 1 y 2), actividades diarias (preguntas 3 y 4), actividades lúdicas (preguntas 5 y 6), trabajo y vida escolar (pregunta 7), relaciones interpersonales (preguntas 8 y 9) y efectos secundarios del tratamiento (pregunta 10). Cada ítem tiene cuatro posibilidades de respuesta según la escala de Likert: 0, no en absoluto/ no relevante; 1, un poco; 2: mucho; 3: muchísimo.

El DLQI se calcula sumando la puntuación de cada pregunta con un resultado máximo de 30 puntos y un mínimo de 0 puntos. Cuanto más alta es la puntuación, mayor es el deterioro de la calidad de vida. El DLQI también puede ser expresado como un porcentaje de la máxima puntuación posible -30-.

Interpretación de las puntuaciones del DLQI: 0-1: No afecta en nada la vida del paciente, 2-5: Pequeño efecto en la vida del paciente, 6-10: Moderado efecto en la vida del paciente, 11-20: Gran efecto en la vida del paciente, 2-30: Extremadamente gran efecto en la vida del paciente (Chren, 2010).

En cuanto al valor de la escala, su consistencia interna de esta escala está entre buena y excelente, y presenta altos coeficientes de correlación de Spearman en su evaluación de fiabilidad test-retest. A pesar de esto, se han encontrado correlaciones altas entre el DLQI y otros instrumentos de calidad de vida excepto en los aspectos mentales y emocionales por lo que se ha sugerido que deba acompañarse de escalas de calidad de vida generales como el SF 36 que complementen las carencias del cuestionario específico. (Lewis y Finlay, 2004) Adicionalmente, esta escala ha presentado un efecto suelo sustancial y una falta de sensibilidad al cambio (Basra *et al.*, 2008), que se traducen en la dificultad que tiene la escala para detectar mejorías en los pacientes y en la incapacidad para detectar cambios en la calidad de vida del paciente cuando se producen cambios clínicos en su estado de gravedad.

SKINDEX-29

El Skindex-29 es un instrumento de calidad de vida específico de enfermedades dermatológicas que mide el impacto de las enfermedades dermatológicas en la calidad de vida de los pacientes. (Abeni *et al.*, 2002)

El instrumento original se componía inicialmente de 61 *ítems*, la versión mejorada que se emplea en la actualidad tiene 29 *ítems* para ser resuelta en 5 minutos, con mayor capacidad de discriminación de evaluación longitudinal y validado para la población española. (Jones-Caballero *et al.*, 2002)

El Skindex 29 se centra en 3 tres escalas de la realidad del paciente: funcional con 12 *ítems* (vida social, aislamiento, sexualidad, trabajo, aficiones), emocional con 10 *ítems* (vergüenza, molestia, depresión, frustración) y sintomática con 7 *ítems* (picor, dolor, irritación). Cada *ítem* tiene una escala de respuesta tipo Likert con posibles opciones, 0 (nunca) a 4 (todo el tiempo). Las preguntas se refieren a las últimas cuatro semanas. Puntajes altos indican peor calidad de vida y se considera un predictor significativo de patología psiquiátrica.

Se trata de una herramienta sensible al cambio, por lo que es adecuada para ser empleada en ensayos clínicos. Skindex-29 es el instrumento que cuenta con la mayor rigurosidad metodológica en su proceso de validación. Presenta poco efecto techo y suelo, y una comprobada interpretabilidad y factibilidad.

Psoriasis Disability Index (PDI)

Es un cuestionario específico de psoriasis autoadministrado y validado en población española. Consta de 15 *ítems*, con 4 opciones de respuesta cada uno; desde 0 (interferencia nula de la psoriasis) y 3 (interferencia máxima de la psoriasis). Referentes a la discapacidad funcional provocada por la enfermedad durante las 4 semanas anteriores en las actividades diarias, trabajo/estudios, relaciones personales, ocio y tratamiento. El PDI se correlaciona con el DLQI pero no mejor que este último con puntuaciones objetivas como el PASI o el SAPASI. Es sensible a los cambios, aunque puede tener un sesgo cultural.

Medidas combinadas de afectación objetiva y calidad de vida

Índice de psoriasis de Salford

Este sistema de medición pretende integrar datos sobre la gravedad clínica del paciente mediante el PASI y por el otro, el impacto psicosocial para el paciente y su historia de gravedad en base a la resistencia al tratamiento previo.

Los datos del PASI se transforman en una escala de 0 a 10 según la siguiente escala de conversión: -PASI 0= 0; PASI 0,1-3=1; PASI 3,1-5= 2; PASI 5,1-8= 3; PASI 8,1-11=4; PASI 11,1-14=5; PASI 14,1-18=6; PASI 18,1-23=7; PASI 23,1-29=8; PASI 29,1-36=9; PASI >36=9. El impacto psicosocial es valorado por el paciente en el momento de la prueba empleando una escala analógica visual del 0 al 10, cuyas puntuaciones se correlacionan bien con el Psoriasis Disability Index.

Por último, la historia de gravedad de la enfermedad se obtiene sumando tres valores: 1 punto por cada tratamiento sistémico; 1 punto más por cada tratamiento sistémico de más de 1 año de duración; 1 punto más si el paciente ha recibido más de 200 sesiones o 1.000 J/cm² de PUVA; 1 punto más por cada 5 ingresos en el hospital, y 1 punto por cada episodio de eritrodermia. De esta forma, se pretende aportar una visión global del paciente, aunque adaptada en cuanto el tercer componente del índice está adaptado al manejo de la psoriasis en Gran Bretaña.

■ 12.3. URTICARIA

La urticaria es una enfermedad cutánea pruriginosa que se caracteriza por la aparición de habones evanescentes (menos de 24 horas), angioedema o ambas. La urticaria se clasifica según su tiempo de evolución en aguda (los síntomas se presentan en menos de seis semanas) y crónica (los síntomas persisten más de seis semanas). A su vez la urticaria crónica se divide en espontánea (UCE) e inducida. La urticaria crónica inducida habitualmente aparece tras estímulos físicos como calor, frío, radiación ultravioleta, roce, presión, entre otras.

La prevalencia estimada de la urticaria crónica espontánea es de 0,5-1 % de la población general, afectándose el sexo femenino aproximadamente dos veces más que el sexo masculino. La duración media de la patología es de 3-5 años y la máxima incidencia se registra entre los 20 y 40 años. (Maurer *et al.*, 2011)

Tabla 1.

PRO EMPLEADOS EN UCE

PRO	Tipo	Dominios	Puntuación	Rango de puntuación	Diferencia mínima significativa	Comentario
UAS	Prospectivo diario	Prurito y habones	Cada pregunta de 0 a 3	0-42 puntos	10 puntos	≤ 6: mínima actividad de la enfermedad. ≥28: severo
UCT	Retrospectivo: 4 ítems de las últimas 4 semanas	Control de la enfermedad	Cada pregunta de 0 a 4	0-16 puntos	3 puntos	≥12: enfermedad bien controlada
USS	12 preguntas	Mejoría de calidad de vida, duración de prurito en hinchazón, medicación	10 preguntas de 0 a 7 (una pesa doble). 2 preguntas de 0 a 8	0-93	-	Mayor puntuación mayor repercusión en calidad de vida
CU-Q2oL	23 preguntas	Prurito, hinchazón, actividades de la vida diaria, sueño, apariencia, limitaciones	Cada pregunta de 1 a 5	23-115 puntos	15 puntos	Mayor puntuación indica mayor afectación de la calidad de vida
AAS	5 preguntas	Afectación física, actividades diarias, impacto cosmético y evaluación global de severidad	Cada pregunta de 0 a 3	0 a 15 (puntuación diaria)	-	Mayor puntuación mayor actividad de la enfermedad
AE-QoL	17 preguntas	Funcionalidad, fatiga/estado anímico, miedo/vergüenza, alimentación	Cada pregunta de 0 a 5	0 a 85	-	Mayor puntuación mayor impacto en la calidad de vida
DLQI	10 preguntas	Síntomas, actividades sociales, trabajo/escuela, relaciones personales problemas del tratamiento	Cada pregunta de 0 a 3	0-30 puntos	2,34-3,1	Mayor puntuación mayor impacto en la calidad de vida

La UCE tiene una repercusión negativa en la calidad de vida de los pacientes debido a los síntomas asociados, impredecibilidad de los brotes y el curso fluctuante de la enfermedad. Así mismo se asocia a alteraciones del sueño, aislamiento social, y conlleva una repercusión económica tanto en costes directos como indirectos (medicación, pruebas de laboratorio, visitas tanto regladas como a los servicios de urgencias, productividad laboral).

En la práctica clínica no se disponen de biomarcadores para medir la actividad de la UCE, con lo que la medición de los PRO tiene una gran relevancia a la hora de evaluar y monitorizar la actividad, gravedad, control de la enfermedad y la calidad de vida.

Se utilizan habitualmente siete PRO en la clínica e investigación de la UCE. Cuatro de ellos son específicos de urticaria: UAS (*Urticaria Activity Score*), UCT (*Urticaria Control Test*); CU-Q20I (*Chronic Urticaria Quality of Life Questionnaire*) y USS (*Urticaria Severity Score*). Dos PRO específicos de angioedema: AAS (*Angioedema Severity Score*) y AE-QoL (*Angioedema Quality of Life Questionnaire*) (Moestrup, Ghazanfar y Thomsen, 2017). Por último, uno es genérico a nivel cutáneo: DLQI (Tabla 1).

Dentro de las mediciones de PRO es el UAS el más extendido y utilizado al ser específico, prospectivo y reportado por el paciente, y evalúa la actividad y la severidad de la respuesta al tratamiento. De hecho, se puede asegurar que hoy en día supone el gold standard, a pesar de tener sus limitaciones, como no recoger la urticaria inducida o angioedema. Dispone de dos componentes: intensidad de prurito (de 0 a 3) y número de habones (de 0 a 3) (Tabla 2).

Los pacientes cumplimentan el cuestionario cada día y la suma resultante corresponde al UAS diario (de 0 a 6). Esta puntuación se suma durante siete días

Tabla 2.

UAS	
¿Cuántos habones han aparecido en las últimas 24 horas?	Puntuación
Ninguno	0
< 20 habones	1
20-50 habones	2
> 50 habones	3
Intensidad del prurito en las últimas 24 horas	Puntuación
Ninguno	0
Leve (presente pero sin resultar molesto o irritante)	1
Moderado (irritante pero no interfiere con actividades diarias o sueño)	2
Severo (interfiere con actividades diarias o sueño)	3

consecutivos y se obtiene el UAS7 que oscila de 0 a 42 puntos. Aunque los cortes no son globalmente aceptados, sí que existe un consenso de catalogar la actividad de la urticaria en los siguientes intervalos (Weller, Zuberbier y Maurer, 2015):

0: total ausencia de habones y prurito.

1-6: mínima actividad de la enfermedad.

7-15: actividad leve.

16-27: actividad moderada.

28-42: actividad severa o grave.

La mínima diferencia clínicamente significativa se ha determinado en la reducción de 10 puntos en el UAS7 (Weller, Zuberbier y Maurer, 2015).

El UAS7 ha sido utilizado en números ensayos clínicos aleatorizados en UCE. En los ensayos de omalizumab se estableció para la inclusión de pacientes un punto de corte un UAS7 de al menos 16 puntos.

Los cuestionarios genéricos de calidad de vida tienen su utilidad para comparar poblaciones con diferentes patologías. El SF-36 y el SF-12 indican una repercusión importante de la UCE en áreas como el sueño, productividad y funcionalidad.

Los cuestionarios de calidad de vida de patologías dermatológicas como el DLQI, indican un impacto similar a la dermatitis atópica e incluso superior a psoriasis y acné (Baiardini, Braido y Brandi, 2006). También se ha utilizado el DLQI en los ensayos clínicos aleatorizados para demostrar la diferencia entre los brazos activos y placebo.

El Skindex-29 ha demostrado que los pacientes con UCE presentan una afectación muy importante en las actividades de la vida diaria como el sueño, ocupación y vida social, con un mayor impacto en el sueño. Así mismo repercute a nivel emocional asociándose con ansiedad, depresión y aparición de miedos (Maurer, Ortonne y Zuberbier, 2009).

El cuestionario específico de urticaria CU-Q2oL proporciona información muy valiosa para determinar la calidad de vida de los pacientes en relación a su enfermedad. Una de sus limitaciones es que es más laborioso que el cálculo del DLQI y que no recoge el dominio del angioedema asociado en algunos pacientes (sería ideal poder implementarlo con el AE-QoL). Sería adecuado utilizarlo en la visita basal de los pacientes y en el seguimiento y tiene una correlación mayor con el UAS7 que el DLQI (Koti *et al.*, 2013).

En resumen, los PRO, a pesar de ser utilizados de forma heterogénea en la UCE, existe un empleo creciente y un esfuerzo de estandarizarlos. Es el UAS7 se considera

el cuestionario de referencia en la evaluación de la actividad de la enfermedad y se ha propuesto que sea el objetivo de éxito terapéutico ($UAS \leq 6$) (Lima *et al.*, 2017).

■ 12.4. DERMATITIS ATÓPICA

La dermatitis atópica (DA), eccema atópico o eccema endógeno es una enfermedad pruriginosa crónica que cursa en brotes. Es una de las enfermedades dermatológicas más frecuentes que puede afectar hasta un 20 % de la población infantil y un 2-8 % de los adultos en la mayoría de países a lo largo del mundo. La prevalencia de la DA ha aumentado en los últimos años. Se considera que menos de un 10 % son formas graves de la enfermedad. La DA suele comenzar en la infancia y los casos más severos pueden permanecer en la edad adulta. No obstante, aproximadamente un tercio de los casos que se presentan en adultos comienzan en la edad adulta. La DA con frecuencia es el primer paso en el desarrollo de otras enfermedades atópicas como la rino conjuntivitis, asma o alergia a alimentos (Wollenberg *et al.*, 2018).

Además de la clara influencia genética, existen hallazgos característicos en su fisiopatología: diferenciación inmunológica hacia linfocitos Th2 con un aumento de IgE; aumento de mediadores inflamatorios; defectos de la función barrera debido a anomalías del metabolismo lipídico y la formación de proteínas estructurales como la filagrina; colonización por *Staphylococcus aureus* o *Malassezia*; e influencias psicósomáticas entre otros.

Tabla 1.

CRITERIOS DIAGNÓSTICOS DE HANIFIN Y RAJKA

Criterios Diagnóstico de Hanifin y Rajka para Dermatitis Atópica

Criterios mayores: Debe tener tres o más:

1. Prurito
2. Distribución y morfología típica
 - Liquenificación o linealidad flexional en adultos
 - Compromiso extensor y facial en infantes y niños
3. Dermatitis crónica o crónicamente recidivante
4. Historia personal o familiar de atopias (asma, rinitis alérgica, dermatitis atópica)

Criterios menores: Debe tener tres o más de:

1. Xerosis
2. Ictiosis, hiperlinealidad palmar, o queratosis pilar
3. Reactividad a pruebas cutáneas inmediata (tipo 1)
4. IgE sérico elevado
5. Comienzo a temprana edad

Tabla 1. (continuación)

CRITERIOS DIAGNÓSTICOS DE HANIFIN Y RAJKA

Criterios Diagnóstico de Hanifin y Rajka para Dermatitis Atópica

Criterios menores: Debe tener tres o más de:

6. Tendencia a infecciones cutáneas (especialmente *S. aureus* y herpes simplex) o defecto de la inmunidad mediada por células.
7. Tendencia a dermatitis en manos y pies inespecíficas
8. Eczema del pezón
9. Queilitis
10. Conjuntivitis recurrente
11. Pliegue infraorbitario de Dennie-Morgan
12. Queratocono
13. Cataratas anterior subcapsular
14. Oscurecimiento orbitario (ojeras)
15. Eritema o palidez facial
16. Pitiriasis alba
17. Pliegues anteriores del cuello
18. Prurito al sudar (hipersudoración)
19. Intolerancia a lanas y solventes orgánicos
20. Acentuación perifolicular
21. Intolerancia a alimentos
22. Curso influenciado por factores ambientales o emocionales
23. Dermografismo blanco

El diagnóstico de la DA es clínico y no existe un biomarcador para su diagnóstico. Son los criterios de Hanifin y Rajka los más comúnmente aceptados y empleados para establecer el diagnóstico (Hanifin y Rajka, 1980) (Tabla 1).

Una vez establecido el diagnóstico la gravedad de la enfermedad se determina evaluando signos objetivos y síntomas subjetivos. Es por ello que es aconsejable emplear para evaluar la gravedad índices compuestos que incluyan la parte subjetiva como el prurito y la alteración del sueño, así como los signos de la enfermedad. El índice compuesto más empleado es el SCORAD (*Scoring of Atopic Dermatitis*) que incluye la extensión de las lesiones, gravedad de los signos (eritema, edema, costras, liquenificación, excoriación, sequedad) y repercusión de la calidad de vida a nivel de alteraciones del sueño y prurito. El máximo SCORAD es de 103. Se considera enfermedad leve cuando un paciente presenta un SCORAD menor de 25; moderada de 25 a 50 y grave mayor de 50.

Un índice de gravedad empleado en gran parte de los ensayos clínicos y en la práctica clínica es el EASI (*Eczema Area and Severity Index*), que se considera el instrumento de medida más adecuado para evaluar los signos de la enfermedad, a pesar de que no recoge los síntomas subjetivos. Es un índice que oscila de 0 a 72 y recoge la extensión de las lesiones, signos (eritema, induración, exoriación y liquefacción). En los criterios de inclusión de los ensayos clínicos de DA de moderada a grave se establece un corte de EASI de al menos 16.

Los investigadores y clínicos también utilizan el índice IGA (*Investigator Global Assessment*) para valorar la gravedad de una forma global y rápida. El IGA oscila de 0 a 4 (0=aclorado; 1=casi aclarado; 2=leve; 3= moderado; 4= grave).

Existe una necesidad creciente del empleo en la DA de los resultados en salud expresados y percibidos por los pacientes, de cara a estandarizar los objetivos terapéuticos a nivel de la obtención de una mejora importante de la calidad de vida así como el gasto eficiente de recursos. Existe una iniciativa denominada HOME (*Harmonising Outcome Measures of Eczema*) cuyo objetivo es homogeneizar los objetivos que han de recogerse en los ensayos clínicos (Schmitt *et al.*, 2011; Chalmers *et al.*, 2016). Uno de los dominios con gran importancia es al que se refiere a los síntomas percibidos por los pacientes y la calidad de vida, para los que se emplean los PRO.

A nivel de síntomas referidos por los pacientes el panel de expertos prefiere por unanimidad el instrumento POEM (*Patient-Oriented Eczema Measure*). Otros instrumentos utilizados son SA-EASI (*Self-Administered Eczema Area and Severity Index*), Itch Severity Scale, PO-SCORAD (*Patient-Oriented SCOring Atopic Dermatitis index*), *Atopic Dermatitis Quickscore* y *Nottingham Eczema Severity Score*.

El instrumento de medida POEM lo cumplimenta el paciente o el familiar y consiste en siete preguntas sobre prurito, sueño, sangrado, humedad, fisuración, descamación y sequedad de la piel en los últimos siete días. Cada pregunta tiene un valor de 0 a 4 y el rango de POEM es de 0 a 28. Se considera significativo una disminución de 4 puntos en el POEM. La severidad de la DA en relación al POEM es:

- 0-2: aclaramiento o casi aclaramiento.
- 3-7: leve.
- 8-16: moderado.
- 17-24: severo.
- 25-28: muy severo

En el dominio de calidad de vida existen dificultades en su definición y su multidimensionalidad. La iniciativa HOME explora fundamentalmente la calidad de vida en la población adulta con DA, ya que la validación de las escalas en esta dimensión en la edad infantil no está correspondientemente validada. Se pone de relieve la importancia de la validación en las diferentes lenguas, así como limitar el número de preguntas para su factibilidad. Las emociones, el impacto del tratamiento y las rela-

ciones personales son esenciales en este dominio. La iniciativa HOME analizó los instrumentos DLQI, QoLIAD (*Quality of Life Index for Atopic Dermatitis*) y Skindex, no estableciendo ningún consenso ni recomendación en su utilización.

Un estudio de la EADV (European Academy of Dermatology and Venereology) TF (Task Force) en calidad de vida, contiene información (validación, experiencia, peculiaridades de determinados grupos, limitaciones y recomendaciones) sobre los instrumentos de evaluación de la calidad de vida relacionada con la salud de los

Tabla 2.

INSTRUMENTOS GENÉRICOS UTILIZADOS EN DERMATITIS ATÓPICA

EuroQoL-5D (EQ-5D)
Short-Form-36 (SF-36)
Short-Form-12 (SF-12)
Quality of Life in Children Revised (KINDL-R)
Quality of Life Scale for Children (AUQEI)
Health Utilities Index Mark 3 (HUI3)
Child Health Questionnaire (CHQ)
DISABKIDS Chronic Generic Measure (DCGM) 37
World Health Organization Quality of Life (WHOQOL)-26
Sickness Impact Profile (SIP)

Tabla 3.

INSTRUMENTOS ESPECÍFICOS DE DERMATOLOGÍA UTILIZADOS EN DERMATITIS ATÓPICA

Children Dermatology Life Quality Index (CDLQI)
Family Dermatology Life Quality Index (FDLQI)
Skindex-29
Skindex-16
Skindex-17
Dermatology Life Quality Index (DLQI)
Impact of Chronic Skin Disease on Daily Life' (ISDL)
Freiburger Quality Of Life Index (FLQA-d)
German instrument for the assessment of quality of life in skin diseases(DIELH)
VQ-Dermato
Adjustment to Chronic Skin Diseases Questionnaire (Subscale 5: Impact on Quality of Life)
Chronic Skin Disease Questionnaire (CSDQ)
Toddler QOL Survey

pacientes con DA y su familia. Esta información ayuda a los dermatólogos e investigadores de cara a elegir el mejor instrumento de medida y evitar errores comunes (Chernyshov *et al.*, 2017).

Tabla 4.

INSTRUMENTOS EN DERMATITIS ATÓPICA

The Infant's Dermatitis Quality of Life Index (IDQoL)
Atopic Dermatitis Burden Scale for Adults (ABS-A)
The Quality of Life Index for Atopic Dermatitis (QoLIAD)
The Dermatitis Family Impact (DFI)
Quality of life in Primary Caregivers of children with Atopic Dermatitis (QPCAD)
DISABKIDS Atopic Dermatitis Module (ADM)
Childhood Atopic Dermatitis Impact Scale (CADIS)
Parents' Index Quality of Life-Atopic Dermatitis (PIQoL-AD)
The Quality of Life in Parents of Children with Atopic Dermatitis
The Childhood Impact of Atopic Dermatitis (CIAD)
The Eczema Disability Index (EDI)

A nivel de los instrumentos de medición calidad de vida relacionada con la salud en los pacientes con DA y familiares se utilizan instrumentos genéricos, específicos de enfermedades dermatológicas y específicos de DA (Tablas 2, 3 y 4).

El instrumento más utilizado para evaluar la calidad de vida en la edad infantil en DA es el CDLQI (*Children's Dermatology Life Quality Index*). Al ser un grupo muy heterogéneo no es fácil establecer en líneas generales un instrumento que sea adecuado para los diferentes escenarios. A nivel del impacto secundario que presentan los familiares de los pacientes afectados de enfermedades dermatológicas el instrumento más utilizado es el FDLQI (*Family Dermatology Life Quality Index*).

Los instrumentos específicos de DA son IDQoL (*Infant's Dermatitis Quality of Life Index*) y DFI (*Dermatitis Family Impact*).

Varios estudios han demostrado la correlación entre la gravedad de la dermatitis atópica a nivel de EASI, SCORAD y PO-SOCRAD y los diferentes instrumentos de calidad de vida como SKINDEX, DLQI, DFI, IDQoL y CDLQI, denotando la importancia de estos últimos.

Algunas de las recomendaciones de la TASK FORCE son:

- Solo se deberían utilizar instrumentos de la medición de la calidad de vida validados.
- En investigación clínica se debería utilizar la combinación de un cuestionario genérico y uno específico dermatológico o de DA.
- En práctica clínica se recomienda un instrumento específico dermatológico y/o de DA.
- Debido a la importancia del impacto familiar de la DA se debería estudiar la calidad de vida de la familia, especialmente en niños pequeños.

Heinl y colaboradores analizaron los PRO empleados en los diferentes ensayos clínicos de DA desde el año 2000 hasta 2014, así como la idoneidad de incluirlos como objetivos primarios (Heinl *et al.*, 2016). En los adultos el DLQI es el instrumento más utilizado; y en niños el CDLQI (*Children Dermatology Life Quality Index*) y IDQOL (*Infant'S Dermatitis Quality of Life Index*). La calidad de vida de los cuidadores de los niños se exploró mediante el cuestionario DFI (*Dermatitis Family Impact*). Fueron conclusiones interesantes del estudio: la gran frecuencia en el empleo de algún PRO (90 %) si bien una minoría se centran en la calidad de vida; el aumento de utilización de PRO en los últimos años; y la escasa presencia de la medición de la calidad de vida de los familiares.

■ 12.5. HIDROSADENITIS SUPURATIVA

La hidrosadenitis supurativa (HS) o hidradenitis es una enfermedad inflamatoria inmunomediada (IMID), progresiva, crónica, dolorosa y recurrente. Se caracteriza por una disregulación inmune que da lugar a inflamación aguda o crónica, causante de daños a tejidos y órganos. Cursa con lesiones inflamatorias, profundas y dolorosas (nódulos, abscesos, fistulas, cicatrices) en las áreas del cuerpo que albergan glándulas apocrinas, siendo las más frecuentes las regiones axilares, inguinales, anogenitales y submarias en mujeres. Cursa en brotes de duración variable que asocian dolor y supuración (Martorell *et al.*, 2015).

En su patogenia están involucrados factores genéticos. Existen antecedentes familiares en un 30-40 % de los casos, con cuadros más tempranos, graves y refractarios al tratamiento. Se han identificado mutaciones en la gamma-secretasa asociadas a la enfermedad. El tabaquismo y la obesidad son factores desencadenantes bien conocidos. La infección bacteriana se desarrolla como una colonización secundaria de las lesiones que puede agravar la enfermedad, pero no es el factor causante.

La prevalencia se estima entre un 1-4 % a nivel mundial y es más frecuente en mujeres, tras la adolescencia, si bien existen casos prepuberales.

El diagnóstico de la enfermedad es clínico, no existiendo una prueba de laboratorio o de imagen válida para establecer el diagnóstico. En general es una

enfermedad poco conocida para los especialistas no dermatólogos y médicos generalistas y por ello existe un retraso en el diagnóstico de varios años que conlleva una merma importante en la calidad de vida de los pacientes y una pérdida de oportunidad de tratamientos precoces de la enfermedad. En nuestro medio se está trabajando entre diversos niveles asistenciales como atención primaria, servicios de urgencias, dermatología, cirugía general, entre otros, de cara a efectuar un diagnóstico precoz. De hecho, muchas instituciones sanitarias entre las que se encuentran la Academia Española de Dermatología y Venereología (AEDV) y la Sociedad Española de Farmacia Hospitalaria (SEFH) han colaborado en la realización del Proyecto Hércules, Iniciativa Estratégica para la HS, para dar respuesta a las necesidades específicas de los pacientes afectados por HS. Los objetivos de este proyecto son:

- Desarrollar, implementar y difundir una ruta clínica asistencial específica para la patología.
- Promover el correcto diagnóstico de la patología.
- Mejorar la detección precoz de la HS.
- Optimizar la coordinación y garantizar la continuidad asistencial.
- Impulsar la evidencia científica para la optimización de los resultados clínicos.

Por su relación con la presencia de artritis, enfermedad inflamatoria intestinal y alteraciones psicológicas/psiquiátricas, así como el aumento de riesgo de eventos endocrino-metabólicos, además de la propia naturaleza de las lesiones, lo apropiado es un abordaje multidisciplinar.

Evaluación de la HS

A pesar de que existen pocos instrumentos para la correcta evaluación de la HS desarrollaremos los más relevantes (Revuz, 2015; Alavi *et al.*, 2015):

Clasificación de Hurley

Hurley, desde una perspectiva fundamentalmente quirúrgica, define los niveles de gravedad en tres:

- Grado I: Absceso/o único o múltiples sin fístulas ni procesos cicatrizales.
- Grado II: Abscesos recidivantes con formación de fístulas y de cicatrices hipertróficas. Lesión única o lesiones múltiples separadas por piel sana.
- Grado III: Afectación difusa o fístulas interconectadas o abscesos sobre toda la extensión de la zona afecta.

Esta clasificación clínica es sencilla y reproducible y es adecuada para determinar el circuito asistencial en HS. No obstante, presenta determinadas limitaciones

al ser estática y por tanto no adecuada para determinar la evolución de la enfermedad ni valorar la respuesta al tratamiento médico. No permite valorar la globalidad de la enfermedad al valorar la afectación de una región determinada.

Puntuación de Sartorius y Sartorius modificada

Es más global y precisa que la clasificación de Hurley. Combina el número de localizaciones, (3 puntos por localización), número y gravedad de las lesiones, distancia de dos lesiones y si existe piel sana entre las lesiones. Se anota complementariamente la percepción del paciente en una escala analógica visual sobre el dolor de la lesión más dolorosa.

La variabilidad interobservador en la puntuación de Sartorius modificada ha demostrado ser baja, y se correlaciona positivamente con la presencia de factores riesgo y con otras mediciones de gravedad (como el DLQI). Sin embargo, su aplicabilidad se encuentra limitada en casos graves, en los que las lesiones separadas acaban confluyendo. Además, aunque se trata de una herramienta más dinámica que el Hurley score, incluye lesiones difícilmente modificables con tratamiento médico (distancia entre 2 lesiones relevantes).

HS-PGA (Physician Global Assessment). Evaluación Global de la Enfermedad

Es uno de los modelos de clasificación actuales más utilizados en ensayos clínicos para evaluar la respuesta terapéutica a tratamientos médicos. Clasifica la gravedad de la enfermedad en categorías, teniendo en cuenta el total de abscesos, fístulas, nódulos inflamatorios y nódulos no inflamatorios presentes (sumando todas las zonas afectadas). El último HS-PGA desarrollado cataloga la gravedad en 6 grados. Es sencilla, rápida de realizar y dinámica. Permite hacer un seguimiento de la progresión de la enfermedad, pero tiene el inconveniente de que las zonas afectadas se valoran de forma conjunta.

HiSCR (Hidradenitis Suppurativa Clinical Response)

Más que un modelo de clasificación es un nuevo parámetro de medición de la respuesta al tratamiento médico para la HS, validado, que pretende cuantificar la gravedad de la enfermedad y determinar un objetivo clínico. El HiSCR se define como una reducción $\geq 50\%$ en el recuento de lesiones inflamatorias (suma de abscesos y nódulos inflamatorios), sin incremento en el número de abscesos o fístulas drenantes respecto a la situación basal.

El HiSCR es, por lo tanto, un objetivo clínico que se basa en el recuento total de lesiones inflamatorias que presenta un paciente afecto de HS en un momento determinado. De esta forma, se puede definir también el porcentaje de reducción de abscesos y nódulos inflamatorios respecto a basal: AN50, AN75 y AN100 (reducción del 50 %, 75 % y 100 % en el número de nódulos inflamatorios frente a la situación basal, respectivamente).

Las evaluaciones del paciente. PRO

Estas evaluaciones son muy utilizadas como criterios secundarios en los estudios y han demostrado su utilidad en la práctica clínica.

A nivel de los PRO, se centra fundamentalmente en la evaluación del dolor: intensidad del dolor en escala analógica visual o escala de 0 a 10; número de días con dolor en un período de tiempo (un mes). Estos PRO se utilizan con frecuencia en los ensayos clínicos.

Los instrumentos de medida de la calidad de vida más empleados en HS son los dermatológicos no específicos de la enfermedad como el DLQI.

Existen estudios en los que se emplean asimismo los instrumentos genéricos como el SF-36. La HS tiene un impacto en la calidad de vida muy alto como han demostrado varios estudios, de hecho se ha evidenciado una repercusión mayor que la dermatitis atópica y la psoriasis. Esta merma viene dada por varios factores como son: la cronicidad y aparición de recurrencias; aparición de lesiones dolorosas e incontrolables; lesiones visibles que distorsionan la imagen corporal al paciente; presencia de cicatrices inestéticas; sensación de vergüenza y miedo a que le vean o huelan las lesiones; afectación de la esfera sexual; limitación de las actividades diarias, el movimiento y el vestido; sensación de dependencia (Esmann y Jemec, 2011). Así mismo los pacientes experimentan con frecuencia problemas económicos al verse alterada la esfera laboral (Deckers y Kimball, 2016).

Kouris et al efectuaron un estudio sobre la calidad de vida y la repercusión psicosocial en una serie de 94 pacientes con HS (Kouris *et al.*, 2016). y encontraron mayor frecuencia de depresión, ansiedad, soledad y aislamiento social, y baja autoestima que los controles. A nivel de DLQI más de un 50 % presentaron cifras de DLQI mayor de 10 (repercusión grave o muy grave de la calidad de vida). El sexo femenino y una gravedad clínica Hurley III se asociaron a peor calidad de vida.

■ REFERENCIAS

- ABENI, D., PICARDI, A., PASQUINI, P., MELCHI, C. F. y CHREN, M. M. (2002). Further evidence of the validity and reliability of the Skindex-29: an Italian study on 2.242 dermatological outpatients. *Dermatology*, 204, pp. 43-49.
- ALAVI, A., ANOOSHIRVANI, N., KIM, W. B., COUTTS, P. y SIBBALD, R. G. (2015). Quality-of-life impairment in patients with hidradenitis suppurativa: a Canadian study. *Am J Clin Dermatol*, 16, pp. 61-65.
- BAIARDINI, I. y BRAIDO, F. y BRANDI, S. (2006). Allergic diseases and their impact on quality of life. *Ann Allergy Asthma Immunol*, 97, pp. 419-428; quiz 429-30, 476.
- BASRA, M. K., FENECH, R., GATT, R. M., SALEK, M. S. y FINLAY, A. Y. (2008). The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results. *Br J Dermatol*, 159(5), pp. 997-1035.
- BRONSARD, V., PAUL, C., PREY, S., PUZENAT, E., GOURRAUD, P. A. y ARACTINGI, S. (2010). What are the best outcome measures for assessing quality of life in plaque type psoriasis? A systematic review of the literature. *J Eur Acad Dermatol Venereol*, 24, pp. 17-22.

- CHALMERS, R. J. (2015). Assessing psoriasis severity and outcomes for clinical trials and routine clinical practice. *Dermatol Clin*, 33, pp. 57-71.
- CHALMERS, J. R., SIMPSON, E., APFELBACHER, C. J., THOMAS, K. S., VON KOBYLETZKI, L., y SCHMITT, J. (2016). Report from the fourth international consensus meeting to harmonize core outcome measures for a topic eczema/dermatitis clinical trials (HOME initiative). *Br J Dermatol*, 175, pp. 69-79.
- CHERNYSHOV, P. V., TOMAS-ARAGONES, L., MANOLACHE, L., MARRON, S. E., SALEK, M. S., POOT, F. *et al.* (2017). Quality of life measurement in atopic dermatitis. Position paper of the European Academy of Dermatology and Venereology (EADV) Task Force on quality of life. *J Eur Acad Dermatol Venereol*, 31, pp. 576-593.
- CHREN, M. M. (2010). Interpretation of quality-of-life scores. *J Invest Dermatol*, 130(5), pp. 1207-1209.
- DAUDÉN, E., PUIG, L., FERRÁNDIZ, C., SÁNCHEZ-CARAZO, J. L. y HERNANZ-HERMOSA, J. M. (2016). Spanish Psoriasis Group of the Spanish Academy of Dermatology and Venereology. Consensus document on the evaluation and treatment of moderate-to-severe psoriasis: Psoriasis Group of the Spanish Academy of Dermatology and Venereology. *J Eur Acad Dermatol Venereol*, 30(Suppl 2), pp.1-18.
- DECKERS, I. E. y KIMBALL, A.B. (2016). The handicap of hidradenitis suppurativa. *Dermatol Clin*, 34, pp. 17-22.
- ESMANN, S. y JEMEC, G. B. (2011). Psychosocial impact of hidradenitis suppurativa: a qualitative study. *Acta Derm Venereol*, 91, pp. 328-332.
- HANIFIN, J. M. y RAJKA, G. (1980). Diagnostic features of atopic dermatitis. *Acta Derm Venereol*, 92(Suppl.), pp. 44-47.
- HEINL, D., CHALMERS, J., NANKERVIS, H. y APFELBACHER, C. J. (2016). Eczema Trials: Quality of Life Instruments Used and Their Relation to Patient-reported Outcomes. A Systematic Review. *Acta Derm Venereol*, 96, pp. 596-601.
- JACOBSON, C. C. y KIMBALL, A. B. (2004). Rethinking the Psoriasis Area and Severity Index: the impact of a reashould be increased. *Br J Dermatol*, 151(2), pp. 381-387.
- JONES-CABALLERO, M., PEÑAS, P. F., GARCÍA-DIEZ, A., CHREN, M. M. y BADÍA, X. (2002). [The Spanish version of Skindex-29. An instrument for measuring quality of life in patients with cutaneous diseases]. *Med Clin (Barc)*, 118, pp. 5-9.
- KOTI, I., WELLER, K., MAKRIS, M. *et al.* (2013). Disease activity only moderately correlates with quality of life impairment in patients with chronic spontaneous urticaria. *Dermatology*, 226, pp. 371-379.
- KOURIS, A., PLATSIDAKI, E., CHRISTODOULOU, C., EFSATHIOU, V., DESSINIOTI, C., TZANETAKOU, V. *et al.* (2016). Quality of Life and Psychosocial Implications in Patients with Hidradenitis Suppurativa. *Dermatology*, 232, pp. 687-691.
- LEWIS, V. y FINLAY, A. Y. (2004). 10 years experience of the Dermatology Life Quality Index (DLQI). *J Invest Dermatol Symp Proc*, 9, pp. 169-180.
- LIMA, H., GOODERHAM, M., DUTZ, J., LYNDE, C., CHAPDELAIN, H, ELLIS, A. *et al.* (2017). Management of chronic spontaneous urticaria (CSU): a treat to target approach using a patient reported outcome. *Allergy ASTHMA CLIN IMMUNOL*, 13, p. 38.
- MARTORELL, A., GARCÍA-MARTÍNEZ, F. J., JIMÉNEZ-GALLO, D., PASCUAL, J. C., PEREYRA-RODRÍGUEZ, J., SALGADO, L., *et al.* (2015). An Update on Hidradenitis Suppurativa (Part I): Epidemiology, Clinical Aspects, and Definition of Disease Severity. *Actas Dermosifiliogr*, 106, pp. 703-715.
- MAURER, M., ORTONNE, J. P. y ZUBERBIER, T. (2009). Chronic urticaria: an internet survey of health behaviours, symptom patterns and treatment needs in European adult patients. *Br J Dermatol*, 160, pp. 633-641.
- MAURER, M., WELLER, K., BINDSLEV-JENSEN, C. *et al.* (2011). Unmet clinical needs in chronic spontaneous urticaria. A GA(2)LEN task force report. *Allergy*, 66, pp. 317-330.

- MOESTRUP, K., GHAZANFAR, M. N. y THOMSEN, S. F. (2017). Patient-reported outcomes (PRO) in chronic urticaria. *Int J Dermatol*, 56, pp. 1342-1348.
- PASCOE, V. L., ENAMANDRAM, M., COREY, K. C., CHENG, C. E., JAVORSKY, E. J., SUNG, S. M. *et al.* (2015). Using the Physician Global Assessment in a Clinical Setting to Measure and Track Patient Outcomes. *JAMA Dermatol*, 151(4), pp. 375-381.
- PUZENAT, E., BRONSARD, V., PREY, S., GOURRAUD, P. A., ARACTINGI, S., BAGOT, M. *et al.* (2010). What are the best outcome measures for assessing plaque psoriasis severity? A systematic review of the literature. *J Eur Acad Dermatol Venereol*, 24(Suppl 2), pp. 10-16.
- REVUZ J. (2015). Clinical evaluation of the severity of hidradenitis suppurativa. *Ann Dermatol Venereol*, 142, pp. 729-735.
- ROBINSON, A., KARDOS, M. y KIMBALL, A. B. (2012). Physician Global Assessment (PGA) and Psoriasis Area and Severity Index (PASI): Why do both? A systematic analysis of randomized controlled trials of biologic agents for moderate to severe plaque psoriasis. *J Am Acad Dermatol*, 66, pp. 369-375.
- SCHMITT, J., LANGAN, S., STAMM, T. y WILLIAMS, H. C. (2011). Core outcome domains for controlled trials and clinical record keeping in eczema: international multiperspective Delphi consensus process. *J Invest Dermatol*, 131, pp. 623-630.
- WELLER, K., ZUBERBIER, T. y MAURER, M. (2015). Clinically relevant outcome measures for assessing disease activity, disease control and quality of life impairment in patients with chronic spontaneous urticaria and recurrent angioedema. *Curr Opin Allergy Clin Immunol*, 15, pp. 220-226.
- WOLLENBERG, A., BARBAROT, S., BIEBER, T., CHRISTEN-ZAECH, S., DELEURAN, M., FINK-WAGNER, A. *et al.* (2018). Consensus-based European guidelines for treatment of atopic eczema (atopic dermatitis) in adults and children: part I. *J Eur Acad Dermatol Venereol*, 32, pp. 657-682.
- WORLD HEALTH ORGANIZATION. (2016). *Global Report on Psoriasis*. World Health Organisation. Disponible en: <http://www.who.int/ncds/management/psoriasis/en/#> (Último acceso 10 de julio de 2018).



13

**RESULTADOS EN SALUD
EN APARATO DIGESTIVO: ENFERMEDAD
INFLAMATORIA INTESTINAL
E INTESTINO IRRITABLE**

DIEGO CASAS DEZA, MANUEL GÓMEZ BARRERA,
SANTIAGO GARCÍA LÓPEZ

En el presente capítulo se van a tratar la repercusión en resultados en salud de dos entidades relevantes en patología digestiva, el síndrome de intestino irritable (SII) y la enfermedad inflamatoria intestinal (EII). En primer lugar se van a comentar resultados económicos y en segundo lugar clínicos o referidos a la calidad de vida de los pacientes.

■ 13.1. RESULTADOS ECONÓMICOS EN EII Y SII

El análisis de resultados en salud desde un punto de vista económico se engloba principalmente en torno a tres figuras: la *evaluación económica*, los *estudios de coste de la enfermedad* o los *estudios de impacto presupuestario*. Las tres dimensiones comentadas son perfectamente compatibles entre sí y en ocasiones se deben contemplar todas ellas para concretar al máximo posible los resultados económicos en salud.

Comenzando con la *evaluación económica*, se debe señalar que su fin último es la selección de aquellas opciones que tengan un impacto sanitario más favorable (Sacristán, 1995). En este sentido, se deben analizar en profundidad los aspectos relacionados con las opciones, incidiendo en dos aspectos, los costes de la aplicación de las diferentes alternativas y los resultados que se esperaría obtener de cada una de ellas. Estos resultados se pueden medir desde diferentes puntos de vista, uno de ellos es clínico, otro está relacionado con la calidad de vida y otro es económico. La evaluación económica de medicamentos puede definirse como la determinación de la eficiencia de un tratamiento farmacológico y su comparación con la de otras opciones, con el fin de seleccionar la alternativa más eficiente, es decir, la que presente un mejor balance entre sus costes y sus resultados (Drummond *et al.*, 2005).

Para realizar una evaluación económica es necesario identificar los costes más relevantes de las opciones que se están considerando. Se pueden establecer además diversos tipos de costes.

Los costes directos generan transacciones económicas o monetarias. Dentro de ellos existen los costes directos sanitarios (cuidados hospitalarios, tratamiento farmacológico) y los costes indirectos no sanitarios (desplazamiento del paciente, cuidados en casa). Los indirectos son los costes producidos por la morbilidad o la mortalidad

prematura asociadas a una enfermedad. Los costes indirectos sanitarios suelen estar relacionados con el consumo adicional de servicios sanitarios a consecuencia de enfermedades que aparecen debido a la intervención evaluada. Por otro lado, los costes indirectos no sanitarios hacen referencia a la pérdida de productividad del paciente debido al tratamiento, es decir, su merma en la capacidad de generar ingresos (Grover *et al.*, 2003).

De las diferentes formas de cuantificar los resultados de las intervenciones analizadas surgen los diferentes tipos de técnicas de evaluación económica, los estudios de minimización de costes, los de coste-efectividad, los de coste-utilidad y los de coste-beneficio (Drummond *et al.*, 2005). El uso de estos estudios ha derivado en que genéricamente se denomine a todos los estudios de evaluación económica como estudios de coste-efectividad, y se asimile su terminología de resultados al resto, si bien es oportuno mostrar las diferencias existentes entre ellos. Los estudios de minimización de costes suponen que las alternativas evaluadas son iguales en cuanto a resultados en salud y se realiza la evaluación únicamente atendiendo a los costes, siendo obviamente la opción preferida será la menos costosa. En los estudios de coste-efectividad los resultados de las alternativas se miden en unidades clínicas, siendo la más aceptada en términos de metodología los años de vida ganados (AVG), pudiéndose emplear otro tipo de variables como puedan ser la proporción de pacientes curados, en todo caso siempre se trata de unidades físicas. En los estudios de coste-utilidad los resultados se miden con una variable subjetiva que son los años de vida ajustados por calidad (AVAC o QALYs), que hacen referencia a los años de vida que le quedan al paciente, tras una intervención sanitaria, pero sujetos a una determinada calidad de vida obteniéndose los datos a partir de cuestionarios de calidad de vida. Finalmente, los estudios de coste-beneficio en los que los resultados se miden en unidades monetarias, son los más rigurosos desde un punto de vista económico, pero los que menos gustan a los profesionales sanitarios, la magnitud de resultados más empleada es la disponibilidad a pagar (DAP o WTP), la cual hace referencia a lo que estaría dispuesto a pagar un paciente por un tratamiento y se obtiene por medio de cuestionarios a pacientes y a sujetos sanos.

Una vez establecidas las alternativas a evaluar, que siempre deben ser al menos dos, se procede a la comparación. Para cada alternativa se puede calcular el coste efectividad medio (CEM) definido como el cociente entre el coste y la efectividad de cada alternativa. Obviamente serán deseables alternativas con un menor valor de CEM. De todos modos, el resultado clave en evaluación económica es el ratio de coste efectividad incremental (RCEI) en estudios de coste efectividad o el ratio de coste utilidad incremental (RCUI) que se define de la siguiente manera (Rovira y Antoñanzas, 1998):

$$\text{RCEI} = (\text{Ca}-\text{Cb})/(\text{Ea}-\text{Eb})$$

Donde Ca y Cb son el coste y Ea y Eb los resultados de las opciones A y B respectivamente. El RCEI informa del coste adicional que supone una unidad adicional de beneficio. Es de señalar que puede tomar valores negativos, en ese caso

siempre existirá una alternativa que, de mejores resultados sanitarios a un menor coste, o positivos, en ese caso existirán alternativas que aporten mejores resultados sanitarios a un mayor coste económico. Por ejemplo, en España, se considera una alternativa más eficiente si es capaz de aportar un AVG adicional frente a otra alternativa menos efectiva con un incremento en el coste inferior a 30.000 euros. En el RCU se consideran utilidades en lugar de efectividades y se asimilan los 30.000 euros/AVG a 30.000 euros/AVAC. Señalar que estos valores están en permanente revisión y siempre deben considerarse en un contexto global económico y sanitario y que se han adaptado del umbral establecido por el National Institute of Health and Care Excellence (NICE), organismo regulador del Reino Unido que fija un umbral de 30.000£/AVAC para considerar que un nuevo tratamiento supone una asignación eficiente de recursos.

Una limitación que presenta la evaluación económica de medicamentos y los estudios económico sanitarios y los estudios económicos en general, es el problema de la generalización de sus conclusiones. Aunque puede parecer tentador, no se deben extrapolar las conclusiones de un estudio económico de forma general (Mason, 1997). Para realizar el estudio de costes y resultados, que toda evaluación supone, se realizan una serie de suposiciones que intentan simplificar la realidad de la práctica clínica, creando lo que se llama un modelo. Esta asunción de suposiciones hace que los modelos no sean habitualmente extrapolables de un lugar a otro lo que dificulta la adopción universal de las conclusiones. Estas suposiciones, que se realizan al hacer evaluaciones económicas, dificultan su adopción general, como se ha comentado. Para solucionar este escollo, es necesario apoyarse en la metodología estandarizada que, aparte de facilitar el trabajo, hace que, si bien las conclusiones no pueden ser generalizables, los modelos sí lo sean y se puedan aplicar en diferentes situaciones. Para ello los modelos deben tener una validez interna, que se la aporta, un punto que toda evaluación económica debe llevar incorporado, el análisis de sensibilidad.

El análisis de sensibilidad surge por un aspecto de capital importancia en estudios económicos como es la incertidumbre (Rubio-Terrés *et al.*, 2004). A la hora de realizar los estudios es fácil apoyarse en los datos disponibles, o accesibles, en el momento de realizar el estudio. La obtención de los datos siempre está sometida a ciertos errores, éstos pueden proceder de la aleatoriedad de los datos, de las asunciones del modelo o de los errores no aleatorios o sesgos que toda recogida de datos puede llevar asociada. El razonamiento básico de los análisis de sensibilidad consiste en someter los resultados del modelo básico, con el que se ha trabajado, a ciertas variaciones en dicho modelo y ver si las conclusiones iniciales obtenidas, a partir de los primeros resultados, se mantienen. El análisis de sensibilidad es de capital importancia, también, para hacer análisis en grupos de población concretos e intentar extrapolar conclusiones de un grupo de población a otro.

Finalmente, un aspecto que se debe comentar es la importancia del punto de vista o perspectiva del estudio. Por ejemplo, si se ciñe un estudio a un punto de vista hospitalario incluirá sólo costes hospitalarios y excluirá otros relevantes como

la pérdida de productividad laboral del paciente. Los posibles puntos de vista a emplear son el de la sociedad, el Ministerio de Sanidad, el gobierno, el centro hospitalario, la institución financiadora o el paciente. Siempre que sea posible se debe aplicar el punto de vista de la sociedad, ya que es el más amplio y el que incluye todos los costes (Eisenberg, 1989). Por otro lado la Canadian Agency for Drugs and Technologies in Health (CADTH) ya recomendaba en el año 1997 utilizar varios puntos de vista, siempre que sea posible (Canadian Coordinating Office for Health Technology Assessment, 1997).

Como se ha comentado, el resultado de los estudios coste efectividad se expresa siempre en términos incrementales (Rovira y Antoñanzas, 1998). Se interpreta como el precio que hay que pagar por curar a una persona más de las que se curarían con la terapia más barata. La pregunta siguiente sería ¿el sistema puede pagar esa cantidad? La respuesta depende de varios factores, tal vez el más relevante sería el número de casos a tratar. Así, por ejemplo, una alternativa barata para tratar una enfermedad con una prevalencia elevada puede ser menos asumible, en términos de coste, que un tratamiento caro de una enfermedad rara, poco frecuente. Partiendo del principio de no negar a nadie un tratamiento, es evidente que no se puede aplicar siempre la alternativa más costosa, ya que a largo plazo sería imposible cumplir otras necesidades de tratamiento, por no disponer de presupuesto o no poder movilizar partidas.

Los estudios de *impacto presupuestario* tratan de ayudar en esta toma de decisiones (Brosa *et al.*, 2005). Suponen un paso adelante con respecto a los estudios de coste-efectividad. Tratan de estimar la población prevalente, o susceptible de tratar, y se intenta analizar el coste y la efectividad esperada, si se generalizan las alternativas a dicha población. Es decir, se pretende conocer el coste global que presenta el tratamiento de todos los pacientes de una enfermedad con el arsenal terapéutico disponible. Es habitual valorar dos escenarios, uno de ellos con los tratamientos disponibles en el momento, y otro con la inclusión de un nuevo tratamiento; la comparación del impacto presupuestario de un escenario con el otro permite conocer el coste económico de la inclusión del nuevo tratamiento y colaborar en la toma de decisiones sobre la inclusión del nuevo tratamiento.

Como conclusión final a esta exposición metodológica, se puede comentar que, de forma similar a la investigación epidemiológica, el estudio de evaluación económica e impacto presupuestario mediante modelos presenta limitaciones en cuanto a la generalización de la evidencia al apoyarse en diferentes supuestos. No obstante, estos modelos deben de ser considerados como una forma válida y necesaria para aportar información relevante a la toma de decisiones en la elección de alternativas terapéuticas.

A continuación, se van a presentar ciertos ejemplos de la aplicación de estos conceptos a las patologías referidas al inicio, la EII y el SII. La selección de los artículos presentados responde únicamente a intereses didácticos y pretende mostrar ejemplos prácticos que permitan una mejor comprensión de los resultados económicos de los tratamientos sanitarios.

Tabla 1.

PRESENTACIÓN DE EJEMPLOS DE RESULTADOS ECONÓMICOS EN SALUD

Referencia	Tipo de estudio	Comparadores	Resultados	Conclusiones
Jean <i>et al.</i> (2018)	Revisión sistemática de estudios de coste efectividad en enfermedad en EII.	No procede.	Las terapias biológicas fueron dominantes en el 23 % de los análisis y fueron rentables de acuerdo con un umbral de 50.000 dólares canadienses por AVAC en el 41 % de los casos y en un 62 % si el umbral se sube hasta 100.000 dólares canadienses por AVAC.	El tratamiento con biológicos, inmunosupresores y 5-ácido amino salicílico puede considerarse una estrategia eficiente.
Buono, Carson y Flores (2017)	Carga de la enfermedad en SII en USA, CVRS y costes indirectos.	Pacientes de SII y controles.	Los pacientes presentaron menor CVRS (media SF-6D: 0,677 frente a 0,741; $p < 0,001$) y mayor ausentismo (5,1 % frente a 2,9 %; $p = 0,004$), presentismo (17,9 % frente a 11,3 %; $p < 0,001$) y pérdida total de productividad laboral (20,7 % frente a 13,2 %; $p < 0,001$), presentaron además un mayor coste indirecto de 2.486\$ (7.008 frente a 4.522; $p < 0,001$).	La presencia de EII se asocia con menor CVRS, mayores alteraciones en trabajo y mayores costes indirectos población no afectada
Fisher <i>et al.</i> (2016)	Coste utilidad de linaclotida en SII con estreñimiento.	Coste utilidad de linaclotida frente a antidepresivos en SII moderado a severo.	En un horizonte temporal de 5 años, los costes incrementales y AVAC generados con linaclotida fueron 659€ y 0,089, lo que implica un RCUI de 7.370€ por AVAC.	Linaclotida puede ser un tratamiento eficiente en el tratamiento de SII en este tipo de pacientes.

Tabla 1. (continuación)

PRESENTACIÓN DE EJEMPLOS DE RESULTADOS ECONÓMICOS EN SALUD

Referencia	Tipo de estudio	Comparadores	Resultados	Conclusiones
Brodzky <i>et al.</i> (2016)	Estudio de impacto presupuestario de la introducción de biosimilares en Bulgaria, República Checa, Hungría, Rumanía y Eslovaquia.	Escenario con la inclusión de biosimilar de infliximab frente a la no inclusión de biosimilar.	Frente al escenario inicial sin biosimilares, se produjo una reducción en el coste de 8,0 millones de € pautando únicamente el biosimilar a nuevos pacientes y de 16,9 millones permitiendo la sustitución de pacientes en tratamiento hasta en un 80 % de los casos. Ello implica que con estos ahorros se puede asumir la terapia con infliximab biosimilar para en 722 o 1.530 pacientes adicionales.	La introducción de infliximab biosimilar para tratar la enfermedad de Crohn puede compensar la inequidad en el acceso a la terapia biológica en enfermedad de Crohn entre los países de Europa Central y Oriental.
Andersson <i>et al.</i> (2011)	Coste efectividad en SII realizado en Estados Unidos.	Terapia de comportamiento cognitivo frente a foro de discusión.	Coste en dólares y resultados según GRS-IBS. RCEI de -16.806.	La terapia del comportamiento cognitivo presenta menor coste y mejores resultados para el paciente.
Juan <i>et al.</i> (2003)	Coste de enfermedad de Crohn en España.	Estudio descriptivo de costes y comparación con población estándar en CVRS.	El coste anual por paciente se estimó en 6.808€ con 2.104€ correspondientes a costes directos médicos y 4.704€ de costes indirectos. El tamaño del efecto aproximado en CVRS fue de 0,8 para salud general, función social, función de funcionamiento físico y vitalidad; 0,5 para el funcionamiento del rol: emocional, dolor corporal y salud mental; y 0,2 para la función física.	La enfermedad de Crohn causa una disminución significativa en la CVRS y un costo económico considerable para los pacientes y la sociedad, principalmente debido a hospitalizaciones y pérdida de productividad.

Notas: AVAC: Años de vida ajustados por calidad; CVRS: Calidad de vida relacionada con la salud; GRS-IBS: *Gastrointestinal Symptom rating scale-IBS version*; Eli: Enfermedad inflamatoria intestinal; SII: Síndrome de intestino irritable; RCEI: Ratio coste efectividad incremental; RCUi: Ratio coste utilidad incremental.

Fuente: Elaboración propia.

■ 13.2. RELEVANCIA EN LA PRÁCTICA CLÍNICA ¿POR QUÉ ENFERMEDAD INFLAMATORIA INTESTINAL Y SÍNDROME DE INTESTINO IRRITABLE?

La patología digestiva es muy importante desde muchos puntos de vista, incluido el económico. Uno de los motivos es su elevada prevalencia, que afecta a gran número de personas, lo que implica una importante afectación a la población, en términos de salud pero también en el coste. Además, algunas enfermedades digestivas son crónicas, lo que multiplica su impacto socioeconómico. Finalmente, algunas enfermedades digestivas son menos frecuentes, pero también crónicas y con un coste directo, no solamente en fármacos, especialmente importante. Son varios los ejemplos de enfermedades digestivas con impacto económico elevado y por tanto en las que es especialmente relevante analizar los resultados en salud. La patología oncológica digestiva y la hepatitis C están actualmente “en boca de todos”, por diversos motivos, pero quizá en especial por el reciente impacto mediático que están teniendo en redes sociales, medios de comunicación y campañas promovidas por las instituciones públicas. Estas dos patologías serán tratadas en otras partes del libro. Nos centraremos a continuación en otras dos patologías en las que creemos especialmente interesante analizar sus resultados en salud desde una perspectiva económica: EII y SII. La EII es una patología relativamente poco frecuente, pero crónica y que afecta a personas jóvenes en plena edad productiva, que requieren tratamientos especialmente costosos desde la llegada de los biológicos, además de cirugías e ingresos, con gastos directos e indirectos muy relevantes. El síndrome de intestino irritable es mucho más frecuente, no condiciona complicaciones importantes pero si afecta a la calidad de vida de las personas, con todo lo que ello conlleva. Su tratamiento médico no es tan costoso proporcionalmente, pero su prevalencia hace que el coste global del manejo de la enfermedad si sea elevado.

■ 13.3. SÍNDROME DE INTESTINO IRRITABLE

El SII es el más representativo de los trastornos funcionales digestivos, y una patología muy prevalente. Es una enfermedad en la que se persigue, una vez considerado que no existe una causa orgánica de su cuadro, no busca tanto una explicación compleja fisiopatológica acerca del origen de sus dolencias, sino una mejoría en sus síntomas y en su calidad de vida. El SII tiene una evolución crónica y se caracteriza por la presencia de molestias o dolores abdominales asociados a alteraciones del ritmo deposicional, con predominio de diarrea, estreñimiento o alternancia entre ambas (Chey, 2011). Como patología funcional que es, no conlleva un incremento en mortalidad y es infrecuente que los pacientes que presentan un SII deban ser ingresados, sometidos a cirugía o requieran de intervenciones agresivas (de hecho cuando esto sucede, no debería ser por el SII). Sin embargo, el SII si provoca afectación en la calidad de vida de los pacientes que lo padecen.

Todo esto hace que la medición de resultados en los pacientes tratados por un SII no deba realizarse a través de una evaluación basada en variables "duras" y objetivas, como sí puede hacerse en otras patologías. Nos referimos a *ítems* como la mortalidad, la necesidad de ingreso, la necesidad de cirugía, etcétera. Por supuesto, estos puntos deben ser controlados como indicadores de calidad de la asistencia, así como la solicitud de pruebas complementarias innecesarias, pero con una atención especializada y experimentada de calidad, todos ellos deberían resultar valores nulos o prácticamente nulos y no nos permitirán evaluar con qué nivel de calidad estamos atendiendo a nuestros pacientes, y cuanto valor estamos generando; al contrario, serían expresión de cuanta iatrogenia estamos cometiendo. La mejor herramienta para evaluar nuestros resultados en el tratamiento de los pacientes con SII, son los propios pacientes y su vivencia personal de la enfermedad. Nuestro principal objetivo en el SII no es otros que el control de la sintomatología, es decir, mejorar en todo lo posible su calidad de vida. La evaluación de nuestros resultados consistirá en una medición de su calidad de vida y de los síntomas subjetivos reportados por el paciente (PRO) (Mearin, Perelló y Perona, 2004). Para realizar la medición de la calidad de vida de la manera más objetiva posible existen varios cuestionarios, algunos de ellos genéricos y otros específicos. A continuación procedemos a describir brevemente los más utilizados en la práctica clínica habitual.

En primer lugar, como cuestionario genérico, el cuestionario de calidad de vida relacionada con la salud *Short Form* de 36 *ítems* (SF-36) es uno de los más utilizados a nivel general en medicina. Desarrollado en los años noventa en EE. UU. es una escala genérica que proporciona información acerca del perfil del estado de salud de pacientes y controles. También existe una versión reducida, el SF-12. La versión completa, el SF-36 está compuesto de 36 preguntas o *ítems* que cubren función física, rol físico, dolor corporal, vitalidad, salud general, función social, rol emocional y salud mental.

En segundo lugar, existen varios cuestionarios de calidad de vida específicos para el síndrome de intestino irritable. Destaca el cuestionario *Irritable Bowel Syndrome-Quality of Life* (IBS-QOL) Este test ha sido la herramienta más analizada y estudiada y es el cuestionario que cuenta con mayor evidencia a su favor. Es un cuestionario validado en castellano consistente en 34 preguntas, acerca de la frecuencia, intensidad, nocturnidad de los síntomas y la afectación física, emocional y de actividades de la vida cotidiana que provoca el SII (Patrick y Drossman, 2002). Otros cuestionarios específicos desarrollados para el SII son el IBSQOL, el FDQOL y el IBS-36, aunque el nivel de evidencia acerca de su utilidad es menor respecto al disponible acerca del IBS-QOL, por lo que la recomendación actual es la utilización de este primero (Drossman *et al.*, 2000).

Por último, Spiegel *et al.*, desarrollaron un cuestionario breve, con cuatro preguntas solamente, que puede ser administrado en menos de un minuto y el cálculo

del score requiere de unos treinta segundos. Este cuestionario se denomina BEST-score y consiste en cuatro preguntas: gravedad percibida de los síntomas, impacto de los síntomas sobre la capacidad de disfrutar de las cosas cotidianas, percepción del enfermo acerca de su padecimiento y repercusión de los síntomas sobre el estado de ánimo y el humor del enfermo. El principal potencial de esta herramienta radica en la posibilidad de ser usada en la práctica clínica habitual, donde los tiempos son más ajustados y no es posible cumplimentar cuestionarios de más de 30 preguntas (Spiegel *et al.*, 2006). Además presenta un elevado grado de concordancia con el cuestionario IBS-QOL.

■ 13.4. ENFERMEDAD INFLAMATORIA INTESTINAL

La EII engloba tres enfermedades diferentes, la colitis ulcerosa (CU), la enfermedad de Crohn (EC) y la colitis indeterminada (CI). Todas ellas presentan una evolución crónica e impredecible, en ocasiones muy tortuosa y con alternancia de periodos de remisión clínica con otros de empeoramiento o brotes (Sairenji, Collins y Evans, 2017).

La prevalencia de estas enfermedades se sitúa entre 1 caso por cada 200 y 1 de cada 300 habitantes en los países desarrollados, con una tendencia alcista. Además, al tratarse de una enfermedad crónica sin posibilidad de curación (el objetivo es mantener al paciente en remisión, no curarlo de su enfermedad), el número de pacientes que se encuentran en seguimiento en las consultas especializadas en EII de los diferentes hospitales es cada vez mayor.

La evaluación de resultados en salud en el caso de la EII es muy diferente al referido en el SII. Cuando hablábamos anteriormente en este mismo capítulo acerca del SII, todo nuestro esfuerzo en la medición de resultados se enfocaba en los PRO, y no en variables “duras” y objetivas. En el caso de la EII este paradigma cambia. Tanto la EC como la CU son enfermedades con importante afectación orgánica y con capacidad de producir complicaciones, potencialmente graves e incluso mortales, como estenosis, patología perianal, abscesos, megacolon tóxico y otras. Esta historia natural potencialmente grave hace que dispongamos de variables duras que evaluar en el seguimiento de nuestros pacientes, tanto en las situaciones de estabilidad clínica y manejo ambulatorio como en aquellos casos en los que nos encontremos en una fase de brote. No obstante, también mediremos síntomas y signos de los pacientes, situación endoscópica como medida objetiva (generalmente en forma de índices compuestos) y, en consonancia con las agencias reguladoras, incluimos la evaluación de los PRO, como una parte muy importante. Por ello en el desarrollo de esta parte del capítulo vamos a separar los resultados a medir en dos grupos: por una parte los resultados percibidos por los observadores, en este caso, los clínicos y por otro, los resultados percibidos por los pacientes.

■ 13.4.1. Resultados comunicados por los clínicos

Aquí es donde vamos a exponer las variables clínicas, los principales *end points*, en los que se basa la medición del éxito del tratamiento de la enfermedad. Como a continuación veremos, en ocasiones son resultados “positivos” que perseguimos y en otros casos son resultados “negativos” que tratamos de evitar. En este grupo de variables tenemos algunas que trabajan bien individualmente, como por ejemplo la necesidad de cirugía, la necesidad de colostomía o la tasa de mortalidad. Otras en cambio son utilizadas en combinación en lo que constituyen los denominados “índices de actividad”. Estos índices se basan en la valoración conjunta de variables analíticas, clínicas, endoscópicas, histológicas y/o radiológicas.

Existen 17 índices que permiten evaluar la actividad en CU, mediante diferentes combinaciones entre los *ítems* antes enumerados. El primer índice que se definió para este fin fueron los criterios de Truelove y Witts en 1995 (Truelove y Witts, 1955), basado en criterios objetivos. A pesar de no haber sido validados, los criterios de Truelove son muy utilizados y constituyeron todo un hito en la investigación médica. Existen varios índices que excluyen las variables analíticas y endoscópicas, permitiendo que sean cumplimentados por el propio paciente. Entre estos índices (8 en total) destacamos el *Simple Clinical Colitis Activity Score*, aunque no ha sido validado. El índice de Mayo, uno de los más empleados hoy en día, incluye valoración endoscópica, clínica y la valoración global (y subjetiva) del médico, existiendo también una versión parcial que excluye la endoscopia del cálculo. Otros índices son el índice de actividad en CU pediátrico, el Rachmilewitz score, el índice de Powell Tuck o el índice de Seo.

En cuanto a los índices en EC, el instrumento más estandarizado es el *Crohn's Disease Activity Index* (CDAI), aunque su cálculo es difícil y no suele usarse en la práctica clínica dada su complejidad. Existe otra versión, abreviada, en la que las preguntas han quedado limitadas a dolor abdominal, frecuencia de las deposiciones y bienestar general. Una de las principales limitaciones del CDAI es el “peso” que en él tienen variables subjetivas y a veces no dependientes de la actividad de la EC. El índice de Harvey-Bradshaw está formado por 5 variables, que son el bienestar general, la gravedad del dolor abdominal, el número de deposiciones líquidas, la presencia de masa abdominal y la presencia de complicaciones (Harvey y Bradshaw, 1980). Sus ventajas son su sencillez y que no requiere recogida prospectiva de los diferentes ítems.

Para la evaluación específica de la actividad perianal existe el *Perineal Disease Activity Index*, ya que aunque el CDAI incluye un apartado de enfermedad perianal, éste solo supone un pequeño peso de la valoración total.

Aparte de los índices combinados antes explicados, ambas patologías disponen de sus propios índices exclusivamente endoscópicos, y por tanto objetivos (relativamente). Estos índices, utilizados principalmente en investigación, permiten una valoración exclusiva de uno de los datos más objetivos que el clínico puede

disponer, como es la curación mucosa. En la colitis ulcerosa destaca el *Mayo Endoscopic Subscore*, una parte del Mayo score antes comentado. El *Ulcerative Colitis Endoscopic Index of Severity* o el *Ulcerative Colitis Colonoscopic Index of Severity* se componen de tres y seis *ítems* respectivamente. Según la Cochrane, existen 19 índices con estudios de validación y otros 18 más que no los tienen. Estos tres que se han descrito aquí son los que presentan mayor grado de evidencia a su favor.

Por su parte, la EC también posee numerosos índices con mayor o menor grado de validación. De todos ellos, los dos más generalizados en el contexto de ensayos clínicos e investigación son el *Cronh's Disease Endoscopic Index of Severity* y el *Simple Endoscopic Score for Crohn's Disease*. A su vez existe un índice específico para valorar la recidiva posquirúrgica de la EC. Se trata del índice de Rutgeerts, que evalúa la recurrencia endoscópica en el neointestino terminal tras una resección ileo-cólica o ileal (Rutgeerts *et al.*, 1990). Es un índice sencillo aunque subjetivo a ojos del examinador y el punto de corte para definir "recurrencia" también es discutido.

Dejando a un lado los índices clínicos, existen otras variables muy importantes en la evolución de los pacientes, y en la historia natural de su enfermedad. Así, como hemos comentado en la introducción, los pacientes con EII son subsidiarios de requerir ingresos, tratamientos con potenciales efectos secundarios graves, cirugías, ostomías o de padecer cáncer de colon. El objetivo primordial en el tratamiento de la EII es mantener al paciente en remisión, pero en ocasiones esa historia natural impredecible que hemos comentado antes hace inevitable la mala evolución del paciente. Si bien esto es cierto en ocasiones, optimizar los recursos de los que el clínico dispone debe permitir reducir al máximo estos casos de evolución tortuosa. La necesidad de colectomía en el caso de la CU (opción que aunque entendida como "curativa" afecta a la calidad de vida del paciente, ya que conlleva secuelas posteriores, como la diarrea permanente) y de cirugía de resección en el caso de la EC suponen el fracaso del tratamiento médico del paciente, aunque en ocasiones sean la mejor alternativa sin duda. Dentro de estas situaciones, la necesidad de una colostomía, con el trastorno vital que ello acarrea para el paciente, también debe considerarse como un ente a evitar.

Por otra parte, el ingreso hospitalario, cada vez menos frecuente gracias al avance de los tratamientos de mantenimiento de la enfermedad, también debe considerarse una situación a evitar. En este caso, una parte de la acción procederá del médico especialista que este al cargo del manejo del enfermo, pero otra pieza fundamental serán los servicios de urgencias, donde la atención a los pacientes con enfermedad inflamatoria es llevada a cabo en muchas ocasiones por personal no tan familiarizado en esta patología y que puede incurrir en ingresos hospitalarios imprevistos en pacientes que podrían haber seguido un manejo ambulatorio.

Por último, no deben obviarse las complicaciones secundarias a los tratamientos empleados por los clínicos. Los fármacos inmunosupresores y biológicos, utilizados de forma muy amplia en ambas entidades presentan una serie de efectos

secundarios, con especial relevancia para las infecciones. Un adecuado *screening* previo, mediante despistaje de infecciones latentes, programa de vacunación y control de dosis y adherencia, permite reducir estos eventos y ofrecer una atención de mayor calidad a los pacientes.

■ 13.4.2. Resultados comunicados por los pacientes

Al igual que en el SII, en la EII también han cobrado gran importancia los PRO con el paso de los años en la evaluación de la respuesta del paciente al tratamiento y en la decisión de la estrategia terapéutica a seguir. La EII comparte con el SII gran parte de la sintomatología digestiva (dolor abdominal, aumento de la frecuencia deposicional con descenso de la consistencia) y también causa o puede causar, una profunda alteración en la calidad de vida del paciente.

El indicador centrado en el paciente más estudiado en EII es la calidad de vida (CV) y existen varios cuestionarios tanto genéricos (ya comentados previamente en este mismo capítulo) como específicos para su valoración. Al igual que en otras enfermedades, los cuestionarios genéricos permiten comparar a los pacientes con controles sanos o con otros grupos de poblaciones afectados por patologías diferentes, como podría tratarse de un estudio que comparase la afectación de la calidad de vida en un grupo de pacientes con EII y otro grupo con EPOC, por ejemplo. En contrapunto, los cuestionarios específicos son más sensibles que los genéricos. Así, como cuestionarios genéricos nuevamente el SF-36 es el más empleado. No obstante, el cuestionario más popular y más extendido en EII es el específico IBDQ (*Inflammatory Bowel Disease Questionnaire*), que consta de 32 ítems, pertenecientes a cuatro categorías / esferas: síntomas sistémicos, síntomas abdominales, función emocional y función social (Guyatt *et al.*, 1989). Existen otras versiones de 36, 10 y 9 ítems. Otros índices destacados son el índice de Manitoba IBD, la escala PRO-2, o el cuestionario IBD-control, este último no validado aún al cuestionario.

Como hemos visto, hay gran número de herramientas para la medición de los PRO en EII, pero cabe la pena detenerse un momento en el proceso que se sigue para determinar qué ítems son los más relevantes en la evaluación del paciente.

Según el profesor Porter, los resultados en salud pueden clasificarse en tres niveles o “tiers”, con dos subgrupos en cada uno a su vez. Estos niveles se definirían en la enfermedad de Crohn, por ejemplo, de la siguiente manera (Porter y Larsson, 2016; Porter, 2010):

- Tier 1 primera casilla (Supervivencia): Probabilidad de sobrevivir a un brote de la enfermedad tras el ingreso en una unidad de inflamatoria especializada.
- Tier 1 segunda casilla (Nivel de salud alcanzado o recuperado): Probabilidad de recuperar una funcionalidad digestiva normal tras un brote, tratado en una unidad especializada.

- Tier 2 primera casilla (Tiempo de recuperación y de regreso a la actividad): Tiempo de retorno a la vida normal tras un ingreso por brote de enfermedad en una unidad especializada.
- Tier 2 segunda casilla (Problemas aparecidos en el proceso terapéutico): Probabilidades de desarrollar una infección oportunista en un paciente en tratamiento con fármacos biológicos.
- Tier 3 primera casilla (sostenibilidad del nivel de salud alcanzado): Probabilidad de requerir intervención quirúrgica en pacientes en tratamiento en una unidad especializada.
- Tier 3 segunda casilla (problemas a largo plazo del proceso terapéutico): Probabilidad de desarrollar linfoma en pacientes en tratamiento con azatioprina en pacientes en seguimiento en una unidad especializada.

Estos casos son solo un ejemplo ilustrativo de cada nivel, pero podrían ser otras variables diferentes, igual de válidas que las que hemos propuesto en este capítulo.

Como complemento a esto, hay que destacar varias iniciativas internacionales que buscan poner al paciente, no solo en el centro de la terapia, sino hacerlo partícipe de cuáles son los objetivos terapéuticos más importantes desde su punto de vista, como receptor de la acción. Estas iniciativas tratan de averiguar cuáles son las prioridades de los pacientes, las cuales en ocasiones coinciden con las planteadas por los facultativos pero que en otras ocasiones pueden ser diferentes.

Una organización que destaca en este empeño es la International Consortium for Health Outcomes Measurement (ICHOM). Este consorcio forma equipos compuestos por pacientes, facultativos e investigadores que definen cuales son los *outcomes* que más importan a los pacientes que viven con diferentes patologías. Han desarrollado una serie de conjuntos para determinadas patologías, entre las que se encuentra la enfermedad inflamatoria intestinal (Kim *et al.*, 2018).

Este conjunto está formado por 12 *outcomes*, clasificados en cuatro categorías:

- Supervivencia y control de la enfermedad: Está compuesto por cuatro *outcomes*, que son anemia, actividad y remisión de la enfermedad (medida mediante el índice Manitoba), la posibilidad de padecer cáncer colorectal y la supervivencia global.
- Síntomas, funcionalidad y calidad de vida: Compuesto por cinco *outcomes* que son la presencia de cambios en los síntomas, el dolor y el disconfort, la capacidad de realizar actividad normales, la fatiga y la falta de energía y el peso. Todas estas manifestaciones se miden mediante el cuestionario IBD-control.

- Problemas en el cuidado: Compuesto por dos *outcomes* que son el uso de corticoides y las complicaciones de la intervención.
- Utilización de los servicios de salud: compuesto por un único *ítem*, hospitalizaciones y visitas a urgencias.

Como podemos ver, muchos de estos *outcomes* ya se encontraban recogidos en los múltiples scores descritos, pero no es así en todos los casos, como por ejemplo el uso de esteroides o las visitas a urgencias. Estos son parámetros que sin duda causan un importante impacto en la calidad de vida del paciente.

■ 13.5. CONCLUSIÓN

Como todos bien sabemos, vivimos actualmente en la era de la Medicina Basada en la Evidencia y en la era de la Medicina Basada en el Paciente (cuyo máximo exponente es el principio de autonomía del paciente). Este enfoque de la Medicina se refleja también en la forma de medir los resultados en salud, como hemos comentado a lo largo de todo el capítulo de dos maneras: Primero, buscando la máxima objetividad y asertividad a la hora de evaluar los resultados, tratando de encontrar medidas que nos den la información de mayor calidad posible acerca de las enfermedades que tratamos, ya sea en investigación o en la práctica diaria en la consulta.

Y en segundo lugar, vemos que no solo buscamos variables clínicas o analíticas, sino que la tendencia actual busca centrar la evaluación en el paciente y en su percepción de la enfermedad y su vivencia de la misma. Como no podía ser de otra manera, las enfermedades digestivas, representadas en este caso por la EII y el SII se han sumado a esta manera de evaluar la información, y lo más probable es que se siga avanzando en esta línea, aumentando el peso del paciente en las decisiones sobre su propia enfermedad.

■ REFERENCIAS

- ANDERSSON, E., LJÓTSSON, B., SMIT, F., PAXLING, B., HEDMAN, E. y LINDEFORS, N. (2011). Cost-effectiveness of internet-based cognitive behavior therapy for irritable bowel syndrome: results from a randomized controlled trial. *BMC Public Health*, 11, p. 215.
- BRODSZKY, V., RENCZ, F., PÉNTÉK, M., BAJI, P., LAKATOS, P. L. y GULÁCSI, L. (2016). A budget impact model for biosimilar infliximab in Crohn's disease in Bulgaria, the Czech Republic, Hungary, Poland, Romania, and Slovakia. *Expert Rev Pharmacoecon Outcomes Res*, 16(1), p. 119-125.
- BROSA, M., GISBERT, R., RODRÍGUEZ, J. M. *et al.* (2005). Principios, métodos y aplicaciones del análisis del impacto presupuestario en el sector sanitario. *Pharmacoecon. Span. Res. Artic*, 2, p. 65.
- BUONO, J. L., CARSON, R. T. y FLORES, N. M. (2017). Health-related quality of life, work productivity, and indirect costs among patients with irritable bowel syndrome with diarrhea. *Health Qual Life Outcomes*, 15(1), p. 35.

- CANADIAN COORDINATING OFFICE FOR HEALTH TECHNOLOGY ASSESSMENT. (1997). *Guidelines for economic evaluation of pharmaceuticals: Canada 2nd ed.* Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA).
- CHEY, W. D. (2011). Irritable bowel síndrome. *Gastroenterology Clinics*, 40(1), pp. 1-264.
- DROSSMAN, D. A., PATRICK, D. L., WHITEHEAD, W. E., TONER, B. B., DIAMANT, N. E., HU, Y., JIA, H. *et al.* (2000). Further validation of the IBSQOL: a disease-specific quality-of-life questionnaire. *Am J Gastroenterol*, 95, pp. 999-1007.
- DRUMMOND, M., SCULPER, M., TORRANCE, G., O'BRIEN, B. y STODDART, G. (2005). *Methods for the economic evaluation of health care programmes*. III Edición (Cap. 2). Oxford: Oxford University Press.
- . (2005). *Methods for the economic evaluation of health care programmes*. III Edición (Cap. 5). Oxford: Oxford University Press.
- EISENBERG, J. M. (1989). Clinical economics. A guide to economic analysis of clinical practices. *JAMA*, 262, pp. 2879-2886.
- FISHER, M., WALKER, A., FALQUÉS, M., MOYA, M., RANCE, M., y TAYLOR, D. (2016). Cost-effectiveness of linaclotide compared to antidepressants in the treatment of irritable bowel syndrome with constipation in Scotland. *Eur J Health Econ*, 17(9), pp. 1091-1100.
- GROVER, S. A., HO, V., LAVOIE, F., COUPAL, L., ZOWALL, H. y PILOTE, L. (2003). The importance of indirect costs in primary cardiovascular disease prevention. *Arch Intern Med*, 163, pp. 333-339.
- GUYATT, G., MITCHELL, A., IRVINE, E. J. *et al.* (1989). A new measure of health status for clinical trials in inflammatory bowel disease. *Gastroenterology*, 96, pp. 804-810.
- HARVEY, R. F. y BRADSHAW, J. M. (1980). A simple index of Crohn's-disease activity. *Lancet*, 1, p. 514.
- JEAN, L., AUDREY, M., BEAUCHEMIN, C. y CONSORTIUM OBOTI. (2018). Economic Evaluations of Treatments for Inflammatory Bowel Diseases: A Literature Review. *Can J Gastroenterol Hepatol*, 2018, p. 7439730.
- JUAN, J., ESTIARTE, R., COLOMÉ, E., ARTÉS, M., JIMÉNEZ, F. J. y ALONSO, J. (2003). Burden of illness of Crohn's disease in Spain. *Dig Liver Dis*, 35(12), pp. 853-861.
- KIM, A. H., ROBERTS, C., FEAGAN, B. G., *et al.* (2018). Developing a standard set of patient-centred outcomes for inflammatory bowel disease: an International, Cross-disciplinary Consensus. *J Crohns Colitis*, 12, pp. 408-418.
- MASON, J. (1997). The generalisability of pharmacoeconomic studies. *Pharmacoeconomics*, 11(6), pp. 503-514.
- MEARIN, F., PERELLÓ, A. y PERONA, M. (2004). Calidad de vida en los pacientes con síndrome del intestino irritable. *Gastroenterol Hepatol*, 27, pp. 24-31. [citado 20 nov. 2018]. Disponible en: <http://www.elsevier.es/es-revista-gastroenterologia-hepatologia-14-articulo-calidad-vida-los-pacientes-con-13058927>
- PATRICK, D. L. y DROSSMAN, D. A. (2002). Comparison of IBS36 and IBS-QOL instruments. *Am J Gastroenterol*, 97, p. 3204.
- PORTER, M. E. (2010). What is value in health care? *N Engl J Med*, 363, pp. 2477-2481. 10.1056/NEJMp1011024.
- PORTER, M. E. y LARSSON, S. (2016). Standardizing patient outcomes measurement. *N Engl J Med*, 374, pp. 504-506.
- ROVIRA, J. y ANTOÑANZAS, F. (1998). Economic Analysis of Health Technologies and Programmes. A Spanish Proposal for Methodological Standardisation. *Pharmacoeconomics*, 8(3), pp. 245-252.
- RUBIO-TERRÉS, C., COBO, E., SACRISTÁN, J. A., PRIETO, L., DEL LLANO, J. y BADÍA, X por el grupo ECOMED. (2004). *Med Clin (Barc)*, 122(17), pp. 668-674.

- RUTGEERTS, P., GEBOES, K., VANTRAPPEN, G., BEYLS, J., KERREMANS, R. e HIELE, M. (1990). Predictability of the postoperative course of Crohn's disease. *Gastroenterology*, 99, pp. 956–963. doi: 10.1016/0016-5085(90)90613-6.
- SACRISTÁN, J. A. (1995). Farmacoeconomía y evaluación económica de medicamentos. En: J. A. SACRISTÁN, X. BADÍA y J. ROVIRA, *Farmacoeconomía: evaluación económica de medicamentos*, (Cap. 1). Madrid: Editores Médicos SA.
- SAIRENJI, T., COLLINS, L. K. y EVANS, D. V. (2017). An Update on Inflammatory Bowel Disease. *Primary Care: Clinics in Office Practice*, 44(4), pp. 673-692.
- SPIEGEL, B. M. R., NALIBOFF, B., MAYER, E. *et al.* (2006). Development and initial validation of a concise point-of-care IBS severity index: the 4-item BEST questionnaire. *Gastroenterology*, 130, p. S1040.
- TRUELOVE, S. C. y WITTS, L. J. (1955). Cortisone in ulcerative colitis; final report on a therapeutic trial. *Br Med J*, 2, pp. 1041–1048.



14

RESULTADOS EN SALUD EN ESCLEROSIS MÚLTIPLE

RICARDO CONSTANTINO GINESTAL LÓPEZ,
PATRICIA SANMARTIN FENOLLERA

■ 14.1. INTRODUCCIÓN

La esclerosis múltiple (EM) es una enfermedad crónica, inflamatoria, desmielinizante y degenerativa del sistema nervioso central. Con una prevalencia superior a 1 en 1.000 en España, es una de las principales causas de discapacidad en los adultos jóvenes en nuestro entorno. Dos fenómenos fisiopatológicos son los responsables de dicha discapacidad. Por un lado, los episodios desmielinizantes agudos de origen inflamatorio, que son responsables de los brotes clínicos de la enfermedad, con recuperación completa o parcial de los déficits adquiridos durante dichos episodios. Por otro, la neurodegeneración, relacionada con la acumulación progresiva de síntomas y signos asociados a la enfermedad.

Es esta dualidad en la forma por la cual los pacientes con esclerosis múltiple pueden acumular discapacidad la que va a definir qué marcadores clínicos y paraclínicos debemos utilizar en un ensayo clínico y en la práctica clínica habitual. No existen marcadores biológicos específicos de la esclerosis múltiple que, utilizados de forma única, nos informen de forma completa sobre la actividad de la enfermedad o de la respuesta al tratamiento. Así, para demostrar la eficacia de una molécula en los ensayos clínicos de formas recurrentes de la EM, debemos medir marcadores clínicos (brotes) y radiológicos (número de lesiones activas en pruebas de resonancia magnética) de inflamación aguda, pero también, y esto reviste aún mayor importancia en ensayos de formas progresivas de la enfermedad, marcadores de neurodegeneración (como la acumulación de discapacidad clínica y de atrofia encefálica cuantificada por resonancia magnética). De la misma manera, en la práctica clínica queremos disponer de medidas que permitan determinar, de forma individualizada, cuál es la mejor opción terapéutica para un paciente y qué respuesta estamos obteniendo del tratamiento instaurado. Se utilizan así los mismos marcadores clínicos y paraclínicos que en investigación, aunque con las lógicas limitaciones logísticas y de financiación de un entorno sanitario de atención universal. Cabe no olvidar, además, que la eficacia/efectividad de un producto sanitario no depende solamente de parámetros clínicos o paraclínicos sino también de la valoración subjetiva que hace del mismo el propio paciente o incluso sus familiares o cuidadores. Entran entonces en la evaluación los resultados en salud reportados por los pacientes (*patient reported outcomes measures*, PROM en su acrónimo en inglés) y por los cuidadores.

Por otro lado, las medidas en esclerosis múltiple no tienen que ver únicamente con la eficacia/efectividad de los medicamentos si no también con la seguridad, la tolerabilidad y el grado de aceptación de los mismos. Estos resultados en salud pueden ser obtenidos, una vez más, por el equipo sanitario, por el paciente o por los cuidadores.

Por último, no se deben olvidar que, en un entorno, como hemos señalado, de atención sanitaria universal de presupuesto limitado, el concepto de eficiencia debe ser también considerado, idealmente desde una perspectiva lo más amplia posible, considerando no únicamente el precio de los medicamentos, también los costes relacionados con el absentismo laboral, el desempleo y las jubilaciones precoces forzadas, tan frecuentes en los pacientes que sufren esclerosis múltiple.

En este capítulo vamos a describir los elementos de medición de resultados en salud más relevantes en una patología compleja como la esclerosis múltiple, tanto en los ensayos clínicos como en la práctica clínica habitual, desde la perspectiva de los agentes involucrados en la prestación sanitaria, del paciente y de su familiar/cuidador.

■ 14.2. RESULTADOS EN ENSAYOS CLÍNICOS, MEDICIÓN DE LA EFICACIA (VAN MUNSTER Y UITDEHAAG, 2017; TUR, MOCCIA Y BARKHOF, 2018)

Los resultados en salud en estudios clínicos de esclerosis múltiple se dividen en medidas de actividad inflamatoria aguda, medidas de progresión de la discapacidad y PROM (resultados en salud reportados por los pacientes) (Tabla 1). Los

Tabla 1.

FORTALEZAS Y DEBILIDADES DE LOS OBJETIVOS PRINCIPALES UTILIZADOS EN ENSAYOS CLÍNICOS DE ESCLEROSIS MÚLTIPLE

Medidas	Fortalezas	Debilidades
Medidas relacionadas con la actividad inflamatoria		
Número de brotes o tasa anualizada de brotes	Fácil de medir y entender	Sólo útil en formas recurrentes, no mide la gravedad de los brotes
Número de brotes severos	Puede reflejar la severidad de los brotes	Gran variabilidad intercentros debido a la ausencia de criterio unificado de tratamiento de los brotes
% pacientes libres de brotes	Útil en ensayos con fármacos de alta eficacia	Muy dependiente de la duración del ensayo
Tiempo hasta brote confirmado	Útil en el síndrome desmielinizante aislado	Sólo útil en formas recurrentes, no mide la gravedad de los brotes
Tiempo hasta fracaso de tratamiento	Mide eficacia, seguridad y tolerabilidad	No específico

Tabla 1. (continuación)

FORTALEZAS Y DEBILIDADES DE LOS OBJETIVOS PRINCIPALES UTILIZADOS EN ENSAYOS CLÍNICOS DE ESCLEROSIS MÚLTIPLE

Medidas	Fortalezas	Debilidades
Medidas de progresión de la discapacidad		
Cambio en EDSS	Fácil de interpretar por los clínicos dedicados a la EM	Baja sensibilidad para los cambios funcionales cognitivos y de miembros superiores. Importante variabilidad intra e interobservadores. Puede reflejar cambios transitorios relacionados con brotes
Cambio en MSFC/PASAT/9HPT/T25FW	No requiere entrenamiento específico, miden funciones concretas no adecuadamente evaluadas por EDSS	Validado para su uso en grupos de pacientes, limita su uso en individuos. Requiere definir el grado significativo de empeoramiento Requiere una población de referencia Fenómeno de aprendizaje (PASAT), ansiedad con PASAT
Cambio en escalas cognitivas	El SDMT no requiere entrenamiento previo. Sensible para la evaluación cognitiva	Las escalas que no son el SDMT requieren entrenamiento previo. Suelen necesitar una población de referencia para poder evaluar los resultados Edad, ansiedad, fatiga, nivel cultural influyen en los resultados
% progresión confirmada en EDSS a 3 ó 6 meses	Medida fácilmente entendible	Sobrestima el % de pacientes que incrementan su discapacidad en el largo plazo, muy dependiente de la longitud del estudio
% libres de progresión confirmada en EDSS a 3 ó 6 meses	Medida fácilmente entendible, útil en ensayos con moléculas de alta eficacia	Infraestima el % de pacientes que incrementan su discapacidad en el largo plazo, muy dependiente de la longitud del estudio
% con mejoría confirmada en EDSS	Útil para valorar mejoras de la discapacidad, generalmente poco estudiados en los ensayos clínicos	Puede ser poco específico en relación con el mecanismo fisiopatológico relacionado con la mejora de la discapacidad
Cambio en medidas combinadas, incluidos NEDA y NEPAD	Mayor sensibilidad que las medidas no combinadas para detectar cambios de discapacidad, menor error tipo I, medida de valor en la práctica clínica	Las medidas combinadas deben incluir medidas no combinadas que tengan relación con el efecto de la medicación a estudio
Patient-Reported Outcome Measures (PROM)		
PROM	La información llega directamente del paciente	La información es subjetiva y puede ser muy variable para el mismo individuo en diferentes momentos

Notas: EDSS: Expanded Disability Status Scale, MSFC: Multiple Sclerosis Functional Composite, PASAT: Paced Auditory Serial Addition Test, 9HPT: Nine hole peg test, T25FW: Timed 25 feet foot walk test, NEDA: No Evidence of Disease Activity, NEPAD: No Evidence of Progression or Active Disease.

Fuente: Elaboración propia.

primeros son predominantes en los ensayos de forma recurrentes-remitentes de la enfermedad. Los segundos adquieren preponderancia en los estudios de las formas progresivas de la EM. Por su parte, los PROM han venido cobrando cada vez mayor importancia en la última década gracias al impulso recibido para su inclusión como objetivos de los ensayos clínicos por las agencias reguladoras.

■ 14.2.1. Marcadores de actividad inflamatoria

■ 14.2.1.1. Marcadores clínicos de actividad inflamatoria

En la actualidad disponemos de 18 principios activos/posologías diferentes para el tratamiento de la forma recurrente-remitente de la esclerosis múltiple. El diseño de los ensayos clínicos en este tipo de forma evolutiva de la enfermedad se basa en utilizar como objetivos primarios y secundarios la reducción de la actividad inflamatoria aguda responsable de los brotes que la caracterizan. Así, se han definido los siguientes criterios clínicos de medición de eficacia: frecuencia de los brotes (generalmente la tasa anualizada de brotes), porcentaje o número de pacientes con brotes respecto a los pacientes sin brotes, tiempo hasta el primer brote tras el inicio de la fase de ensayo, y criterios combinados.

Un brote es definido como aquel síntoma de nueva aparición o el empeoramiento de un síntoma previo que aparece al menos 30 días tras un brote previo, que dura al menos 24 horas y que no se da en el contexto de fiebre o infección. En la mayoría de los ensayos clínicos se exige que el brote sea confirmado objetivamente por un neurólogo, habitualmente ciego para el tratamiento recibido, y medido por un incremento de puntuación en la escala EDSS (*Expanded Disability Status Scale*) y/o en alguno de sus sistemas funcionales. La medida de eficacia más comúnmente utilizada en relación con los brotes es la tasa anualizada. En más de la mitad de los más de 40 ensayos clínicos en EM recurrente-remitente llevados a cabo en las últimas décadas, la tasa anualizada de brotes fue el objetivo primario. Es un parámetro fácilmente medible y computable y tiene una gran significación clínica en formas recurrentes. Sin embargo, puede resultar poco específico porque no tiene en consideración la gravedad clínica del brote. Por este motivo, se ha intentado introducir la tasa anualizada de brotes severos, definidos por la necesidad de tratamiento corticoideo y/o de hospitalización del paciente. No obstante, la diferencia entre los procedimientos de los diferentes centros de los ensayos clínicos multicéntricos a estos dos respectos ha hecho poco estandarizable esta medida. La discriminación entre pacientes con y sin brotes está cobrando creciente valor como objetivo de eficacia fundamentalmente en los ensayos clínicos que evalúan medicamentos de alta eficacia, fármacos con los cuales se puede alcanzar el estatus de libre de actividad en brotes. Con un objetivo tan exigente, el factor tiempo es de gran importancia, siendo muy difícil mantener el estatus de libre de enfermedad en los ensayos clí-

nicos más prolongados. El objetivo clínico de tiempo hasta el primer brote es más frecuentemente utilizado en la fase de la esclerosis múltiple llamada síndrome desmielinizante aislado. Tiene particular relevancia en esta fase porque representa lo que tardará el paciente en tratamiento en convertirse en una esclerosis múltiple clínicamente definida al sufrir un segundo brote de la enfermedad. Así, en ensayos clínicos en síndrome desmielinizante aislado, el tiempo hasta el primer brote suele ser el objetivo primario del estudio. Finalmente, como medida combinada de eficacia en brotes, el parámetro más relevante utilizado en ensayo clínico es el tiempo hasta fracaso del tratamiento, el cual se define por la ocurrencia de cualquiera de uno de los dos siguientes eventos: brote o abandono del tratamiento por cualquier motivo.

■ 14.2.1.2. Imagen por Resonancia Magnética (Sormani y Bruzzi, 2013)

La cuantificación de lesiones focales por medio de estudios de resonancia magnética es un objetivo primario en ensayos de fase II y un objetivo secundario en los ensayos de fase III realizados en esclerosis múltiple. Es una evaluación de especial importancia en las formas más inflamatorias de la enfermedad, esto es, en el síndrome desmielinizante aislado y la forma recurrente-remitente, aunque también se computa en las formas progresivas. Las dos mediciones más frecuentes son la presencia de lesiones que realzan contraste con Gadolinio en secuencias ponderadas en T1 y el incremento de lesiones en secuencias ponderadas en T2, ya que reflejan la presencia de actividad inflamatoria de la enfermedad. Las lesiones ponderadas en T1 que realzan con Gadolinio representan una dehiscencia de la barrera hematoencefálica en el momento en que se está realizando la resonancia y, por lo tanto, hablan de inflamación aguda. Dada la escasa duración (semanas) de estas lesiones en T1 realzadas por contraste, lo que limita su posibilidad de detección con resonancias poco frecuentes, se utiliza la comparación entre el número y volumen de las lesiones en T2 entre resonancias magnéticas sucesivas para determinar, en ausencia de inflamación aguda en T1, si han aparecido nuevas lesiones o estas han crecido en el intervalo entre dos resonancias. Esto significaría que entre ambos estudios radiológicos ha habido actividad inflamatoria, representada por dichas nuevas (o mayores en volumen) lesiones en T2. De hecho, un objetivo combinado de resonancia magnética que permite determinar con mayor precisión la actividad inflamatoria de la enfermedad es la cuantificación de lesiones combinadas activas (*combined unique active*, CUA), definidas por la suma de las lesiones que realzan con contraste en T1 en la última resonancia y las lesiones T2 nuevas o que han incrementado su tamaño entre las dos últimas resonancias.

La gran ventaja de los objetivos medidos por resonancia magnética es su objetividad y, en el caso de las formas inflamatorias, su sencillez en la evaluación y gran sensibilidad a los cambios. Además, los brotes clínicos pueden ser predichos de forma razonablemente precisa por los estudios de resonancia magnética, de tal forma que se ha postulado considerar, en el futuro, la utilización de los parámetros de resonancia magnética como objetivo primario de ensayos en esclerosis múltiple

en situaciones concretas como en la evaluación de biosimilares o genéricos de medicamentos con un bien conocido mecanismo de acción, o en EM pediátrica en la evaluación de fármacos ya aprobados para su uso en adultos.

■ **14.2.2. Marcadores de discapacidad (Lavery, Verhey y Waldman, 2014; y Cohen *et al.*, 2012)**

■ **14.2.2.1. Marcadores clínicos de discapacidad**

La medición de la progresión de la discapacidad suele ser un objetivo secundario en los ensayos clínicos en esclerosis múltiple recurrente-remitente pero el principal objetivo primario de los estudios en fases progresivas. Tras numerosos ensayos clínicos previos en formas progresivas con diferentes moléculas que no habían conseguido demostrar beneficio sobre la discapacidad, los recientes estudios ORATORIO (Ocrelizumab en EM primariamente progresiva) y EXPAND (Siponimod en EM secundariamente progresiva) han demostrado eficacia de los productos a estudio utilizando este objetivo primario. Los objetivos definidos para evaluar la progresión de la discapacidad pueden ser clasificados en cinco grupos: discapacidad como un valor continuo (empeoramiento en una escala de discapacidad, generalmente la EDSS), pacientes con/sin progresión confirmada de la discapacidad, pacientes con/sin mejoría confirmada de la discapacidad, tiempo hasta progresión confirmada, y criterios de medición combinados. El parámetro más frecuentemente utilizado es el cambio en la escala EDSS. El problema de la escala EDSS es que el incremento de la discapacidad no es lineal a lo largo de toda la escala (no es lo mismo, por ejemplo, cambiar de 1.0 a 1.5 que de 6.5 a 7.0, ya que en el primer caso el cambio es funcionalmente inapreciable y en el segundo, siendo también un empeoramiento de medio punto en la escala, el paciente pasa de caminar con dos muletas a no poder caminar y tener que desplazarse en una silla de ruedas). Además, la escala EDSS no mide adecuadamente variaciones de función de los miembros superiores, de la ambulación a corta distancia, o de la función cognitiva de los pacientes, por lo que pierde sensibilidad. Por último, la escala EDSS, sobre todo en las puntuaciones más bajas, se presta a una gran variabilidad intra e interobservadores. A pesar de todas estas limitaciones, la EDSS sigue siendo la escala más utilizada tanto en el pasado como en el presente para medir el cambio de discapacidad de los pacientes en los ensayos clínicos de EM. Para intentar compensar las limitaciones de esta escala, se han introducido otras como el test de los 25 pies (*Timed 25 feet footwalk test*, mide la velocidad en recorrer dicha distancia), el test de los nueve agujeros (*Nine hole peg test*, evalúa la destreza bimanual) y el PASAT (*Paced Auditory Serial Addition Test*, mide la atención, concentración, memoria de trabajo y aritmética). Otros test cognitivos empleados en ensayos clínicos para evaluar la progresión de los déficits, infrarrepresentados en la escala EDSS, son el SDMT (*Symbol Digit Modalities Test*)

o el más completo BRN-B (*Brief Repeatable Battery of Neuropsychological Tests*), entre otros.

La medición binaria de la progresión de la discapacidad, clasificando en pacientes que progresan y pacientes que no lo hacen, se ha utilizado de forma muy variable en los diferentes ensayos clínicos en EM. La progresión confirmada de la discapacidad se define como un empeoramiento de la puntuación en la escala EDSS de 1.5 puntos si partíamos de 0, de 1.0 punto si nos encontrábamos inicialmente entre 1.0 y 5.5, y de 0.5 punto si el paciente presentaba un EDSS superior a 5.5 antes del incremento. La confirmación del incremento en la puntuación se hace generalmente a los tres o seis meses de la primera medición, para intentar asegurarnos de que no es un incremento transitorio de la discapacidad (como podría ocurrir, por ejemplo, tras un brote). La confirmación de la progresión a tres o seis meses, no obstante, puede sobreestimar el incremento de la discapacidad en el largo plazo por lo que sería deseable que la confirmación de la discapacidad se hiciera a 12 o incluso a 24 meses. Aunque siguen sobreestimando el incremento de la discapacidad a largo plazo en un 20 % y un 11 % respectivamente, son medidas más precisas. La medición de la progresión de la discapacidad confirmada a 12 y 24 meses no se ha implementado aún en los ensayos clínicos. Otra forma de medir la progresión de la discapacidad es con el *Multiple Sclerosis Functional Composite* (MSFC) que resulta de la aplicación combinada del test de los 25 pies, del test de los nueve agujeros y del PASAT, todos ellos nombrados anteriormente. Parece que el MSFC podría ser más sensible que la EDSS para medir cambios en la discapacidad. Sin embargo, su uso no ha reemplazado a la EDSS en los ensayos clínicos y aún menos en la práctica clínica debido a una mayor complejidad para interpretar los resultados desde un punto de vista clínico (Meyer-Moock *et al.*, 2014). Muestra de ello es que, en vez de abandonarla y sustituirla por otra, se siguen realizando propuestas de mejora de la escala EDSS como la de la evaluación de la agudeza visual mediante el uso de láminas de bajo contraste. Utilizadas en sustitución de las clásicas láminas de Snellen o de láminas de visión cercana, frecuentemente manejadas para la EDSS, dichas láminas de bajo contraste parecen tener validez, consistencia, reproducibilidad y ser clínicamente significativas en la detección de cambios en la agudeza visual de los pacientes con esclerosis múltiple (Balcer *et al.*, 2017).

Entre los parámetros que se utilizan para medir la mejoría de la discapacidad, el más utilizado es la mejoría confirmada evaluada mediante la escala EDSS. Tal y como el empeoramiento de la discapacidad, su mejoría también debe ser confirmada en el tiempo, al menos a los tres meses de la primera medición.

Finalmente, en cuanto a las medidas combinadas de medición de la discapacidad, se ha visto que podría ser más sensible definir la progresión como el cumplimiento de uno de los tres eventos siguientes: progresión de la EDSS (1.0 punto si la puntuación de origen era de 5.5 o menos, 0.5 punto si era de 6.0 o más), incremento de más del 20 % en el tiempo para recorrer 25 pies, o incremento de 20 % en el test de los nueve agujeros.

■ 14.2.2.2. *Imagen por resonancia magnética (Sormani y Bruzzi, 2013)*

La atrofia encefálica mide la neurodegeneración, proceso íntimamente relacionado con la acumulación de discapacidad irreversible en pacientes con esclerosis múltiple. En este sentido, se puede medir el volumen encefálico global o el de una fracción del cerebro como pueden ser el volumen de la sustancia blanca o el volumen cortical. En los ensayos clínicos de esclerosis múltiple recurrente-remitente, el efecto del tratamiento sobre la atrofia cerebral se correlaciona con la progresión de la discapacidad a dos años y es totalmente independiente de su efecto sobre las lesiones activas detectadas por resonancia magnética. Las dos medidas más frecuentemente utilizadas para medir la atrofia cerebral son la fracción de parénquima cerebral y el cambio en porcentaje del volumen cerebral. Las metodologías más universales para la determinación de fracciones de volumen cerebral concretas generalmente evalúan el volumen de la sustancia gris (corteza, tálamo, ...) o de la sustancia blanca. Es posible que la medición de cambios en el volumen de la sustancia gris sea más sensible a los cambios que la progresión de la atrofia cerebral global, y que se correlacione mejor que ésta tanto con la discapacidad global como con la discapacidad cognitiva de los pacientes con esclerosis múltiple.

La atrofia medular se produce a una ratio de hasta 2.2 %/año en las formas secundariamente progresivas de la enfermedad lo que parece apuntar a que sería un marcador sensible y clínicamente significativo de discapacidad. Generalmente se mide a nivel cervical y se asocia con la acumulación de discapacidad motora a largo plazo tanto en formas progresivas como en el síndrome desmielinizante aislado y la EM recurrente-remitente. En ensayos de formas primaria y secundariamente progresivas, la medición de la atrofia medular cervical se ha utilizado como objetivo secundario. Tiene ciertas limitaciones en cuanto a la estandarización de la técnica y su posibilidad de replicación entre centros.

■ 14.2.2.3. *Tomografía de Coherencia Óptica (OCT, Optical Coherence Tomography) (Saidha et al., 2015)*

Capaz de medir la pérdida neuronal y axonal en la vía óptica anterior, se correlaciona con la clínica visual pero también con el proceso de neurodegeneración, sobre todo en las formas progresivas de la enfermedad. Por este motivo se ha incluido como objetivo exploratorio en ensayos clínicos de neuritis óptica pero también de EM sin clínica de neuritis óptica. Es una técnica de fácil realización y consistente, pero se necesitan datos a más largo plazo para poder estandarizarla como objetivo principal de ensayos clínicos.

■ 14.2.3. Combinación de resultados clínicos y de resonancia magnética para la medición global de la eficacia (actividad inflamatoria y progresión de la discapacidad) (Giovannoni *et al.*, 2017; Parks *et al.*, 2017; Wolinsky *et al.*, 2017)

El concepto de una medida combinada clínico-radiológica para medir la actividad de la esclerosis múltiple data de 2006. En el ensayo clínico fase III de Natalizumab, uno de los parámetros más precisos para predecir ausencia de progresión en el seguimiento de los pacientes fue la combinación de tres elementos: ausencia de brotes sumado a ausencia de progresión en la EDSS y a ausencia de lesiones nuevas/mayores en T2 más ausencia de lesiones que realcen con Gadolinio en T1. Nace así el concepto de NEDA (*No Evidence of Disease Activity*). Algunos autores han propuesto añadir a esta definición la ausencia de progresión anormalmente rápida de la atrofia cerebral (por encima de 0.4 % al año), lo que se conoce como NEDA-4, dada la correlación ya comentada entre progresión radiológica anormal de la atrofia y mayor velocidad de acumulación de discapacidad física y cognitiva.

Un concepto aún más reciente es el de NEPAD (*No Evidence of Progression or Active Disease*). Es la medida combinada que cumple todos los siguientes criterios: no hay progresión confirmada a las 12 semanas de la discapacidad de 1.0/0.5 puntos en la EDSS si la EDSS basal era de ≤ 5.5 / >5.5 respectivamente, no evidencia confirmada a las 12 semanas de la discapacidad medida por un incremento de 20 % en el test de los nueve agujeros, no evidencia confirmada a las 12 semanas de la discapacidad medida por un incremento de 20 % en el test de los 25 pies, ausencia de lesiones nuevas/mayores en T2 + ausencia de lesiones que realcen con Gadolinio en T1, y ausencia de brotes. En el ensayo clínico ORATORIO (Ocrelizumab en EM primariamente progresiva), los pacientes que recibieron el tratamiento activo tuvieron tres veces más probabilidades de cumplir criterios NEPAD a las 120 semanas que los que recibieron placebo. Esto parece indicar que, aunque extraordinariamente exigente, la medida combinada NEPAD podría ser útil para confirmar la ausencia de elementos clínicos y radiológicos de actividad inflamatoria y de progresión en la forma primariamente progresiva de la esclerosis múltiple.

■ 14.2.4. Resultados en salud reportados por el paciente (*Patient-reported outcome measures, PROM*) (Khurana *et al.*, 2017)

Utilizados para evaluar el impacto de la EM en las actividades de la vida diaria, la funcionalidad de los pacientes y su calidad de vida, están cada vez más presentes entre los objetivos de los ensayos clínicos en esta enfermedad (Tabla 2). Los PROM pueden ser medidos con escalas genéricas (como la SF-36) o específicas de la enfermedad que estemos evaluando. Algunos ejemplos de escalas específicas para la EM son los siguientes: *Multiple sclerosis impact scale* (MSIS:29), *Patient-Reported Indices in Multiple Sclerosis* (PRIMUS), *Multiple Sclerosis Quality of Life 54 items* (MSQoL-54), *Multiple Sclerosis Quality of Life* (MUSIQoL), *Functional Assessment*

of Multiple Sclerosis (FAMS), Leeds Multiple Sclerosis Quality of Life (LMSQoL) y Hamburg Quality of Life Questionnaire in Multiple Sclerosis (HAQUAMS). Otras escalas están dirigidas a evaluar síntomas concretos como la capacidad ambulatoria o la muy prevalente fatiga de los pacientes con EM. Algunos ejemplos son la MSWS-12 (*Multiple Sclerosis Walking Scale-12 ítems*), la *Fatigue Impact Scale* (FIS) y la *Fatigue Severity Scale* (FSS). La MSWS-12 ha servido para evaluar la respuesta a Fampridina en ensayos clínicos, considerándose un cambio clínicamente significativo una pérdida de 4 a 6 puntos en la escala.

Tabla 2.

PRINCIPALES ESCALAS DE OBJETIVOS REPORTADOS POR LOS PACIENTES UTILIZADAS EN ENSAYOS CLÍNICOS DE FASE III EN ESCLEROSIS MÚLTIPLE

Medida	Características
Arminindex	Escala que describe la funcionalidad del miembro superior
PRIMUS	Calidad de vida y funcionalidad
EQ-5D o MSQoL-54	Calidad de vida
FIS	Fatiga
MSWS-12	Calidad de la marcha
MSIS-29	Impacto físico y psicológico de la enfermedad
SF-36	Calidad de Vida
TSQM	Satisfacción con el tratamiento

Notas: PRIMUS: *Patient-Reported Indices for Multiple Sclerosis*, EQ-5D: EuroQoL 5 dimensiones, MSQoL-54: *Multiple Sclerosis Quality of Life (54 ítems)*, FIS: *Fatigue Impact Scale*, MSWS-12: *Multiple Sclerosis Walking Scale (12 ítems)*, MSIS-29: *Multiple Sclerosis Impact Scale (29 ítems)*, SF-36: *Short Form 36 Health Survey*, TSQM: *Treatment Satisfaction Questionnaire for Medicines*.

Fuente: Elaboración propia.

■ 14.2.5. Métodos de expresión de los resultados de eficacia en ensayos clínicos

En los ensayos en esclerosis múltiple recurrente-remitente, la evidencia se expresa generalmente en forma de riesgo/proporción de evento, reducción del riesgo, reducción del riesgo relativo, razón de riesgo, y reducción relativa del número de lesiones en la resonancia magnética. Una forma muy clínica de expresar los resultados es por medio de los NNT (*Number Needed to Treat*). A pesar de las limitaciones metodológicas y las precauciones que hay que observar en la interpretación de los resultados (Okwuokenye *et al.*, 2017), resulta fácilmente comprensible para el médico clínico interpretar los valores NNT para tener una idea de la amplitud del efecto y así poder aplicar este conocimiento a la práctica individualizada con los pacientes. Se han realizado algunos estudios de revisión sistemática aplicando esta metodología a los ensayos clínicos en esclerosis múltiple recurrente-remitente, con resultados no siempre esperados (Mendes, Alves y Batel-Marques, 2016).

■ 14.3. RESULTADOS EN SALUD EN LA PRÁCTICA CLÍNICA

■ 14.3.1. Evaluación de la EFECTIVIDAD (Tur *et al.*, 2018; Karabudak *et al.*, 2015)

Los ensayos clínicos permiten conocer la eficacia de una molécula en un grupo poblacional bien definido y delimitado por unas características muy concretas, durante un periodo de seguimiento generalmente corto. Además, en los ensayos clínicos, dicha búsqueda de eficacia obliga al uso de medidas que se caractericen ante todo por una alta sensibilidad que permita discriminar entre las diferentes opciones terapéuticas. Contrariamente, en la práctica clínica las medidas usadas se emplean para monitorizar la evolución de la enfermedad durante largos periodos de tiempo (décadas), evaluar la respuesta al tratamiento y guiar las decisiones terapéuticas en un solo individuo, el paciente, procurándose por encima de todo que sean medidas de alta especificidad, aunque sea a expensas de cierta pérdida de sensibilidad. De esta forma, puede ser deseable utilizar algunos parámetros de los ensayos clínicos en la práctica habitual pero otros factores, entre ellos los PROM (aceptación del tratamiento y percepción del riesgo por parte del paciente, por ejemplo) van a cobrar gran importancia en este último contexto.

■ 14.3.1.1. Resultados clínicos

En la práctica clínica se siguen utilizando criterios de medición de brotes y de progresión de la discapacidad, ya que son los que caracterizan a la esclerosis múltiple (Tabla 3).

La medición del número de brotes en un periodo de tiempo dado, generalmente entre 6 meses y 1 año, es el elemento utilizado en la práctica clínica habitual para evaluar la actividad en brotes. La presencia de brotes, sobre todo en los primeros años de la enfermedad, se ha asociado tanto con una mayor posibilidad de alcanzar un EDSS de 6.0 (limitación severa de la movilidad) como de pasar a una forma secundariamente progresiva (Tremlett *et al.*, 2009).

La discapacidad se cuantifica generalmente gracias a la escala EDSS. Esta escala permite interpretar fácilmente a qué estado clínico corresponde cada valor numérico, sobre todo por encima de 4.0, valor a partir del cual la influencia de la deambulación en la puntuación es mayor. Como ya hemos señalado anteriormente, la EDSS peca de poca reproducibilidad intra e inter-evaluadores y tiene poca sensibilidad al cambio para datos de elevada importancia clínica como la destreza de miembros superiores y, sobre todo, la afectación cognitiva. Mucho menos utilizada en la práctica clínica, tal y como sucede en los ensayos clínicos, la escala combinada MSFC tiene la ventaja respecto a la EDSS de que evalúa bien la función

Tabla 3.

MEDIDAS CLÍNICAS Y RADIOLÓGICAS PRINCIPALES UTILIZADAS EN LA PRÁCTICA CLÍNICA

Objetivos clínicos y radiológicos principales utilizados en la práctica clínica

Objetivos clínicos

Número de brotes en un periodo de tiempo.

Cambio de EDSS en la evolución o puntuación al final de un periodo de seguimiento.

Timed 25 feet Walk Test (T25FW): z-score al final del periodo de seguimiento o cambio en el tiempo.

9-Hole Peg Test (9-HPT): z-score al final del periodo de seguimiento o cambio en el tiempo.

Paced Auditory Serial Addition Test (PASAT): z-score al final del periodo de seguimiento o cambio en el tiempo.

Symbol Digit Modalities Test (SDMT): número de aciertos en 90 segundos o z-score al final del periodo de seguimiento o cambio en el tiempo.

Fatigue Impact Scale (FIS), *Fatigue Severity Scale (FSS)*: puntuación al final del periodo de seguimiento.

Objetivos radiológicos

Lesiones encefálicas en T2: número de lesiones nuevas o que han aumentado de volumen entre dos estudios por resonancia magnética.

Lesiones encefálicas en T1 que realzan con Gadolinio: número de lesiones por estudio de resonancia magnética.

Lesiones encefálicas en T1 que no realzan con Gadolinio: número de lesiones por estudio de resonancia magnética.

Número de lesiones medulares en T2: número de lesiones por estudio de resonancia magnética.

Nota: EDSS: *Expanded Disability Status Scale*.

Fuente: Elaboración propia.

de miembros superiores y la cognitiva, pero como desventajas que necesita una población sana de referencia para comparar los valores y que obliga a definir qué es un cambio clínicamente significativo para que se confirme un deterioro funcional. Además, el PASAT presenta un fenómeno de aprendizaje por el cual pacientes que repiten las evaluaciones realizan éstas mejor, minimizando la capacidad de la herramienta de detectar posibles empeoramientos cognitivos en la evolución. Para intentar evitar este efecto de aprendizaje y, sobre todo, hacer una evaluación cognitiva más completa de los pacientes con esclerosis múltiple, se ha propuesto el uso de la batería BICAMS (*Brief International Cognitive Assessment for Multiple Sclerosis*). Esta batería se compone del *Symbol Digit Modalities Test (SDMT)*, que mide la atención y la velocidad de procesamiento de la información, el *California Verbal Learning Test* (segunda edición) y el *Brief Visuospatial Memory Test* (versión revisada). Un test aún más completo, pero que requiere disponer de una hora y media para cada paciente y que es más difícil de interpretar (generalmente requiere personal especializado tal como un Neuropsicólogo), es el BRB-N. Por último, debido a que la realización de test cognitivos a pacientes con esclerosis múltiple es difícil por la complejidad de los mismos pero, sobre todo, por las limitaciones de tiempo que la

práctica clínica impone, herramientas informatizadas como el CogState®, que son realizadas online y autoadministradas por el paciente, pueden ser una solución de compromiso en algunos casos, al menos como herramienta de triaje de pacientes con sospecha de deterioro cognitivo.

En la práctica clínica son muy útiles los PROM mencionados anteriormente, midiendo elementos como calidad de vida, depresión, ansiedad, fatiga y otros elementos como la funcionalidad en la vida personal y profesional, grado de satisfacción con el tratamiento, etc. El problema de estas medidas es su interpretación dada la gran variabilidad inter e intraindividual, que se ve compensada en los ensayos clínicos por el tamaño muestral pero no en la práctica clínica a nivel individualizado. Además, requieren un tiempo importante para ser completados, hay muchas escalas disponibles, ninguna ha demostrado una clara superioridad respecto a las otras, y dan una gran cantidad de información que requiere de tiempo y esfuerzo para convertirla en datos útiles para la práctica clínica. En cualquier caso, parece claro que las prioridades de los pacientes y de los clínicos no siempre son las mismas, siendo necesario medir las prioridades de los enfermos para tener toda la información necesaria respecto a la evolución de la esclerosis múltiple y su respuesta al tratamiento.

Los PROM, tal y como se ha descrito previamente, pueden valorar la calidad de vida del paciente o síntomas específicos de la enfermedad como la fatiga o la capacidad de caminar. Sin embargo, existen también escalas que permiten determinar el impacto funcional que la esclerosis múltiple tiene en diversos aspectos de la vida del enfermo. Entre estos, uno muy importante por su relevancia en adultos jóvenes es el impacto a nivel laboral. La *Multiple Sclerosis Work Difficulties Questionnaire* (MSWDQ-23) es una herramienta autoadministrada para definir los problemas en el trabajo relacionados con la esclerosis múltiple y abarca tres dominios: físico, psicológico/cognitivo y barreras externas. Su versión en español, validada recientemente, se muestra fácil de realizar, exacta y sensible para utilizarla como un PROM que nos indique las limitaciones que la esclerosis múltiple impone a nuestros pacientes en el ámbito laboral (Martínez-Ginés *et al.*, 2018).

Se ha realizado algún intento adicional de ampliar las escalas de impacto de la esclerosis múltiple no sólo al enfermo si no también al cuidador. Con estas escalas se pretende ver dicho impacto de la enfermedad en el cuidador/acompañante principal del paciente. La escala *caregiver health-related quality of life in multiple sclerosis* (CAREQOL-MS) es una escala específica para esta enfermedad, compuesta de 24 *ítems*, creada y validada en España. Su uso, no obstante, es marginal en la práctica clínica habitual (Benito-León *et al.*, 2011).

■ 14.3.1.2. Imagen por resonancia magnética (Sormani y Bruzzi, 2013)

Aunque no está perfectamente establecido, algunas recomendaciones internacionales sugieren la realización de resonancia magnética encefálica +/- medular

con una frecuencia anual en los pacientes con esclerosis múltiple, sobre todo si se encuentran en tratamiento. Si el tratamiento es el Natalizumab, y los pacientes son seropositivos para el virus John Cunningham, dicha frecuencia debe ser mayor. De la misma forma, se recomienda la realización de una resonancia magnética encefálica y/o medular ante la presencia de síntomas inesperados o atípicos. La resonancia, cuando se utiliza para la monitorización de la evolución de un paciente y para determinar la respuesta a un tratamiento, debe ser realizada idealmente utilizando la misma máquina y el mismo protocolo de recogida de imágenes que el de la resonancia magnética basal con la que se pretende comparar.

La medida de resonancia magnética más frecuentemente utilizada para determinar la evolución de la enfermedad y la respuesta a tratamiento es la presencia de lesiones ponderadas en T2 nuevas o de mayor tamaño. Su presencia se asocia a un peor pronóstico en los pacientes en tratamiento y puede indicar la necesidad de un cambio de tratamiento. La presencia de tres lesiones nuevas en T2 tras el primer año de tratamiento con Interferón β se asocia con un riesgo del 27 % de fracaso de tratamiento (incremento confirmado de EDSS o necesidad de cambio de tratamiento por falta de respuesta) y de un 22 % de riesgo de empeoramiento del EDSS en los siguientes tres años.

Otra medida de resonancia magnética habitualmente utilizada es la presencia de lesiones que realzan Gadolinio en secuencias ponderadas en T1. Esta secuencia permite la demostración de actividad inflamatoria sin la necesidad de comparar con una resonancia magnética previa, como ocurre con la determinación de nuevas lesiones en T2, manteniendo el valor predictivo de éstas en cuanto a mala evolución futura de la enfermedad. Como contrapartida, la fugacidad de las lesiones en T1 realzadas con Gadolinio, solamente de tres a seis semanas, hace que sea más difícil encontrarlas que las nuevas en T2 por lo que la determinación de éstas es preferible si no se pueden hacer las dos secuencias radiológicas. Además, el coste de la administración del contraste y los muy poco frecuentes, pero potencialmente graves, efectos adversos asociados a su uso son elementos a tener en cuenta al utilizar secuencias con Gadolinio.

La medición de la atrofia en la práctica clínica sigue siendo controvertida. Como hemos señalado, la progresión de la atrofia se asocia a mayor discapacidad física y cognitiva en los grupos de pacientes a estudio en los ensayos clínicos. Sin embargo, su aplicación como medida de uso común en la práctica clínica se ve limitada por diversos factores como la ausencia de claros valores normativos de pérdida de volumen cerebral en controles sanos de la misma edad y sexo, variaciones fisiológicas transitorias del volumen cerebral (debidas, por ejemplo, a deshidratación), y el uso de tratamientos inmunomoduladores que pueden modificar el volumen cerebral de forma transitoria al disminuir drásticamente la inflamación y, por lo tanto, el contenido de agua en el parénquima cerebral (concepto llamado pseudoatrofia).

■ 14.3.1.3. Medidas combinadas (clínicas + resonancia magnética)

No existe una evidencia muy sólida que soporte el uso del concepto NEDA (ausencia de brotes + ausencia de progresión en la EDSS + ausencia de lesiones nuevas/mayores en T2 y ausencia de lesiones que realcen con Gadolinio en T1) en la práctica clínica. Un estudio en 219 pacientes con EM mostró que aquellos pacientes que conseguían el estatus NEDA tras dos años de tratamiento tenían una alta probabilidad (78.3 %) de no progresar en la EDSS más de 0.5 puntos al cabo de siete años de seguimiento. Sin embargo, otro estudio similar realizado posteriormente no mostró beneficio del estatus NEDA al segundo año en cuanto al pronóstico en el seguimiento a 10 años. Más recientemente, un estudio retrospectivo de los pacientes del estudio fase III de Interferón β 1b confirmó que no tener brotes ni progresión de la discapacidad (NEDA clínico) tras dos años de tratamiento fue predictor de no fallecimiento, no progresión de la discapacidad a EDSS>6.0, no necesidad de silla de ruedas, y no paso a forma secundariamente progresiva en el seguimiento a 21 años ($p=0.0029$) (Goodin *et al.*, 2018). Los criterios radiológicos de NEDA no añadieron, en este estudio, un mejor valor predictivo. La conclusión que se puede extraer de los diferentes estudios realizados al respecto hasta el momento es que el estatus NEDA probablemente tenga un valor predictivo positivo alto pero un valor predictivo negativo bajo: ser NEDA es probablemente indicativo de un buen pronóstico a largo plazo, pero no serlo no necesariamente indica un mal pronóstico. La implementación en la práctica clínica del concepto NEDA-4, que añade la no progresión anormalmente elevada de la atrofia cerebral medida por resonancia magnética, tiene los mismos problemas de aplicación en la práctica clínica que el NEDA y no se ha estandarizado aún para valorar pacientes de forma individualizada.

A día de hoy, todavía no se han publicado datos sobre la aplicación del concepto NEPAD en la práctica clínica.

■ 14.3.2. Bases de datos de vida real para la obtención de resultados en salud

Como consecuencia de la actividad asistencial y la variedad de medidas de resultados en salud expuestas, el sector sanitario dispone de numerosas fuentes de datos heterogéneas con gran cantidad de información relacionada con los pacientes, las enfermedades y los centros sanitarios. Son enormes cantidades de datos sobre salud recogidos en historias clínicas electrónicas, dispositivos de telemedicina, pruebas clínicas, e incluso en dispositivos móviles.

Son datos de vida real (Real World Data) que ISPOR (International Society for Pharmacoeconomics and Outcomes Research) define como datos usados para tomar decisiones sobre medicamentos o procedimientos médicos que se recogen fuera de los ensayos clínicos aleatorizados. Del análisis riguroso de esta información se extrae la denominada evidencia de vida real (Real World Evidence). Podría-

mos decir que la evidencia en el mundo real es el equivalente médico de *big data*, pues hace uso de conjuntos de datos muy grandes para determinar cómo funcionan las intervenciones sanitarias más allá del alcance de los ensayos clínicos.

Así pues, el análisis de esa gran cantidad de información puede ayudar a tomar decisiones tanto a los médicos como a los gestores de los centros sanitarios, con el objetivo de obtener mejores resultados en salud, además en el campo de la investigación se muestra como una herramienta con enorme potencialidad.

Iniciativas como el proyecto ICHOM (*International Consortium for Health Outcomes Measures*) (<https://www.ichom.org/>) tiene por objetivo definir los resultados de valor para el paciente en distintas condiciones médicas. Para alcanzar sus metas necesitan utilizar grandes volúmenes de información. La misión concreta de ICHOM es descubrir el potencial de la atención médica basada en el valor, definiendo conjuntos estándar globales de medidas de resultado que realmente importan a los pacientes, promoviendo la adopción y el informe de estas medidas en todo el mundo. ICHOM de momento en el campo de la neurología no han abordado la EM, y si la demencia y la enfermedad de Parkinson.

A modo de ejemplo, las siguientes iniciativas de análisis de *big data*, todas ellas patrocinadas tanto por entidades públicas como privadas, tienen establecido como objetivo el avance en el conocimiento de la esclerosis múltiple:

The Big Multiple Sclerosis Data (BMSD) Network: una iniciativa que reúne datos de registros franceses, daneses, italianos y suecos. Un gran volumen de datos longitudinales que responde a una necesidad científica en EM. En 2017 disponían de datos de 138.148 pacientes con EM (Iaffaldano *et al.*, 2017).

The OPTIMIZE Project (<https://www.optimize-ms.org/>): proyecto de ámbito europeo que propone una solución frente al panorama de desarrollo de múltiples bases de datos de investigación clínica en EM durante las últimas décadas, con acceso limitado y problemas de integración directa, que además no estandarizan o personalizan eficazmente los datos relevantes para la farmacovigilancia y la investigación de efectividad clínica. Se plantea el desarrollo de una herramienta flexible y eficiente de registro de datos de clínicos y pacientes, una plataforma abierta, con herramientas asociadas (aplicaciones de teléfonos inteligentes) para la captura de datos de biomarcadores clínicos, datos reportados por el paciente, e imágenes. OPTIMIZE está diseñado para usuarios médicos, investigadores y pacientes.

The Big Data for Better Outcomes (<http://bd4bo.eu/>): un programa de investigación europeo cuyo fin es la transformación del sistema de atención médica mediante el uso de *big data*. La iniciativa desarrollará plataformas para integrar y analizar diversos conjuntos de datos, enfocándose en los resultados que importan a los pacientes, y generando ideas significativas que ayudarán a mejorar la atención médica. Entre las entidades participantes que se han sumado a este proyecto se encuentra la European Multiple Sclerosis Platform.

■ 14.3.3. Factores determinantes de los resultados en salud consecuencia de la farmacoterapia

Por tener la EM características de enfermedad crónica, los resultados clínicos obtenidos como consecuencia de los tratamientos farmacológicos se encuentran influenciados por factores como la seguridad (efectos adversos, interacciones farmacológicas) y la adherencia y persistencia al tratamiento, así como el grado satisfacción y de formación del paciente, relativos a su enfermedad y su tratamiento.

■ 14.3.3.1. Seguridad

Los fármacos o tratamientos modificadores de la enfermedad (TME) actualmente disponibles difieren con respecto a propiedades farmacológicas, mecanismos de acción, eficacia, perfil de riesgo y efectos secundarios. Para la evaluación de seguridad, los ensayos controlados aleatorios no son completamente capaces de identificar eventos adversos raros. La fuente más rica de datos de seguridad se encuentra en la fase posterior a la comercialización. La evidencia de vida real y los estudios observacionales reflejan la utilidad de los medicamentos en la práctica asistencial, y tienen la capacidad de descubrir efectos adversos poco comunes o reacciones raras a medicamentos. El perfil de seguridad de los fármacos es un factor que influye en la adherencia y persistencia al tratamiento. En la cohorte italiana de 520 pacientes seguida por Ferraro *et al.* (2018) durante un año, los efectos adversos fueron la razón más frecuente de interrupción de los TME, tanto orales (62,24 %) como inyectables (60,25 %). El seguimiento de los efectos adversos por parte del equipo clínico permite tomar medidas a tiempo encaminadas a resolverlos o minimizarlos. Ello repercute positivamente en la adherencia al tratamiento, la satisfacción del paciente y mejora de su calidad de vida.

■ 14.3.3.2. Interacciones farmacológicas

Los fármacos modificadores del curso de la enfermedad con frecuencia se acompañan de otros tratamientos farmacológicos (tratamientos sintomáticos, medicamentos sin receta, terapias alternativas iniciadas al margen del tratamiento médico convencional y sin control médico). Además, la EM, a pesar de debutar más frecuentemente en adultos jóvenes, se caracteriza por ser de larga evolución, de forma que los pacientes con el paso de los años pueden presentar comorbilidades que requerirán tratamiento farmacológico específico. Esto nos conduce a pacientes en los cuales la pluripatología y la polimedicación se añaden a la cronicidad que caracteriza la esclerosis múltiple. Es necesario tener en cuenta las interacciones farmacocinéticas y farmacodinámicas, debido a su potencial influencia en la efectividad y seguridad de los tratamientos. Así, encontramos la recomendación de no administrar vacunas atenuadas vivas con fármacos inmunomoduladores e inmunosupresores. En cuanto

a interacciones farmacocinéticas, a modo de ejemplo, citamos el incremento significativo de la exposición al Fingolimod por su administración conjunta con inhibidores del CYP3A4. Otro ejemplo lo encontramos en Teriflunomida, que inhibe entre otros el CYP2C8 in vivo, de forma que los medicamentos metabolizados por el CYP2C8 se deben utilizar con precaución (Bainbridge, 2015).

También pueden ser relevantes las interacciones fármaco-alimento. Por ejemplo, la recomendación tras la administración de dimetilfumarato de evitar durante la primera hora, beber cantidades mayores de 50ml de bebidas alcohólicas fuertes (con volumen de alcohol >30 %), por riesgo de gastritis (Ficha técnica Tecfidera®, 2015).

■ 14.3.3.3. Adherencia y persistencia

La falta de adherencia a los fármacos modificadores del curso de la enfermedad puede conducir a una progresión de la enfermedad más rápida e incapacitante y a un mayor consumo de recursos sanitarios. Se ha puesto de manifiesto a través de estudios observacionales tasas de adherencia al tratamiento por debajo del 80 % en proporciones relevantes de pacientes con EM. Así Burks, Marshall y Ye (2017) observaron una media de adherencia del 72,6 % en pacientes tratados con TME orales e inyectables, con porcentajes de pacientes no adherentes del 41,4 % y 38,6 % respectivamente. Otros autores arrojan cifras de 24,9 % de pacientes no adherentes (nivel de adherencia <80 %) a TME inyectables (Evans *et al.*, 2017). Se han identificado múltiples factores que pueden influir en el cumplimiento terapéutico: percepción de pérdida de efectividad, efectos adversos, así como factores neuropsiquiátricos como el olvido, la fatiga y la depresión como principales factores determinantes. Otros factores como EDSS, problemas con el dispositivo de inyección y frecuencia de la dosificación también han sido relacionados con la adherencia (Bartolomé-García *et al.*, 2017).

La persistencia se define como el tiempo durante el cual el paciente se administra la medicación desde el inicio hasta la interrupción por cualquier causa. La medida de la persistencia es relevante, entre otras razones porque es necesario un tiempo mínimo de duración del tratamiento para que los fármacos puedan desarrollar su efecto. Se han identificado como factores de discontinuación la edad al diagnóstico, edad del primer TME y EDSS basal (Bartolomé-García *et al.*, 2017; Correia *et al.*, 2016)

Es especialmente importante identificar a los pacientes que no siguen las recomendaciones de administración de sus tratamientos, analizar las causas y desarrollar estrategias individualizadas de tipo cognitivo, conductual, emocional y organizativo, para modificar estas actitudes o comportamientos, y con ello mejorar la efectividad de los tratamientos. La monitorización de la adherencia deberá ser continua y las estrategias para su mejora adaptables, por tratarse de un proceso dinámico.

Dado que el paciente con EM puede llegar a padecer un alto grado de discapacidad es necesario eliminar barreras físicas que faciliten el acceso a los fármacos y contribuyan a la mejora de la adherencia, por lo tanto, es importante desarrollar sistemas alternativos de dispensación (por ejemplo: dispensación a domicilio o *home delivery*).

■ 14.3.3.4. Formación/información al paciente

Un paciente informado está en condiciones de hacerse con las riendas del control de su enfermedad, por ello el empoderamiento del paciente es clave a la hora de lograr resultados en salud. Para alcanzar ese estatus, el paciente necesita de información precisa e imparcial. Este conocimiento puede ayudar a prevenir las expectativas poco realistas, por ejemplo, en los efectos de las terapias, y sirve como recurso para ayudar a superar barreras (efectos secundarios y comorbilidades). Sin embargo, la evidencia que sustenta estas afirmaciones es de momento débil. A pesar de ello, las guías clínicas sobre EM han reconocido la necesidad de una información equilibrada, unido a la participación del paciente en la toma de decisiones. El "Código de Buenas Prácticas" sobre los derechos de las personas con EM considera que la entrega de información clara y concisa de alta calidad, desde el diagnóstico en adelante, es necesaria para empoderar al paciente y así poder autogestionar su enfermedad en la medida de lo posible (European Multiple Sclerosis Platform, 2015).

En este sentido, la MSIF (Multiple Sclerosis International Federation) ha publicado los siete principios para mejorar la calidad de vida del paciente con EM (Multiple Sclerosis International Federation, 2016) (Tabla 4). Concretamente su primer punto proclama: *El empoderamiento, la independencia y el papel central de las personas afectadas por la EM a la hora de tomar decisiones que les afecten.*

Tabla 4.

SIETE PRINCIPIOS PARA MEJORAR LA CALIDAD DE VIDA

1	El empoderamiento, la independencia y el papel central de las personas afectadas por la EM a la hora de tomar decisiones que les afecten.
2	Acceso a tratamientos completos y eficaces, y prestación de las necesidades cambiantes relacionadas con la salud física y mental que surgen con la EM.
2	Apoyo a la red de familiares, amigos, seres queridos y cuidadores no remunerados.
4	Oportunidades de trabajo, voluntariado, educación y ocio accesibles y flexibles.
5	Lugares, tecnologías y transporte públicos y privados que sean accesibles.
6	Recursos económicos que se ajusten a las necesidades cambiantes y al coste de vida que supone la EM.
7	Actitudes, políticas y prácticas que promuevan la igualdad y planten cara al estigma y la discriminación.

Fuente: Multiple Sclerosis International Federation (2016).

■ 14.3.3.5. *Satisfacción del paciente*

Doyle, Lennox y Bell, 2013, en un metaanálisis que incluye una amplia gama de grupos demográficos y patologías, encuentran evidencia de relación entre la experiencia positiva del paciente y la efectividad clínica. Las preferencias de los pacientes es uno de los factores a considerar para lograr un elevado grado de satisfacción con el tratamiento. Un estudio internacional realizado sobre 246 pacientes con EM (IMPrESS, 2018) (Kanavos *et al.*, 2016) muestra que los pacientes otorgan más valor a aspectos como la comodidad y la seguridad que los médicos. La incorporación de nuevos fármacos al arsenal terapéutico ha aumentado la complejidad del manejo farmacológico de la EM, pero a su vez, ha permitido adaptar los tratamientos a las diferentes características de los pacientes. Así, las preferencias de los pacientes son uno de los atributos a tener en cuenta a la hora de seleccionar el tratamiento, además de la efectividad, la seguridad y la eficiencia.

Al incorporar las preferencias de los pacientes en la toma de decisiones, el paciente se está corresponsabilizando y participando en la gestión de su propia salud. Para tomar decisiones el paciente necesita disponer de información suficiente para entender la enfermedad y su tratamiento. La adquisición de este conocimiento, le permitirá convertirse en un paciente con capacidad para decidir y resolver problemas, y con control sobre su vida, es decir, un paciente con capacidad de autogestión. Proporcionar información veraz, en lenguaje sencillo es responsabilidad de los profesionales sanitarios que atienden al paciente. Actuaciones encaminadas a promocionar la autogestión del paciente, tendrán como consecuencia mejoras en los resultados en salud (Jadad *et al.*, 2010).

■ 14.3.4. **Modelo de atención sanitaria multidisciplinar y resultados en salud**

El trabajo multidisciplinar centrado en el paciente crónico constituye una herramienta para la mejora de los resultados en salud. El modelo de atención sanitaria centrada en los pacientes que hace un mejor uso de las habilidades de los distintos profesionales sanitarios intervinientes es el basado en equipos multidisciplinarios. Las interacciones entre sus miembros no solo mejoran el resultado final, además optimizan los recursos empleados en el proceso.

Este modelo de atención es el más adecuado para atender las necesidades del paciente con EM, por su carácter de enfermedad crónica. Los pilares fundamentales en la atención al paciente de EM están constituidos por la atención médica, farmacéutica y de enfermería, así como rehabilitadores, neuropsicólogos, asistentes sociales. Alcanzaremos resultados de salud óptimos actuando coordinadamente e incorporando la perspectiva de los pacientes.

La integración del farmacéutico hospitalario (FH) en el equipo asistencial que atiende al paciente con EM permite la realización de actividades de educación sobre

medicamentos y evaluación de su uso correcto en el inicio ya lo largo toda la terapia. Además, el contacto más frecuente del FH con el paciente, ligado generalmente, pero no exclusivamente, al acto de dispensación, permite fomentar la adherencia, el uso adecuado de medicamentos, la monitorización continua de la seguridad, y proporciona una vía de comunicación del paciente con el resto del equipo. Esto facilita el acceso del paciente al sistema, incrementando también su confianza. La mejora de los resultados del paciente es el objetivo prioritario de este modelo de atención que se caracteriza por estar centrada en el paciente. Existen evidencias de que este modelo beneficia clínicamente al paciente con EM, y como consecuencia, contribuye a aumentar la productividad en el trabajo, y disminuir la tasa de consumo de otros recursos de atención médica. Distintas experiencias (Hanson *et al.*, 2014 y Taddei-Allen *et al.*, 2017) desde diferentes modelos de provisión de atención sanitaria muestran resultados satisfactorios.

Desde la Sociedad Española de Farmacia Hospitalaria se ha iniciado una reingeniería de procesos de atención al paciente con tratamiento autoadministrado en domicilio denominada proyecto MAPEX. El proyecto se ha subdividido en patologías prevalentes, entre las cuales se encuentra la EM. Como consecuencia de esta iniciativa, ha surgido un nuevo grupo de trabajo entorno al paciente con EM (grupo GEAFEN), que agrupa a los farmacéuticos hospitalarios dedicados a la atención de los pacientes con enfermedades neurodegenerativas en todos los ámbitos de atención hospitalaria.

Del proyecto MAPEX también ha surgido un nuevo enfoque de atención farmacéutica, el denominado modelo CMO. Este modelo incorpora como novedades más destacadas la estratificación de pacientes y la entrevista motivacional como herramienta fundamental para alcanzar los objetivos de resultados en salud del paciente. En consecuencia, se postula que el paciente afectado de EM con tratamiento autoadministrado en domicilio, por su complejidad, encontrará en este modelo de atención farmacéutica un abordaje pleno de sus necesidades de atención.

■ 14.4. RESULTADOS DE EFICIENCIA EN ESCLEROSIS MÚLTIPLE

■ 14.4.1. Datos de eficacia y efectividad

Con un número creciente de fármacos modificadores del curso de la enfermedad para el tratamiento de la esclerosis múltiple, la necesidad de datos comparativos de eficacia y efectividad es cada vez más relevante. Sin embargo, los datos provenientes de ensayos clínicos o de la práctica clínica no son totalmente satisfactorios en este sentido. En los ensayos clínicos, la escasa duración de estos, la escasez de estudios comparativos entre diferentes alternativas, y lo restrictivo de las poblaciones evaluadas son algunos de los factores que limitan la utilidad de los datos. En la vida real, la falta de grupos control, la heterogeneidad de los grupos

de estudio, el alto número de pacientes que se pierden en los seguimientos, y un diseño frecuentemente retrospectivo tampoco permiten la extracción de datos concluyentes en muchas ocasiones.

Por estos motivos, cuando se intenta hacer una evaluación de la eficiencia de los medicamentos para el tratamiento de la esclerosis múltiple, una enfermedad de larga duración (más de 20 años de media desde el diagnóstico) se recurre a los datos de eficacia/efectividad en brotes, discapacidad, datos clínicos y radiológicos tanto de ensayos clínicos como de estudios observacionales, pero, para reducir el grado de incertidumbre se recurre a modelizaciones que permitan tomar decisiones.

■ 14.4.2. Costes

Tal y como sucede en cualquier otra patología, la atención al paciente con esclerosis múltiple acarrea unos costes económicos que se pueden dividir en costes directos, indirectos y costes intangibles. Desde hace casi dos décadas (Henriksson *et al.*, 2001) sabemos que el coste global del cuidado de los pacientes aumenta en paralelo al incremento de la discapacidad de los enfermos. Los fármacos modificadores del curso de la enfermedad representaron el mayor peso dentro de los gastos directos, en pacientes con EDSS ≤ 3.0 . Sin embargo, los costes directos, exceptuando dichos fármacos, y los recursos no sanitarios fueron mayores en pacientes con discapacidades severas (EDSS ≥ 7.0). Este comportamiento en cuanto a los costes es reproducible de forma muy semejante en todos los países europeos, con ligeras diferencias regionales (Kobelt *et al.*, 2017). En España, el coste anual por paciente con esclerosis múltiple se sitúa en los siguientes valores: 20.600 euros para un paciente con EDSS entre 0 y 3, 48.500 euros para un paciente con EDSS entre 4.0 y 6.5, 68.700 euros para un paciente con EDSS entre 7 y 9 de EDSS. Dicho incremento en los costes no sólo no es paralelo al del tratamiento inmunomodulador si no inverso, ya que éste es de 12.644, 15.598 y 7.419 para cada uno de los tres grupos, respectivamente. Son fundamentalmente los costes informales los que incrementan la tarifa en los pacientes más discapacitados (Oreja-Guevara *et al.*, 2017).

■ 14.4.3. Modelización

No habiendo dudas en cuanto a la efectividad de una intervención precoz en la EM, por su impacto en una menor acumulación de secuelas ¿se puede considerar que nuestras decisiones terapéuticas son coste-efectivas? Desde un punto de vista global/social, tratar pronto supone un gasto en fármacos costosos (impacto presupuestario, coste directo) en el paciente nada o levemente discapacitado que se compensaría con creces más adelante evitando una acumulación de discapacidad asociada a unos mayores costes globales (costes directos + costes indirectos +

costes de cuidado informal), como se ha mostrado anteriormente (Henriksson *et al.*, 2001; Kobelt *et al.*, 2017; Oreja-Guevara *et al.*, 2017).

Los estudios de eficiencia en esclerosis múltiple suelen ser de tipo coste-efectividad (generalmente basados en medidas de brotes y discapacidad medida por la escala EDSS) o de coste-utilidad (años de vida ajustados por la calidad, AVAC). Deben basarse en los objetivos clínicos y radiológicos descritos anteriormente para los ensayos clínicos y en los costes conocidos para las diferentes fases de la enfermedad, utilizando una modelización a un tiempo no inferior a 20 años ya que es el periodo mínimo de duración de la esclerosis múltiple, desde su diagnóstico. En este sentido, se han publicado recomendaciones específicas para la modelización de estudios de eficiencia de los tratamientos de la esclerosis múltiple (Thompson, Abdolahi y Noyes, 2013).

■ 14.4.4. Gestión económica farmacoterapéutica de la esclerosis múltiple y resultados en salud

Los tratamientos utilizados en esclerosis múltiple pueden modificar la historia natural de la enfermedad, pero tienen un elevado impacto económico. Los sistemas sanitarios se enfrentan a una crisis de sostenibilidad que, unido a la necesidad de incorporar la innovación, obliga a los gestores a un cambio de enfoque y a un cambio de lenguaje que ha sustituido el término gasto por el término inversión. Es preciso conocer los resultados obtenidos con los recursos económicos consumidos, y así avanzar en la consecución de los resultados en salud más eficientes. Herramientas clásicas como estudios de impacto presupuestario, o el establecimiento y monitorización de indicadores económicos y de actividad, unidos a nuevas herramientas como los indicadores de eficiencia y el empleo de nuevos modelos de financiación, son necesarias para este cambio de enfoque en la gestión.

Es imprescindible estimar el impacto presupuestario al incorporar una innovación farmacológica. La prevalencia de la EM y el número de pacientes a tratar con la nueva alternativa terapéutica, son aspectos determinantes, junto con el coste de la innovación. Así, por ejemplo, el Servicio Catalán de Salud realizó un análisis del impacto presupuestario debido a la incorporación de teriflunomida, dimetil fumarato e interferón pegilado beta-1a. Dado que el coste anual de estos fármacos era similar al de los fármacos disponibles en ese momento para la primera línea de la EM-RR, se concluyó que su incorporación no conllevaría un aumento significativo del impacto presupuestario (Servicio Catalán de Salud, 2014). Al tiempo, es imprescindible establecer mecanismos y políticas sanitarias que faciliten y agilicen la incorporación de fármacos genéricos y biosimilares al arsenal terapéutico asistencial, para así disminuir costes que amortiguarán el impacto financiero derivado de la incorporación de la innovación farmacológica.

Los indicadores de calidad se utilizan cada vez más para la evaluación comparativa, como una herramienta de auditoría, o para medir el efecto de las interven-

ciones. El establecimiento por parte de la administración sanitaria de indicadores ligados a los resultados en salud, además de los clásicos indicadores económicos, de actividad y de proceso, constituyen una herramienta de ayuda a la gestión eficiente (Calvo, 2017). Algunos ejemplos de indicadores se muestran en la siguiente tabla (Tabla 5).

Tabla 5.

EJEMPLOS DE INDICADORES RELACIONADOS CON LA FARMACOTERAPIA EN EM

Tipo indicadores	Ejemplos
Indicadores económicos	Coste tratamiento por paciente EM.
Indicadores de actividad	Número de pacientes con EM en TME.
Indicadores de proceso	Porcentaje de prescripciones de TME que se ajustan al protocolo institucional.
Indicadores de resultados en salud	Proporción de pacientes en tratamiento con TME sin progresión en la escala EDSS en un período fijado. Proporción de pacientes con tasa anual de brotes ≤ 1 .

Nota: EM: Esclerosis Múltiple, TME: Tratamiento Modificador de la Enfermedad, EDSS: *Expanded Disability Status Scale*

Fuente: Elaboración propia.

Es necesario recurrir a nuevos modelos de financiación con el enfoque de pago por resultados, como los acuerdos de riesgo compartido ya sean financieros o de resultados en salud. Estos acuerdos deben ser sensibles a las modificaciones que pueda experimentar un mercado tan cambiante como el farmacéutico. Entre los acuerdos financieros se encuentra el techo de gasto por patología, que podría ser por ejemplo la sumatoria de los techos de gasto de cada uno de los TME utilizados en la EM. Para los acuerdos de resultado en salud, es preciso concretar las indicaciones y establecer las características del paciente candidato, seleccionar las variables de resultado a monitorizar, que deben estar estandarizadas y ser fácilmente medibles, y además fijar la frecuencia de medida (Espín, Oliva y Rodríguez-Barrios, 2010). Pueden seleccionarse variables basadas en las medidas clínicas como número de brotes en un periodo de tiempo, cambio de EDSS en la evolución, cambio en el tiempo del resultado del Test T25FW, o también, variables basadas en medidas radiológicas. Existen experiencias de acuerdos de pagos por resultados a nivel internacional, algunos controvertidos, que disponen de resultados de eficiencia publicados (Palace *et al.*, 2015; Slof, Ruiz y Vila, 2015). Ejemplos de estos acuerdos se muestran en la Tabla 6.

Son todos ellos ejemplos de modelos de colaboración entre gestores de salud e industria innovadora basados en una retribución por los resultados en salud conseguidos. Se trata de soluciones cuyo objetivo es contribuir a la sostenibilidad del sistema sanitario, facilitando la incorporación de nuevas tecnologías en un escena-

Tabla 6.

EJEMPLOS DE ACUERDOS DE PAGO POR RESULTADOS

País/año del acuerdo	Fármaco	Descripción del acuerdo de pago por resultados
Reino Unido/2002 Department of Health, 2002	Interferón beta y acetato de glatiramero	Los precios de los medicamentos se establecieron cerca de los niveles comerciales, pero se reducirían si los resultados de los pacientes fueran inferiores a los requeridos para cubrir un costo por QALY de 36000€.
Italia/2013 Agenzia Italiana del Farmaco, 2014	Nabiximols (Sativex®)	Acuerdo de participación en el coste (<i>costsharing</i>), con 50 % de reembolso del coste de 6 semanas o 3 semanas gratuitas para todos los pacientes elegibles, y un acuerdo de pago por resultados, con el reembolso del 100 % del coste en los pacientes que no responden a las 4 semanas (sin mejoría del 20 % en NRS- <i>Numerical Ranting Scale</i> para espasticidad)
España/2013 Ministerio de Sanidad Servicios Sociales e Igualdad, 2013	Fampridina (Fampyra®)	Las dos primeras semanas de tratamiento sin coste, tras este periodo, si se consigue una mejoría de la marcha ≥ 20 % evidenciada mediante la aplicación del test de los 25 pies y la escala de 12 <i>ítems</i> , el paciente continúa la terapia financiándolo el Sistema Nacional de Salud

Nota: EM: Esclerosis Múltiple, TME: Tratamiento Modificador de la Enfermedad, EDSS: *Expanded Disability Status Scale*

Fuente: Elaboración propia.

rio cambiante debido al continuo desarrollo que está experimentando el abordaje terapéutico de la EM.

14.5. CONCLUSIONES

- La esclerosis múltiple (EM) es una enfermedad crónica, inflamatoria, desmielinizante y degenerativa del sistema nervioso central que se diagnostica en torno a los 30 años de edad.
- Se caracteriza por la presencia de episodios desmielinizantes agudos de origen inflamatorio, responsables de los brotes clínicos del fenotipo más frecuente de la enfermedad.
- Por otro lado, existe un proceso neurodegenerativo que, independientemente de los brotes, se relaciona íntimamente con la acumulación progresiva de síntomas y signos asociados con la enfermedad.
- Tanto en los ensayos clínicos como en la práctica clínica, el efecto de un fármaco para la esclerosis múltiple se evalúa midiendo sus resultados sobre la actividad inflamatoria y sobre el acúmulo de discapacidad.

- La actividad inflamatoria se determina con parámetros clínicos (número, frecuencia, gravedad de los brotes) y paraclínicos (lesiones que realzan contraste con Gadolinio en secuencias ponderadas en T1 y el incremento del número o el tamaño de las lesiones en secuencias ponderadas en T2).
- Lo mismo sucede con la discapacidad, aunque los parámetros de medición son diferentes (empeoramiento en una escala de discapacidad, generalmente la EDSS, como parámetro clínico; atrofia cerebral, como ejemplo de parámetro radiológico).
- Unos objetivos de eficacia/efectividad más exigentes son las medidas combinadas ya que valoran conjuntamente la actividad inflamatoria y la progresión de la discapacidad mediante parámetros clínicos y radiológicos: NEDA (*No Evidence of Disease Activity*); NEPAD (*No Evidence of Progression or Active Disease*).
- En la práctica clínica, en una enfermedad crónica y compleja como la esclerosis múltiple, cabe señalar la importancia de las bases de datos de vida real para la obtención de resultados en salud.
- Además, por ser una enfermedad crónica, los resultados clínicos obtenidos en la esclerosis múltiple como consecuencia del uso de medicamentos se encuentran influenciados por factores como los efectos adversos y las interacciones farmacológicas de los mismos, la adherencia y la persistencia al tratamiento, la formación/información de los pacientes que los usan y el grado de satisfacción de éstos con su utilización.
- En cuanto los resultados de eficiencia de los tratamientos para la esclerosis múltiple, los estudios al respecto deben basarse en los objetivos clínicos y radiológicos descritos anteriormente para los ensayos clínicos y en los costes conocidos para las diferentes fases de la enfermedad, utilizando una modelización a un tiempo no inferior a 20 años ya que es el periodo mínimo de duración de la esclerosis múltiple, desde su diagnóstico.

■ REFERENCIAS

- AGENZIA ITALIANA DEL FARMACO. (2014). [Internet]. Disponible en: http://www.agenziafarmaco.gov.it/sites/default/files/43_Sativex_18.02.2014.zip [acceso 19 de noviembre de 2018].
- BAINBRIDGE, J. L. (2015). Putting the Action in Drug-Interaction: Disease Modifying Therapies (DMTs) in the Aging Adult. En: *2015 The Consortium of Multiple Sclerosis Centers Annual Meeting*. Indianapolis, 7-30 mayo, 2015. Disponible en: https://cdn.ymaws.com/www.ms-care.org/resource/resmgr/2015AM/S15_Bainbridge_MS_Symptoms_D.pdf [acceso 19 de noviembre de 2018].
- BALCER, L. J., RAYNOWSKA, J., NOLAN, R. *et al.* (2017). Validity of low-contrast letter acuity as a visual performance outcome measure for multiple sclerosis. *MultScler*, 23(5), pp. 734-747.
- BARTOLOMÉ-GARCÍA, E., USARRALDE-PÉREZ, Á., SANMARTÍN-FENOLLERA, P. y PÉREZ ENCINAS, M. (2017). Persistence and adherence to interferon and glatiramer acetate in patients with multiple sclerosis *Eur J Hosp Pharm*. Published online first: 31 August 2017. doi: 10.1136/ejpharm-2017-001286.

- BENITO-LEÓN, J., RIVERA-NAVARRO, J., GUERRERO, A. L. *et al.* (2011). Caregiver quality of life in multiple sclerosis (CAREQOL-MS) Study Group. The CAREQOL-MS was a useful instrument to measure caregiver quality of life in multiple sclerosis. *J Clin Epidemiol*, 64, pp. 675–686.
- BURKS, J., MARSHALL, T. S. y YE, X. (2017). Adherence to disease-modifying therapies and its impact on relapse, health resource utilization, and costs among patients with multiple sclerosis. *Clinicoecon Outcomes Res*, 9, pp. 251-260.
- CALVO, M. J. (2017). Visión y alternativas desde la gestión en las Comunidades Autónomas. Medicamentos, precios y salud. En: *62 Congreso Nacional SEFH*, Madrid, 18-21 octubre. . Disponible en: https://www.sefh.es/eventos/62congreso/ponencias/maria_jose_calvo_SEFH_2017_DEFINITIVA.pdf [acceso 20 de noviembre de 2018].
- COHEN, J. A., REINGOLD, S. C., POLMAN, C. H. y WOLINSKY, J. S. (2012). International Advisory Committee on Clinical Trials in Multiple Sclerosis. Disability outcome measures in multiple sclerosis clinical trials: current status and future prospects. *Lancet Neurol*, 11(5), pp. 467-476.
- CORREIA, I., MARQUES, I. B., SOUSA, M. *et al.* (2016). Predictors of first-line treatment persistence in a Portuguese cohort of relapsing-remitting multiple sclerosis. *J Clin Neurosci*, 36, pp. 73–78.
- DEPARTMENT OF HEALTH. (2002). Cost effective provision of disease modifying therapies for people with multiple sclerosis. London: Health Service Circular (2002/004), Stationery Office.
- DOYLE, C., LENNOX, L. y BELL, D. (2013). A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open*, 3(1), pii: e001570. doi: 10.1136/bmjopen-2012-001570
- ESPÍN, J., OLIVA, J. y RODRIGUEZ-BARRIOS, J. M. (2010). Esquemas innovadores de mejora del acceso al Mercado de nuevas tecnologías: los acuerdos de riesgo compartido. *Gac Sanit*, 24(6), pp. 491–497.
- EUROPEAN MULTIPLE SCLEROSIS PLATFORM. (2015). *The European Code of Good Practice in MS*. Disponible en: <http://www.emsp.org/wp-content/uploads/2015/06/140824-Revised-Code-of-Good-Practice.pdf> [acceso 19 de noviembre de 2018].
- EVANS, C., MARRIE, R. A., ZHU, F. *et al.* (2017). Adherence to disease-modifying therapies for multiple sclerosis and subsequent hospitalizations. *Pharmacoepidemiol Drug Saf*, 26(6), pp. 702-711.
- FERRARO, D., CAMERA, V., BALDI, E. *et al.* (2018). First-line disease-modifying drugs in relapsing-remitting multiple sclerosis: an Italian real-life multicenter study on persistence. *Curr Med Res Opin*, 12, pp. 1-5.
- FICHA TÉCNICA TECFIDERA® 2015 [acceso 10 de noviembre de 2018]. Disponible en: http://www.ema.europa.eu/docs/es_ES/document_library/EPAR_-_Product_Information/human/002601/WC500162069.pdf [acceso 1 de junio de 2016].
- GIOVANNONI, G., TOMIC, D., BRIGHT, J. R. y HAVRDOVÁ, E. (2017). "No evident disease activity": The use of combined assessments in the management of patients with multiple sclerosis. *Mult Scler*, 23(9), pp. 1179-1187.
- GOODIN, D. S., REDER, A. T., TRABOULSEE, A. L. *et al.* (2018). IFNB Multiple Sclerosis Study Group and the 16- and 21-Year LTF Investigators. Predictive validity of NEDA in the 16- and 21-year follow-up from the pivotal trial of interferon beta-1b. *Mult Scler*, 2018, 1352458518773511.
- HANSON, R. L., HABIBI, M., KHAMO, N., ABDU, S. y STUBBINGS, J. (2014). Integrated clinical and specialty pharmacy practice model for management of patients with multiple sclerosis. *American Journal of Health-System Pharmacy*, 71(6), pp. 463-469.
- HENRIKSSON, F., FREDRIKSON, S., MASTERMAN, T. y JÖNSSON, B. (2001). Costs, quality of life and disease severity in multiple sclerosis: a cross sectional study in Sweden. *Eur J Neurol*, 8, pp. 27-35.
- IAFFALDANO, P., BUTZKUEVEN, H., HILLERT, J. *et al.* (2017). Big Multiple Sclerosis Data network: data sharing among five large MS registries. En: *ACTRIMS-ECTRIMS 2017*. París, 26 de octubre, 2017; abstract200393. Disponible en: <https://onlinelibrary.ectrims-congress.eu/ectrims/2017/ACTRIMS->

- ECTRIMS2017/200393/pietro.iaffaldano.big.multiple.sclerosis.data.network.data.sharing.among.five.html [acceso 20 de noviembre de 2018].
- JADAD, A., CABRERA, A., LYONS, R. F., MARTOS, F. y SMITH, R. (2010). When people live with multiple chronic diseases: A collaborative approach to an emerging global challenge. *EASP*, 5, pp. 117-138.
- KANAVOS, P., TINELLI, M., EFTHYMIADOU, O., VISINTIN, E., GRIMACCIA, F. y MOSSMAN, J. (2016). *Towards better outcomes in multiple sclerosis by addressing policy change: The International MultiPIE Sclerosis Study (IMPrESS)*. The London School of Economics and Political Science, 2016 Disponible en: <http://www.lse.ac.uk/LSEHealthAndSocialCare/research/LSEHealth/MTRG/IMPRESS-Report-March-2016.pdf> [acceso 19 de noviembre de 2018].
- KARABUDAK, R., DAHDALEH, M., ALJUMAH, M. *et al.* (2015). Functional clinical outcomes in multiple sclerosis: Current status and future prospects. *Mult Scler Relat Disord*, 4(3), pp. 192-201.
- KHURANA, V., SHARMA, H., AFROZ, N., CALLAN, A. y MEDIN, J. (2017). Patient-reported outcomes in multiple sclerosis: a systematic comparison of available measures. *Eur J Neurol*, 24(9), pp. 1099-1107.
- KOBELT, G., THOMPSON, A., BERG, J., GANNEDAHL, M. y ERIKSSON, J. (2017). MSCOI Study Group; European Multiple Sclerosis Platform. New insights into the burden and costs of multiple sclerosis in Europe. *Mult Scler*, 23(8), pp. 1123-1136.
- LAVERY, A. M., VERHEY, L. H. y WALDMAN, A. T. (2014). Outcome measures in relapsing-remitting multiple sclerosis: capturing disability and disease progression in clinical trials. *MultSclerInt*, 2014, p. 262350.
- MARTÍNEZ-GINÉS, M. L., GARCÍA-DOMÍNGUEZ, J. M., FORERO, L. *et al.* (2018). Spanish validation of a specific measure to assess work-related problems in people with multiple sclerosis: The Multiple Sclerosis Work Difficulties Questionnaire (MSWDQ-23). *Mult Scler Relat Disord*, 22, pp. 115-119.
- MENDES, D., ALVES, C. y BATEL-MARQUES, F. (2016). Benefit-risk of therapies for relapsing-remitting multiple sclerosis: testing the Number Needed to Treat to Benefit (NNTB), Number Needed to Treat to Harm (NNTH) and the Likelihood to be Helped or Harmed (LHH): a systematic review and meta-analysis. *CNS Drugs*, 30(10), pp. 909-929.
- MEYER-MOOCK, S., FENG, Y. S., MAEURER, M., DIPPEL, F. W. y KOHLMANN, T. (2014). Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis. *BMC Neurol*, 14, p. 58.
- MINISTERIO DE SANIDAD SERVICIOS SOCIALES E IGUALDAD. (2013). *Nomenclátor*, suplemento septiembre 2013. Real e Ilustre Colegio de Farmacéuticos de Sevilla [Internet] Disponible en: <http://profesionales.farmaceticosdesevilla.es/opencms/export/sites/default/Proyecto/proyecto/RICOFs/Circulares/Circulares2013381.pdf> [acceso 20 de noviembre de 2018].
- MULTIPLE SCLEROSIS INTERNATIONAL FEDERATION. (2016). *Seven principles to improve quality of life*. Disponible en: <https://www.msif.org/living-with-ms/what-influences-quality-of-life/seven-principles-to-improve-quality-of-life/> [acceso 19 de noviembre de 2018].
- VAN MUNSTER, C. E. y UITDEHAAG, B. M. (2017). Outcome measures in clinical trials for multiple sclerosis. *CNS Drugs*, 31(3), pp. 217-236.
- OKWUOKENYE, M., ZHANG, A., PACE, A. y PEACE, K. E. (2017). Number Needed to Treat in multiple sclerosis clinical trials. *Neurol Ther*, 6(1), pp. 1-9.
- OREJA-GUEVARA, C., KOBELT, G., BERG, J., CAPSA, D. y ERIKSSON, J. (2017). European Multiple Sclerosis Platform. New insights into the burden and costs of multiple sclerosis in Europe: results por Spain. *Mult Scler*, 23(2_suppl), pp. 166-178.
- PALACE, J., DUDDY, M., BREGENZER, T. *et al.* (2015). Effectiveness and cost-effectiveness of interferon beta and glatiramer acetate in the UK Multiple Sclerosis Risk Sharing Scheme at 6 years: a clinical cohort study with natural history comparator. *Lancet Neurol*, 14(5), pp. 497-505.
- PARKS, N. E., FLANAGAN, E. P., LUCCHINETTI, C. F. y WINGERCHUCK, D. M. (2017). NEDA treatment target? No evident disease activity as an actionable outcome in practice. *J Neurol Sci*, 383, pp. 31-34.

- SAIDHA, S., AL-LOUZI, O., RATCHFORD, J. N. *et al.* (2015). Optical coherence tomography reflects brain atrophy in multiple sclerosis: a four-year study. *Ann Neurol*, 78(5), pp. 801-813.
- SERVICIO CATALÁN DE SALUD. (2014). *Informe técnico CAMHDA dimetil-fumarato-teriflunomida-peginterfero-beta1a*. Disponible en: http://catsalut.gencat.cat/web/.content/minisite/catsalut/proveidors_professionals/medicaments_farmacia/harmonitzacio/informes/dimetilfumarat/Informe-tecnic-CAMHDA-dimetil-fumarat-teriflunomida-i-peginterfero-beta1a.pdf [acceso 20 de noviembre de 2018].
- SLOF, J., RUIZ, L. y VILA, C. (2015). Cost-effectiveness of Sativex in multiple sclerosis spasticity: new data and application to Italy. *Expert Rev Pharmacoecon Outcomes Res*, 15(3), pp. 379-391.
- SORMANI, M. P. y BRUZZI, P. (2013). MRI lesions as a surrogate for relapses in multiple sclerosis: a meta-analysis of randomised trials. *Lancet Neurol*, 12(7), pp. 669-676.
- TADDEI-ALLEN, P., KING, M., OELOFSEN, M. y PAGE, R. (2017). Implementing a Pharmacist-Led Multiple Sclerosis Specialty Pharmacy Clinical Outreach Program. Poster NASP.
- THOMPSON, J. P., ABDOLAH, A. y NOYES, K. (2013). Modelling the cost effectiveness of disease-modifying treatments for multiple sclerosis: issues to consider. *Pharmacoeconomics*, 31(6), pp. 455-469.
- TREMLET, H., YOUSEFI, M., DEVONSHIRE, V., RIECKMANN, P. y ZHAO, Y. (2009). UBC Neurologists. Impact of multiple sclerosis relapses on progression diminishes with time. *Neurology*, 73(20), pp.1616-1623.
- TUR, C., MOCCIA, M., BARKHOF, F. y cols. (2018). Assessing treatment outcomes in multiple sclerosis trials and in the clinical setting. *Nat Rev Neurol*, 14(2), pp. 75-93.
- WOLINSKY, J., MONTALBAN, X., KAPPOS, L. *et al.* (2017). Evaluation of no evidence of progression active disease (NEPAD) in patients with primary progressive multiple sclerosis in the ORATORIO trial. *J Neurol Neurosurg Psychiatry*, 88(5), pp. e31-e32.



15

BENEFICIO CLÍNICO EN ONCOLOGÍA: UNA COMPLEJA PREGUNTA

GERARDO CAJARAVILLE ORDOÑANA,
JOSÉ IGNACIO CHACÓN LÓPEZ-MUÑIZ,
ANA ROSA RUBIO SALVADOR, RUTH VERA GARCÍA

■ 15.1. JUSTIFICACIÓN DEL CAPÍTULO

El cáncer es uno de los mayores problemas de salud actualmente y el desarrollo de nuevos fármacos para su tratamiento es, en este momento, uno de los aspectos más relevantes de la compleja situación que rodea el abordaje terapéutico de esta enfermedad, tanto en términos de eficacia, efectividad, seguridad y eficiencia como en términos de sostenibilidad y acceso.

Un elemento esencial en este contexto es la necesidad de dinamismo que requiere el proceso de desarrollo de un nuevo fármaco, desde la investigación básica de la farmacodinamia de una nueva molécula y la confirmación de su eficacia clínica en la población diana, hasta su incorporación a la práctica clínica de manera rutinaria, con un lugar en terapéutica bien definido y con mecanismos de acceso rápidos, eficientes y equitativos.

En este contexto de desarrollo clínico, ágil por necesidad, la optimización de la eficiencia económica emerge como parte esencial de un proceso continuo de toma de decisiones de atención médica, que pretende alcanzar altos niveles de calidad de atención y seguridad con los recursos disponibles.

Algunas herramientas analíticas que se utilizan para determinar el valor fármaco-económico de una intervención sanitaria son los análisis de coste-beneficio, coste-efectividad, coste-utilidad y minimización de costes. Estas herramientas fundamentan su análisis en tres conceptos principales: eficacia, efectividad y eficiencia. La eficacia es la medida de los resultados obtenidos en condiciones ideales y controladas, normalmente en el contexto de estudios experimentales (ensayos clínicos). Su validez es universal siempre que se mantengan las condiciones que originaron el resultado de la aplicación de la intervención.

La efectividad mide los resultados obtenidos en condiciones reales de práctica clínica habitual, lejos de la situación controlada de los estudios experimentales, y proporcionan generalmente un mayor potencial de validez externa, ya que los resultados son en mayor grado extrapolables a la población real.

La evaluación de una intervención sanitaria no debería limitarse únicamente a medir la eficacia de las técnicas, sino que sería además necesario asegurar que se utilizan de forma óptima, es decir, produciendo el máximo beneficio posible. La estimación de medidas de eficacia y efectividad son, en principio, complementarias,

ya que es imprescindible conocer los resultados de una intervención sanitaria en condiciones ideales para obtener su aprobación administrativa, pero también es necesario conocer su efecto real en condiciones de utilización habitual en la población general, por lo que ambas informan sobre aspectos distintos pero adicionales, y la medida de una no exime la necesidad de estimación de la otra. Por tanto, y junto con la eficacia de las tecnologías sanitarias es cada vez más importante estudiar si esas tecnologías aportan también ventajas desde un punto de vista no solo de la práctica clínica real, sino también desde el punto de vista económico y social.

Las medidas de eficiencia, por su parte, incorporan aspectos económicos asociados a la intervención, y aportan información sobre la oportunidad de asignación de recursos económicos a una determinada intervención sanitaria, especialmente importante en el contexto de recursos sanitarios escasos.

Estas herramientas, englobadas dentro de la disciplina Evaluación Económica de Intervenciones Sanitarias, aunque útiles para proporcionar datos relativos al valor económico en relación con los resultados en salud que aporta una determinada tecnología sanitaria, no son suficientes para configurar el valor real en terapéutica, su beneficio clínico en el contexto terapéutico en el que son activos. De aquí la necesidad de desarrollar nuevas metodologías que aporten información de la relación precio-resultados en salud desde una óptica más conectada con el valor real que aportan las diferentes alternativas terapéuticas en cada situación clínica.

■ 15.2. ESCALAS DE BENEFICIO CLÍNICO

El valor de una nueva estrategia terapéutica (o de un nuevo tratamiento) viene determinado por la magnitud de su beneficio clínico en función de su coste (Porter, 2010). Esta definición, aparentemente tan breve, encierra en realidad una gran complejidad, por muchos motivos, que van desde la misma definición del beneficio clínico, que no es universal, hasta las herramientas para medirlo, pasando por los estándares de las determinaciones empleadas.

Además, mientras que las organizaciones europeas supranacionales, como ESMO, ponen el foco en el valor del resultado final de la estrategia terapéutica analizada, las organizaciones americanas, esencialmente ASCO, pusieron inicialmente el acento en el coste de la atención sanitaria, aunque, como veremos, más tarde se focalizaron en el valor.

De modo que, en la actualidad, existen dos grandes tendencias mundiales para analizar el beneficio clínico de los tratamientos del cáncer: la liderada por la europea y supranacional ESMO (European Society of Medical Oncology) y la liderada por la americana ASCO (American Society of Clinical Oncology). Ambas sociedades, de forma independiente, han creado grupos de trabajo y realizado extensos esfuerzos para encontrar definiciones de consenso y desarrollar herramientas de referencia que permitan medir y evaluar el beneficio clínico de las diferentes estrategias

terapéuticas que se proponen para el tratamiento del cáncer. Las dos sociedades publicaron sus trabajos iniciales el mismo año, 2015, lo que refleja la preocupación común por el problema.

En este capítulo vamos a analizar estos dos procedimientos de trabajo y sus resultados. Describiremos también el proceso por el que ambas organizaciones han emprendido el camino del encuentro con el fin de diseñar una única escala a nivel mundial que permita evaluar el beneficio clínico de las estrategias terapéuticas que se diseñan para enfrentarse al cáncer.

■ 15.2.1. ESMO-MCBS

MCBS son las siglas de *Magnitude of Clinical Benefit Scale*, que es como ESMO ha denominado a la herramienta diseñada para valorar la magnitud del beneficio clínico de los nuevos tratamientos o estrategias. Las principales características de esta herramienta son que es reproducible, está validada y tiene una vocación dinámica, para ir evolucionando en el tiempo a medida que lo hagan los tratamientos oncológicos (Cherny *et al.*, 2015).

Para el diseño de la herramienta tuvieron en cuenta varios principios:

- Siempre que fue posible, utilizaron para la validación de un tratamiento o estrategia ensayos clínicos fase III o metaanálisis bien diseñados y estructurados.
- Dado que no existía una herramienta estándar para medir el beneficio clínico, consensuaron una serie de los potenciales beneficios que debía tener un nuevo tratamiento para considerarse como productor de beneficio clínico, y que mostramos en la Tabla 1. Como objetivos principales de valoración del beneficio clínico no contaron solo con el principio de “Vivir más” (prolongar la supervivencia), sino también con el de “Vivir mejor” (valoración de la calidad de vida).
- Después de una extensa serie de comparaciones consiguieron determinar los umbrales mínimos para cada escenario a partir de los cuales un beneficio clínico podía considerarse como tal.
- En relación con la regla anterior, y como complemento, también decidieron crear dos escenarios de puntuación del beneficio clínico independientes: el de los tratamientos con posibilidad de ser curativos; y los tratamientos de situaciones que por definición no son curativos. Para estos dos escenarios (tratamientos curativos vs no curativos) crearon dos sistemas de puntuación independientes que después veremos.
- Con toda esta información fueron capaces de crear una tabla de resultados mínimos para considerar el beneficio clínico como aceptable dependiendo

del escenario terapéutico del que se trate. Estos principios se resumen en la Tabla 2.

- Todas estas escalas de beneficio clínico se emplearon para valorar estudios en tumores sólidos.

Tabla 1.

POTENCIALES BENEFICIOS DE UN NUEVO TRATAMIENTO SELECCIONADOS PARA VALORAR EL BENEFICIO CLÍNICO DE UN ESTUDIO

Potenciales beneficios de un nuevo tratamiento

Prolongar la vida

Mejorar la SG (OS)

Mejorar el subrogado de la SG (OS):

SLE (DFS) (cuando los datos no son suficientemente maduros en el entorno adyuvante)

Mejorar la SLP (PFS)

Vivir mejor

Mejorar la calidad de vida

Mejorar el subrogado de la calidad de vida: Mejorar la SLP (PFS)

Reducir la toxicidad

Notas: ESG: Supervivencia global (OS; Overall survival); SLE: Supervivencia Libre de Enfermedad (DFS; Disease Free Survival); SLP: Supervivencia Libre de Progresión (PFS; Progression Free Survival)

Fuente: Cherny *et al.* (2015).

Tabla 2.

BENEFICIOS MÍNIMOS QUE DEBEN OBSERVARSE PARA OBTENER LAS PUNTUACIONES MÁXIMAS

Tratamientos con intención curativa

Mejora >5% de la supervivencia a los 3 años

Mejoras tan sólo en la SLE de HR <0.60 (objetivo principal) en los estudios sin datos maduros de supervivencia

Tabla 2. (continuación)

BENEFICIOS MÍNIMOS QUE DEBEN OBSERVARSE PARA OBTENER LAS PUNTUACIONES MÁXIMAS

Tratamientos sin intención curativa

Cuando el objetivo principal es la SG (OS)

Control <12 meses*

HR <0.65 Y ganancia de >3 meses de supervivencia O

Incremento de la supervivencia a los dos años >10%

Control >12 meses*

HR <0.70 Y ganancia de >5 meses de supervivencia O

Incremento de la supervivencia a los 3 años >10%

Cuando el objetivo principal es la SLP (PFS)

Control <6 meses*

HR <0.65 Y ganancia de >1.5 meses de SLP

Control >6 meses*

HR <0.65 Y ganancia >3 meses

Notas: *Tiempos de SG o de SLP de los controles con los que se comparará el nuevo tratamiento que se está evaluando.

Fuente: Cherny *et al.* (2015).

Con estos puntos de partida, el equipo de trabajo diseñó dos sistemas de medida del beneficio clínico dependiendo de la intención del tratamiento:

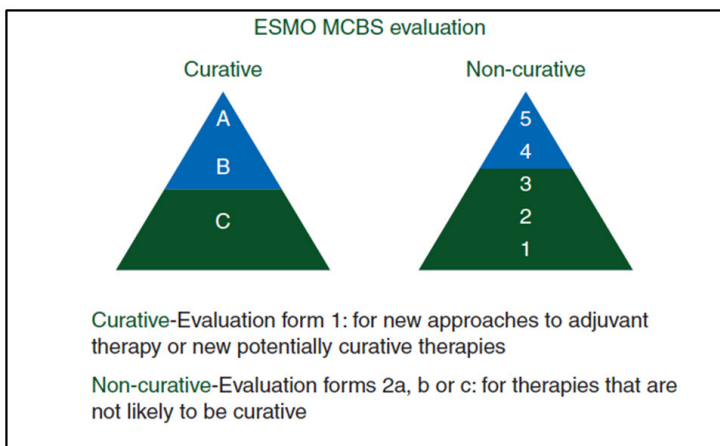
- **Tratamientos con intención curativa.** Es el llamado *form 1*. Se incluyen aquí los tratamientos adyuvantes y neoadyuvantes, y para la enfermedad local o metastásica tratada con intención curativa. La escala produce tres grados: A, B y C. Los dos primeros se consideran como un beneficio sustancial del tratamiento analizado (Figura 1). En los tratamientos adyuvantes analizados cuando los datos de supervivencia no están suficientemente maduros se pueden emplear para el análisis del beneficio clínico los datos de supervivencia libre de enfermedad (SLE; DFS en inglés). En los tratamientos neoadyuvantes no se puede emplear la tasa de remisión completa patológica (pRC) como criterio del beneficio clínico porque existe mucha controversia sobre si la pRC es un factor subrogado válido de la supervivencia global (Bianchini y Gianni, 2011; Burki, 2014; Cortazar *et al.*, 2014).
- **Tratamientos sin intención curativa.** El beneficio clínico de estos estudios se valora de 1 a 5. En esta escala, los valores de 4 y 5 son los que representan un alto nivel de beneficio clínico demostrado (Figura 1). En este grupo, que ellos llaman *form 2* contemplan además tres posibles escenarios de análisis:

- **Form 2a.** Se incluyen aquí los estudios que tienen la supervivencia global (SG;OS, en inglés) como variable principal. Este grupo se estratifica por una mediana de la SG en el brazo control por encima y por debajo de los 12 meses de supervivencia global. En este análisis se incluyen datos de la calidad de vida (QoL), que pueden incrementar la puntuación final en 1 punto si los datos de calidad de vida son favorables. La puntuación final de 5 sólo se consigue incorporando en el análisis datos de calidad de vida favorables.
- **Form 2b.** Esta versión se utiliza para los tratamientos que emplean la SLP (PFS) como objetivo principal. Este grupo se estratifica para una mediana de PFS de seis meses. Los estudios que evalúen QoL y tengan resultados positivos a ese respecto elevarán también 1 punto la puntuación final en este apartado.
- **Form 2c.** Con este sistema se evalúan los estudios de no inferioridad y aquellos en los que el objetivo principal reside en la QoL, la toxicidad o la tasa de respuestas.

Para constatar que el planteamiento era correcto, llevaron a cabo el análisis del beneficio clínico de los tratamientos modernos en una extensa serie de tumores sólidos: cáncer de mama, cáncer de pulmón, cáncer de próstata, cáncer colo-rectal, cáncer de ovario, carcinoma de células renales, sarcomas, melanoma, cáncer de páncreas, y cáncer gastro-esofágico. Lo que supone, además de una extensa

Figura 1.

ESCALAS DE EVALUACIÓN PARA LOS TRATAMIENTOS CON INTENCIÓN CURATIVA Y NO CURATIVA



Fuente: Tomado de Cherny *et al.* (2015).

Figura 2.

TABLA DE LA PUNTUACIÓN FINAL DE ESMO-MCBS PARA LOS ESTUDIOS DE CÁNCER DE COLON. LA COLUMNA ESMO-MCBS ES LA QUE CONTIENE LA PUNTUACIÓN FINAL DEL SISTEMA QUE VALORA EL NIVEL DE CALIDAD DEL ESTUDIO)

Table 6. Field testing ESMO-MCBS v1.0: colorectal cancer

Medication	Trial name	Setting	Primary outcome	PFS control	PFS gain	PFS HR	OS control	OS gain	OS HR	QoL	Toxicity	ESMO-MCBS	Ref.
FOLFOLX ± panitumumab	PRIME	First-line metastatic (post hoc KRAS, NRAS BRAF WT)	PFS	7.9 months	2.3 months	0.72 (0.58–0.90)	20.2 months	5.8 months	0.78 (0.62–0.99)			4	[74]
Panitumumab + mFOLFOLX versus bevacizumab + mFOLFOLX	PEAK	First-line metastatic (KRAS-WT)	PFS			NS	24.3 months	9.9 months	0.62 (0.44–0.89)			4 ^a	[75]
FOLFIRI ± cetuximab	CRYSTAL	First-line metastatic stratified for KRAS-WT (post hoc KRAS, NRAS WT)	PFS	8.4 months	3.0 months	0.56 (0.41–0.76)	20.2 months	8.2 months	0.69 (0.54–0.88)			4	[76]
Cetuximab versus best supportive care		Refractory metastatic KRAS-WT	OS	1.9 months	1.8 months	0.4 (0.30–0.54)	4.8 months	4.7 months	0.55 (0.41–0.74)			4	[77]
FOLFOLX ± panitumumab	PRIME	First-line metastatic KRAS-WT	PFS	8 months	1.6 months	0.80 (0.66–0.97)	19.4 months	4.4 months	0.83 (0.70–0.98)			3	[78, 79]
FOLFIRI ± cetuximab	CRYSTAL	First-line metastatic stratified for KRAS-WT	PFS	8.4 months	1.5 months	0.70 (0.56–0.87)	20 months	3.5 months	0.80 (0.67–0.95)			3	[80, 81]
ILF ± bevacizumab		First-line metastatic	OS	3.9 months	2 months	0.73 (0.59–0.90)	15.6 months	4.7 months	0.66 (0.54–0.81)			3	[82]
FOLFIRI ± panitumumab		Second-line metastatic KRAS-WT	PFS									3	[83]
FOLFOLX ± bevacizumab	E3200	Second-line metastatic after FOLFIRI	OS	7.3 weeks	5 weeks	0.45 (0.34–0.59)	10.8 months	2.1 months	0.75 (0.63–0.89)			2	[84]
versus bevacizumab alone		Third-line metastatic stratified for KRAS	PFS									2	[85]
Panitumumab, versus best supportive care		First-line metastatic	PFS	9.7 months	2.4 months	0.75 (0.62–0.90)			NS			2	[86]
FOLFIRI bevacizumab versus FOLFOXIRI bevacizumab		Third-line metastatic	PFS									2	[87]
TAS-102 versus placebo	CONCOURSE	Third-line or beyond metastatic	OS	5.3 months	1.8 months	0.68 (0.658–0.81)	5 months	1.4 months	0.77 (0.64–0.94)			1	[88]
Regorafenib versus placebo	CORRECT	Third-line metastatic	OS	9.6 months	1.5 months	0.81 (0.69–0.94)						1	[89]
Second-line chemotherapy ± bevacizumab	ML18147	Second line beyond progression on bevacizumab	OS									1	[90]
FOLFIRI ± aflibercept	VELOUR	Second line after oxaliplatin-based treatment	OS	4.7 months	2.2 months	0.76 (0.66–0.87)	11.7 months	1.6 months	0.84 (0.73–0.97)			1	[91]
FOLFIRI ± Ramucirumab	RAISE	Second-line metastatic after bevacizumab, oxaliplatin, fluoropyrimidine	OS									1	[91]

^aUnbalanced crossover.

Fuente: Tomado de Cherny et al. (2015).

Figura 3.

TABLA DE LA PUNTUACIÓN FINAL DE ESMO-MCBS PARA LOS ESTUDIOS DE CÁNCER DE PULMÓN. LA COLUMNA ESMO-MCBS ES LA QUE CONTIENE LA PUNTUACIÓN FINAL DEL SISTEMA QUE VALORA EL NIVEL DE CALIDAD DEL ESTUDIO

Table 3. Field testing ESMO-MCBS v1.0: lung cancer

Lung cancer Medication (new versus control)	Trial name	Setting	Primary outcome	PFS control	PFS gain	PFS HR	OS control	OS gain	OS HR	Qual.	Toxicity	ESMO-MCBS	Ref.
Erlotinib versus carboplatin/gemcitabine	OPTIMAL CTONG-082	First-line stage IIIB or IV non-squamous, with EGFR mutation	PFS	4.6 months	8.5 months	0.16 (0.10-0.26)					1.2% less serious adverse events	4	[43]
Erlotinib versus platinum-based chemotherapy doublet	EUREAC	First-line stage IIIB or IV non-squamous, with EGFR mutation	PFS (crossover allowed)	5.2 months	4.5 months	0.37 (0.25-0.54)	19.5 months	NS			1.8% less severe adverse reactions	4	[44]
Gefitinib versus carboplatin + pemetrexed	IPASS	First-line stage IIIB or IV adenocarcinoma, with EGFR mutation	PFS (crossover allowed)	6.3 months	3.3 months	0.48 (0.34-0.67)				Improved	Reduced toxicity	4	[45, 46]
Afinib versus Cisplatin + pemetrexed	LUX-Lung 3	First-line stage IIIB or IV adenocarcinoma, with EGFR mutation (D419/ L858R)	PFS (crossover allowed)	6.9 months	4.2 months	0.58 (0.43-0.78)				Improved	Improved	4	[47, 48]
Cisrotinib versus chemotherapy		First-line stage IIIB or IV non-squamous, with ALK mutation	PFS (crossover allowed)	3.0 months	4.7 months	0.49 (0.37-0.64)				Improved	Improved	4	[49]
Cisrotinib versus cisplatin + pemetrexed		First-line stage IIIB or IV non-squamous, with EGFR mutation	PFS	2.0 months	3.9 months	0.45 (0.35-0.60)				Improved	Improved	4	[50]
Pemetrexed versus placebo		Stage IIIB or IV disease maintenance after responding to four cycles platinum doublet	PFS (crossover for biologic disease maintenance (non-squamous))	2.6 months	1.9 months	0.47 (0.37-0.60)	10.3 months	5.2 months	0.70 (0.56-0.88)			4	[51]
Cisplatin pemetrexed versus cisplatin gemcitabine		First-line stage IIIB or IV (non-squamous)	OS (non-inferiority)		10.4 months	1.4 months	0.81 (0.70-0.94)				Less grade 3+ toxicity neutropenia anaemia thrombocytopenia	4	[52]
Chemotherapy ± palliative care		Stage IV non-small-cell ECOG <2	Qual.		8.9 months	2.7 months	HR for death in control arm 1.7 (1.14-2.54)			Improved		4	[53]
Paclitaxel/ carboplatin ± bevacizumab		First-line stage IIIB or IV, non-squamous	OS	10.3 months	2.0 months	0.79 (0.67-0.92)						2	[54]
Erlotinib versus placebo	SATURN	Stage IIIB or IV disease maintenance after responding to four to six cycles platinum doublet	PFS	11.1 weeks	1.2 weeks	0.71 (0.62-0.82)	11.0 months	1.0 months	0.81 (0.70-0.95)			1	[55]

Fuente: Tomado de Cherny et al. (2015).

prueba de que la herramienta funciona, el análisis del beneficio clínico de prácticamente todos los estudios importantes que abarca la gran mayoría de la patología oncológica del adulto.

Como ejemplo del resultado final del sistema, en las Figura 2 y 3 se muestran las tablas de valoración como resultado del análisis ESMO-MCBS para cáncer de colon y de pulmón que se publican en el artículo de referencia.

■ 15.2.2. ASCO

El planteamiento del grupo de trabajo de ASCO, denominado en 2013 *Value in Cancer Care Task Force*, centra el esfuerzo del análisis en el valor de los tratamientos, más que en el beneficio clínico como ESMO lo entendió. Para determinar este valor crean un sistema de comparación que tiene en cuenta el beneficio clínico como tal, la toxicidad y el coste de los tratamientos. Toda la definición de esta estructura se publica en 2015 (Schnipper *et al.*, 2015), con el nombre de *ASCO Value Framework*, y de forma prácticamente simultánea a la publicación del sistema de ESMO que antes hemos analizado.

■ 15.2.2.1. Definición del valor en la asistencia sanitaria

Dado que todo el sistema de ASCO se basa en el valor del tratamiento, su inicial esfuerzo se centra en definirlo. Para ello, recurren a los seis elementos que el Institute of Medicine (IOM) considera que definen la calidad de un acto de asistencia sanitaria: la seguridad, la efectividad, la eficiencia, que esté centrado en el paciente, que sea oportuno en el tiempo y la equidad (Institute of Medicine, 2001 y 2013). Sin embargo, argumentando que los tres últimos elementos (que esté centrado en el paciente, la oportunidad temporal y la equidad) son muy difíciles de medir y que rara vez se encuentran datos sobre ellos en los ensayos clínicos, basan la definición del valor del acto sanitario en los tres primeros, que son más fácil y fiablemente medibles: el beneficio clínico (la eficacia), la toxicidad (seguridad) y el coste (eficiencia).

Sobre estos tres parámetros ASCO creó tres grupos de trabajo, cada uno de ellos centrado en definir los parámetros esenciales de las tres áreas de interés, para después incorporarlos, a modo de puzzle, al esquema global del valor (*the value framework*).

Al igual que sucede en el esquema de ESMO, establecen dos escenarios de puntuación distintos: para los tratamientos potencialmente curativos (neoadyuvantes y adyuvantes) y para el tratamiento del cáncer avanzado.

Los parámetros de beneficio clínico y toxicidad se combinan para generar una nueva puntuación, el beneficio sanitario neto (o Net Health Benefit, NHB). Finalmente, este valor del NHB se confronta con el coste directo del tratamiento, para proporcionar una valoración global final del tratamiento analizado. A continuación, explicaremos la definición de los tres elementos que conforman el esquema de valor.

Beneficio clínico. Las valoraciones son ligeramente diferentes para la enfermedad avanzada y para los tratamientos con intención curativa.

En el primer caso, el de la enfermedad avanzada, el beneficio clínico se mide con una puntuación de 1 a 5 sobre la base del incremento que el nuevo tratamiento supone en la mediana de la supervivencia global (SG; OS en inglés) con respecto al tratamiento estándar con el que se le compara. Para ello se utiliza la mediana de la *hazard ratio* (HR). Si no se dispone de los datos de SG, se utilizan los de la supervivencia libre de progresión (SLP; PFS en inglés); cuando no se dispone de ninguna de estas dos medidas de la supervivencia, o el estudio analizado es de un solo brazo, son comparados, se utiliza la tasa de respuestas (RR en inglés). El valor de 1 a 5 asignado a la SG se multiplica x16 (se eligió este múltiplo para expresar que hasta 80 [16x5] de 100 puntos pueden atribuirse a la SG, como expresión de la importancia de este parámetro). Para la SLP y la RR los múltiplos de corrección son, respectivamente, 11 y 8, expresando así su menor importancia en el valor del beneficio clínico que se está calculando.

Para el análisis de los tratamientos con intención curativa se asigna la misma puntuación (1 a 5) a la diferencia de la SG entre el estudio valorado y el brazo control (diferencia expresada por la mediana de la HR), y esa puntuación se multiplica por 16. Cuando no se dispone de la SG se utiliza la supervivencia libre de enfermedad (SLE; DFS en inglés), en ese caso, se multiplica por 15. De nuevo se da a la SG un valor relativo de 80 sobre 100, expresando así la importancia que este parámetro tiene para evaluar el valor de un tratamiento.

Toxicidad. En los dos escenarios (intención curativa y enfermedad avanzada) el valor de la toxicidad se calcula como toxicidad relativa del nuevo tratamiento en comparación con el estándar. Para ello, asigna puntuaciones entre -20 y +20 en función de la incidencia de toxicidades grados 3-5 (de acuerdo con las definiciones de los CTCs, o *Common Terminology Criteria for Adverse Events*¹).

Bonus points. El sistema contempla la posibilidad de conceder puntos de bonificación (*bonus points*) a los tratamientos evaluados en dos supuestos: *palliative care bonus*. El sistema contempla conceder puntos adicionales si el tratamiento en evaluación consigue una mejora estadísticamente significativa en algún síntoma relacionado con el tumor; y *bonus* asociados con el intervalo libre de tratamiento (*treatment-free interval bonus*). El nuevo tratamiento obtendrá puntos adicionales si

¹ https://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_8.5x11.pdf

demuestra ventaja estadísticamente significativa sobre el tratamiento de control en el tiempo libre de tratamiento. Esta puntuación adicional se otorga porque la prolongación del tiempo sin tratamiento se asume asociada a estabilización de la enfermedad y ausencia de efectos tóxicos, lo que es asimilable a una mejoría en la calidad de vida.

Coste del tratamiento. El coste del tratamiento es un factor fundamental del sistema de cálculo del valor del nuevo tratamiento. Se incorporan tanto los costes de adquisición de los fármacos como los costes en los que incurre el paciente durante su tratamiento y sus cuidados. El análisis de estos costes es complejo, y no siempre se puede llevar a cabo con fiabilidad. Además, son estudios de costes muy extensos. Por ejemplo, como parte del coste del tratamiento se incluyen también los fármacos de soporte (p. ej., los antieméticos). Todo ello dificulta el análisis fiable, y sobre todo reproducible, del coste de los tratamientos antineoplásicos.

NHB (Net Health Benefit). Es el otro elemento esencial, junto con el coste, para calcular el valor de la intervención terapéutica. Como se ha dicho antes, su valor se obtiene combinando los valores del beneficio clínico y la toxicidad con la metodología antes reseñada. Ampliando un poco este concepto, no siempre fácil de manejar para los clínicos, el NHB ha sido definido por el Intitute for Clinical and Economic Review como el “balance entre los beneficios clínicos, por un lado, y los riesgos y los efectos adversos, por otro. El resultado de este balance se utiliza para valorar la magnitud de la diferencia entre un agente terapéutico y su comparador estándar (Ollendorf y Pearson, 2013).

Valoración final. Como conclusión de la valoración del procedimiento se proporciona la información sobre el NHB y el coste del tratamiento. El valor del tratamiento se infiere a partir de la relación entre el NHB y el coste de este, de la que se obtiene el grado de beneficio para el paciente.

■ 15.2.2.2. *Aplicación de la técnica de valoración a los escenarios clínicos*

Para la publicación de la técnica se probó en cuatro escenarios clínicos en los que existen múltiples tratamientos que se han comparado con los estándares de tratamiento: la primera línea de cáncer no microcítico de pulmón metastásico, el tratamiento del mieloma múltiple avanzado, el tratamiento del cáncer de próstata metastásico resistente a la castración y el tratamiento adyuvante del cáncer de mama HER-2 positivo. Para todos estos escenarios, en el artículo de referencia se muestran las gráficas del análisis final, que incluye los valores para el brazo del control y el que se está valorando del beneficio clínico, la toxicidad, el NHB y el coste. En la Figura 4 se muestra, como ejemplo, la valoración del tratamiento adyuvante del cáncer de mama HER-2 positivo cuando se añade trastuzumab, comparándolo con el tratamiento control.

15.2.3. Actualizaciones

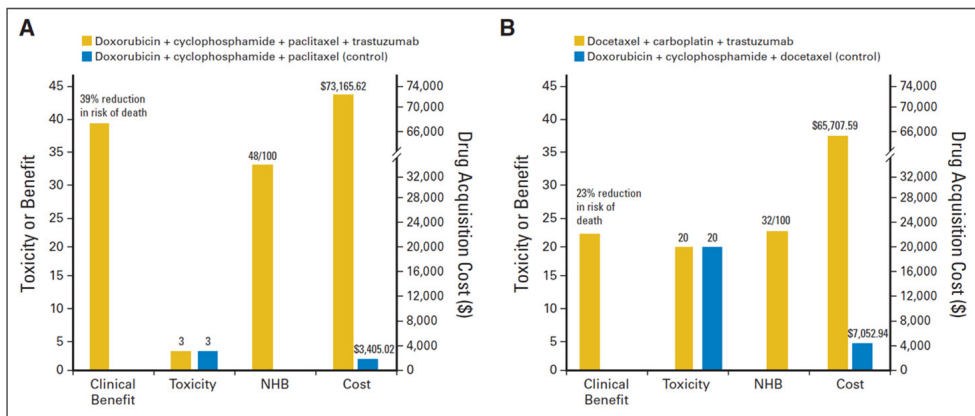
Tanto el sistema americano (*ASCO Clinical Oncology Value Framework*) como el europeo (*ESMO-MCBS*) nacen con la clara vocación de ser sistemas actualizables. Fieles a esa idea, ambos han publicado actualizaciones: ASCO en 2016 (Schnipper *et al.*, 2016) y ESMO en 2017 (Cherny *et al.*, 2017). En ambos casos, las modificaciones recogidas proceden de sugerencias y comentarios del colectivo clínico y se ocupan de modificaciones metodológicas y algunas conceptuales puntuales. Los autores expresan en ambos casos su confianza en que los sistemas de valoración ganen en calidad con las modificaciones efectuadas y sigan siendo evolutivos.

15.2.4. ESMO-MCBS vs ASCO-Value Framework: un intento de convergencia

En 2018 se ha publicado un extenso análisis que evalúa la posible convergencia de los dos sistemas (Cherny *et al.*, 2018). Para ello, los creadores del sistema de ESMO evalúan los 102 estudios que utilizaron para crearlo con el sistema de evaluación de ASCO. La conclusión a la que llegan es que hay una convergencia

Figura 4.

EVALUACIÓN DEL BENEFICIO CLÍNICO, EL NHB Y EL COSTE EN EL TRATAMIENTO ADYUVANTE DEL CÁNCER DE MAMA HER-2 POSITIVO CON QUIMIOTERAPIA Y TRASTUZUMAB COMPARADO CON EL TRATAMIENTO CON QUIMIOTERAPIA SIN EL ANTI-HER2



Fuente: Tomado de Schnipper *et al.* (2015).

entre los dos sistemas de 0.68; o, dicho de otro modo, en casi el 70 % de los estudios los dos sistemas de evaluación, aparentemente muy distintos, coinciden en la valoración de los estudios. En los 37 estudios discordantes, las discrepancias se centraron con diferentes aproximaciones a la evaluación de la ganancia absoluta y relativa de la supervivencia (tanto global como libre de progresión) y en la valoración de la toxicidad. Los autores concluyen que existen suficientes puntos de convergencia entre los dos sistemas para seguir avanzando en la construcción de un sistema común.

■ 15.3. CALIDAD DE VIDA Y RESULTADOS REPORTADOS POR LOS PACIENTES

La supervivencia y la calidad de vida son las dos medidas de resultados más importantes en los ensayos clínicos en oncología. Por ese motivo, los primeros cuestionarios para medir la calidad de vida fueron desarrollados a finales del siglo pasado para su uso en la investigación clínica. En la actualidad disponemos de una gran variedad. En su elaboración se sigue un proceso laborioso y complejo que culmina con la validación, que debe repetirse cada vez que se traduce a un idioma diferente. Algunos son genéricos y pueden utilizarse frente a cualquier situación patológica (EQ-5D), mientras que otros se orientan a situaciones o pacientes concretos. Entre los específicos de oncología, los más usados son los de la EORTC (European Organisation for Research and Treatment of Cancer) y la organización americana FACIT (Functional Assessment of Chronic Illness Therapy). La EORTC creó su grupo de trabajo en calidad de vida en 1980 y sigue muy activo. Ha desarrollado un cuestionario general (EORTC-Q30) que se puede complementar con otros específicos para cada tipo de tumor o incluso para algunos síntomas o situaciones clínicas concretas (caquexia, astenia, etc.). Paralelamente, FACIT presentó el FACT-G, su cuestionario general de cáncer, en la década de los 90. Al igual de los de la EORTC, se complementa con cuestionarios específicos. Muchas otras organizaciones e instituciones han desarrollado diferentes cuestionarios. En ocasiones, puede resultar difícil seleccionar la herramienta más apropiada para cada caso. Por otra parte, la comparación de resultados obtenidos con diferentes cuestionarios plantea problemas importantes.

En la actualidad, la calidad de vida se engloba dentro de los llamados “resultados reportados por el paciente” (PRO, por su acrónimo en inglés). Según la EMA, un resultado reportado por el paciente es cualquier resultado evaluado por el propio paciente basado en su percepción sobre la enfermedad y su tratamiento (European Medicines Agency, 2016) Incluye evaluaciones del estado de salud, calidad de vida, síntomas, satisfacción con los cuidados recibidos y adherencia a la medicación.

Recientemente se ha acuñado un concepto más amplio denominado “Datos de salud generados por el paciente” que incluye cualquier dato relacionado por la salud y que ha sido creado, registrado, reunido o inferido por o desde el paciente

o sus representantes o cuidadores para afrontar un problema de salud (Locklear *et al.*, 2015). Este término incluye los PRO pero también datos cuyo origen son biosensores u otro tipo de dispositivos.

La medida de la calidad de vida en los ensayos clínicos se considera un componente esencial pero adquiere más importancia cuando la variable principal no es la supervivencia global, sino una variable subrogada, cuando se comparan alternativas terapéuticas con resultados de efectividad semejante (p.e. el cáncer de próstata precoz), cuando el beneficio clínico esperado es limitado, cuando las diferencias en supervivencia son mínimas y se espera toxicidad o cuando se espera mayor paliación de síntomas con uno de los tratamientos.

A pesar de la dilatada experiencia con los cuestionarios de calidad de vida la realidad es que la utilización de los PRO en la investigación clínica no se considera óptima.

Una publicación reciente (Marandino *et al.*, 2018) en la que se han revisado los 446 ensayos clínicos fase III publicados entre 2012 y 2016 en las principales revistas oncológicas ha puesto de manifiesto que tan solo un 47 % de los EC incluyen la calidad de vida como variable de estudio, aunque sea secundaria o exploratoria. Esta cifra sigue siendo baja incluso en los ensayos dirigidos a población metastásica (40 %). Entre aquellos ensayos en los que la medida de la calidad de vida está incluida entre las variables de estudio (43 % del total incluido en el estudio), en el 39 % de los mismos no se incluye en la publicación inicial, sino que otra específica que puede acumular demoras importantes, en el mejor de los casos, o incluso no llega a publicarse. A esto se debe añadir las dificultades metodológicas para el diseño, interpretación y redacción de los resultados. Las guías consort y spirit han incluido sendas extensiones para el manejo de los PRO.

Las agencias reguladoras apoyan el uso de PRO como un elemento clave en el proceso de aprobación de nuevos fármacos o indicaciones. Lo consideran el *Gold estándar* para la evaluación de síntomas subjetivos. Para promover su utilización, han elaborado recomendaciones para su incorporación al proceso regulatorio (European Medicines Agency, 2016; Department of Health and Human Services Food and Drug Administration, 2009).

Tradicionalmente, los investigadores recogen la toxicidad en los ensayos clínicos utilizando el sistema estandarizado CTCAE (*Common Terminology Criteria for Adverse Events*) desarrollado por el Instituto Nacional del Cáncer americano (NCI). Recientemente, el propio NCI ha desarrollado un sistema paralelo utilizando los PRO. La nueva herramienta, denominada PRO-CTCAE incorpora los 78 *items* del CTCAE que se manifiestan en forma de síntomas y ha sido validada en varios idiomas, incluyendo el español y se considera un complemento muy útil a la valoración de los clínicos.

En los últimos años ha surgido un interés creciente por la aplicación de la medida de la calidad de vida y otros resultados reportados por el paciente (PRO) en

el proceso asistencial. En la actualidad son considerados un componente esencial de la atención oncológica de alta calidad y centrada en el paciente y del nuevo concepto de Porter denominado “Value Health System” que promueve la transformación del Sistema de Salud a través de la medición sistemática del valor que aporta cada intervención (Porter, 2010).

Según establece el modelo de sistema sanitario basado en el valor (*Value Health System*) la única manera de hacer el sistema sanitario sostenible y a la vez de mejorarlo es medir los resultados relevantes para el paciente así como los recursos económicos que se utilizan para conseguirlos. El adjetivo de “relevantes” supone un cambio conceptual revolucionario ya que deja en un segundo plano muchos resultados a los que los clínicos conceden importancia. Este nuevo enfoque obliga a mirar con los ojos del paciente y los PRO adquieren una gran importancia puesto que serán la única fuente de información para la mayoría de los resultados “relevantes” para el paciente. Sin embargo, uno de los grandes retos de la actualidad es identificar cuales son concretamente los resultados relevantes para el paciente que hay que medir. Especialmente, si tenemos en cuenta que estos varían con las diferentes enfermedades o incluso en fase diferentes de una misma enfermedad.

Muy alineado con la filosofía de Porter, uno de sus fundadores, el International Consortium for Health Outcomes Measurement (ICHOM) es una institución sin ánimo de lucro que mediante un proceso colaborativo internacional en el que participan todas las partes implicadas (profesionales, pacientes, pagadores...) se define el conjunto de resultados relevantes para cada proceso patológico, así como las herramientas para medirlos y los intervalos de tiempo en los que se han de aplicar. En el campo de la oncología se han definido set de resultados para el cáncer de pulmón, de mama, colorectal, próstata localizada y próstata avanzada. Los resultados propuestos suelen incluir la supervivencia, medidas de morbilidad del tratamiento y un gran número de variables relacionadas con síntomas y calidad de vida. Los PRO tienen un papel fundamental en las propuestas de ICHOM. El objetivo último es que se aplique la misma metodología en diferentes instituciones y los resultados sean públicos de manera que se utilicen como herramienta de mejora (benchmarking) y que el paciente pueda elegir donde prefiere ser tratado

Los cuestionarios en formato papel requieren la introducción posterior de los datos y retrasa de forma importante el procesamiento de los mismos. La utilización de los PRO en el proceso asistencial solo ha sido posible con la llegada de los dispositivos electrónicos que permiten recoger las respuestas de los cuestionarios y procesarlos a tiempo real a través de páginas WEB o incluso con APPs instaladas en dispositivos portátiles como teléfonos o tabletas. Esta modalidad de PRO recibe el nombre de ePROs.

El beneficio de utilizar los PRO en el proceso asistencial ha sido confirmado por Basch *et al.* (2017) en un ensayo aleatorizado en el que participaron 766 pacientes con tumores metastásicos de mama, pulmón, ginecológicos o genitourinarios en tratamiento con quimioterapia en el Memorial Sloan Kettering Cancer Center

de Nueva York. En el brazo experimental se utilizó una herramienta de ePRO para monitorizar 12 síntomas comunes. Los pacientes recibían recuerdos semanales para reportar los síntomas. Cuando un síntoma empeoraba o alcanzaba un valor alto, el sistema generaba un correo electrónico a los profesionales sanitarios. Los pacientes del brazo control fueron monitorizados con el procedimiento habitual. El objetivo principal fue la calidad de vida. Un mayor porcentaje de los pacientes del grupo de intervención mejoró la puntuación a los seis meses (34 % v 18 %) y un menor porcentaje empeoró (38 % v 53 %) en comparación con el grupo control ($P < 0.001$). Sin embargo, el mayor impacto de este estudio es que también se pudo evidenciar una mejora de la supervivencia global con una diferencia de 5 meses en las medianas de supervivencia.

La integración de los PRO en el proceso asistencial oncológico, que se caracteriza por su complejidad, saturación y limitación de recursos, plantea muchas dificultades. No es suficiente la utilización de cuestionarios electrónicos vía web o app. Se requiere además un procesamiento a tiempo real y una integración de los resultados en la HCE con una presentación visual y longitudinal (en el tiempo), de manera que no suponga una carga de trabajo mayor para los clínicos, sino que, por el contrario, permitan orientar y optimizar el proceso de la entrevista clínica.

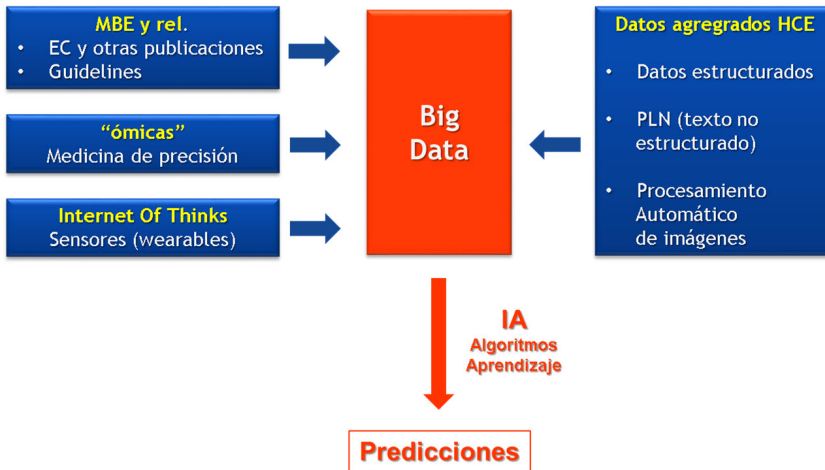
■ 15.4. EXPERIENCIAS DE AGREGACIÓN DE DATOS Y *BIG DATA*

La gran implantación de la historia clínica electrónica (HCE) en los sistemas de salud, el desarrollo de la genómica, y en general, de todas las “ómicas” y la posibilidad de recoger datos de salud a través de dispositivos móviles portátiles ha hecho crecer el interés por la utilización del big data en sanidad. La creciente capacidad de los sistemas informáticos para procesar grandes volúmenes de datos a gran velocidad crea el entorno adecuado para hacer posible el análisis de datos agregados de orígenes muy diversos a tiempo real. Este nuevo enfoque puede suponer una aportación definitiva a la investigación complementando los datos de los ensayos clínicos con datos de la vida real y abriendo una nueva dimensión a la medicina basada en la evidencia. A través de este mecanismo se puede confirmar datos de eficacia en la vida real o establecer relaciones con características genéticas, o plantear nuevas hipótesis de investigación, entre otras muchas utilidades. Y en el campo de la práctica clínica, puede ser la base de potentes sistemas de soporte a la decisión clínica basándose no solo en guías clínicas sino en modelos predictivos elaborados con técnicas de inteligencia artificial.

En la actualidad existe un acuerdo absoluto en que el análisis de datos agregados con técnicas de *big data* generará un conocimiento que cambiará de forma radical la forma de trabajo en sanidad, transformándose en un sistema de aprendizaje rápido basado en los datos de vida real y mejorando previsiblemente los resultados actuales.

Figura 5.

ESQUEMA DE LA APLICACIÓN DEL BIG DATA PARA LA TOMA DE DECISIONES CLÍNICAS



Fuente: Elaboración propia.

Sin embargo, a medida que se han intentado poner en marcha este tipo de herramientas, las grandes expectativas iniciales se han ido modulando hacia planteamientos más realistas, con objetivos más modestos, reconociendo que el camino será más lento que lo esperado debido a las limitaciones presentes a día de hoy. Los mejores resultados se consiguen cuando se orienta la herramienta a responder una pregunta concreta y precisa, y se dispone de datos para el aprendizaje. Este es el caso, por ejemplo, de aplicaciones para el diagnóstico por imagen en radiología o en Anatomía patológica. A medida que aumenta la complejidad y variabilidad del proceso, como en el caso de la oncología, disminuye la probabilidad de éxito a corto-medio plazo.

Solo una pequeña parte de la información de la HCE está almacenada en forma estructurada (como campos de una base de datos). La mayor parte de la misma está en formato texto formando parte de evolutivos o informes diversos, o incluso documentos escaneados. A pesar del indudable avance de los sistemas de procesamiento de lenguaje natural, el desarrollo actual de los mismos no es suficiente para procesar eficazmente esta información. En el caso de la oncología, a diferencia del de otros escenarios clínicos, un alto porcentaje de la información más valiosa no está estructurada y por tanto esta limitación tiene mayor impacto. Pero incluso el tratamiento de la información estructurada plantea algunos retos ya que para agregar información de orígenes distintos hay que hacer un gran esfuerzo para uniformizar y estandarizar los contenidos. Este problema se ve agravado porque en

oncología está pendiente un amplio consenso sobre como expresar los complejos conceptos que definen el proceso oncológico.

Por otra parte, para poder extraer la máxima información se hace necesario conectar los datos procedentes de la HCE con los de otras bases de datos, por ejemplo las de secuenciación genómica, mortalidad o utilización de recursos. Esto supone una dificultad añadida ya que los diferentes aplicativos no suele ser capaces de interoperar y por otra parte, la necesaria protección de la privacidad de los datos y la anonimización de los mismos representan un gran obstáculo adicional difícil de vencer.

A pesar de estas dificultades, existen experiencias concretas que han sido capaces de obtener algunos logros. Entre ellas caben destacar algunos proyectos de I+D+i financiados por la Unión Europea (proyecto HARMONY), iniciativas de Sociedades profesionales (CANCERLINQ de la Sociedad Americana de Oncología Médica) o de agrupaciones de Hospitales (ORIEN) o de empresas privadas (FLATIRON HEALTH o IBM Watson Oncology).

La Alianza HARMONY (*Healthcare Alliance for resourceful medicines offensive against neoplasms in hematology*)² es una Red de Excelencia Europea público-privada, establecida en enero de 2017 y liderada por dos hospitales españoles. Su objetivo es desbloquear y difundir conocimiento valioso sobre neoplasias malignas hematológicas entre todas las partes interesadas, incluyendo centros académicos, hospitales, sociedades profesionales, compañías farmacéuticas, agencias reguladoras y representantes de pacientes, con el objetivo de aprovechar y extraer *big data* para acelerar el desarrollo de mejores tratamientos para los pacientes y estrategias de tratamiento más eficaces. Actualmente incorpora 53 socios y 32 miembros asociados de 22 países, incluidas 8 compañías farmacéuticas de la Federación Europea de Industrias Farmacéuticas y Asociaciones (EFPIA).

HARMONY es el programa de mayor envergadura del proyecto IMI de la Unión Europea sobre aprovechamiento de *big data* para la obtención de mejores resultados en salud.

Aloja múltiples tipos de datos, incluidos los datos clínicos relevantes, síntomas, bioquímica y exámenes físicos, y otra información recopilada sobre tratamiento, supervivencia, omics, calidad de vida y utilización de recursos. Para garantizar que la información descriptiva, comparativa y predictiva generada por los análisis realizados en la plataforma de datos sea fiable, se presta especial atención a determinar la calidad de los datos que ingresan en la Plataforma. Se verifica que la información de entrada está estandarizada, anonimizada, completa y correcta.

Los datos transferidos a HARMONY se reciben en el mismo formato que la fuente, lo que implica formatos, organizaciones y codificación dispares. Esto obliga a conocer perfectamente estas características para garantizar la integridad de los

² HARMONY (Healthcare Alliance for resourceful medicines offensive against neoplasms in hematology) <https://www.harmony-alliance.eu/>

datos y la armonización en toda la plataforma. El proceso de estandarización y armonización en HARMONY se logra al convertir los datos al Modelo de datos común de la Asociación de Resultados Médicos Observacionales (OMOP), que utiliza terminologías estándar internacionales como SNOMED o LOINC para referirse a la información contenida en la fuente. Eventualmente, todas las fuentes se vuelven estructuralmente y semánticamente homogéneas.

Aunque toda la información se almacena en una sola base de datos, solo los subconjuntos apropiados están disponibles para su análisis. Las fracciones de los datos recuperados corresponden a los datos esenciales que se utilizarán como entrada para el análisis diseñado para una pregunta de investigación específica. El acceso a los subconjuntos se otorgará a un grupo limitado de usuarios durante un período de tiempo específico. Además de esto, el hecho de que solo los datos anónimos ingresen a la plataforma y de que tenga la certificación ISO27001 agrega mayor seguridad con respecto a la privacidad y la seguridad de los datos.

Siguiendo las directrices de la Asociación Médica Americana de poner en marcha “rapid learning Health systems” para mejorar los resultados en salud, ASCO creó Cancerlinq (*Cancer learning intelligence network for quality*) en 2012 como una extensión de su programa de calidad (QOPI). Tras validar el prototipo, en 2016 se incorporan las primeras instituciones. En la actualidad participan en el programa más de 100 y se han incorporado más de 1.000.000 de registros. Existe un repositorio único anonimizado con un modelo de datos interno para uniformizar los datos procedentes de los aplicativos de cada institución. El desarrollo está siendo mucho más lento que el esperado. ASCO reconoce las dificultades actuales para incorporar datos no estructurados así como la falta de un modelo de datos común para los registros oncológicos. En la actualidad aporta sobre todo un sistema de monitorización de indicadores de calidad analizando directamente los datos de la HCE y un sistema de visualización incluyendo una presentación de los eventos en una línea temporal. Actualmente se está desarrollando el módulo para analizar datos para investigación denominado “CancerLinQ Discovery”.

El desarrollo de *Watson Oncology* ha sido también mucho más limitado que el esperado. La herramienta dispone de módulos para facilitar el reclutamiento de pacientes para ensayos clínicos en función de la información de la historia clínica, para generar informes basados en la información genética sobre posibles tratamientos que actúan en relación a la mutación, para dar soporte a la toma de decisiones en los comités de tumores o para visualización de una historia clínica, entre otros. Pero está lejos de las alcanzar las expectativas iniciales de sistemas de soporte a la decisión clínica. Por ahora, el sistema no ha sido capaz de aprender por sí mismo sino que el conocimiento le ha sido transmitido a través de los profesionales con los que IBM ha tenido acuerdos (Memorial Sloan Kettering Centre de NY o MD Anderson). El hecho de que IBM esté detrás del proyecto hace pensar en un mayor desarrollo futuro.

Flatiron Health ha desarrollado uno de los sistemas más prometedores. Se trata de una empresa que ha tenido un crecimiento muy rápido. Fue fundada en 2012 con un capital de 313 millones de dólares y comprada por Roche a principios de 2018 por 1.900 millones. Tienen varios módulos: historia clínica oncológica, facturación, portal del paciente y Onco-Analytics, que permite extraer información de todos los pacientes almacenados. En la actualidad de utiliza en más de 250 instituciones y el repositorio anonimizado único incluye más de 1.000.000 de pacientes. Flatiron es capaz de incorporar no solo la información estructurada sino también la no estructurada. Para ello actualmente utiliza un sistema de extracción manual asistida por ordenador, a la espera de que los programas de procesamiento de lenguaje natural sean capaces de realizar esta labor de forma automática. En los dos últimos años se han realizado diversas publicaciones con datos extraídos de su repositorio y en ese sentido son la plataforma más avanzada. Estas publicaciones aportan datos de resultados en vida real y permiten la comparación con los datos de los ensayos clínicos y potencialmente la comparación intercentros.

En la medida en que los resultados reportados por los pacientes sean incorporados a la HCE o se almacenen en aplicativos interconectados, las plataformas de *big data* podrán incorporar este componente esencial de la medida de resultados. Estos avances, junto con el mayor desarrollo de las herramientas de procesamiento de lenguaje natural y de las de análisis de *big data* permitirán incorporar el conocimiento de Real Word Data en la investigación y en la toma de decisiones clínicas. El ritmo de los avances tecnológicos hace concebir la esperanza de que este nuevo escenario no se demore mucho en el tiempo.

■ REFERENCIAS

- BASCH, E., DEAL, A. M., AMYLOU, C. *et al.* (2017). Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. *JAMA*, 318(2), pp. 197–198. doi: 10.1001/jama.2017.7156
- BIANCHINI, G. y GIANNI, L. (2011). Surrogate markers for targeted therapy-based treatment activity and efficacy. *J Natl Cancer Inst Monogr*, 2011, pp. 91–94.
- BURKI, T. K. (2014). Pathological complete response is no surrogate for survival. *Lancet Oncol*, 15, p. e111.
- CHERNY, N. I., DAFNI, U., BOGAERTS, J. *et al.* (2017). ESMO-Magnitude of Clinical Benefit Scale version 1.1. *Annals of Oncology*, 28, pp. 2340–2366.
- CHERNY, N. I., SULLIVAN, R., DAFNI, U., KERST, J. M., SOBRERO, A., ZIELINSKI, C., DE VRIES, E. G. y PICCART, M. J. (2015). A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). *Ann Oncol*, 26(8), pp. 1547-1573. doi: 10.1093/annonc/mdv249.
- CHERNY, N. I., DE VRIES, E., DAFNI, U. *et al.* (2018). Comparative Assessment of Clinical Benefit Using the ESMO-Magnitude of Clinical Benefit Scale Version 1.1 and the ASCO Value Framework Net Health Benefit Score. *J Clin Oncol*, 37, pp. 336-349.
- CORTAZAR, P., ZHANG, L., UNTCH, M. *et al.* (2014). Pathological complete response and long term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet*, 384, pp. 164–172.

- DEPARTMENT OF HEALTH AND HUMAN SERVICES, FOOD AND DRUG ADMINISTRATION. (2009). Guidance for Industry on Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims; Availability, 65132-65133 [E9-29273]. Disponible en: <https://regulations.justia.com/regulations/fedreg/2009/12/09/E9-29273.html>
- EUROPEAN MEDICINES AGENCY. (2016). Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man - the use of patient-reported outcome (PRO) measures in oncology studies. Disponible en: <https://www.ema.europa.eu/en/appendix-2-guideline-evaluation-anticancer-medicinal-products-man-use-patient-reported-outcome-pro>
- INSTITUTE OF MEDICINE. (2001). *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press.
- . (2013). *Delivering High-Quality Cancer Care: Charting a New Course for a System in Crisis*. Washington, DC: National Academies Press.
- LOCKLEAR, T. *et al.* (2015). Reaching consensus on patient-centered definitions: a report from the Patient-Reported Outcomes PCORnet Task Force. *NIH Collaboratory*. Disponible en: https://www.nihcollaboratory.org/Products/Reaching%20Consensus_April_9_2015.pdf.
- MARANDINO, L., LA SALVIA, A., SONETTO, C. *et al.* (2018). Deficiencies in health-related quality-of-life assessment and reporting: a systematic review of oncology randomized phase III trials published between 2012 and 2016. *Ann Oncol*, 29(12), pp. 2288-2295. doi: 10.1093/annonc/mdy449.
- OLLENDORF, D. y PEARSON, S. D. (2013). *ICER Evidence Rating Matrix: A User's Guide*. Disponible en: <http://www.icer-review.org/wpcontent/uploads/2013/04/Rating-Matrix-User-Guide-Exec-Summ-FINAL.pdf>
- PORTER, M. E. (2010). What is value in health care? *N Engl J Med*, 363, pp. 2477-2481.
- SCHNIPPER, L. E., DAVIDSON, N. E., WOLLINS, D. S. *et al.* (2015). American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options. *J Clin Oncol*, 33, pp. 2563-2577.
- . (2016). Updating the American Society of Clinical Oncology Value Framework: Revisions and Reflections in Response to Comments Received. *Journal of Clinical Oncology*, 34, pp. 2925-2934.



16

EVALUACIÓN DE RESULTADOS EN SALUD: HEMATOLOGÍA

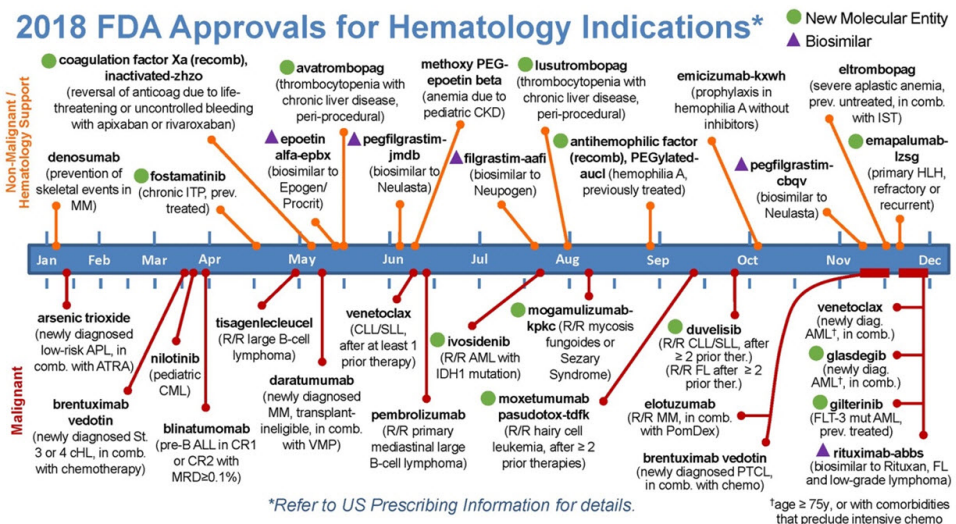
RAÚL CÓRDOBA MASCUÑANO, JAVIER BÉCARES MARTÍNEZ

16.1. INTRODUCCIÓN

Tradicionalmente, la investigación en hematología se ha centrado en su mayoría en la búsqueda de nuevos tratamientos para los pacientes con neoplasias hematológicas, aunque recientemente se ha visto un aumento de nuevos fármacos para tratar algunos tipos de anemias y algunos problemas de la coagulación como la hemofilia (Figura 1). Pero si hablamos de grandes avances en fármacos innovadores, en las últimas dos décadas han aparecido un gran número de fármacos antineoplásicos orales, en ocasiones más eficaces que los tratamientos quimioterápicos convencionales, y que vienen a introducir un elemento más en la cadena del proceso asistencial: la adherencia al tratamiento. En este aspecto, es por primera vez cuándo se le da al paciente un papel más activo en el tratamiento de su enfer-

Figura 1.

FÁRMACOS APROBADOS EN 2018 POR LA FOOD AND DRUG ADMINISTRATION (FDA) AMERICANA PARA INDICACIONES EN ENFERMEDADES DE LA SANGRE



Fuentes: FDA Oncology on Twitter.

edad. Y en parte, la adherencia o más bien, la falta de la misma, se debe a la aparición de efectos adversos no deseados que hace que el paciente no cumpla con el tratamiento tal y como se ha prescrito, conllevando en ocasiones pérdida de eficacia de los mismos. A raíz de este cambio en el modo de tratar las enfermedades de la sangre, vamos a involucrar al paciente en los resultados de eficacia que podemos conseguir, y con ello se abre un nuevo escenario en el que el paciente participa de forma activa en todo el proceso asistencial. De ahí que los *Patient Reported Outcomes* (PRO) en hematología cobren más relevancia tras la irrupción de estas nuevas terapias (Efficace, Gaidano y Lo-Coco, 2017).

Está claro que tenemos que medir los resultados en salud comunicados por los pacientes, pero la cuestión es cómo hacerlo. Uno de los problemas a los que nos enfrentamos es la falta de estandarización a la hora de recoger los resultados y luego analizarlos. Los PRO no solo pretenden analizar los efectos adversos experimentados por los pacientes, sino que pretenden ir un paso más allá: analizar el impacto que tienen en la calidad de vida y en la capacidad funcional de los pacientes (De Muro *et al.*, 2013).

Respecto a la aparición de efectos adversos, es muy común que los hematólogos registren menos efectos secundarios en comparación con lo que experimentan los pacientes, tal y como se recogen en diferentes trabajos (Di Maio *et al.*, 2016). Las razones por las que esto puede suceder no están bien explicadas, pero pueden verse justificadas por el escaso tiempo que se les dedica a los pacientes en las consultas, por no preguntar de forma activa y específica por determinados efectos adversos, por el bloqueo del paciente en la consulta en la que no se le olvida todo lo que le quería comentar al médico, o incluso en ocasiones, por el miedo de que el hematólogo le interrumpa el tratamiento debido a los efectos adversos y se quede sin opciones terapéuticas. Y no solo ocurre en el porcentaje de pacientes que comunican efectos secundarios, sino en la severidad de los mismos. Los médicos tienden a minimizar o trivializar algunos efectos secundarios que son vividos por los pacientes como algo que impacta negativamente en su calidad de vida, como ocurre por ejemplo en pacientes con determinadas formas de leucemia crónica que precisan tratamientos antineoplásicos orales de forma crónica y continuada (Efficace, Rosti y Aaronson, 2014). Un efecto secundario leve considerado por el hematólogo, pero mantenido en el tiempo de forma indefinida, puede convertirse en un problema relevante en la calidad de vida de los pacientes y que tenga más importancia de la que se le puede dar desde el punto de vista de la toxicidad farmacológica.

Finalmente, la Asociación Europea de Hematología (EHA, en inglés *European Hematology Association*), ha incluido dentro de sus líneas estratégicas el diseño de tratamientos mejores tolerados, en especial para los pacientes mayores (Engert *et al.*, 2016). Una de las herramientas diseñadas para tal fin es la HM-PRO (del inglés, *Hematologic Malignancy – Patient Reported Outcome*), presentada por primera vez en el congreso anual de la EHA en Madrid en 2017, ante la falta de una herramienta específica para pacientes con neoplasias hematológicas (Goswami, Khatib y Salek, 2018).

■ 16.2. EVALUACIÓN DE RESULTADOS EN ONCOHEMATOLOGÍA

■ 16.2.1. Leucemia linfática crónica

La leucemia linfática crónica (LLC) ha experimentado grandes avances en la última década, no solo con la incorporación de nuevos anticuerpos monoclonales, sino con los tratamientos moleculares dirigidos como los inhibidores de tirosina quinasa de Bruton (BTK), de PI3Kdelta o de BCL2. Con respecto a los nuevos anticuerpos monoclonales, obinutuzumab, un anticuerpo anti-CD20 tipo 2, en combinación con clorambucilo, ha demostrado incrementar la supervivencia libre de progresión y tiempo hasta el próximo tratamiento cuando se administra en pacientes no tratados previamente y que presentaran comorbilidades severas y/o no eran candidatos a recibir esquemas basados en fludarabina (Goede *et al.*, 2014), en comparación con rituximab. Este incremento de eficacia se vio asociado a un incremento en efectos secundarios, como mayor porcentaje de pacientes que experimentaron neutropenia, trombopenia y de reacciones infusionales. No obstante, la aparición de estos efectos adversos no impactó negativamente en la calidad de vida de los pacientes. Además, en un estudio realizado en España, la combinación de obinutuzimab con clorambucilo resultó ser fármaco-eficiente, con un incremento de los años de vida ajustados a calidad (AVAC) si se comparaba con la combinación rituximab con clorambucilo (Casado *et al.*, 2016).

Con respecto a los nuevos tratamientos moleculares dirigidos orales, ibrutinib ha demostrado mejorar no solo los síntomas sino el estado funcional de los pacientes con LLC en comparación con brazo control (Barrientos *et al.*, 2018; Cramer *et al.*, 2018). En cambio, en estudios de vida real, el porcentaje de discontinuación de los nuevos inhibidores como ibrutinib e idelalisib es más elevado de los comunicados inicialmente en los ensayos clínicos (Mato *et al.*, 2018), y será algo en lo que debemos de trabajar.

Recientemente, se ha publicado un estudio donde se analizan un total de 17 estudios sobre calidad de vida e impacto económico en esta enfermedad en la era de los nuevos fármacos, realizados entre 1 enero 2000-2 junio 2019. Los pacientes reportaron un deterioro en varios dominios de la calidad de vida en comparación con los controles sanos, incluida la fatiga, la ansiedad, el funcionamiento físico, el funcionamiento social, la depresión, los trastornos del sueño y la interferencia del dolor. Los factores clave asociados con un impacto negativo fueron el sexo femenino, el aumento de la gravedad de la enfermedad y el inicio de múltiples líneas de tratamiento. Se evaluó la carga económica para los pacientes con LLC según el estado de la enfermedad y el régimen de tratamiento recibido. Los principales generadores de costes relacionados con la LLC fueron los costes relacionados con la hospitalización y los pacientes ambulatorios, principalmente como resultado de la quimioinmunoterapia, los eventos adversos y la progresión de la enfermedad. El tratamiento con agentes dirigidos, es decir, ibrutinib y venetoclax, se asoció con

costes médicos más bajos que la quimioinmunoterapia, aunque ibrutinib se asoció con algunos costes aumentados de eventos adversos relacionados con toxicidades cardíacas. Todo indica que las terapias dirigidas pueden reducir la carga económica de la CLL, pero se necesitan datos de seguimiento más prolongados (Waweru *et al.*, 2020).

■ 16.2.2. Linfoma

Los pacientes con linfoma, a diferencia de otras neoplasias hematológicas, en lugar de tratamientos crónicos y prolongados, a menudo reciben estrategias limitadas en el tiempo con una intención curativa. Existe algunos subtipos de linfomas potencialmente curables y, por tanto, los pacientes se convierten en largos supervivientes. En este grupo de pacientes se están centrando gran cantidad de esfuerzos al haberse identificado la persistencia de efectos adversos a muy largo plazo, como, por ejemplo, la fatiga, que puede extenderse hasta más allá de los 10 años desde el diagnóstico (Oerlemans *et al.*, 2013). La presencia de enfermedad estadio IV y de comorbilidades favorecerían la persistencia de la fatiga a tan largo plazo.

Debido a estas características, se ha diseñado una estrategia para identificar e intervenir de forma precoz en la aparición de los síntomas comunicados por los pacientes con linfoma. Es el ensayo LIVE, donde próximamente tendremos resultados sobre la relevancia y la necesidad de tener un acceso precoz a los datos comunicados por los pacientes, así como diseñar estrategias que reduzcan o eliminen los efectos secundarios que experimentan (Arts *et al.*, 2017).

En un estudio reciente, la Organización Europea para el Tratamiento e Investigación del Cáncer (EORTC) ha evaluado el cuestionario de calidad de vida (QLQ-C30) en pacientes con linfoma de Hodgkin (LH) en estadio avanzado y supervivientes de LH a 5 años. El QLQ-C30 es el cuestionario específico de cáncer más utilizado para evaluar la calidad de vida relacionada con la salud. Los pacientes informaron peores puntuaciones de calidad de vida que los supervivientes en la mayoría de las escalas funcionales y escalas de síntomas. Estas puntuaciones variaron en función del género, pero no de la edad. Los informes de calidad de vida de los supervivientes fueron comparables a los de la población general. Estos valores proporcionan un marco de evaluación para la comparación e interpretación de las puntuaciones del QLQ-C30 en pacientes con LH en estadio avanzado (Vachon *et al.*, 2021).

■ 16.2.3. Mieloma múltiple

El mieloma múltiple ha sufrido una revolución con la irrupción de nuevos fármacos, como los inhibidores de proteosoma (bortezomib, carfilzomib, ixazomib), inmunomoduladores (talidomida, lenalidomida, pomalidomida) y anticuerpos monoclonales (daratumumab, elotuzumab). Todos los ensayos de estos nuevos agentes

incluyeron como objetivos secundarios la calidad de vida, y si bien, muchos de estos fármacos en combinaciones de tripletes son más eficaces, también presentan más toxicidad. Pero quizás en parte por una mayor rapidez de respuesta, los síntomas se controlan antes y, por tanto, los pacientes comunican mejores resultados de calidad de vida (Stewart, Dimopoulos y Masszi, 2016).

Dado que los pacientes con mieloma múltiple experimentan un gran número de síntomas, en ocasiones relacionadas con la propia enfermedad, y en otras relacionadas con las toxicidades producidas por los tratamientos, un grupo de expertos nacionales ha elaborado unas recomendaciones para la evaluación de la salud por parte de los pacientes con mieloma múltiple, entre los que se incluyen escalas como la *European Organisation for the Research and Treatment of Cancer (EORTC); Quality of Life Questionnaire (QLQ); Core questionnaire 30 (C30); EORTC-QLQ-Multiple Myeloma (MY20)* y *EORTC-QLQ-Breast Cancer (BR23)*, escala visual analógica para evaluación del dolor, así como preguntas específicas sobre preferencias del paciente y grado de satisfacción (Blade *et al.*, 2018).

■ 16.2.4. Leucemias agudas

En las leucemias agudas, se han obtenido aprobaciones de nuevos fármacos con mecanismo de acción novedosos. En las leucemias agudas linfoblásticas (LAL), blinatumumab, un anticuerpo biespecífico anti-CD19-CD3, y inotuzumabozogamizina, un anticuerpo anti-CD19 conjugado con un agente quimioterápico, han sido aprobados para LAL en recaída o refractarias. En el caso de blinatumumab, el ensayo clínico demostró prolongar la supervivencia global de los pacientes y, además, mejorar el control de síntomas y aspectos funcionales (Kantarjian, Stein y Gökbuget, 2017).

Con respecto a las terapias con células T modificadas genéticamente con receptor antigénico quimérico, también llamadas células T CAR, la aprobación de tisagenlecleucel ha venido a revolucionar el armamentario disponible para niños y adultos hasta 25 años con LAL refractaria. Los estudios preliminares de calidad de vida demuestran cómo mejoran el control de síntomas a los 28 días de su infusión, y que siguen mejorando a los tres y seis meses (Vairy *et al.*, 2018).

Con respecto a las leucemias agudas mieloblásticas (LAM), la FDA ha aprobado en 2018 diferentes fármacos para pacientes con LAM: ivosidenib, para pacientes con mutaciones en IDH1; venetoclax; glasdegib; y gilterinib, para pacientes con LAM FLT3 positivos. La aparición de estos nuevos fármacos viene a revolucionar el tratamiento de una forma de leucemia en la que en los últimos 30 años no había experimentado ningún avance terapéutico. El diferente perfil de seguridad de los mismos hará necesario un control estrecho del impacto de los mismos en la calidad de vida de los pacientes, al ser todos ellos fármacos orales y con exposición prolongada a los mismos, pudiendo existir en ocasiones aparición de eventos adversos en

tratamientos prolongados, no comunicados inicialmente en los ensayos clínicos de registro que llevaron a la aprobación de los mismos.

Un caso especial es el de la leucemia aguda promielocítica. La incorporación del trióxido de arsénico a la primera línea en combinación con ATRA en pacientes de riesgo bajo o intermedio ha venido a revolucionar el tratamiento de una leucemia con elevadísimos porcentajes de curación, pero con elevada morbilidad en el primer mes de tratamiento. Este esquema libre de quimioterapia ha demostrado además mejorar los PRO en pacientes tratados post-inducción con este esquema (Efficace, Mandelli y Avvisati, 2014).

■ 16.2.5. Síndromes mielodisplásicos

Los pacientes con síndromes mielodisplásicos suelen tener un gran número de síntomas que frecuentemente impactan negativamente en su calidad de vida y en la adherencia a los tratamientos (Patel y Gerds, 2017). Además, el diagnóstico en pacientes de edad avanzada, y frecuentemente con otras comorbilidades, hace necesario la correcta identificación de los síntomas para poder abordarlos de manera multidisciplinar.

■ 16.2.6. Leucemia mieloide crónica

Probablemente sea la leucemia mieloide crónica la que más nos ha enseñado sobre los resultados en salud comunicados por los pacientes desde la irrupción en el mercado de imatinib, el primer inhibidor tirosina quinasa indicado para tratar esta forma de leucemia. Más de 20 años después han llevado a aprender a manejar los efectos secundarios derivados de los mismos, pero también nos ha enseñado cómo una mala tolerancia de efectos adversos considerados por los hematólogos como leves (grados 1-2 según la CTCAE), llevaba a la discontinuación del tratamiento de forma unilateral por parte del paciente. Y gracias a eso se pudo descubrir que hasta un 40 % de los pacientes en respuestas moleculares profundas no recaían tras la discontinuación del tratamiento, lo cual llevó a diseñar ensayos clínicos de discontinuación, no solo con imatinib, sino con los otros inhibidores de tirosina quinasa de segunda generación como nilotinib y dasatinib, también indicados en la primera línea de tratamiento. Cada inhibidor tiene un perfil de toxicidad diferente, lo cual permite seleccionar uno frente al otro en base a las características individuales de cada paciente. El planteamiento de un tratamiento de forma indefinida hace necesario evaluar el impacto en la calidad de vida de los pacientes, tal y como ha quedado demostrado en numerosos estudios al respecto (Efficace y Cannella, 2016).

■ 16.2.7. Otras neoplasias mieloproliferativas crónicas

La mielofibrosis representa el área donde los PRO han contribuido más al diseño de nuevas estrategias de seguimiento y evaluación de la respuesta de los tratamientos. Se han desarrollado herramientas específicas como el *Myelofibrosis Symptom Assessment Form* en el ensayo comparando ruxolitinib frente a placebo (Mesa, Gotlib y Gupta, 2013) y frente a mejor terapia de soporte (Deisseroth, Kaminskas y Grillo, 2012) donde se observó una mejoría en el control de los síntomas, incluyendo la fatiga.

■ 16.3. EVALUACIÓN DE RESULTADOS EN TRASPLANTE DE PROGENITORES HEMATOPOYÉTICOS

Los pacientes sometidos a un trasplante de progenitores hematopoyético (TPH) sufren a menudo de alteraciones a largo plazo, difícilmente cuantificables. Esta modalidad de tratamiento se ofrece en la mayoría de las ocasiones como estrategias de consolidación de tratamiento, y frecuentemente con intención curativa, que dependiendo de la enfermedad hematológica oscila entre 40 %-70 %. Los pacientes sometidos a un TPH y que superan el periodo precoz pos-TPH sin morbilidad importante, pasan a formar parte del grupo de largos supervivientes, en ocasiones con síntomas no recuperados a largo plazo y que impactan negativamente en su calidad de vida como miedo a recaer, fatiga, infertilidad, falta de autoestima por secuelas, ... Una identificación precoz e intervención para que ayuda a controlar mejor los síntomas tales como las alteraciones neuropsicológicas (Majhail *et al.*, 2018). Un plan de cuidados individualizado pos-TPH se hace necesario para la identificación precoz de los síntomas y realizar su abordaje de forma más eficiente (Denzen *et al.*, 2018).

■ 16.4. EVALUACIÓN DE RESULTADOS EN SALUD EN HEMOPATÍAS BENIGNAS: EL EJEMPLO DE LA ANEMIA DE CÉLULAS FALCIFORMES

En Estados Unidos, la anemia de células falciformes es la alteración hereditaria de la sangre más frecuente, con unos 100.000 adultos afectados que causa unas 113.000 hospitalizaciones anuales (Brousseau *et al.*, 2010). Lo más característico de esta enfermedad son las crisis de dolor agudo, responsables de la mayoría de las visitas a urgencias y de los ingresos. Además, los pacientes padecen dolor crónico, afectación de órganos y déficits neurológicos. Lo que antes era una enfermedad de la infancia ahora es una enfermedad de adultos. En el año 2002 el National Heart, Lung, and Blood Institute (*NHLBI*) en Estados Unidos, organizó un grupo de trabajo con pacientes, médicos y científicos, que concluyeron que la necesidad más importante era desarrollar una herramienta para conocer la calidad de vida relacionada con la salud de los pacientes afectados de ACF. Este grupo de trabajo desarrolló

el proyecto ASCQ-Me (*Adult Sickle Cell Quality-of-Life Measurement System*). Una herramienta para analizar los siguientes aspectos de la vida de los pacientes: Dolor, salud mental, satisfacción con la atención sanitaria, oportunidades de formación y empleo, carga que supone el tratamiento, discriminación y estigmatización, e independencia para la vida diaria (Marsha *et al.*, 2014).

El cuestionario ASCQ-Me¹ consiste en siete áreas que valoran:

- Impacto emocional: cinco preguntas.
- Episodios de dolor: cinco preguntas que miden el dolor agudo debido a las crisis veno-oclusivas.
- Impacto del dolor: Son cinco cuestiones sobre el dolor en los últimos siete días.
- Checklist de la historia clínica de la ACF: nueve preguntas
- Impacto en el sueño: cinco cuestiones.
- Impacto en la funcionalidad social: también cinco preguntas.
- Efecto sobre la frialdad o rigidez: cinco preguntas.

Para medir los resultados reportados por los pacientes pediátricos, se exploró una adaptación del cuestionario general PEDS-QL a la anemia de células falciformes (Peds-QL SCD). Se trata de un formulario de 43 preguntas relacionadas con el dolor, la preocupación por la enfermedad, las emociones, el tratamiento y la comunicación (Panepinto *et al.*, 2013).

Se ha tratado de comparar varios cuestionarios para evaluar cómo afecta esta enfermedad a los pacientes adultos y niños. En adultos se compara el cuestionario ASCQ-Me con el PROMIS de medida específica del dolor en adultos, en el que se concluye que el cuestionario ASCQ-Me es válido, factible y más sensible que el PROMIS para evaluar los resultados en salud en un momento dado, pero ninguno de ellos sirve para evaluarlo a lo largo del tiempo. En niños, se ha comparado el PROMIS de dolor pediátrico con el módulo *PedsQL SCD*. Ambos son factibles y válidos para responder a los cambios en el dolor agudo en niños (Curtis y Brandon, 2017).

Con la llegada de nuevos fármacos y el futuro de la terapia génica abre las puertas a explorar la evaluación de síntomas por los pacientes como algo necesario para monitorizar la respuesta al tratamiento, y no utilizar únicamente parámetros analíticos como el nivel de Hb o la necesidad transfusional.

¹ www.ascq-me.org/ASCQ-Me

■ 16.5. EVALUACIÓN DE RESULTADOS EN SALUD EN TROMBOSIS Y HEMOSTASIA

■ 16.5.1. Anticoagulantes orales

Los pacientes que reciben tratamiento con anticoagulantes orales precisan habitualmente de un tratamiento prolongado que, en ocasiones, suele ser de por vida. Uno de los retos a los que se enfrentan los médicos prescriptores es la adherencia y satisfacción por parte del paciente. Existen varios cuestionarios para medir la adherencia y satisfacción de los pacientes con los anticoagulantes orales (Barret *et al.*, 2014):

- *Anti-Clot Treatment Scale (ACTS)*:
 - Es un cuestionario específico para medir la satisfacción frente a la anticoagulación.
 - Basado en el DASS, pero modificado para ser más aplicable a mayor variedad de pacientes y lenguajes.
 - No limitado a unas condiciones específicas ni a la vía de administración del anticoagulante.
 - Contiene quince preguntas, de dos campos, uno es una escala de cargas o responsabilidades con doce cuestiones y el otro una escala de satisfacción con tres.
 - Dos notas: impacto positivo o negativo del tratamiento en la calidad de vida.
 - El periodo de recuerdo es de cuatro semanas.
- *Duke anticoagulation Satisfaction Scale (DASS)*:
 - Es un cuestionario específico de anticoagulación para medir la satisfacción con el tratamiento.
 - No limitado a la vía de administración del anticoagulante.
 - Contiene 25 preguntas en dos áreas principales, que son los aspectos positivos y negativos del tratamiento anticoagulante.
 - No especifica el tiempo de recuerdo del paciente en sus respuestas.
- *Perception of Anticoagulant Treatment Questionnaire (PACT-Q)*:
 - Mide las expectativas y la satisfacción del paciente frente al tratamiento.
 - No limitado a unas condiciones específicas ni a la vía de administración del anticoagulante.

- Contiene 27 preguntas, originalmente propuestas en cuatro aspectos: Expectativas del tratamiento; Conveniencia del mismo; Carga de la enfermedad y del tratamiento y satisfacción con el tratamiento anticoagulante.
- No se especifica el periodo de recuerdo.
- **Sawicki Questionnaire:**
 - Es un cuestionario específico de anticoagulación desarrollado en Alemania, para comparar la anticoagulación en el domicilio y en el hospital.
 - Contiene 32 *items*, incluyendo cinco temas:
 - Satisfacción general con el tratamiento.
 - Molestias diarias.
 - Relaciones sociales tensas.
 - Angustia.
 - No se especifica el periodo de recuerdo.
- **Deep Venous Thrombosis Quality of Life (DVTQOL):**
 - Es un cuestionario específico de trombosis venosa profunda (TVP) para medir los resultados en calidad de vida (incluye las demandas prácticas y psicológicas del tratamiento anticoagulante oral) en relación con un evento primario de TVP.
 - Contiene 29 preguntas, agrupadas en seis dimensiones:
 - Alteración emocional.
 - Síntomas: dolor, hinchazón de tobillos, calambres, hematomas.
 - Limitación de la actividad física.
 - Dificultades con la monitorización de la anticoagulación.
 - Alteraciones del sueño.
 - Problemas con la dieta.
 - No se especifica el periodo de recuerdo.

Un ejemplo de la utilización del cuestionario ACTS es el trabajo realizado en el que se comparó rivaroxabán con enoxaparina seguida de antivitamina K en el tratamiento del embolismo pulmonar agudo y sintomático. Participaron 2.397 voluntarios de siete países (Canadá, Francia, Alemania, Italia, Holanda, Reino Unido, Canadá y Estados Unidos). Se trata de una parte del estudio EINSTEIN PE. Además del cuestionario ACTS, se empleó el cuestionario para la satisfacción con el tratamiento ver-

sión II (TSQM II). La satisfacción mostrada por los pacientes fue estadísticamente más favorable al tratamiento con rivaroxaban (Prins *et al.*, 2012).

■ 16.5.2. Hemofilia

Los últimos resultados reportados por pacientes nos indican que, a pesar del tratamiento recibido, los pacientes no dejan de padecer dolor agudo-crónico y hemorragias frecuentes en músculos o articulaciones. Sin embargo, a medida que hay un aumento de la esperanza de vida en los pacientes hemofílicos, es común que dicha patología se acompaña de comorbilidades habituales. Entre estas cabe destacar que en la edad adulta los pacientes pueden sufrir osteoporosis, obesidad, hipertensión y enfermedad cardiovascular (World Federation of Hemophilia, 2018). En la era actual de la profilaxis del sangrado es necesario medir los resultados percibidos por los pacientes, en el campo asistencial y de investigación. En la hemofilia se han empleado herramientas para medir PRO de otras enfermedades crónicas (SF36, EQ-5D, IPAQ) pero ninguno de ellos discrimina o detecta los cambios más relevantes para los pacientes con hemofilia (Recht *et al.*, 2016). En un estudio se compararon la fiabilidad de cinco herramientas para medir PRO en el campo de la hemofilia: EQ-5D-5L, *Brief Pain Inventory* (BPI), *International Physical Activity Questionnaire* [IPAQ], *Short Form 36 Health Survey v2* (SF-36v2), y *Hemophilia Activities List* (HAL). Todas ellas mostraron tasas similares de fiabilidad, por lo que, según los autores del trabajo, la elección del cuestionario se debe hacer pensando en sus características, más que en su fiabilidad (Kempton *et al.*, 2017).

■ 16.6. EVALUACIÓN DE RESULTADOS EN TRATAMIENTOS DE SOPORTE

Los cuidados continuos o cuidados de soporte son tratamientos que tienen como objetivo la mejora de los síntomas a través de un cuidado total, activo y continuado cubriendo sus necesidades físicas, psicológicas, emocionales, emocionales y sociales.

■ 16.6.1. Soporte transfusional

Los pacientes con determinadas enfermedades hematológicas como los síndromes mielodisplásicos, pueden tener una dependencia transfusional. Los síntomas atribuibles a la anemia en ocasiones no se correlacionan con el nivel de hemoglobina, al ser pacientes con anemia crónica que generan mecanismos compensatorios para mejorar sus síntomas. Existen trabajos en los que se analiza el impacto de la calidad de vida en pacientes dependientes de transfusiones, en los que claramente existe un beneficio en el control de síntomas, si bien el punto de corte varía y no debería utilizarse el de los pacientes en situaciones de anemia aguda (Jansen *et al.*, 2003).

■ 16.6.2. Quelación

La terapia de quelación consiste en la administración de agentes quelantes para eliminar los metales pesados tóxicos del cuerpo. Es una terapia que requiere una estricta supervisión médica y precaución, ya que puede llegar a ocasionar la muerte del paciente. La mayoría de las enfermedades hematológicas requieren transfusiones en función del estado y patología del paciente, por lo que se produce una acumulación de hierro, elemento necesario para el correcto funcionamiento del organismo, pudiendo generar una toxicidad en el receptor. La quelación del hierro es el único tratamiento farmacológico eficaz para tratar la sobrecarga férrica en pacientes con síndromes mielodisplásicos, siendo la enfermedad hematológica más común en adultos mayores de 60 años y precisando constantes transfusiones sanguíneas, lo que produce una sobrecarga de hierro, empeorando la calidad de vida y llegando a ocasionar incluso la muerte. La anemia afecta al 80 % de los pacientes con síndromes mielodisplásicos quienes necesitarán soportes transfusionales de forma periódica, para evitar los síntomas y problemas que condiciona la anemia. Estas transfusiones originan, antes o después, una sobrecarga férrica que precisa administrar un tratamiento quelante, único tratamiento farmacológico eficaz. El tratamiento de quelación del hierro en la actualidad es un tratamiento oral crónico, donde la adherencia del paciente es fundamental (Escudero-Vilaplana *et al.*, 2016). En un trabajo basado en PRO sobre las preferencias del paciente frente a dos formulaciones de deferasirox, comprimidos recubiertos frente a dispersables, se emplearon tres cuestionarios: cuestionario de satisfacción con el tratamiento de quelación del hierro modificado; cuestionario de palatabilidad y diario de síntomas gastrointestinales. La información obtenida de los tres cuestionarios fue concordante (Taher *et al.*, 2018).

■ 16.7. E-RES SALUD, UN EJEMPLO DE PROGRAMA DE EVALUACIÓN DE RESULTADOS EN SALUD EN HEMATOLOGÍA

E-Res Salud es un programa de la evaluación de resultados en salud, un programa de medicina basada en el valor en el que medimos lo que verdaderamente importa al paciente. Tradicionalmente se ha medido la actividad asistencial con actos (consultas realizadas al año, tratamientos puestos en hospital de día, estancia media, etc), sin analizar si lo que se realiza aporta valor. Por eso, se ha comenzado a preguntarle a los pacientes si lo que se hace en los actos asistenciales, a ellos les importa. Y para eso se dispone de cuestionarios PROM, de resultados reportados por los pacientes, y PREM, de experiencia del paciente.

Con lo presentado anteriormente, se pone de manifiesto que la percepción que tienen los médicos y los pacientes de un mismo problema de salud puede ser muy diferente, y lo que uno considera leve o poco relevante, al otro le genera un alto impacto. De ahí la importancia de dar voz al paciente para que cuente cómo afecta una enfermedad o tratamiento a su calidad de vida, el impacto que provoca

en ella y los posibles efectos adversos. Existen escalas de medición de estos resultados validadas y que se están empezando a utilizar en oncohematología, como el PRO-CTCAE del Instituto Nacional del Cáncer Americano, y el HM-PRO, de la Asociación Europea de Hematología.

Tradicionalmente, los cuestionarios de calidad de vida y percepción de salud reportados por los pacientes se han realizado en papel. La transformación digital ha permitido que dichos cuestionarios sean ahora cumplimentados a través de nuevas herramientas, como webs o aplicaciones.

Este programa se ha implementado por la Unidad de Innovación Clínica y Organizativa (UICO) en los hospitales universitarios Fundación Jiménez Díaz, Rey Juan Carlos, Infanta Elena y General de Villalba. Los principales beneficiados son los pacientes. En E-Res Hematología se han identificado algunos síntomas de alerta sobre los que se puede actuar casi en tiempo real. Cuando se detectan estos síntomas, de forma proactiva se puede llamar a los pacientes para preguntarles cómo están, confirmar el síntoma que ha generado la alerta y dar respuesta y resolverlo. Con este sistema se ha optimizado el consumo de recursos, al poder gestionar tratamientos en el Hospital de Día, realizar consultas no presenciales o cursar ingresos programados sin necesidad de acudir a Urgencias.

El programa en Hematología comenzó en diciembre de 2019, y ya se han incluido a más de 600 pacientes, con un grado de cumplimentación que ronda el 40 %, lo cual está en consonancia con lo que reportan otros centros internacionales. Respecto a la capacidad de identificar síntomas de alerta, se han reportado casi 2.000, permitiendo dar respuesta de forma rápida y eficaz a las mismas.

El siguiente reto es analizar el impacto de este sistema de evaluación de resultados en salud con el consumo de recursos, con el fin de poder realizar un consumo más eficiente de los mismo. Hasta ahora se tienen ejemplos de una mejor planificación y programación de actividades ambulatorias en el Hospital de Día, sin tener que consumir recursos tales como los ingresos en urgencias o las hospitalizaciones no programadas.

■ 16.8. CONCLUSIONES

- Los pacientes con neoplasias hematológicas, a diferencia de los que tienen tumores sólidos, se van a caracterizar por una experiencia muy prolongada de síntomas con los que tienen que convivir, bien por ser enfermedades de curso lento, con supervivencias muy prolongadas como ocurre con algunos linfomas indolentes, que pueden tener supervivencias de 15-18 años; bien por la “cronificación” de la enfermedad con las nuevas estrategias de tratamiento, como la leucemia mieloide crónica, que ha pasado de tener una supervivencia menor de 5 años a tenerla igual a la población general; o bien por tener un elevado porcentaje de pacientes curados que deben de sufrir las

secuelas del tratamiento y sus síntomas el resto de su vida, como muchos pacientes con linfomas agresivos o leucemias agudas.

- Una adecuada valoración de identificación permitirá diseñar estrategias individualizadas, encaminadas siempre a mejorar la calidad de vida de los pacientes, y en los pacientes que requieren medicación de forma crónica y por tiempo indefinido, garantizar la adherencia al tratamiento y mejorar por ello los resultados de eficacia de los mismos.
- Existe una falta de homogeneización a la hora de comunicar resultados en salud en pacientes con enfermedades hematológicas. Sería aconsejable diseñar y validar una herramienta común para poder ser explorada en todos los pacientes de forma homogénea.

■ REFERENCIAS

- ARTS, L. P. J., VAN DE POLL-FRANSE, L. V., VAN DEN BERG, S. W., PRINS, J. B., HUSSON, O., MOLS, F., BRANDS-NIJENHUIS, A. V. M., TICK, L. y OERLEMANS, S. (2017). Lymphoma InterVENTion (LIVE) - patient-reported outcome feedback and a web-based self-management intervention for patients with lymphoma: study protocol for a randomised controlled trial. *Trials*, 18(1), p. 199.
- BARRET, A. *et al.* (2014). *Literature Review of Patient-Reported Outcome Measures Assessing Anticoagulant Therapy*. Disponible en: <https://www.rtihs.org/sites/default/files/25509%20Barrett%202014%20Ispor%20Lit%20Review.pdf>
- BARRIENTOS, J. C., O'BRIEN, S., BROWN, J. R., KAY, N. E., REDDY, N. M., COUTRE, S., TAM, C., MULLIGAN, S., JAEGER, U., DEVEREUX, S., POCOCK, C., ROBAK, T., SCHUSTER, S. J., SCHUH, A., GILL, D., BLOOR, A., DEARDEN, C., MORENO, C., CULL, G., HAMBLIN, M., JONES, J. A., ECKERT, K., SOLMAN, I. G., SUZUKI, S., HSU, E., JAMES, D. F., BYRD, J. C. y HILLMEN, P. (2018). Improvement in Parameters of Hematologic and Immunologic Function and Patient Well-being in the Phase III RESONATE Study of Ibrutinib Versus Ofatumumab in Patients With Previously Treated Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma. *Clin Lymphoma Myeloma Leuk*, 18(12), pp. 803-813.e7.
- BLADE, J., CALLEJA, M. Á., LAHUERTA, J. J., POVEDA, J. L., DE PAZ, H. D. y LIZÁN, L. (2018). Defining a set of standardised outcome measures for newly diagnosed patients with multiple myeloma using the Delphi consensus method: the IMPORTA project. *BMJ Open*, 8(2), p. e018850.
- BROUSSEAU, D. C., PANEPINTO, J. A., NIMMER, M. y HOFFMANN, R. G. (2010). The number of people with sickle-cell disease in the United States: National and state estimates. *Am J Hematol*, 85(1), pp. 77-78.
- CASADO, L. F., BURGOS, A., GONZÁLEZ-HABA, E., LOSCERTALES, J., KRIVASI, T., OROFINO, J., RUBIO-TERRES, C. y RUBIO-RODRÍGUEZ, D. (2016). Economic evaluation of obinutuzumab in combination with chlorambucil in first-line treatment of patients with chronic lymphocytic leukemia in Spain. *Clinicoecon Outcomes Res*, 8, pp. 475-484.
- CRAMER, P., FRASER, G., SANTUCCI-SILVA, R., GROSICKI, S., DILHUYDY, M. S., JANSSENS, A., LOSCERTALES, J., RULE, S., GOY, A., TRAINA, S., CHAN, E. K. H., DIELS, J., SENGUPTA, N., MAHLER, M., SALMAN, M., HOWES, A. y CHANAN-KHAN, A. (2018). Improvement of fatigue, physical functioning, and well-being among patients with severe impairment at baseline receiving ibrutinib in combination with bendamustine and rituximab for relapsed chronic lymphocytic leukemia/small lymphocytic lymphoma in the HELIOS study. *Leuk Lymphoma*, 59(9), pp. 2075-2084.

- CURTIS, S. y BRANDON, A. M. (2017). Responsiveness of Patient-Reported Outcome Measurement Information System (PROMIS) pain domains and disease-specific patient-reported outcome measures in children and adults with sickle cell disease. *Hematology Am Soc Hematol Educ Program*, 2017(1), pp. 542–545.
- DEISSEROTH, A., KAMINSKAS, E. y GRILLO, J. (2012). U.S. Food and Drug Administration approval: ruxolitinib for the treatment of patients with intermediate and high-risk myelofibrosis. *Clin Cancer Res*, 18(12), pp. 3212-3217.
- DENZEN, E. M., PREUSSLER, J. M., MURPHY, E. A., BAKER, K. S., BURNS, L. J., FOSTER, J., IDOSSA, L., MOORE, H. K., PAYTON, T. J., HAVEN, D., JAHAGIRDAR, B., KAMANI, N., RIZZO, J. D., SALAZAR, L., SCHATZ, B. A., SYRJALA, K. L., WINGARD, J. R. y MAJHAIL, N. S. (2018). Tailoring a Survivorship Care Plan: Patient and Provider Preferences for Recipients of Hematopoietic Cell Transplantation. *Biol Blood Marrow Transplant*. Recuperado de: pii: S1083-8791(18)30614-1. doi: 10.1016/j.bbmt.2018.10.005. [Epub ahead of print].
- EFFICACE, F. y CANNELLA, L. (2016). The value of quality of life assessment in chronic myeloid leukemia patients receiving tyrosine kinase inhibitors. *Hematology*, 2016(1), pp. 170-179.
- EFFICACE, F., GAIDANO, G. y LO-COCO, F. (2017). Patient-reported outcomes in hematology: is it time to focus more on them in clinical trials and hematology practice? *Blood*, 130(7), pp. 859-866.
- EFFICACE, F., MANDELLI, F. y AVVISATI, G. (2014). Randomized phase III trial of retinoic acid and arsenic trioxide versus retinoic acid and chemotherapy in patients with acute promyelocytic leukemia: health-related quality-of-life outcomes. *J Clin Oncol*, 32(30), pp. 3406-3412.
- EFFICACE, F., ROSTI, G. y AARONSON, N. (2014). Patient- versus physician-reporting of symptoms and health status in chronic myeloid leukemia. *Haematologica*, 99(4), pp. 788-793.
- ENGERT, A., BALDUINI, C., BRAND, A., COIFFIER, B., CORDONNIER, C., DÖHNER, H., DE WIT, T. D., EICHINGER, S., FIBBE, W., GREEN, T., DE HAAS, F., IOLASCON, A., JAFFREDO, T., RODEGHIERO, F., SALLES, G. y SCHURINGA, J. J. (2016). EHA Roadmap for European Hematology Research. The European Hematology Association Roadmap for European Hematology Research: a consensus document. *Haematologica*, 101(2), pp. 115-208.
- ESCUDERO-VILAPLANA, V., GARCÍA-GONZÁLEZ, X., OSORIO-PRENDES, S., ROMERO-JIMÉNEZ, R. M. y SANJURJO-SÁEZ, M. (2016). Impact of medication adherence on the effectiveness of deferasirox for the treatment of transfusional iron overload in myelodysplastic syndrome. *J Clin Pharm Ther*, 41, pp. 59–63.
- GOEDE, V., FISCHER, K., BUSCH, R., ENGELKE, A., EICHHORST, B., WENDTNER, C. M., CHAGOROVA, T., DE LA SERNA, J., DILHUYDY, M. S., ILLMER, T., OPAT, S., OWEN, C. J., SAMOYLOVA, O., KREUZER, K. A., STILGENBAUER, S., DÖHNER, H., LANGERAK, A. W., RITGEN, M., KNEBA, M., ASIKANIUS, E., HUMPHREY, K., WENGER, M. y HALLEK, M. (2014). Obinutuzumab plus chlorambucil in patients with CLL and coexisting conditions. *N Engl J Med*, 370(12), pp. 1101-1110.
- GOSWAMI, P., KHATIB, Y. y SALEK, M. S. (2018). Haematological malignancy: are we measuring what is important to patients? A systematic review of quality of life instruments. *Eur J Haematol*. doi: 10.1111/ejh.13203. [Epub ahead of print].
- JANSEN, A. J., ESSINK-BOT, M. L., BECKERS, E. A., HOP, W. C., SCHIPPERUS, M. R. y VAN RHENEN, D. J. (2003). Quality of life measurement in patients with transfusion-dependent myelodysplastic syndromes. *Br J Haematol*, 121(2), pp. 270-274.
- KANTARJIAN, H., STEIN, A. y GÖKBÜGET, N. (2017). Blinatumomab versus chemotherapy for advanced acute lymphoblastic leukemia. *N Engl J Med*, 376(9), pp. 836-847.
- KEMPTON, C. L. et al. (2017). Reliability of patient-reported outcome instruments in US adults with hemophilia: the Pain, Functional Impairment and Quality of life (P-FiQ) study. Published on: Dovepress.com (September 19th, 2017).
- DI MAIO, M., BASCH, E., BRYCE, J. y PERRONE, F. (2016). Patient-reported outcomes in the evaluation of toxicity of anticancer treatments. *Nat Rev Clin Oncol*, 13(5), pp. 319-325.

- MAJHAIL, N. S., MURPHY, E., LAUD, P., PREUSSLER, J. M., DENZEN, E. M., ABETTI, B., ADAMS, A., BESSER, R., BURNS, L. J., CERNY, J., DREXLER, R., HAHN, T., IDOSSA, L., JAHAGIRDAR, B., KAMANI, N., LOREN, A., MATTILA, D., MCGUIRK, J., MOORE, H., REYNOLDS, J., SABER, W., SALAZAR, L., SCHATZ, B., STIFF, P., WINGARD, J. R., SYRJALA, K. L. y BAKER, K. S. (2018). Randomized controlled trial of individualized treatment summary and survivorship care plans for hematopoietic cell transplantation survivors. *Haematologica*, Recuperado de: pii:haematol.2018.203919. doi: 10.3324/haematol.2018.203919. [Epub ahead of print].
- MARSHA, J. T., HASSELL, K., LEVINE, R. y KELLER, S. (2014). Adult Sickle Cell Quality-of-Life Measurement Information System (ASCQ-Me): Conceptual Model Based on Review of the Literature and Formative Research. *Clin J Pain*, 30(10), pp. 902–914.
- MATO, A. R., SAMP, J. C., GAUTHIER, G., TERASAWA, E. y BRANDER, D. M. (2018). Drivers of treatment patterns in patients with chronic lymphocytic leukemia stopping ibrutinib or idelalisib therapies. *Cancer Biol Ther*, 19(7), pp. 636–643.
- MESA, R. A., GOTLIB, J. y GUPTA, V. (2013). Effect of ruxolitinib therapy on myelofibrosis-related symptoms and other patient-reported outcomes in COMFORT-I: a randomized, double-blind, placebo-controlled trial. *J Clin Oncol*, 31(10), pp. 1285–1292.
- DEMURO, C., CLARK, M., DOWARD, L., EVANS, E., MORDIN, M. y GNANASAKTHY, A. (2013). Assessment of PRO label claims granted by the FDA as compared to the EMA (2006–2010). *Value Health*, 16(8), pp. 1150–1155.
- OERLEMANS, S., MOLS, F., ISSA, D. E., PRUIJT, J. H., PETERS, W. G., LYBEERT, M., ZIJLSTRA, W., COEBERGH, J. W. y VAN DE POLL-FRANSE, L. V. (2013). A high level of fatigue among long-term survivors of non-Hodgkin's lymphoma: results from the longitudinal population-based PROFILES registry in the south of the Netherlands. *Haematologica*, 98(3), pp. 479–486.
- PANEPINTO, J. A. *et al.* (2013). PedsQL™ Sickle Cell Disease Module: Feasibility, Reliability and Validity. *Pediatr Blood Cancer*, 60(8), pp. 1338–1344.
- PATEL, S. S. y GERDS, A. T. (2017). Patient-Reported Outcomes in Myelodysplastic Syndromes and MDS/MPN Overlap Syndromes: Stepping Onto the Stage with Changing Times. *Curr Hematol Malig Rep*, 12(5), pp. 455–460.
- PRINS, M., BAMBER, L., CANO, S., WANG, M. y LENSING, A. W. A. (2012). Lensing and Rupert Bauersachs. Patient-Reported Treatment Satisfaction with Oral Rivaroxaban Versus Standard Therapy in the Treatment of Acute Symptomatic Pulmonary Embolism. *Blood*, 120, p. 1163.
- RECHT, M., KONKLE, B. A., JACKSON, S., NEUFELD, E. J. y PIPE, S. (2016). Recognizing the need for personalization of haemophilia patient-reported outcomes in the prophylaxis era. *Haemophilia*, 22, pp. 825–832.
- STEWART, A. K., DIMOPOULOS, M. A. y MASSZI, T. (2016). Health-related quality of life results from the open-label, randomized, phase III ASPIRE trial evaluating carfilzomib, lenalidomide, and dexamethasone versus lenalidomide and dexamethasone in patients with relapsed multiple myeloma. *J Clin Oncol*, 34(32), pp. 3921–3930.
- TAHER *et al.* (2018). Patient-reported outcomes from a randomized phase II study of the deferasirox film-coated tablet in patients with transfusion-dependent anemias *Health and Quality of Life Outcomes*, 16, p. 216.
- VACHON, H., MIERZYNSKA, J., TAYE, M., PE, M., COENS, C., MARTINELLI, F., FORTPIED, C., FLECHTNER, H. H., VESTMOE MARALDO, M., HUTCHINGS, M., MEIJNDERS, P., ALEMAN, B., LUGTENBURG, P., SPINA, M., ANDRÉ, M., HERTZBERG, M., BRIONES, J. y BOTTOMLEY, A. (2021). EORTC Quality of Life Group, Lymphoma Group. Reference values for the EORTC QLQ-C30 in patients with advanced stage Hodgkin lymphoma and in Hodgkin lymphoma survivors. *Eur J Haematol*. doi: 10.1111/ejh.13601. Epub ahead of print. PMID: 33570765.

- VAIRY, S., LOPES GARCÍA, J., TEIRA, P. y BITTENCOURT, H. (2018). CTL019 (tisagenlecleucel): CAR-T therapy for relapsed and refractory B-cell acute lymphoblastic leukemia. *Drug Des Devel Ther*, 12, pp. 3885–3898.
- WAWERU, C., KAUR, S., SHARMA S. y MISHRA, N. (2020). Health-related quality of life and economic burden of chronic lymphocytic leukemia in the era of novel targeted agents. *Curr Med Res Opin*, 36(9), pp.1481-1495. doi: 10.1080/03007995.2020.1784120. Epub 2020 Jul 23. PMID: 32634056.
- WORLD FEDERATION OF HEMOPHILIA. (2018). *Tratamiento de la hemofilia* - [Internet]. Wfh.org. 2018. Disponible en: <https://www.wfh.org/es/page.aspx?pid=781> [cited 21 December 2018].



17

RESULTADOS EN SALUD: PATOLOGÍA CARDIOVASCULAR

INMACULADA ROLDÁN RABADÁN

■ 17.1. INTRODUCCIÓN

■ 17.1.1. Una aproximación inicial a la medida de resultados

¿Por qué es imprescindible la medida de resultados en Cardiología?

La medición y el seguimiento de resultados es un elemento clave para el desarrollo adecuado y coherente de la estrategia de las organizaciones sanitarias. La introducción de los sistemas de clasificación de pacientes y ajuste de riesgo, desde los años ochenta, constituyeron la base fundamental para la incorporación de la medición de la actividad. La complejidad de la práctica médica exige medir los resultados para conseguir detectar los problemas e introducir medidas de corrección y mejora. La exigencia de sistemas de medida dirigidos a determinar los resultados clínicos en términos de efectividad y eficiencia constituye un rasgo fundamental para garantizar la calidad y seguridad.

La incorporación de sistemas integrales de medidas de resultados en cardiología constituye por tanto un elemento fundamental en la organización y planificación de la gestión clínica cardiológica. El análisis de los resultados de un Servicio, Área, o Unidad de Gestión de Cardiología es un instrumento básico para proporcionar el soporte adecuado a la mejora de la práctica clínica. Además, permite aumentar el conocimiento clínico sobre los procesos cardiológicos y compartir las mejores prácticas.

¿Cuáles son la áreas y campos de la medida de resultados?

Habitualmente los resultados se han venido estructurando en unas dimensiones bien definidas: actividad en términos cuantitativos, morbilidad atendida, satisfacción de los pacientes, adecuación de los procedimientos, efectividad clínica, eficiencia y seguridad de los procedimientos diagnósticos y terapéuticos. En cada uno de estos apartados se estructuran una serie de indicadores que permiten el análisis, la comparación y evaluación en el tiempo. La AHRQ (Agency for Healthcare Research and Quality) en el campo de los indicadores de calidad de la hospitalización recoge indicadores de mortalidad hospitalaria, de utilización de procedimientos susceptibles de utilización inadecuada, y volumen de procedimientos relacionados con menor mortalidad.

En nuestro medio sin embargo, no existen indicadores preespecificados iguales para todo el territorio español.

La Central de Resultados de Cataluña recoge datos de adecuación, efectividad, seguridad, eficiencia, satisfacción, datos económicos, docencia, y tecnologías de la información para el análisis de la red hospitalaria. Así, en el ámbito cardiológico para la evaluación se recogen indicadores como:

1. *Adecuación*: tiempo de entrada en sala y tiempo de angioplastia primaria en pacientes con código infarto. Además, porcentaje de pacientes en código infarto atendidos en menos de 120 minutos.
2. *Efectividad*: Mortalidad al alta y a los 30 días de pacientes con infarto agudo de miocardio con el elevación del segmento ST (IAMCEST).
3. *Eficiencia*: estancia media para pacientes con insuficiencia cardiaca (IC).

El Observatorio de Resultados del Servicio Madrileño de Salud (SERMAS) presenta indicadores referentes a actividad clínica, seguridad del paciente, eficiencia, atención al paciente, docencia e investigación. Así, en el área de efectividad clínica y seguridad del paciente cardiológico introduce la mortalidad por infarto agudo de miocardio (IAM) en los hospitales del SERMAS. Además incluye el ratio de mortalidad estandarizada como el cociente entre los casos observados y los casos esperados en función de las características de los pacientes (RME). También se considera la mortalidad por IC que se expresa como en el caso del IAM, con el RME. Analizan asimismo los reingresos por IC en los primeros 30 días después del alta.

Andalucía recoge a través del Servicio Andaluz de Salud indicadores referentes a la equidad, eficiencia, accesibilidad, efectividad, experiencia del paciente, seguridad, coordinación, capacidad del sistema e innovación.

Datos de actividad como indicadores de la eficiencia de un Servicio de Cardiología

Los indicadores habituales de ingresos, estancias, estancia media, índice de ocupación, índice de rotación, índice de mortalidad, urgencias, urgencias ingresadas, presión de urgencias o bien los referidos a quirófanos y consultas externas son indicadores de funcionamiento, en la actualidad bien recogidos, permiten una aproximación inicial. Sin embargo es necesario profundizar y para ello es preciso la identificación de los procesos atendidos.

Morbilidad y sistemas de clasificación de pacientes por grupos de diagnósticos relacionados

Las altas hospitalarias, de hospital día y cirugía mayor ambulatoria recogidas en el Conjunto Mínimo Básico de Datos (CMBD) nos permiten una información clave sobre los procesos atendidos cuantitativa y cualitativamente y una evaluación de las tendencias en frecuentación por problemas de salud cardiológicos de la población

atendida. La codificación de los procesos atendidos en cardiología, en el momento actual a través de la CIE-10, y su posterior agregación mediante la AP-GRDs (*All Patients related diagnoses groups*) en grupos homogéneos de isoconsumo de recursos (GRDs) permite un análisis específico de la morbilidad atendida. Efectivamente para cada GRD podemos establecer la estancia media, y calcular la estancia media ajustada por la casuística del estándar, así como la estancia media ajustada por el funcionamiento del estándar y determinar el índice de complejidad. Estos indicadores para cada GRD cardiológico nos permiten ya un análisis más riguroso y la posibilidad de establecer comparaciones y revisar las actuaciones y procedimientos realizados en cada proceso.

El cuadro de mando como herramienta para integrar y resumir las medidas de resultados

La medición y seguimiento de los resultados de una Unidad de Cardiología se plasma en el cuadro de mando integral. Este permite evaluar el cumplimiento de objetivos, identificar las áreas de mejora, comparar resultados con los mejores (*Benchmarking*), establecer medidas correctoras, etc.

Medidas de la calidad de vida de los resultados en salud en cardiología

El resultado real de las acciones médicas y de los servicios sanitarios es su impacto en la calidad y cantidad de vida y también en la modificación de actitudes y hábitos de salud de la población. Las medidas de resultados en calidad de vida han sido tempranamente aplicadas desde la década de los años noventa en la IC, uno de los primeros procesos donde se utilizaron, mediante instrumentos genéricos como la aplicación de cuestionarios validados el SF-36 (*Short Form 36 Health Survey*) y NHP (Perfil de Salud de Nottingham) para medir el impacto en la calidad de vida. También en IC se desarrollaron instrumentos específicamente diseñados para medir de forma sensible los efectos de la enfermedad en la calidad de vida del paciente como el *Minnesota Living with Heart Failure Questionnaire* (cuestionario de calidad de vida de IC de Minnesota). Estos cuestionarios se han ido extendiendo a otros procesos cardiológicos como por ejemplo a la angina de pecho con el APQLQ (*Angina Pectoris Quality of Life Questionnaire*).

Los sistemas de medida de la efectividad clínica

La aplicación de los QALY (años ganados ajustados por calidad de vida, AVACS en inglés) como instrumento de medida de la efectividad clínica de las diferentes alternativas diagnósticas y terapéuticas está siendo utilizada de forma habitual en los estudios de coste efectividad de los diferentes procesos diagnósticos y terapéuticos así por ejemplo la agencia NICE utiliza Qaly de forma sistemática para la realización de sus recomendaciones. En este campo cabe destacar en cardiología los estudios de coste efectividad en cirugía coronaria, intervencionismo coronario y en fármacos anti trombóticos por ejemplo.

■ 17.1.2. Evolución de los sistemas de medida de los resultados

A través de los trabajos de Donald Berwick (Berwick, 2017) conocemos como las prácticas clínicas inapropiadas suponen el 25-30 % de los presupuestos sanitarios. En este sentido se ha comunicado que determinados procesos clínicos concentran una proporción exagerada de prácticas inapropiadas. Así según señala el "Right Care Inform 1" (Brownlee *et al.*, 2017) el 30 % de las coronariorafías y el 22 % del ICP (intervencionismo coronario percutáneo) realizadas en Italia corresponden a procesos no indicados o justificados. En USA el 1.1 % del ICP primario no está indicado y el 11 % de los procedimientos en procesos no agudos tampoco estaban justificados con una variación interhospitalaria del 6-16.7 %. Por ello se han dirigido importantes líneas de trabajo para dotar a la práctica clínica y en concreto a la cardiología de indicadores de resultados que permitan realizar una asistencia de excelencia e identificar las oportunidades de mejora. En este campo son bien conocidos por todos los profesionales los trabajos del NICE (The National Institute for Health and are Excellence) británico. En nuestro medio se han realizado también diferentes iniciativas para analizar la adecuación de procedimientos como la MAPAC (mejora de la adecuación de la práctica asistencial y clínica), el compromiso por la Calidad de las sociedades científicas en España, la Agencia de Evaluación y Calidad Sanitaria de Cataluña, el programa de indicadores de calidad en cardiología (INCARDIO) y SEC-excelente de la Sociedad Española de Cardiología (SEC).

Es interesante aquí repasar el concepto de valor incorporado por Michael Porter (Porter, 2010) a la práctica clínica. Con este término señala la necesidad de establecer una metodología que permita calcular y determinar el valor añadido que los procedimientos clínicos aportan en cada proceso. De acuerdo con su planteamiento el valor se determina mediante una fracción en la cual los resultados de un proceso constituyen el numerador y los costes incurridos en cada proceso el denominador. En esta fórmula el elemento fundamental es el numerador, los resultados que hay que determinar desde nuestra perspectiva clínica como cardiólogos. Para su medición hay que definir cuáles son los resultados específicos para cada proceso mediante un análisis en el cual los clínicos desarrollan un trabajo de reflexión de acuerdo con el conocimiento clínico.

■ 17.1.3. Evolución actual e incorporación de la experiencia del paciente

Al enfoque tradicional basado en la medida de resultados clínicos se ha incorporado una visión de carácter integral donde se incluye la experiencia del paciente. Así, es imprescindible redefinir los procesos clínicos incorporando a los indicadores tradicionales aquellos que preocupan a los pacientes como recomienda el consorcio internacional para la medida de resultados de salud (ICHOM).

En nuestro medio por ejemplo, el hospital Vall de Hebrón ha desarrollado tres tipos de resultados: a) medida de los resultados clínicos: supervivencia, complica-

ciones, reingresos, etc...b) medida de los resultados de los que informa el paciente. Son básicamente aquellos datos que afectan a la calidad de vida del paciente y c) medidas de las experiencias que aporta el paciente. Se refiere aquí a la experiencia del paciente en el curso de la atención.

Como podemos observar este esquema de trabajo aporta dos elementos que no son habituales en la medida de resultados clínicos, ambos referidos al paciente.

■ 17.1.4. Objetivos del trabajo

Este capítulo tiene como objetivo definir en primer lugar, los indicadores de resultados generales en la práctica de la cardiología en la actualidad, en segundo lugar exponer un ejemplo concreto de mejoría de la calidad asistencial en la enfermedad cardiovascular: definición del proceso fibrilación auricular de excelencia.

■ 17.2. INDICADORES DE RESULTADOS EN CARDIOLOGÍA

A continuación revisaremos el proceso de modelización de resultados en cardiología. La mayoría de los datos de la literatura en este contexto se refieren a la práctica clínica hospitalaria y no de consultas externas, cuya actividad es difícil de codificar y medir con fiabilidad en el momento actual. Recientemente la SEC elaboró un informe con los indicadores de calidad de los cuidados cardiológicos. El informe elige tres categorías de indicadores: mortalidad, reingreso y volumen. En la categoría de mortalidad se incluyen: 1) mortalidad hospitalaria a los 30 días de la realización de un ICP, ajustada a riego; 2) Tasa de mortalidad hospitalaria en los 30 días siguientes a la realización de un bypass aorto-coronario (Ao-C); 3) Mortalidad hospitalaria a los 30 días tras cirugía de reemplazo valvular aórtico, y 4) Mortalidad a 30 días de la realización combinada de bypass Ao-C y reemplazo valvular aórtico. En la categoría de reingreso se definen: 1) Reingreso urgente por cualquier causa en los 30 días siguientes al alta hospitalaria de ICP; 2) Reingreso por cualquier causa a los 30 días al alta de un episodio de bypass Ao-C. En la categoría de volumen se incluye: el número total de ICP realizadas. Así mismo la ACC/AHA (American College of Cardiology/American Heart Association) ha revisado en 2016, 2017 y 2018 las medidas de desempeño y calidad del proceso fibrilación auricular (Heidenreich *et al.*, 2016) del Síndrome coronario Agudo (Jneid *et al.*, 2017) y de la rehabilitación cardiaca (Thomas *et al.*, 2018).

En nuestro entorno europeo disponemos también de una propuesta rigurosa de nuestro país que es la definida por la Sociedad Española de Cardiología que en noviembre de 2015 publicó un excelente trabajo en la *Revista Española de Cardiología* y poco después en la *Revista Europea de Cardiología*, coordinado y dirigido por Lopez-Sendón (Lopez-Sendón *et al.*, 2015) donde después de una rigurosa metodología de consenso de un grupo de trabajo en el que participaron 13 profesionales con

elevada experiencia en varios campos de la cardiología, se proponen los principales indicadores para medir la calidad de los resultados en la práctica clínica cardiológica hospitalaria y las medidas de rendimiento asociadas a los mejores resultados de práctica clínica, que constituyen los indicadores de práctica asistencial.

Los autores señalan como objetivo principal: la elaboración para el Servicio Nacional de Salud (SNS) español de un conjunto de indicadores para medir la calidad de los resultados cardiológicos. En su planteamiento se propusieron identificar los indicadores de resultado mejores y más representativos y en segundo término establecer los parámetros de medida. Para llevar a cabo estos objetivos, como señalamos anteriormente, se constituyó un grupo de trabajo llamado INCARDIO formado por expertos seleccionados por la SEC y la Sociedad de Cirugía Torácica y Cardiovascular (SECTCV).

En la metodología aplicada se tuvo en cuenta la gradación de las indicadores en base a cuatro aspectos:

- Relevancia clínica y práctica.
- Fuentes de información: origen y dificultad de la recogida uniforme de los datos.
- Facilidad para auditoría y verificación de los indicadores elegidos.
- Evidencia existente en la literatura de los indicadores propuestos.

Los autores consideraron también el tipo y características de los hospitales que clasifican en tres niveles de complejidad: baja, intermedia y alta, dependiendo de la presencia o no de cirugía cardíaca, electrofisiología, Unidad de Cardiología intervencionista, y Unidad de Cuidados Críticos cardiológicos.

■ 17.2.1. Indicadores de calidad

Finalmente, de acuerdo con el análisis y reflexión efectuado señalan los cinco principales indicadores por su alta relevancia para la medida de la calidad en la práctica clínica habitual :

- Mortalidad por cualquier causa. La relevancia de la mortalidad por cualquier causa es evidente en si misma y constituye el principal parámetro evaluable en los ensayos clínicos.
- Duración de la estancia hospitalaria y tasa de reingreso. Siendo variables claras de calidad no siempre pueden medirse de forma rigurosa como es deseable. Los reingresos se estiman en los 30 días siguientes al alta hospitalaria.

- IAM durante la hospitalización o después del alta. Se trata de una variable imprescindible de objetivo primario de los ensayos clínicos pero puede ser difícil de estimar en poblaciones amplias durante los primeros días de ingreso por la pobre estandarización del diagnóstico.
- Ictus. El ictus invalidante tiene una relevancia obvia y se relaciona con la iatrogenia del intervencionismo, la cirugía o el tratamiento antitrombótico. Sin embargo, las formas menores de ictus son a veces difíciles de diagnosticar y puede ser un parámetro confuso.
- Seguridad: hemorragia mayor, infecciones graves, taponamiento cardiaco post ICP, etc. No obstante, el diagnóstico uniforme y comparable es complejo de evaluar.

Estos indicadores deberían ser adoptados y utilizados por todos los hospitales de nuestro país y publicados anualmente para conocimiento de todos los profesionales. Si bien la premisa clave para la correcta evaluación y comparación de los resultados de calidad es la recogida universal y rigurosa del CMBD (Conjunto mínimo básico de datos) nacional.

Se incluye además en este trabajo, el ajuste de riesgo y se analiza la repercusión según la estrategia utilizada: edad y sexo, puntuaciones de riesgo específicas como la escala europea "EuroSCORE", la escala del grupo de estudio "TIMI" (*The Thrombolysis in Myocardial Infarction*), etc... Se suma también la evaluación comparativa con diferentes hospitales y las tendencias de la mortalidad en entidades específicas como por ejemplo la intervención de sustitución de válvula aórtica. La idea es obtener con el ajuste de riesgo poblaciones comparables homogéneas. El mejor método es comparar los GRDs de hospitales de semejante complejidad. Esto exige una correcta codificación de los diagnósticos de forma similar en todos los hospitales. En la Tabla 1 se señalan los principales diagnósticos para evaluar la calidad de los resultados de práctica clínica y los valores de referencia recomendados.

Tabla 1.

INDICADORES PARA EVALUAR LA CALIDAD GENERAL DE LOS RESULTADOS EN LA PRÁCTICA CLÍNICA CARDIOLÓGICA	
Indicador	Valor de referencia propuesto
Mortalidad^a	
Mortalidad por IAMCEST(excluidos los pacientes en Killip IV o que han recibido reanimación cardiovascular)	< 5 % (1)
Mortalidad por SCASEST(excluidos los pacientes en Killip IV o que han recibido reanimación cardiovascular)	< 3 % (1)

Tabla 1. (continuación)

INDICADORES PARA EVALUAR LA CALIDAD GENERAL DE LOS RESULTADOS EN LA PRÁCTICA CLÍNICA CARDIOLÓGICA

Indicador	Valor de referencia propuesto
Mortalidad^a	
Mortalidad por ICP en etapas	< 1 % (1)
Mortalidad por TAVI	< 6 % (1)
Mortalidad por TV después de IAM y otras intervenciones de ablación percutánea complejas	< 3 % (1)
Mortalidad por implante de marcapasos, DAI, TRC	< 1 % (1)
Mortalidad por insuficiencia cardíaca	< 7 % (1)
Mortalidad por primera intervención quirúrgica de SVA programada (excluido el TAVI)	< 5 % (1) < 7 % (2)
Mortalidad por primera intervención quirúrgica de sustitución de válvula mitral programada	< 7 % (1) < 9 % (2)
Mortalidad por primera intervención quirúrgica de reparación de válvula mitral programada	< 3 % (1) < 5 % (2)
Mortalidad por primer CABG programada (sin otra intervención quirúrgica combinada)	< 3 % (1) < 5 % (2)
Mortalidad por primer CABG + SVA combinado	< 6 % (1) < 8 % (2)
Trasplante cardíaco	< 15 % (1) (3)
Hospitalización^b	
Número de días de estancia en el hospital por IAMCEST	< 10
Número de días de estancia en el hospital por SCASEST	< 10
Número de días de estancia en el hospital por insuficiencia cardíaca	< 9
Número de días de estancia en el hospital por primera intervención quirúrgica aórtica, mitral o de CABG en etapas	< 15
Rehospitalización post SCA, insuficiencia cardíaca o cirugía previa ^c	Menor que el valor medio en los registros nacionales

Notas: CABG: cirugía de revascularización aortocoronaria; DAI: desfibrilador automático implantable; IAM: infarto agudo de miocardio; IAMCEST: infarto agudo de miocardio con elevación del segmento ST; SCA: síndrome coronario agudo; SCASEST: síndrome coronario agudo sin elevación del segmento ST; SVA: sustitución de válvula aórtica; TAVI: implante percutáneo de válvula aórtica; TRC: terapia de desincronización cardíaca; TV: taquicardia ventricular.

Los valores de referencia se presentan a modo de orientación. Para la evaluación comparativa, se recomienda decididamente un objetivo de valor de referencia menor que la mediana de los hospitales participantes.

a. Se prefiere la mortalidad por cualquier causa a 30 días a la mortalidad antes del alta hospitalaria solo si se puede obtener datos fiables (registros específicamente dedicados a este fin y auditables). (1): mortalidad observada (valor medio); (2): mortalidad esperada, corregida por el EuroSCORE logístico para esa población; (c): mortalidad o retrasplante.

b. Se prefiere el número de días en cualquier hospital durante los primeros 30 días siguientes a la hospitalización índice al número de días entre la hospitalización y el alta.

c. Reingreso no programado por cualquier causa en cualquier hospital de agudos en los 30 días siguientes al alta hospitalaria.

■ 17.2.2. Indicadores de práctica asistencial

Se refieren en la publicación a continuación los indicadores de calidad relacionados con mejores resultados en la práctica clínica (indicadores de práctica asistencial) y se definen como los indicadores de procesos con conocida influencia positiva en resultados deseables.

Se estructuran en bloques delimitados:

- Cardiología clínica.
- Diagnóstico por la imagen cardíaca.
- Cuidados críticos cardiológicos.
- Cardiología intervencionista
- Electrofisiología y arritmias complejas
- Insuficiencia cardíaca
- Rehabilitación cardíaca
- Cirugía cardíaca

La descripción detallada de todos los indicadores recogidos en los apartados referidos excede el propósito de este capítulo y definiremos a modo de ejemplo los referidos a la insuficiencia cardíaca, entidad que hoy día sigue aumentando su complejidad y precisa de medidas de excelencia para mejorar su pronóstico. Los indicadores que hay que evaluar incluyen la mortalidad y los reingresos hospitalarios y según los autores se exponen en la Tabla 2 las recomendaciones que se deben cumplir en todos los hospitales. La Sociedad Canadiense de Cardiología en su informe *Best practices for developing cardiovascular quality indicators* incluye los siguientes indicadores de calidad para la insuficiencia cardíaca definidos en forma de fracción (McKelvie *et al.*, 2016):

1. Electrolitos en sangre, urea y creatinina. Numerador: número de pacientes con medida diaria de Na, K, urea y creatinina. Denominador: número de pacientes con IC que reciben tratamiento intravenoso.
2. Radiografía de tórax. Numerador: número de pacientes con radiografía de tórax al ingreso o 24 horas después. Denominador: número de pacientes con IC aguda atendidos en urgencias o en hospitalización.
3. Utilización en el ingreso de un inhibidores de la enzima de conversión de la angiotensina (IECA) o un antagonistas del receptor de la angiotensina II (ARA II). Numerador: número de pacientes que reciben un IECA o ARA II. Denominador: número de pacientes hospitalizados con IC.

Tabla 2.

INDICADORES DE CALIDAD ASISTENCIAL EN INSUFICIENCIA CARDIACA RELACIONADOS CON MEJORES RESULTADOS EN LA PRÁCTICA CLÍNICA

Indicador	Recomendación
Volúmenes del hospital	Número de pacientes con insuficiencia cardiaca dados de alta del hospital.
Tecnología deseable	<ul style="list-style-type: none"> • Péptidos natriuréticos. • Hospitales de tipo II y III: ecocardiografía siempre disponible. Clínica ambulatoria multidisciplinaria de insuficiencia cardiaca Tratamiento de DAI y TRC. • Hospitales de tipo III: UC intensiva, dispositivos de asistencia circulatoria.
Dotación de personal	<ul style="list-style-type: none"> • Hospitales de tipo II y III: cardiólogos asignados al manejo de la insuficiencia cardiaca. • Hospitales de tipo III: cardiólogos con acreditación asignados a un programa de insuficiencia cardiaca avanzada. • Hospitales de tipo III: enfermeras especializadas asignadas al tratamiento de la insuficiencia cardiaca. • Consulta ambulatoria de enfermería.
Acreditación	<ul style="list-style-type: none"> • Hospitales de tipo III: programa multidisciplinario de insuficiencia cardiaca acreditado, inclusión de cardiólogos, especialistas en medicina interna, oncología, rehabilitación, médicos generales, otros. • Hospitales de tipo III: cardiólogo con acreditación en insuficiencia cardiaca avanzada.
Servicios prestados a los pacientes	<ul style="list-style-type: none"> • Hospitales de tipo III: clínica ambulatoria de insuficiencia cardiaca. • Hospitales de tipo III: insuficiencia cardiaca durante el programa de manejo intrahospitalario. • Todos los hospitales: disponibilidad en el centro o acceso a rehabilitación, unidades de insuficiencia cardiaca avanzada, trasplante cardiaco, unidades de hipertensión pulmonar compleja y unidades de cuidados paliativos.
Proceso Asistencial	
Duración de la estancia	Recomendado < 9 días; 8,6 días de media + 1 desviación estándar de los últimos 5 años en la base de datos nacional del sistema de salud de España.
Prestación de asistencia al alta	<p>Recomendaciones por escrito para el autotratamiento.</p> <p>Vía definida para el seguimiento en el momento del alta hospitalaria.</p> <p>Primera visita después del alta en < 2 semanas.</p>
Resultados	Resultados en poblaciones especiales como referidos en Tabla 1.

Tabla 2. (continuación)

INDICADORES DE CALIDAD ASISTENCIAL EN INSUFICIENCIA CARDIACA RELACIONADOS CON MEJORES RESULTADOS EN LA PRÁCTICA CLÍNICA

Indicador	Recomendación
Controles de calidad	
Adherencia a las recomendaciones de las guías de ESC/AHA/ACC	Instrucciones al alta. Recomendado: 100 %.
	Visita programada después del alta. Recomendado: 100 %.
	Evaluación de la función ventricular. Recomendado: 100 %.
	Consejo para dejar de fumar. Recomendado: 100 %.
	Uso de bloqueadores beta/(ivabradina si la FC es > 70 lpm), IECA o ARA-II, antagonistas de la aldosterona en pacientes con FEVI < 40 % y sin contraindicaciones para el alta hospitalaria. Recomendado > 90 %.
	Uso de DAI/TRC en clase I-A: recomendado menos que el valor medio en los registros nacionales.

Notas: ACC: American College of Cardiology; AHA: American Heart Association; ARAII: antagonistas del receptor de la angiotensina II; DAI: desfibrilador automático implantable; ECO: ecocardiografía; ESC: Sociedad Europea de Cardiología; FC: frecuencia cardiaca; FEVI: fracción de eyección del ventrículo izquierdo; IECA: inhibidores de la enzima de conversión de la angiotensina; SU: servicio de urgencias, TRC: terapia de resincronización cardiaca; UCIC: unidad de cuidados intensivos coronarios.

- Evaluación de la función ventricular izquierda (FEVI). Numerador: número de pacientes con valoración de la FEVI por cualquier técnica de imagen. Denominador: número de pacientes con IC.
- Tasa de reingreso a los 30 días. Numerador: número de pacientes con reingreso en 30 días. Denominador: número de pacientes con IC dados de alta.
- Educación de pacientes. Numerador: número de pacientes con IC que han sido formados en la enfermedad. Denominador: número de pacientes ingresados con IC.

La SEC en su programa de calidad dentro del apartado denominado SEC-excelente, recoge también los indicadores referidos al proceso insuficiencia cardiaca. Llama la atención que salvo por algunos indicadores que se repiten en las distintas publicaciones, ninguno de los registros son similares.

17.2.3. Datos de morbilidad atendida

Interesa por último señalar, que los datos específicos de morbilidad atendida hospitalaria clasificados a través de los GRD y analizados en función de su estancia

media, severidad y mortalidad son también una herramienta de extraordinaria utilidad en la medida de la calidad de los resultados. En ese sentido y como ejemplo exponemos la realidad práctica de un Servicio de Cardiología y los indicadores utilizados para evaluar la morbilidad atendida. Se exponen los resultados de los GRD del Servicio de Cardiología del Hospital Universitario La Paz correspondientes al año 2015.

Cardiología Hospitalización excluida la Unidad Coronaria

1. El peso medio de los GRD atendidos durante el año 2015 es de 2,52.
2. La estancia media (EM) es de 5,52 mejor que la del estándar en 2,35 puntos. El índice de EM ajustada (IEMA) es también mejor que el estándar.
3. Se han evitado 1.281 estancias que representan un ahorro de 232 ingresos en los GDR más eficientes, pero se han perdido 1.037 en los GRD menos eficientes. Por tanto hemos evitado 244 ingresos este año.

Todo ello indica que el funcionamiento del servicio, una vez realizados los ajustes de estancia media por la casuística y funcionamiento, es un 2,47 % mejor que el estándar.

Más del 80 % de la patología ingresada en cardiología hospitalización general, corresponde a los 14 GDR representados en la tabla superior.

<i>Principales GRD en cardiología hospitalización (excluida la Unidad Coronaria)</i>				
GRD	Altas	EM hospital	EM estándar	Ingresos evitados
112 Procedimientos cardiovasculares percutáneos sin IAM o IC.	288	1,71	2,47	-88
853 Procedimiento cardiovascular percutáneo con Stent liberador de fármaco con IAM o shock.	191	5,28	4,59	29
124 Trastorno circulatorio, excepto IAM, con cateterismo & diagnóstico complejo.	149	3,66	6,81	-69
125 Trastorno circulatorio excepto IAM, con cateterismo sin diagnóstico complejo.	141	2,16	3,48	-54
854 Procedimiento cardiovascular percutáneo con Stent liberador de fármaco sin IAM o shock.	123	3,32	3,47	-5
550 Otros procedimientos vasculares con complicación mayor.	121	12,26	11,35	10
127 IC & Shock.	112	7,69	6,75	15
143 Dolor torácico.	89	2,75	3,16	-12
808 Procedimientos cardiovasculares percutáneos con IAM, IC o shock.	58	5,41	4,71	9
138 Arritmia o trastorno conducción y complicación	55	4,95	4,8	2

Principales GRD en cardiología hospitalización (excluida la Unidad Coronaria) (continuación)				
GRD	Altas	EM	EM	Ingresos evitados
		hospital	estándar	
140 Angina de pecho.	55	3,13	4,64	-18
121 IAM complicado.	52	6,54	8,19	-10
544 IC y arritmia con complicación mayor.	48	12,29	9,51	14
543 Enfermedad cardiovascular, no IAM, no IC, no endocarditis, con complicación mayor.	40	7,3	10,1	-11

■ 17.3. MEDIDAS DE DESEMPEÑO E INDICADORES DE CALIDAD DEL PROCESO DE EXCELENCIA DE FIBRILACIÓN AURICULAR

Las medidas de desempeño e indicadores de calidad del proceso fibrilación auricular (FA), son herramientas que miden la calidad del manejo clínico de la FA identificando las oportunidades de mejora. Su aplicación por los profesionales implicados permite trasladar a la clínica con fiabilidad las evidencias científicas en este campo, evitando la variabilidad clínica intercentro e interprofesional (Heidenreich *et al.*, 2016). Además, los indicadores y medidas de desempeño permiten la comparación de los cuidados y resultados clínicos entre hospitales y son aplicables al continuo de la atención del proceso FA, desde la situación ambulatoria a la hospitalaria, desde la atención primaria a la especializada.

Los indicadores señalados en este capítulo son los adoptados por la Sociedad Española de Cardiología (SEC) dentro de su programa SEC-excelente y en concreto SEC excelente FA (Ruiz *et al.*, 2019). Se han incorporado en este proceso los indicadores que están en concordancia con las guías de práctica clínica de nuestro entorno y por tanto, preferentemente europeas (Hindricks *et al.*, 2020; Cowan *et al.*, 2014). Por ello se han considerado razones basadas en las evidencias científicas que explican su inclusión en las guías clínicas (Hindricks *et al.*, 2020; January *et al.*, 2014; Cox *et al.*, 2016). Se han tenido en cuenta también los indicadores ya establecidos por la ACC/AH (January *et al.*, 2014), la Sociedad Canadiense de Cardiología (Cox *et al.*, 2016), NICE AF quality care (Buthany *et al.*, 2015), publicaciones sobre indicadores en FA en nuestro país y posicionamiento terapéutico oficial de la SEC sobre prescripción de anticoagulación (Barón-Esquivias *et al.*, 2016; Roldán y Marín, 2016).

Los indicadores o estándares de evaluación del desempeño, calidad y resultados valoran diferentes parámetros de efectividad clínica, seguridad del paciente o coordinación y comunicación del proceso asistencial y serían los necesarios para acreditar la gestión del proceso FA como excelente en nuestro país. Dividimos los indicadores deseables en estándares de estructura, indicadores de desempeño del proceso, de calidad y de resultados:

■ 17.3.1. Estándares de estructura

No disponemos de referencias bibliográficas específicas, ni españolas ni europeas que nos indiquen con exactitud cuáles son los estándares de estructura que se deben adoptar. Disponemos de algunas indicaciones de las recientes guías europeas (Hindricks *et al.*, 2020; Cox *et al.*, 2016), sobre el cuidado integral y multidisciplinar del proceso FA que incluyen una estructura organizativa de cuidado y seguimiento del proceso con clase de indicación IIa y evidencia B, teniendo presente el papel central en el mismo, del paciente. La ACC/AHA, la Sociedad Canadiense y NICE (Heidenreich *et al.*, 2016; Cox *et al.*, 2016; y Buthany *et al.*, 2015), están de acuerdo en señalar la necesidad de los mismos pero sin coincidencias claras en la selección de los indicadores básicos. No obstante, los estándares de estructura son imprescindibles pues tienen como objetivo, facilitar la adhesión a las guías clínicas, y disminuir la hospitalización y mortalidad del proceso (January *et al.*, 2017) y como señala el proceso SEC excelente FA deben incluir (Martín-Ruiz *et al.*, 2019).

- Vía clínica, escrita y publicada electrónicamente, común y con acceso de todos los profesionales implicados. Debe estar basada en las guías de práctica clínica del entorno y revisable al menos cada año por un equipo multidisciplinar. Debe incluir: las rutas de acceso del paciente y el flujo interprofesional de los cuidados, cuándo enviar de forma electiva o urgente al cardiólogo, el establecimiento de realización de un ecocardiograma, vigilancia de la función renal, la decisión de anticoagulación y cómo, en base al riesgo trombótico y consideración del riesgo hemorrágico, el control de síntomas, factores precipitantes y comorbilidad, decisión de estrategia de control de ritmo o frecuencia, uso de la cardioversión, ablación, cierre de la orejuela o cirugía y actitud de manejo de la anticoagulación periprocedimiento y pericirugía. Además, debe contemplar la toma de decisiones terapéuticas en situaciones complejas por un *Heart-team* (Equipo del Corazón).
- Historia electrónica en todos los casos con ECG diagnóstico y analítica con función renal.
- Accesible a todos los niveles.
- Check-list en cada escalón asistencial con las tareas de desempeño.
- Accesibilidad electrónica a los INRs a todos los niveles.
- Programa de formación continuada profesional en el proceso.
- Formación del paciente experto en FA.
- Base de datos electrónica común, con un responsable de su seguimiento, que incluya los resultados clínicos para la evaluación y establecimiento de mejoras en el proceso.

■ 17.3.2. Indicadores de desempeño

Incluyen medidas bien soportadas por la evidencia científica y con clase de indicación alta en las guías de práctica clínica. Se comunican públicamente y en determinados escenarios sanitarios se adoptan para el pago por programas de coste por proceso o pago por programas de desempeño.

■ 17.3.3. Indicadores de calidad

Son parámetros utilizados para mejorar la calidad de los cuidados localmente. Como refieren las guías de la ACC (Heidenreich *et al.*, 2016), en ocasiones su implementación demuestra mejoras importantes clínicas y son incorporados a los indicadores de desempeño.

Indicadores de resultados

Miden de forma clara y objetiva el impacto clínico de las medidas de mejora adoptadas en el proceso.

Los autores del documento SEC-excelente FA seleccionan un listado amplio de indicadores deseables, pero tras realizar una revisión bibliográfica de los indicadores publicados relacionados con el proceso FA y considerar las indicaciones de clase I de las recientes guías de la Sociedad Europea de Cardiología sobre el manejo de la FA, validadas por un grupo de expertos de la Sociedad Española de Cardiología (Arribas *et al.*, 2017) y los inconvenientes actuales para la evaluación de todos ellos en nuestro entorno, se considera un subconjunto de indicadores de mayor relevancia como imprescindibles. Serían los necesarios para monitorizar el proceso de fibrilación auricular en los diferentes centros hospitalarios y extrahospitalarios de nuestro SNS. La Tabla 3 señala los siete indicadores básicos para el proceso FA, 3 de desempeño, 3 de calidad y uno de resultado. Son los que el comité de SEC excelente FA determinó utilizando el siguiente sistema de selección: 1) que sean derivados de recomendaciones de clase I de las últimas guías de la Sociedad Europea de Cardiología sobre el manejo de FA y aceptadas por el grupo de expertos de la SEC que revisó dichas guías; 2) que sean importantes en el manejo práctico de la FA; y 3) que sean evaluables y cuantificables en una historia clínica convencional y/o en los sistemas electrónicos de información de pacientes empleados en nuestro país. En la Tabla 4 se puede observar la correcta definición de los mismos.

Recientemente también se han publicado los indicadores clave del proceso FA estimados por consenso de un grupo de expertos en la profilaxis de la tromboembolia por FA, reunidos con el propósito de establecer un procesos de atención integrada para estos pacientes en la Comunidad de Madrid (Ortiz *et al.*, 2018). Se consensuaron los siguientes indicadores de proceso: 1. *Indicadores de estructura*: a) existencia de vía clínica común compartida entre atención especializada y primaria; b) existencia de documentación clínica con el diagnóstico de FA electrocar-

Tabla 3.

INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FIBRILACIÓN AURICULAR

Indicadores seleccionados

- Documentación y pruebas básicas en fibrilación auricular.
- Estratificación de riesgo tromboembólico en fibrilación auricular.
- Anticoagulación en pacientes de alto riesgo tromboembólico.
- Porcentaje de pacientes con identificación de la calidad deficiente de anticoagulación (TTR < 65 %).
- Porcentaje de pacientes con paso de dicumarínicos a ACOD cuando los INRs son lábiles.
- Ablación en pacientes sintomáticos refractarios a antiarrítmicos.
- Necesidad de atención en urgencias, hospitalización o muerte.

diográficamente, analítica con función renal y hepática, ecocardiograma, escalas de riesgo trombótico y hemorrágico, motivo de la indicación de anticoagulación y existencia clara del INR con el rango establecido si el anticoagulante elegido es el acenocumarol. 2. *Indicadores de proceso*: a) proporción de pacientes en tratamiento con anticoagulantes orales en los que consta en la historia clínica el balance riesgo/beneficio; b) proporción de pacientes con indicación de anticoagulación que lo están correctamente con indicación en historia clínica de porqué se ha indicado acenocumarol o anticoagulante directo. 3. *Indicadores de resultados*: a) número de pacientes anticoagulados que presentan complicaciones tromboembólicas severas o hemorragia grave que precisa hospitalización o transfusión sanguínea; b) número de controles de INR en pacientes tratados con antivitamina k en rango terapéutico. Los indicadores elegidos en este caso, son casi superponibles a los adoptados por la SEC referidos anteriormente y demuestran la necesidad de adoptar medidas que avalen la calidad de la atención al proceso FA en nuestro entorno inmediato y el conjunto de nuestro país.

Tabla 4.

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4. 1. Indicador de calidad "documentación de FA".

Nombre	Documentación y pruebas básicas en fibrilación y flutter auricular
Descripción	Porcentaje de pacientes que tienen una historia clínica una anamnesis centrada en el aparato cardiovascular, una analítica que incluya al menos hemograma, función renal y tiroidea, un ECG diagnóstico de fibrilación auricular y un ecocardiograma transtorácico.
Tipo de indicador	<i>Calidad.</i>
Numerador	Pacientes con fibrilación auricular que tienen en su historia clínica una analítica centrada en el aparato cardiovascular, con una analítica que incluya al menos un recuento de eritrocitos, hemoglobina sérica o hematocrito, creatinina sérica o aclaramiento de creatinina, y determinación de TSH y T4, un ECG diagnóstico de fibrilación auricular y un ecocardiograma transtorácico.
Denominador	Todos los pacientes fibrilación auricular.
Exclusiones del numerador y denominador (documentadas)	<p>Pacientes cuya fibrilación auricular responde a causas transitorias o corregidas.</p> <p>Pacientes que fallecen durante el estudio.</p> <p>Pacientes que son transferidos a otro centro durante el periodo de aplicación.</p> <p>Pacientes que solicitan el alta voluntaria del centro durante el periodo de aplicación.</p> <p>Pacientes que rechazan cualquier actitud diagnóstica y/o terapéutica.</p>
Excepciones	En centros cuya historia clínica electrónica no permita el almacenamiento de electrocardiogramas externos no será de aplicación el requisito de la presencia de un documento electrocardiográfico de la fibrilación auricular, a condición de que este electrocardiograma se encuentre adecuadamente informado y descrito en la historia clínica.
Momento de aplicación	La anamnesis se debe haber realizado en el momento de la primera evaluación del paciente en el centro/unidad. La analítica se debe haber obtenido antes de tres meses de la primera evaluación del paciente en el centro/unidad. El ecocardiograma transtorácico se deberá haber realizado antes de 6 meses de la primera evaluación del paciente en el centro/unidad. En pacientes que vienen diagnosticados ya de otros centros/unidades se admitirán analíticas y ecocardiogramas previos a condición de que tengan menos de 12 meses antigüedad y de queden adecuadamente descritos (cuantitativamente) en la historia clínica junto a la fecha de la determinación.

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.1. Indicador de calidad "documentación de FA" (continuación)

Nombre	Documentación y pruebas básicas en fibrilación y flutter auricular
Fuente de datos	Historia clínica.
Ámbito aplicación	Hospitalaria y extrahospitalaria.
Justificación	Los pacientes con FA deben tener una historia clínica completa, exploración rigurosa, valoración de las enfermedades concomitantes (clase de recomendación I, nivel de evidencia C), ECG diagnóstico de FA (clase de recomendación I, evidencia B) y ecocardiograma (clase de recomendación I, evidencia C).
Referencias	<p>ARRIBAS, F.; SEC WORKING GROUP FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: ARIBAS, F., ROLDÁN, I., LUIS MERINO, J., ROLDÁN, V., ARENAL, Y., TAMARGO, J., RUIZ-CRANELL, R., MONT, L., ANGUITA, M., MARIN, F.; REVIEWERS FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: BARRIOS ALONSO, V., BARÓN, G., ESQUIVIAS, COSÍN SALES, J., DÍAZ INFANTE, E., ESCOBAR CERVANTES, C., FERREIRO GUTIÉRREZ, J. L., GUERRA RAMOS, J. M., JIMÉNEZ CANDIL, F. J., PÉREZ CASTELLANO, N., SÁMBOLA AYALA, A., TELLO MONTOLIÚ, A.; SEC GUIDELINES COMMITTEE: SAN ROMÁN, A., ALFONSO, F., ARIBAS, F., EVANGELISTA, A., FERREIRA, I., JIMÉNEZ, M., MARIN, F., DE ISLA, L. P., RODRÍGUEZ PADIAL, L., SÁNCHEZ, P. L., SIONIS, A. Y VÁZQUEZ, R. (2017). Comments on the 2016 ESC Guidelines for the Management of Atrial Fibrillation. <i>Rev Esp Cardiol</i>, 70, pp. 2-8.</p> <p><i>Atrial Fibrillation/Flutter Quality Indicators Chapter. The Canadian Cardiovascular Society Quality Indicators E-Catalogue</i> FINAL – March 2019.</p> <p>COX, J. L., DAI, S., GONG, Y., MCKELVIE, R., MCMURTRY, M. S., OKAES, G. H., SKANES, A., VERMA, A., WILTON, S. B. y WYSE, D. G. (2016). The Development and Feasibility Assessment of Canadian Quality Indicators for Atrial Fibrillation. <i>Can J Cardiol</i>, 32, pp. 1566-1569.</p> <p>HART, R. G., EIKELBOOM, J. W., BRIMBLE, K. S., MCMURTRY, M. S. e INGRAM, A. J. (2013). Stroke prevention in atrial fibrillation patients with chronic kidney disease. <i>Can J Cardiol</i>, 29, pp. 571-78.</p> <p>DONAL, E., LIP, G. Y., GALDERISI, M., GOETTIE, A., SHAH, D., MARMAN, M., LEDERLIN, M., MONDILLO, S., EDVARDSEN, T., SITGES, M., GRAPSA, J., GARRI, M., SENIOR, R., GIMELLI, A., POTPARA, T. S., VAN GELDER, I. C., GORENEK, B., MABO, P., LANCELOTTI, P., KUCK, K. H., POPESCU, B. A., HINDRICKS, G., HABIB, G., COSYNS, B., DELGADO, V., HAUGAA, K. H., MURARU, D., NIEMAN, K. Y COHEN, A. (2016). EACVI/EHRA Expert Consensus Document on the role of multi-modality imaging for the evaluation of patients with atrial fibrillation. <i>EurHeart J Cardiovasc Imaging</i>, 17, pp. 355-383.</p>

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.1. Indicador de calidad "documentación de FA" (continuación)

Nombre	Documentación y pruebas básicas en fibrilación y flutter auricular
Referencias	<p>HEIDENREICH, P. A., SOLIS, P., ESTES, N. A. 3^{er}, FONAROW, G. C., JURGENS, C. Y., MARINE, J. E., McMANUS, D. D. y McNAMARA, R. L. (2016). 2016 ACC/AHA Clinical Performance and Quality Measures for Adults With Atrial Fibrillation or Atrial Flutter: A Report of the American College of Cardiology/American Heart Association Task Force on Performance Measures. <i>J Am CollCardiol</i>, 68, pp. 525-568.</p> <p>KIRCHHOF, P., AURICCHIO, A., BAX, J., CRUINS, H., CAMM, J., DIENER, H. C., GOETTE, A., HINDRICKS, G., HOHNLOSER, S., KAPPEMBERGER, L., KUICK, K. H., LIP, G. Y., OLSSON, B., MEINERTZ, T., PRIORI, S., RAVENS, U., STEINBECK, G., SVERNHAGE, E., TUISSEN, J., VINCENT, A. y BREITHARDT, G. (2007). Outcome parameters for trials in atrial fibrillation: executive summary: Recommendations from a consensus conference organized by the German Atrial Fibrillation Competence NETwork (AFNET) and the European Heart Rhythm Association (EHRA). <i>EurHeart J</i>, 28, pp. 2803-2817.</p> <p>KIRCHHOF, P., BENUSSI, S., KOTCHA, D., AHLSSON, A., ATAR, D., CASADEI, B., CASTELLA, M., DIENER, H. C., HEIDBUCHEL, H., HENDRIKS, J., HINDRICKS, G., MANOLIS, A. S., OLDGREN, J., POPESCU, B. A., SCHOTTEN, U., VAN PUTTE, B., VARDAS, P., AGEWALL, S., CAMM, J., BARON, ESQUIVAS, G., BUDTS, W., CARERI, S., CASSELLMAN, F., COCA, A., DE CATERINA, R., DEFETEROS, S., DOBREV, D., FERRO, J. M., FLIPPATOS, G., FITZSIMONS, D., GORENEK, B., GUENOUN, M., HOHNLOSER, S. H., KOUH, P., LIP, G. Y., MANOLIS, A., McMURRAY, J., PONIKOWSKI, P., ROSENHEK, R., RUSCHITZKA, F., SÁVELLEVA, I., SHARMA, S., SUWALSKI, P., TAMARGO, J. L., TAYLOR, C. J., VAN GELDER, I. C., VOORS, A. A., WINDECKER, S., ZAMORANO, J. L. y ZEPPENFELD, K. (2016). 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. <i>Europace</i>, 18, pp. 1609-1678.</p> <p>ORTIZA, M., MASJUAN, J., EGOCHEGA, M. I., MARTIN, A., SUÁREZ, C., ROLDÁN, I., SALGADO, R., MIRAL, J. J. y LLAMAS, P. (2018). <i>Diseño del proceso asistencial de profilaxis de la tromboembolia en pacientes con fibrilación auricular</i>. https://doi.org/10.1016/j.jhqr.2018.02.004.</p>

4.2. Indicador de desempeño "estratificación del riesgo tromboémico"

Descripción	Porcentaje de pacientes con fibrilación auricular que tienen en su historia clínica documentado su riesgo tromboembólico mediante la escala CHA ₂ DS ₂ -VASC.
Tipo de indicador	<i>Desempeño (performance)</i> .
Numerador	Pacientes con fibrilación auricular cuyo riesgo tromboembólico se ha estimado mediante la aplicación correcta de la escala CHA ₂ DS ₂ -VASC (edad, sexo, presencia de hipertensión, diabetes, insuficiencia cardiaca, enfermedad vascular coronaria o periférica, y antecedentes de ictus, accidentes isquémicos transitorios o embolismos sistémicos) y documentado en la historia clínica.
Denominador	Todos los pacientes con fibrilación auricular.

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.2. Indicador de desempeño “estratificación del riesgo tromboótico” (continuación)

Nombre	Documentación y pruebas básicas en fibrilación y flutter auricular
Exclusiones del numerador y denominador (documentadas)	Pacientes con estenosis mitral y prótesis valvulares mecánicas.
Excepciones	Ninguna.
Momento de aplicación	Primera evaluación médica.
Fuente de datos	Historia clínica.
Ámbito aplicación	Hospitalaria y extrahospitalaria.
Justificación	Se recomienda emplear la escala CHA ₂ DS ₂ -VASc para la estratificación de riesgo tromboembólico de los pacientes con fibrilación auricular (clase de recomendación I, nivel de evidencia A).
Referencias	<p>ARRIBAS, F.; SEC WORKING GROUP FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: ARIBAS, F., ROLDÁN, I., LUIS MERINO, J., ROLDÁN, V., ARENAL, Y., TAMARGO, J., RUIZ-GRANELL, R., MONT, L., ANGLITA, M., MARÍN, F.; REVIEWERS FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: BARRIOS ALONSO, V., BARÓN, G., ESQUIVIAS, COSIN SALES, J., DIAZ INFANTE, E., ESCOBAR CERVANTES, C., FERREIRO GUTIÉRREZ, J. L., GUERRA RAMOS, J. M., JIMÉNEZ CANDIL, F. J., PÉREZ CASTELLANO, N., SAMBOLA AYALA, A., TELLO MONTOLIÚ, A.; SEC GUIDELINES COMMITTEE: SAN ROMÁN, A., ALFONSO, F., ARIBAS, F., EVANGELISTA, A., FERREIRA, I., JIMÉNEZ, M., MARÍN, F., DE ISLA, L. P., RODRÍGUEZ PADIAL, L., SÁNCHEZ, P. L., SIONIS, A. Y VÁZQUEZ, R. (2017). Comments on the 2016 ESC Guidelines for the Management of Atrial Fibrillation. <i>Rev Esp Cardiol</i>, 70, pp. 2-8.</p> <p><i>Atrial Fibrillation/Flutter Quality Indicators Chapter</i>. The Canadian Cardiovascular Society Quality Indicators E-Catalogue FINAL – March 2019.</p> <p>Cox, J. L., Dai, S., Gong, Y., McKemie, R., McMurry, M. S., Oakes, G. H., Skanes, A., Verma, A., Wilton, S. B. y Wyse, D. G. (2016). The Development and Feasibility Assessment of Canadian Quality Indicators for Atrial Fibrillation. <i>Can J Cardiol</i>, 32, pp. 1566-1569.</p> <p>FRIBERG, L., ROSENQVIST, M. Y LIP, G. Y. (2012). Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish Atrial Fibrillation cohort study. <i>EurHeart J</i>, 33, pp. 1500–1510.</p> <p>HEIDENREICH, P. A., SOLIS, P., ESTES, N. A. 3RD, FOMAROW, G. C., JURGENS, C. Y., MARINE, J. E., McMANUS, D. D. Y McNAMARA, R. L. (2016). 2016 ACC/AHA Clinical Performance and Quality Measures for Adults With Atrial Fibrillation or Atrial Flutter: A Report of the American College of Cardiology/American Heart Association Task Force on Performance Measures. <i>J Am CollCardiol</i>, 68, pp. 525-568.</p>

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.2. Indicador de desempeño “estratificación del riesgo trombotico” (continuación)

Nombre	Documentación y pruebas básicas en fibrilación y flutter auricular
Referencias	<p>KIRCHHOF, P., BENUSSI, S., KOTECHEA, D., AHLSSON, A., AYAR, D., CASADEI, B., CASTELLA, M., DIENER, H. C., HEIDRUCHEL, H., HENDRIKS, J., HINDRICKS, G., MANOLIS, A. S., OLDOGREN, J., POPESCU, B. A., SCHOTTEN, U., VAN PUTTE, B., VARDAS, P., ACEWALL, S., CAMM, J., BARON ESQUIVIAS, G., BUDITS, W., CARERI, S., CASSELLMAN, F., COCA, A., DE CATERINA, R., DEFERRIOS, S., DOBREV, D., FERRO, J. M., FILIPPATOS, G., FITZSIMONS, D., GORENEK, B., GUENOUN, M., HOHNLOSER, S. H., KOHL, P., LIP, G. Y., MANOLIS, A., McMURRAY, J., PONIKOWSKI, P., ROSENHEK, R., RUSCHITZKA, F., SAVELIEVA, I., SHARMA, S., SUWALSKI, P., TAMARGO, J. L., TAYLOR, C. J., VAN GELDER, I. C., VOORS, A. A., WINDECKER, S., ZAMORANO, J. L. Y ZEPPENFELD, K. (2016). 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. <i>Europace</i>, 18, pp. 1609-1678.</p> <p>LIP, G. Y., NIEUWLAAT, R., PISTERS, R., LANE, D. A. Y CRUINS, H. J. (2010). Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. <i>Chest</i>, 137, pp. 263–272.</p> <p>OLESEN, J. B., LIP, G. Y., HANSEN, M. L., HANSEN, P. R., TOLSTRUP, J. S., LINDHARDSEN, J., SELMER, C., AHLEHOFF, O., OLSEN, A. M., GISLASON, G. H. Y TORP-PEDERSEN, C. (2011). Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study. <i>BMJ</i>, 342, p. d1124.</p> <p>ORTIZA, M., MASJUAN, J., EGOCHAGA, M. I., MARTIN, A., SUÁREZ, C., ROLDÁN, I., SALGADO, R., MIRAL, J. J. Y LLANAS, P. (2018). <i>Diseño del proceso asistencial de profilaxis de la tromboembolia en pacientes con fibrilación auricular</i>. https://doi.org/10.1016/j.jhq.2018.02.004.</p>

4.3. Indicador de desempeño “anticoagulación indicada”

Nombre	Anticoagulación en pacientes de alto riesgo tromboembólico
Descripción	<p>Porcentaje de pacientes con fibrilación auricular que tienen un alto riesgo tromboembólico (varones con $CHA_2DS_2-VASc \geq 2$ y mujeres con $CHA_2DS_2-VASc \geq 3$) a los que se les ha prescrito anticoagulantes orales.</p>
Tipo de indicador	<p><i>Desempeño (performance)</i>.</p>
Numerador	<p>Pacientes con fibrilación auricular que tienen un alto riesgo tromboembólico (varones con $CHA_2DS_2-VASc \geq 2$ y mujeres con $CHA_2DS_2-VASc \geq 3$) a los que se les ha prescrito anticoagulantes orales (acenocumarol, warfarina, dabigatán, rivaroxabán, apixabán o edoxabán).</p>

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.3. Indicador de desempeño “anticoagulación indicada” (continuación)

Nombre	Anticoagulación en pacientes de alto riesgo tromboembólico
Denominador	Todos los pacientes con fibrilación que tienen un alto riesgo tromboembólico (varones con $CHA_2DS_2-VASc \geq 2$ y mujeres con $CHA_2DS_2-VASc \geq 3$).
Exclusiones del numerador y denominador (documentadas)	<p>Pacientes cuya fibrilación responde a causas transitorias, evitables o corregidas.</p> <p>Pacientes con estenosis mitral y prótesis valvulares mecánicas.</p> <p>Pacientes que fallecen antes del momento de aplicación.</p> <p>Pacientes que son transferidos a otro centro antes del momento de aplicación.</p> <p>Pacientes que solicitan el alta voluntaria del centro antes del momento de aplicación.</p> <p>Pacientes que rechazan la anticoagulación a pesar de la recomendación médica.</p>
Excepciones	<p>Pacientes en los que se ha considerado la anticoagulación y finalmente no se ha recomendado en base a algún motivo suficientemente justificado, razonado y documentado en la historia clínica.</p> <p>Pacientes anticoagulados con heparinas por un motivo razonado.</p> <p>Pacientes portadores de un dispositivo de cierre de orejuela.</p>
Momento de aplicación	Al alta hospitalaria o en la primera consulta médica.
Fuente de datos	Historia clínica.
Ámbito aplicación	Hospitalaria y extrahospitalaria.
Justificación	Los pacientes varones con $CHA_2DS_2-VASc \geq 2$ y mujeres con $CHA_2DS_2-VASc \geq 3$ que han presentado algún episodio de fibrilación auricular debieran estar anticoagulados salvo que la taquiarritmia se haya debido a una causa transitoria, evitable, o corregida, o exista una contraindicación para la anticoagulación (clase de recomendación I, nivel de evidencia A). El motivo para no anticoagular deberá quedar razonado en la historia clínica.

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.3. Indicador de desempeño "anticoagulación indicada" (continuación)

Nombre	Anticoagulación en pacientes de alto riesgo tromboembólico
Referencias	<p>ARRIBAS, F.; SEC WORKING GROUP FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: ARIBAS, F., ROLDÁN, I., LUIS MERINO, J., ROLDÁN, V., ARENAL, Y., TAMARGO, J., RUIZ-GRANELL, R., MONT, L., ANGUITA, M., MARÍN, F.; REVIEWERS FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: BARRIOS ALONSO, V., BARÓN, G., ESQUIVIAS, COSÍN SALES, J., DÍAZ INFANTE, E., ESCOBAR CERVANTES, C., FERREIRO GUTIÉRREZ, J. L., GUERRA RAMOS, J. M., JIMÉNEZ CANDIL, F. J., PÉREZ CASTELLANO, N., SAMBOLA AYALA, A., TELLO MONTOLIÚ, A.; SEC GUIDELINES COMMITTEE: SAN ROMÁN, A., ALFONSO, F., ARIBAS, F., EVANGELISTA, A., FERREIRA, I., JIMÉNEZ, M., MARÍN, F., DE ISLA, L. P., RODRÍGUEZ PADIAL, L., SANCHEZ, P. L., SIONIS, A. Y VÁZQUEZ, R. (2017). Comments on the 2016 ESC Guidelines for the Management of Atrial Fibrillation. <i>Rev Esp Cardiol</i>, 70, pp. 2-8.</p> <p><i>Atrial Fibrillation/Flutter Quality Indicators Chapter. The Canadian Cardiovascular Society Quality Indicators E-Catalogue FINAL</i> – March 2019.</p> <p>COX, J. L., DAI, S., GONG, Y., MCKELVIE, R., MCMURTRY, M. S., OAKES, G. H., SKANES, A., VERMA, A., WILTON, S. B. y WYSE, D. G. (2016). The Development and Feasibility Assessment of Canadian Quality Indicators for Atrial Fibrillation. <i>Can J Cardiol</i>, 32, pp. 1566-1569.</p> <p>FRIBERG, L., ROSENQVIST, M. y LIP, G. Y. (2012). Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish Atrial Fibrillation cohort study. <i>EurHeart J</i>, 33, pp. 1500–1510.</p> <p>HEIDENREICH, P. A., SOLIS, P., ESTES, N. A. 3th, FONAROW, G. C., JURGENS, C. Y., MARINE, J. E., McMANUS, D. D. y McNAMARA, R. L. (2016). 2016 ACC/AHA Clinical Performance and Quality Measures for Adults With Atrial Fibrillation or Atrial Flutter: A Report of the American College of Cardiology/American Heart Association Task Force on Performance Measures. <i>J Am CollCardiol</i>, 68, pp. 525-568.</p> <p>KIRCHHOFF, P., BENUSSI, S., KOTECHE, D., AHLSSON, A., ATAR, D., CASADEI, B., CASTELLA, M., DIENER, H. C., HEIDRICH, H., HENDRIKS, J., HINDRICKS, G., MANOLIS, A. S., OLDGREN, J., POPESCU, B. A., SCHOTTEN, U., VAN PUTTE, B., VARDAS, P., ACEVALL, S., CAMM, J., BARON ESQUIVIAS, G., BUDIS, W., CAREI, S., CASSELMAN, F., COCA, A., DE CATERINA, R., DEFTEREOS, S., DOBREV, D., FERRO, J. M., FILIPPATOS, G., FITZSIMONS, D., GORENEK, B., GUENOUN, M., HOHNLOSER, S. H., KOHL, P., LIP, G. Y., MANOLIS, A., McMURRAY, J., PONIKOWSKI, P., ROSENHEK, R., RUSCHITZKA, F., SÁVELLEVA, I., SHARMA, S., SUWALSKI, P., TAMARGO, J. L., TAYLOR, C. J., VAN GELDER, I. C., VOORS, A. A., WINDECKER, S., ZAMORANO, J. L. y ZEPPENFELD, K. (2016). 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. <i>Europace</i>, 18, pp. 1609-1678.</p> <p>LIP, G. Y., NIEUWLAAAT, R., PIETERS, R., LANE, D. A. y CRIJNS, H. J. (2010). Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. <i>Chest</i>, 137, pp. 263–272.</p>

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.3. Indicador de desempeño "anticoagulación indicada" (continuación)

Nombre	Anticoagulación en pacientes de alto riesgo tromboembólico
Referencias	<p>OLESEN, J. B., LIP, G. Y., HANSEN, M. L., HANSEN, P. R., TOLSTRUP, J. S., LINDHARDSEN, J., SELMER, C., AHLEHOFF, O., OLSEN, A. M., GISLASON, G. H. y TORP-PEDERSEN, C. (2011). Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study. <i>BMJ</i>, 342, p. 124.</p> <p>ORTIZ, M., MASJUAN, J., EGOCHAGA, M. I., MARTIN, A., SUÁREZ, C., ROLDÁN, I., SALGADO, R., MIRAL, J. J. y LLAMAS, P. (2018). <i>Diseño del proceso asistencial de profilaxis de la tromboembolia en pacientes con fibrilación auricular</i>. https://doi.org/10.1016/j.jhqr.2018.02.004.</p>

4.4. Indicador de desempeño "calidad de la anticoagulación"

Nombre	Mala calidad de la anticoagulación con AVK
Descripción	Porcentaje de pacientes con FA con identificación deficiente de la calidad de la anticoagulación con AVK, considerada como la presencia de un TTR <65 % medido por el método de Rosendaal.
Tipo de indicador	<i>Desempeño (performance)</i> .
Numerador	Pacientes con fibrilación auricular tratados con AVK durante >3 meses que tienen un TTR <65 %.
Denominador	Todos los pacientes con fibrilación auricular tratados con AVK durante > 3 meses.
Exclusiones del numerador y denominador (documentadas)	<p>Pacientes que fallecen antes del momento de aplicación.</p> <p>Pacientes que son transferidos a otro área antes del momento de aplicación.</p> <p>Pacientes que rechazan la anticoagulación a pesar de la recomendación médica.</p> <p>Pacientes en los que se ha considerado la anticoagulación y finalmente no se ha recomendado en base a algún motivo suficientemente justificado, razonado y documentado en la historia clínica.</p> <p>Pacientes anticoagulados con heparinas por un motivo razonado.</p> <p>Pacientes portadores de un dispositivo de cierre de orejuela.</p>
Momento de aplicación	En consulta sucesiva tras 3 meses de tratamiento con AVK.
Fuente de datos	Historia clínica.
Ámbito aplicación	Hospitalaria en ingresos sucesivos al diagnóstico inicial y extrahospitalaria.

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.4. Indicador de desempeño "calidad de la anticoagulación" (continuación)

Nombre	Anticoagulación en pacientes de alto riesgo tromboembólico
Justificación	<p>Los pacientes con FA anticoagulados con AVK deben tener los INRs en el rango terapéutico adecuado establecido y la calidad de la anticoagulación, medida por el TTR por el método Rosendaal, el más correcto, debe ser >65 % para evitar eventos tromboembólicos y mejorar la supervivencia. Ciras de TTR<65 % se asocian a mal pronóstico con aumento de eventos trombóticos y hemorrágicos y mortalidad.</p>
Referencias	<p>ARRIBAS, F.; SEC WORKING GROUP FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: ARIBAS, F., ROLDÁN, I., LUIS MERINO, J., ROLDÁN, V., ARENAL, Y., TAMARGO, J., RUIZ-GRANELL, R., MONT, L., ANGUITA, M., MARIN, F.; REVIEWERS FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: BARRIOS ALONSO, V., BARÓN, G., ESQUIVIAS, COSÍN SALES, J., DÍAZ INFANTE, E., ESCOBAR CERVANTES, C., FERREIRO GUTIÉRREZ, J. L., GUERRA RAMOS, J. M., JIMÉNEZ CANDIL, F. J., PÉREZ CASTELLANO, N., SÁMBOLA AYALA, A., TELLO MONTOLIÚ, A.; SEC GUIDELINES COMMITTEE: SAN ROMÁN, A., ALFONSO, F., ARIBAS, F., EVANGELISTA, A., FERREIRA, I., JIMÉNEZ, M., MARÍN, F., DE ISLA, L. P., RODRÍGUEZ PADIAL, L., SÁNCHEZ, P. L., SIONIS, A. y VÁZQUEZ, R. (2017). Comments on the 2016 ESC Guidelines for the Management of Atrial Fibrillation. <i>Rev Esp Cardiol</i>, 70, pp. 2-8.</p> <p><i>Atrial Fibrillation/Flutter Quality Indicators Chapter. The Canadian Cardiovascular Society Quality Indicators E-Catalogue FINAL</i> – March 2019.</p> <p>BONDE, A. N., STÆRK, L., LEE, C. J. Y. <i>et al.</i> (2018). Outcomes Among Patients With Atrial Fibrillation and Appropriate Anticoagulation Control. <i>J Am Coll Cardiol</i>, 18, pp. 1357-1365.</p> <p>COX, J. L., DAI, S., GONG, Y., MCKELVIE, R., MCMURTRY, M. S., OAKES, G. H., SKANES, A., VERMA, A., WILTON, S. B. y WYSE, D. G. (2016). The Development and Feasibility Assessment of Canadian Quality Indicators for Atrial Fibrillation. <i>Can J Cardiol</i>, 32, pp. 1566-1569.</p> <p>JANUARY, C. T., WANIN, L. S., CALKINS, H., FIELD, M. E., CHEN, L. Y., FURIE, K. L., CIGARROA, J. E., HEIDENREICH, P. A., CLEVELAND, JR. J. C., MURRAY, K. T., ELLINOR, P. T., SHEA, J. B., EZEKOWITZ, M. D., TRACY, C. M. y YANG, C. W. (2019). 2019 AHA/ACC/HRS Focused Update of the 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation. <i>Heart Rhythm</i>, doi:https://doi.org/10.1016/j.hrthm.2019.01.024.</p> <p>KIRCHHOF, P., BENUSSI, S., KOTECHA, D., AHLSSON, A., ATAR, D., CASADEI, B., CASTELLA, M., DIENER, H. C., HEIDBUCHEL, H., HENDRIKS, J., HINDRICKS, G., MANOLIS, A. S., OLDFREN, J., POPESCU, B. A., SCHOTTEN, U., VAN PUTTE, B., VARDAS, P., ACEWALL, S., CAMM, J., BARON ESQUIVIAS, G., BUDIS, W., CARERI, S., CASSELMAN, F., COCA, A., DE CATERINA, R., DEFETEROS, S., DOBREV, D., FERRO, J. M., FLIPPATOS, G., FITZSIMONS, D., GORENEK, B., GUENOUN, M., HOHNLOSER, S. H., KOHL, P., LIP, G. Y., MANOLIS, A., MCMURRAY, J., PONIKOWSKI, P., ROSENHEK, R., RUSCHITZKA, F., SÁVELLEVA, I., SHARMA, S., SUWALSKI, P., TAMARGO, J. L., TAYLOR, C. J., VAN GELDER, I. C., VOORS, A. A., WINDECKER, S., ZAMORANO, J. L. y ZEPPENFELD, K. (2016). 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. <i>Europace</i>, 18, pp. 1609-1678.</p>

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.4. Indicador de desempeño "calidad de la anticoagulación" (continuación)

Nombre	Anticoagulación en pacientes de alto riesgo tromboembólico
Referencias	<p>ORITZA, M., MASJUAN, J., EGOCHEGA, M. I., MARTÍN, A., SUÁREZ, C., ROLDÁN, I., SALGADO, R., MIRAL, J. J. Y LLANAS, P. (2018). <i>Diseño del proceso asistencial de proflaxis de la tromboembolia en pacientes con fibrilación auricular</i>. https://doi.org/10.1016/j.jhq.2018.02.004.</p> <p>RIVERA-CARAVACA, J. M., ROLDÁN, V., ESTEVE-PASTOR, M. A. et al. (2018). Reduced Time in Therapeutic Range and Higher Mortality in Atrial Fibrillation Patients Taking Acenocoumarol. <i>Clin Ther</i>, 40, pp. 114-122.</p>

4.5. Indicador calidad "cambio de AVK a ACOD directo con INRs lábiles"

Nombre	Cambio de AVK a ACOD si INRS lábiles
Descripción	Porcentaje de pacientes con FA y tratamiento con AVC con mala calidad de anticoagulación medida por TTR <65 % (INRs lábiles), a los que se les prescribe paso de dicumarínicos a ACOD.
Tipo de indicador	Calidad
Numerador	Pacientes con fibrilación auricular tratados con AVK y mal control de la anticoagulación (TTR<65 %) medida a los tres meses que pasan a recibir un ACOD en lugar del AVK
Denominador	Todos los pacientes con fibrilación auricular tratados con AVK y mala calidad de la anticoagulación medida a los 3 meses con TTR<65 %.
Exclusiones del numerador y denominador (documentadas)	<p>Pacientes que fallecen antes del momento de aplicación.</p> <p>Pacientes que son transferidos a otra área antes del momento de aplicación.</p> <p>Pacientes que rechazan la anticoagulación a pesar de la recomendación médica.</p>
Excepciones	Pacientes con clara contraindicación para ACOD.
Momento de aplicación	Ingreso hospitalario o en consulta tras >3 meses de tratamiento con AVK.
Fuente de datos	Historia clínica.
Ámbito aplicación	Hospitalaria y extrahospitalaria.

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.5. Indicador calidad "cambio de AVK a ACOD directo con INRs lábiles" (continuación)

Nombre Anticoagulación en pacientes de alto riesgo tromboembólico

Justificación Los pacientes con FA tratados con AVK y mala calidad de la anticoagulación medida a los tres meses del tratamiento tienen mal pronóstico que solo mejora con el cambio preceptivo a ACOD si no existe contraindicación absoluta (Clase IIb A).

ARRIBAS, F.; SEC WORKING GROUP FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: ARIBAS, F., ROLDÁN, I., LUIS MERINO, J., ROLDÁN, V., ARENAL, Y., TAMARGO, J., RUIZ-GRANELL, R., MONT, L., ANGUITA, M., MARÍN, F.; REVIEWERS FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: BARRIOS ALONSO, V., BARÓN, G., ESQUIVIAS, COSÍN SALES, J., DÍAZ INFANTE, E., ESCOBAR CERVANTES, C., FERREIRO GUTÉRREZ, J. L., GUERRA RAMOS, J. M., JIMÉNEZ CANDIL, F. J., PÉREZ CASTELLANO, N., SAMBOLA AYALA, A., TELLO MONTOLIU, A.; SEC GUIDELINES COMMITTEE: SAN ROMÁN, A., ALFONSO, F., ARIBAS, F., EVANGELISTA, A., FERREIRA, I., JIMÉNEZ, M., MARÍN, F., DE ÍSLA, L. P., RODRÍGUEZ PADIAL, L., SÁNCHEZ, P. L., SIONIS, A. Y VÁZQUEZ, R. (2017). Comments on the 2016 ESC Guidelines for the Management of Atrial Fibrillation. *Rev Esp Cardiol*, 70, pp. 2-8.

BONDE, A. N., STAERK, L., LEE, C. J. Y. *et al.* (2018). Outcomes Among Patients With Atrial Fibrillation and Appropriate Anticoagulation Control. *J Am Coll Cardiol*, 18, pp. 1357-1365.

JANUARY, C. T., WANN, L. S., CALKINS, H., FIELD, M. E., CHEN, L. Y., FURIE, K. L., CIGARROA, J. E., HEIDENREICH, P. A., CLEVELAND, JR. J. C., MURRAY, K. T., ELLINOR, P. T., SHEA, J. B., EZEKOWITZ, M. D., TRACY, C. M. Y YANG, C. W. (2019). 2019 AHA/ACC/HRS Focused Update of the 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation. *Heart Rhythm*, doi:https://doi.org/10.1016/j.hrthm.2019.01.024.

KIRCHHOF, P., BENJUSI, S., KOTCHA, D., AHLSSON, A., ATAR, D., CASADEI, B., CASTELLA, M., DIENER, H. C., HEIDBUCHEL, H., HENDRIKS, J., HINDRICKS, G., MANOLIS, A. S., OLDGREN, J., POPESCU, B. A., SCHOTTEN, U., VAN PUTTE, B., VARDAS, P., ACEWALL, S., CAMM, J., BARON ESQUIVIAS, G., BUDITS, W., CARERI, S., CASSELMAN, F., COCA, A., DE CATERINA, R., DEFTEREOS, S., DOBREV, D., FERRO, J. M., FILIPPATOS, G., FITZSIMONS, D., GORENEK, B., GUENOUN, M., HOHNLOSER, S. H., KOHL, P., LIP, G. Y., MANOLIS, A., McMURRAY, J., PONIKOWSKI, P., ROSENHEK, R., RUSCHITZKA, F., SÁVELLEVA, I., SHARMA, S., SUWALSKI, P., TAMARGO, J. L., TAYLOR, C. J., VAN GELDER, I. C., VOORS, A. A., WINDECKER, S., ZAMORANO, J. L. y ZEPPENFELD, K. (2016). 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Europace*, 18, pp. 1609-1678.

RIVERA-CARAVACA, J. M., ROLDÁN, V., ESTEVE-PASTOR, M. A. *et al.* (2018). Reduced Time in Therapeutic Range and Higher Mortality in Atrial Fibrillation Patients Taking Acenocoumarol. *Clin Ther*, 40, pp. 114-122.

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.6. Indicador calidad "ablación en pacientes refractarios a antiarrítmicos"

Nombre	Ablación en pacientes sintomáticos refractarios a antiarrítmicos
Descripción	Porcentaje de pacientes con fibrilación auricular paroxística sintomática recurrente a pesar de fármacos antiarrítmicos que son remitidos para ablación en un centro con experiencia en este procedimiento (recomendación clase I, nivel de evidencia A).
Tipo de indicador	<i>Calidad</i>
Numerador	Pacientes con fibrilación auricular paroxística sintomática recurrente a pesar de dosis adecuadas de fármacos antiarrítmicos (200 mg/día de flecaínida, 450 mg/día de propafenona, 200 mg/día de amiodarona, 240 mg/día de sotalol, 800 mg/día dronedarona) que son remitidos para ablación en un centro con experiencia en este procedimiento.
Denominador	Todos los pacientes con fibrilación auricular paroxística sintomática recurrente a pesar de dosis adecuadas de fármacos antiarrítmicos (200 mg/día de flecaínida, 450 mg/día de propafenona, 200 mg/día de amiodarona, 240 mg/día de sotalol, 800 mg/día dronedarona).
Exclusiones del numerador y denominador (documentadas)	<p>Pacientes cuya fibrilación responde a causas transitorias, evitables o corregidas.</p> <p>Pacientes que fallecen antes del momento de ablación.</p> <p>Pacientes que son transferidos a otro centro antes del momento de aplicación.</p> <p>Pacientes que solicitan el alta voluntaria del centro antes del momento de aplicación.</p> <p>Pacientes que rechazan la ablación a pesar de la recomendación médica.</p>
Excepciones	<p>Pacientes en los que se ha considerado la ablación y finalmente no se ha recomendado en base a algún motivo suficientemente justificado, razonado y documentado en la historia clínica.</p> <p>Pacientes con ablación previa de fibrilación auricular.</p> <p>Pacientes con corta esperanza de vida, comorbilidades significativas o contraindicaciones para la ablación.</p>
Momento de aplicación	En ausencia de factores precipitantes evitables, corregibles o tratables, se considera que una fibrilación auricular paroxística es refractaria a antiarrítmicos cuando se producen dos recurrencias sintomáticas en un periodo de seis meses. Desde este momento el paciente deberá ser referido a ablación de fibrilación auricular en un centro con experiencia en un plazo de seis meses.

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.6. Indicador calidad “ablación en pacientes refractarios a antiarrítmicos” (continuación)

Nombre	Ablación en pacientes sintomáticos refractarios a antiarrítmicos
Fuente de datos	Historia clínica.
Ámbito aplicación	Hospitalaria y extrahospitalaria.
Justificación	Los pacientes con FA paroxística sintomática refractaria a fármacos antiarrítmicos deben ser remitidos para ablación de FA en un centro con experiencia en este procedimiento (clase de recomendación I, nivel de evidencia A).
Referencias	<p>ARBELO, E., BRUGADA, J., HINDRICKS, G., MAGGIONI, A. P., TAVAZZI, L., VARDAS, P., LAROCHE, C., ANSELME, F., INAMA, G., JAIS, P., KALARUS, Z., KAUTZNER, J., LEWALTER, T., MAIRESSE, G. H., PEREZ-VILLACASTIN, J., RIAHI, S., TABORSKY, M., THEODORAKIS, G., TRINES, S. A. Y ATRIAL FIBRILLATION ABLATION PILOT STUDY INVESTIGATORS (2014). The atrial fibrillation ablation pilot study: a European Survey on Methodology and results of catheter ablation for atrial fibrillation conducted by the European Heart Rhythm Association. <i>EurHeart Jm</i>, 35, pp. 1466–1478.</p> <p>ARRIBAS, F.; SEC WORKING GROUP FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: ARIBAS, F., ROLDÁN, I., LUIS MERINO, J., ROLDÁN, V., ARENAL, Y., TAMARGO, J., RUIZ-GRANELL, R., MONT, L., ANGLITA, M., MARIN, F.; REVIEWERS FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: BARRIOS ALONSO, V., BARÓN, G., ESQUINAS, COSÍN SALES, J., DJAZ INFANTE, E., ESCOBAR CERVANTES, C., FERREIRO GUTIÉRREZ, J. L., GUERRA RAMOS, J. M., JIMÉNEZ CANDIL, F. J., PÉREZ CASTELLANO, N., SAMBOLA AYALA, A., TELLO MONTOLIÚ, A.; SEC GUIDELINES COMMITTEE: SAN ROMÁN, A., ALFONSO, F., ARIBAS, F., EVANGELISTA, A., FERREIRA, I., JIMÉNEZ, M., MARIN, F., DE ISLA, L. P., RODRIGUEZ PADIAL, L., SÁNCHEZ, P. L., SIONIS, A. Y VÁZQUEZ, R. (2017). Comments on the 2016 ESC Guidelines for the Management of Atrial Fibrillation. <i>Rev Esp Cardiol</i>, 70, pp. 2-8.</p> <p>CALKINS, H., REYNOLDS, M. R., SPECTOR, P., SONDHI, M., XU, Y., MARTIN, A., WILLIAMS, C. J. Y SLEDGE, I. (2009). Treatment of atrial fibrillation with antiarrhythmic drugs or radiofrequency ablation: two systematic literature reviews and meta-analyses. <i>Circ Arrhythm Electrophysiol</i>, 2, pp. 349–361.</p> <p>CAPPATO, R., CALKINS, H., CHEN, S. A., DAVIES, W., IESAKA, Y., KALMAN, J., KIM, Y. H., KLEIN, G., NATALE, A., PACKER, D., SKANES, A., AMBROGI, F. Y BIGANZOLI, E. (2010). Updated worldwide survey on the methods, efficacy, and safety of catheter ablation for human atrial fibrillation. <i>Circ Arrhythm Electrophysiol</i>, 3, pp. 32–38.</p> <p>COSEDIS, N. J., JOHANNESSEN, A., RAATIKAINEN, P., HINDRICKS, G., WALFRIDSSON, H., KONGSTAD, O., PEHRSON, S., ENGLUND, A., HARTIKAINEN, J., MORTENSEN, L. S. Y HANSEN, P. S. (2012). Radiofrequency ablation as initial therapy in paroxysmal atrial fibrillation. <i>N Engl J Med</i>, 367, pp.1587–1595.</p>

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.6. Indicador calidad “ablación en pacientes refractarios a antiarrítmicos” (continuación)

Nombre	Ablación en pacientes sintomáticos refractarios a antiarrítmicos
Referencias	<p>KIRCHHOF, P., BENUSSI, S., KOTÉCHA, D., AHLSSON, A., AVIAR, D., CASADEI, B., CASTELLA, M., DIENER, H. C., HEIDBUCHHEL, H., HENDRIKS, J., HINDRICKS, G., MANOLIS, A. S., OUDGREN, J., POPESCU, B. A., SCHOTTEN, U., VAN PUTTE, B., VARDAS, P., AGEWALL, S., CAMM, J., BARON ESQUIVIAS, G., BUDITS, W., CARERI, S., CASSELMAN, F., COCA, A., DE CATERINA, R., DEFETEROS, S., DOBREV, D., FERRO, J. M., FILIPPATOS, G., FITZSIMONS, D., GORENEK, B., GUENOUN, M., HOHNLOSER, S. H., KOH, P., LIP, G. Y., MANOLIS, A., McMURRAY, J., PONIKOWSKI, P., ROSENHEK, R., RUSCHITZKA, F., SÁVELIEVA, I., SHARMA, S., SUWALSKI, P., TAMARGO, J. L., TAYLOR, C. J., VAN GELDER, I. C., VOORS, A. A., WINDECKER, S., ZAMORANO, J. L. y ZEPPENFELD, K. (2016). 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. <i>Europace</i>, 18, pp. 1609-1678.</p> <p>WILBER, D. J., PAPPONE, C., NEUZIL, P., DE PAOLA, A., MARCHLINSKI, F., NATALE, A., MACLE, L., DAQUID, E. G., CALKINS, H., HALL, B., REDDY, V., AUGELLO, G., REYNOLDS, M. R., VINEKAR, C., LIU, C. Y., BERRY, S. M. y BERRY, D. A. (2010). ThermoCool AF Trial Investigators. Comparison of antiarrhythmic drug therapy and radiofrequency catheter ablation in patients with paroxysmal atrial fibrillation: a randomized controlled trial. <i>JAMA</i>, 303, pp. 333–340.</p>

4.7. Indicador de resultado “atención urgente, hospitalización o muerte”

Nombre	Necesidad de atención en urgencias, hospitalización o muerte
Descripción	Porcentaje de pacientes que acuden a urgencias, son ingresados o fallecen por un evento relacionado con fibrilación auricular.
Tipo de indicador	<i>Resultador.</i>
Numerador	Pacientes que acuden a urgencias o son ingresados en cualquier centro hospitalario, o fallecen, por un evento relacionado con fibrilación auricular en un año natural.
Denominador	Pacientes con fibrilación auricular evaluados en el centro en el año objeto de estudio y en el año anterior.
Exclusiones del numerador y denominador (documentadas)	<p>Pacientes cuya fibrilación responde a causas transitorias, evitables o corregidas.</p> <p>Pacientes que son transferidos a otro centro antes del comienzo del momento de aplicación.</p> <p>Pacientes que fallecen por causas no relacionadas con fibrilación auricular antes del comienzo del momento de aplicación.</p>
Excepciones (del numerador y denominador)	Pacientes con información insuficiente o dudosa en los sistemas de información

Tabla 4. (continuación)

DEFINICIÓN DE LOS INDICADORES SELECCIONADOS POR EL COMITÉ SEC-EXCELENTE-FA

4.7. Indicador de resultado "atención urgente, hospitalización o muerte" (continuación)

Nombre	Ablación en pacientes sintomáticos refractarios a antiarrítmicos
Momento de aplicación	Año natural (enero a diciembre).
Fuente de datos	Sistemas electrónicos de información de pacientes.
Ámbito aplicación	Hospitalaria y extrahospitalaria.
Justificación	<p>Dados los sistemas de información disponibles actualmente y la planificación sanitaria en España, se considera que la combinación de atención en urgencias u hospitalización puede guardar el mejor equilibrio entre la capacidad de evaluar el resultado la atención médica en fibrilación auricular y la capacidad de medición. La selección de otros índices propuestos en documentos de otros países como la tasa anual de ictus o la tasa anual de hemorragias está sujeta a limitaciones importantes de información que, unido a la baja tasa esperada de estos eventos, restan viabilidad a estos indicadores.</p>
Referencias	<p><i>Atrial Fibrillation/Flutter Quality Indicators Chapter</i>. The Canadian Cardiovascular Society Quality Indicators E-Catalogue FINAL – March 2019.</p> <p>COX, J. L., DAI, S., GONG, Y., MCKELVIE, R., MCMURTRY, M. S., OAKES, G. H., SKANES, A., VERMA, A., WILTON, S. B. y WYSE, D. G. (2016). The Development and Feasibility Assessment of Canadian Quality Indicators for Atrial Fibrillation. <i>Can J Cardiol</i>, 32, pp. 1566-1569.</p> <p>HEIDENREICH, P. A., SOLIS, P., ESTES, N. A. 3^{er}, FONAROW, G. C., JURGENS, C. Y., MARINE, J. E., McMANUS, D. D. y McNAMARA, R. L. (2016). 2016 ACC/AHA Clinical Performance and Quality Measures for Adults With Atrial Fibrillation or Atrial Flutter: A Report of the American College of Cardiology/American Heart Association Task Force on Performance Measures. <i>J Am Coll Cardiol</i>, 68, pp. 525-568.</p> <p>ORTIZ, M., MASUANI, J., ECOCHEGA, M. I., MARTÍN, A., SUÁREZ, C., ROLDÁN, I., SALGADO, R., MIRAL, J. J. y LLANAS, P. (2018). <i>Diseño del proceso asistencial de profilaxis de la tromboembolia en pacientes con fibrilación auricular</i>. https://doi.org/10.1016/j.jhqr.2018.02.004.</p>

■ 17.4. CONCLUSIONES

- Para alcanzar la excelencia en la práctica clínica de la cardiología es imprescindible la medida de los resultados.
- La SEC ha definido para el ámbito nacional indicadores básicos de resultados de práctica clínica que permiten la comparación de hospitales: mortalidad por cualquier causa. Duración de la estancia hospitalaria y tasa de reingreso. IAM durante la hospitalización o después del alta. Ictus y Hemorragia mayor posintervencionismo.
- La premisa clave para la correcta evaluación y comparación de los resultados de calidad es la recogida universal y rigurosa del CMBD nacional.
- No tenemos indicadores cualitativos relevantes de consultas externas, pues no disponemos de datos fiables codificados en el CMBD de los procesos atendidos en este ámbito.
- Cada servicio de cardiología debe estimar y comunicar anualmente sus resultados asistenciales cuantitativos y cualitativos, mediante el análisis de la morbi-mortalidad referida por los GRD atendidos y compararlos con los de hospitales del mismo grupo.
- Disponemos de indicadores de excelencia de varios procesos cardiológicos claves como la IC o la FA. Los referidos a este último proceso se han descrito con detalle en este capítulo y corresponden a tres de desempeño, tres de calidad y uno de resultado.

■ REFERENCIAS

- ARRIBAS, F.; SEC WORKING GROUP FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: ARIBAS, F., ROLDÁN, I., LUIS MERINO, J., ROLDÁN, V., ARENAL, A., TAMARGO, J., RUIZ-GRANELL, R., MONT, L., ANGUITA, M., MARÍN, F.; REVIEWERS FOR THE 2016 ESC GUIDELINES FOR THE MANAGEMENT OF ATRIAL FIBRILLATION: BARRIOS ALONSO, V., BARÓN, G., ESQUIVIAS, COSÍN SALES, J., DÍAZ INFANTE, E., ESCOBAR CERVANTES, C., FERREIRO GUTIÉRREZ, J. L., GUERRA RAMOS, J. M., JIMÉNEZ CANDIL, F. J, PÉREZ CASTELLANO, N., SAMBOLA AYALA, A., TELLO MONTOLIU, A.; SEC GUIDELINES COMMITTEE: SAN ROMÁN, A., ALFONSO, F., ARIBAS, F., EVANGELISTA, A., FERREIRA, I., JIMÉNEZ, M., MARÍN, F., DE ISLA, L. P., RODRÍGUEZ PADIAL, L., SÁNCHEZ, P. L., SIONIS, A. y VÁZQUEZ, R. (2017). Comments on the 2016 ESC Guidelines for the Management of Atrial Fibrillation. *Rev Esp Cardiol*, 70, pp. 2-8.
- BARÓN-ESQUIVIAS, G., GÓMEZ, S., BRUFAU, H., GARCÍA, L., AMO, C., GUTIÉRREZ, J. M. *et al.* (2016). Indicadores asistenciales en pacientes con fibrilación auricular: evaluación del manejo de los problemas clínicos y de las diferencias por sexo. *Rev Esp Cardiol*, 69(4), pp. 384–391.
- BERWICK, D. M. (2017). Avoiding overuse—the next quality frontier. *The Lancet*, 390, pp. 102-104.
- BROWNLEE, S., CHALKIDOU, K., DOUST, J. *et al.* (2017). Evidence for overuse of medical services around the world. *Right Care 1. The Lancet*, 390, pp. 156-168.
- BUTHANY, G., BOSTOCK, J., BROMLEY, H., CHOWHAN, H., DE LA MOTTE, A., DUNN, P. *et al.* (2015). *Atrial Fibrillation. Quality Standard*. Disponible en: Nice.org.uk/guidance/qs93

- COWAN, C., CAMPBELL, J., CHEONG, V. L., CHUNG, G., FAY, M., FITZMAURICE, D. *et al.* (2014). Atrial fibrillation: Management. *NICE Clinical guideline*. Disponible en: nice.org.uk/guidance/cg180
- COX, J. L., DAI, S., GONG, Y., MCKELVIE, R., MCMURTRY, M. S., OAKES, G. H. *et al.* (2016). The Development and Feasibility Assessment of Canadian Quality Indicators for Atrial Fibrillation. *Canadian Journal of Cardiology*, 2016: pp. 1-4.
- HEIDENREICH, P. A., SOLIS, P., EATEA, M. N. A. III *et al.* (2016). 2016 ACC/AHA Clinical Performance and Quality Measures for Adults with Atrial Fibrillation or Atrial Flutter. *JACC*, 68(5), pp. 525-568. doi: 10.1016/j.jacc.2016.03.521.
- HINDRICKS, G., POTPARA, T., DAGRES, N., *et al.* (2020). 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). *European Heart Journal*, 00, pp. 1-125. doi:10.1093/eurheartj/ehaa612.
- ICHOM. Why measure outcomes? Disponible en: www.ichom.org
- JANUARY, C. T., WANN, L. S., ALPERT, J. S., CALKINS, H., CIGARROA, J. E., CLEVELAND, J. C. *et al.* (2014). 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol*, 64(21), pp. 2246–2280.
- JNEID, H., ADDISON, D., BHATT, D. L. *et al.* (2017). 2017 AHA/ACC Clinical Performance and Quality Measures for Adults with ST-Elevation and Non-ST-Elevation Myocardial Infarction. *JACC*. doi: 10.1016/j.jacc.2017.06.032.
- LÓPEZ-SENDÓN, J., GONZÁLEZ-JUANATEY, J. R., PINTO, F. *et al.* (2015). Indicadores de calidad en cardiología. Principales indicadores para medir la calidad de los resultados (indicadores de resultados) y parámetros de calidad relacionados con mejores resultados en la práctica clínica (indicadores de práctica asistencial). INCARDIO (Indicadores de Calidad en Unidades Asistenciales del Área del Corazón): Declaración de posicionamiento de consenso de SEC/SECTCV. *Rev Esp Cardiol*, 68(11), pp. 976–995.
- MCKELVIE, R. S. y HECKMAN, G. A. (2016). Blais C et al. Canadian Cardiovascular Society Quality Indicators For Heart Failure. *CJC*, 32, pp. 1023, e5-1038.69.
- ORTIZ, M., MASJUÁN, J., EGOCHEGA, M. I., MARTÍN, A., SUÁREZ, C., ROLDÁN, I. *et al.* (2018). Diseño del proceso asistencial de profilaxis de la tromboembolia en pacientes con fibrilación auricular. *J. Healthc. Qual. Res.* Disponible en: <https://doi.org/10.1016/j.jhqr.2018.02.004>
- PORTER, M. (2010). What is value in Health Care. *N England J Med*, 363(26), pp. 2477-2481.
- Proceso Fibrilación Auricular SEC-excelente: www.secardiologia.es.
- ROLDÁN, I. y MARÍN, F. (2016), en representación del Grupo de Trombosis Cardiovascular de la Sociedad Española de Cardiología y el Comité elaborador de la propuesta. En el camino de un mejor uso de los anticoagulantes en la fibrilación auricular no valvular. Propuesta de modificación del posicionamiento terapéutico UT/V4/23122013. *Rev Esp Cardiol*, 69(6), pp. 551–553.
- THOMAS, R. J., BALADY, G., BANKA, G. *et al.* (2018). ACC/AHA Clinical Performance and Quality Measures for Adults for Cardiac Rehabilitation. *JACC*, 71(16). doi: 10.1016/j.jacc.2018.01.004.



18

RESULTADOS EN SALUD A EVALUAR EN ENFERMEDADES RARAS

MARÍA TERESA CABALLERO MOLINA, MARÍA PILAR AIBAR ABAD,
MARÍA REYES ABAD SAZATORNIL

■ 18.1. CONCEPTO DE ENFERMEDAD RARA Y MEDICAMENTO HUÉRFANO

■ 18.1.1. Enfermedad rara

El concepto de enfermedad rara (ER) está basado en criterios de prevalencia y gravedad. Las enfermedades raras (EE. RR.) o poco frecuentes son enfermedades de carácter crónico y discapacitante, que presentan una prevalencia tan baja que requieren esfuerzos especiales dirigidos a prevenir la morbilidad, la mortalidad temprana o la reducción de la calidad de vida de las personas que las padecen y sus familias (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2013). En la Unión Europea (UE) se considera que una enfermedad es rara cuando afecta a menos de una persona de cada 2.000 habitantes (Parlamento Europeo, 2000).

Las EE. RR. presentan una amplia variedad de trastornos, signos y síntomas clínicos y producen un gran impacto sanitario. Se estima que un 80 % de las EE. RR. tienen una etiología genética y la mayoría son incurables. Aproximadamente el 75 % afecta a los niños y de ellos el 30 % de los casos no viven a los 5 años. Se han identificado entre 5.000 y 8.000 EE. RR. y se estima que cada semana se describen cinco nuevas EE. RR. en la literatura médica. Aunque cada una de ellas afecta a pocos pacientes, entre el 6 % y el 8 % de la población mundial está afectado por una ER, lo que supone 30 millones en la UE y otros 30 millones de personas en Estados Unidos (EE. UU.) (Parlamento Europeo, 2000; Benjamin *et al.*, 2017).

■ 18.1.2. Medicamento huérfano

El Reglamento (CE) nº 141/2000 del Parlamento Europeo y del Consejo, de 16 de Diciembre de 1999, define medicamento huérfano (MH) (Parlamento Europeo, 2000), como aquel que:

- se destina al diagnóstico, prevención o tratamiento de una afección que ponga en peligro la vida o conlleve una incapacidad crónica y que no afecte a más de cinco personas por cada diez mil habitantes en la UE o,
- se destina al diagnóstico, prevención o tratamiento de una afección que ponga en peligro la vida o conlleve una incapacidad crónica y que, sin incen-

tivos, la comercialización de dicho medicamento en la UE no generaría suficientes ingresos para justificar la inversión necesaria, y

- que no existe ningún método satisfactorio alternativo o, que de existir, el tratamiento propuesto aportaría un beneficio considerable.

En 1983 en EE. UU., y en 2000 en la UE se aprobaron sendas regulaciones que incluían una serie de incentivos regulatorios para potenciar el desarrollo de los medicamentos huérfanos (MM. HH.). Concretamente en Europa el Reglamento (CE) nº 141/2000 estableció que la designación de un medicamento como MH conlleva los siguientes incentivos:

- Asistencia de la European Medicine Agency (EMA) en la elaboración de protocolos de investigación.
- Procedimiento centralizado de autorización.
- Reducción de tasas para las solicitudes.
- Exclusividad en el mercado durante 10 años (en el caso de que el medicamento se comercialice). Solamente se aprobará otro MH similar para la misma indicación si el titular de la autorización de comercialización del MH inicial ha dado su consentimiento al segundo solicitante o no puede suministrar suficiente cantidad del medicamento o bien si el segundo solicitante puede demostrar, que su medicamento, aunque similar al MH ya autorizado, es más seguro, más eficaz o clínicamente superior en otros aspectos. Este periodo de exclusividad se alarga dos años más si con el medicamento se ha realizado un programa de investigación pediátrica acordado con la EMA, independientemente de que, al final, se obtenga o no la indicación en niños.
- Facilidad para acceder a financiación pública para la investigación.
- Proporciona además incentivos adicionales para pequeñas y medianas empresas.

Estos incentivos han resultado positivos puesto que han favorecido la autorización de MM. HH. en la UE, pasando de 3 en 2001 a 28 en 2018.

Los pacientes con EE. RR. presentan una serie de necesidades clínicas no resueltas relacionadas con el acceso al cuidado clínico y a la información (Benjamin *et al.*, 2017; EURORDIS, 2009). Frecuentemente se producen retrasos o errores en el diagnóstico por falta de acceso a especialistas con conocimiento de la enfermedad. Incluso una vez diagnosticados correctamente se producen dificultades en el acceso a fármacos específicos y a otras intervenciones terapéuticas, bien por la inexistencia de fármacos o por su acceso restringido debido, entre otras causas, a su elevado coste. Los pacientes y sus familias también se quejan del desconocimiento de su enfermedad y de la limitación que ello supone para su acceso a las opciones de cuidado sanitario o al cuidado informado. Algunas de estas enfermedades incapacitan al paciente para realizar las actividades de la vida diaria, con la con-

siguiente dependencia de la familia o los cuidadores y la disminución de la calidad de vida relacionada con la salud (CVRS). Todo esto tiene un efecto importante sobre la familia y los cuidadores. Así, las familias indican preocupaciones psicosociales, sentimientos de aislamiento y depresión.

■ 18.2. DIFICULTADES ASOCIADAS A LOS ENSAYOS CLÍNICOS EN ENFERMEDADES RARAS

Generalmente, el desarrollo de un medicamento es un proceso largo, complejo, arriesgado y costoso. En el caso de los MM. HH., las dificultades inherentes al desarrollo de fármacos se ven incrementadas por el escaso conocimiento de la fisiopatología de la enfermedad, la falta de modelos preclínicos específicos, el infradiagnóstico, el reducido número de expertos clínicos y la dificultad de realizar ensayos clínicos por el reducido tamaño muestral o la falta de variables clínicas validadas, lo que eleva tanto la inversión necesaria para desarrollar un nuevo tratamiento como la incertidumbre asociada a su eficacia. Además, las empresas deberán recuperar la inversión realizada en I+D+i con ventas a un número muy reducido de pacientes antes de que expire la correspondiente patente.

En relación a los ensayos de fármacos para EE. RR. el problema más obvio es el reclutamiento de pacientes, adecuados y en número suficiente, por lo que a menudo se requiere colaboración multicéntrica y multinacional (Sun, Zheng y Simeonov, 2017; Logviss, Krievins y Purvina, 2018). La aprobación de un fármaco, generalmente se basa en un ensayo en fase III, doble ciego, aleatorizado y controlado, considerado con frecuencia el estándar de referencia, sin embargo puede ser particularmente difícil utilizar este tipo de ensayos en MM. HH. Así un estudio puso de manifiesto que los ensayos en fase III fueron la base para la aprobación por la Food and Drug Administration (FDA) en el 45 % de los MM. HH. frente al 73 % de los no huérfanos, en indicaciones oncológicas (Richey *et al.*, 2009).

Un estudio reciente analizó los ensayos clínicos de MM. HH. realizados en Letonia, registrados en la base de datos EU Clinical Trials Register (www.clinicaltrialsregister.eu) en el periodo 2004-2016 y los comparó con los ensayos para medicamentos no huérfanos. En el mismo, se puso de manifiesto que los ensayos clínicos de medicamentos para EE. RR. tienen menos probabilidades de ser ensayos controlados aleatorizados (62,7 % vs. 83,3 %; $p = 0,008$), en relación con el enmascaramiento, tenían menos probabilidades de ser doble ciego (45,1 % vs. 63,7 %; $p = 0,035$), también utilizaron un comparador activo con menos frecuencia (36,4 % vs. 58,8 %) e incluyeron menos participantes (media 18,3 % vs. 40,2 % sujetos) (Logviss, Krievins y Purvina, 2018).

Algunos autores revisaron la evidencia de los 64 MM. HH. autorizados por la EMA en 2012 para 78 indicaciones en EE. RR. mediante el análisis de los 117 estudios principales de dichos fármacos. De estos estudios, menos del 20 % inclu-

ieron variables robustas, sólo el 64,8 % fueron aleatorizados y el 68,5 % presentó un brazo control. La mitad fueron ciegos e incluyeron estudios de búsqueda de dosis de manera previa al ensayo pivotal (55 %), una minoría (26,9 %) incluyeron variables de calidad de vida. El tamaño muestral fue insuficiente en más del 50 % de los estudios analizados. La mayoría de los MM. HH. aprobados por la EMA demostraron una calidad de la evidencia moderada e incluso algunos habían recibido la aprobación de comercialización en base sólo a un ensayo en fase II no controlado, un estudio retrospectivo o la revisión de literatura. La duración de los ensayos de MM. HH. fue a menudo demasiado corta en relación con la historia natural de la enfermedad (Picavet *et al.*, 2013).

En un análisis de los MM. HH. comercializados en España en el periodo 2010-2015 (Núñez *et al.*, 2017) se observó que en el 49,9 % de los ensayos clínicos el comparador era placebo, mientras que el 27 % no disponían de comparador activo. El tamaño muestral más frecuente fue de 50 a 249 pacientes (41,6 %), siendo el tamaño muestral medio de 260 pacientes. El 12,5 % de los ensayos tuvieron menos de 50 pacientes.

Preocupa de manera importante que la muestra objeto de estudio no refleje adecuadamente la población a la cual va dirigido el fármaco, seleccionando los casos con síntomas atípicos (menos graves) y excluyendo mujeres, niños o pacientes más graves lo que puede originar sesgo en la interpretación de los resultados.

Por estas razones los datos clínicos aportados por las compañías farmacéuticas en la solicitud de aprobación de los MM. HH. pueden ser menos robustos que en el caso de medicamentos para enfermedades comunes, lo que conduce a la autorización condicional y bajo circunstancias excepcionales. La aprobación es condicional cuando los datos no están todavía completos y la compañía está obligada a la realización de estudios adicionales. Esta autorización se debe renovar cada año hasta que se hayan completado los estudios y, entonces pasaría a tener una autorización normal. Este tipo de autorización únicamente se concede para necesidades médicas no cubiertas con el objetivo de proporcionar un acceso temprano al medicamento. La aprobación bajo circunstancias excepcionales se da cuando no es posible suministrar datos completos sobre la eficacia y seguridad del medicamento para el que se solicita la autorización. Suele estar motivado por la rareza de la enfermedad a la que está destinado, el limitado conocimiento científico en el área de que se trate o por consideraciones éticas involucradas en la recopilación de esos datos. La información se revisa anualmente para reevaluar el balance beneficio-riesgo (Hernández, 2016).

De los 110 MM. HH. que tenían autorizada su comercialización en la Unión Europea en febrero de 2019, 13 presentan una aprobación condicional (11,8 %) y 14 (12,7 %) bajo circunstancias excepcionales. Este tipo de autorizaciones se han convertido en una herramienta relevante para acelerar el acceso a la "innovación terapéutica" según la EMA.

■ 18.3. RESULTADOS A MEDIR EN ENFERMEDADES RARAS

■ 18.3.1. Biomarcadores

Los biomarcadores o marcadores biológicos son aquellas sustancias que permiten medir de forma fiable un estado de salud o una enfermedad (se pueden usar para detectar una enfermedad, un cambio fisiológico, una respuesta al tratamiento o un trastorno psíquico).

Son una herramienta emergente en el desarrollo clínico de los fármacos (Sun, Zheng y Simeonov, 2017), y se distinguen cuatro tipos: subrogado, farmacodinámico, predictivo y pronóstico.

Los objetivos principales del uso de biomarcadores son la mejora de los procesos de desarrollo de fármacos, dado que la finalidad de los ensayos clínicos es medir las respuestas de los pacientes a un tratamiento. Si no es posible medir la respuesta directamente, los biomarcadores pueden ser un método alternativo para medir un resultado y la personalización del tratamiento.

La investigación en el campo de los biomarcadores está contribuyendo a mejorar la predicción del riesgo de padecer una enfermedad, la evolución de una enfermedad una vez diagnosticada y la respuesta de una persona a un tratamiento farmacológico. Esto permitirá tomar decisiones más seguras y eficaces por lo que respecta al tratamiento.

La identificación de los biomarcadores adecuados puede mejorar el desarrollo de fármacos, reduciendo el tiempo y el tamaño muestral en los ensayos clínicos. Al usarse como variable de resultado subrogada los biomarcadores pueden utilizarse para predecir resultados en salud y ayudar en la toma de decisiones regulatorias.

En situaciones donde la determinación de variables clínicas no es posible, tales como enfermedades muy lentamente progresivas o con daño irreversible en el momento del inicio de los síntomas o en enfermedades con eventos muy poco frecuentes, los cambios en un biomarcador pueden ser una alternativa a una variable clínica.

Los biomarcadores son a menudo más baratos, más fáciles y más rápidos de medir y con frecuencia pueden ser sorprendentemente más aceptados por pacientes y médicos que los resultados clínicos. Los biomarcadores no miden directamente el beneficio clínico, pero pueden ser medidas analíticas, pruebas de imagen y resultados clínicos intermedios.

En enfermedades graves o que pongan en peligro la vida, la aprobación acelerada o condicional de un medicamento puede basarse en un biomarcador con un menor grado de evidencia siempre que los reguladores consideren que predice

de manera razonable el beneficio de los pacientes, lo que puede conducir a que el medicamento esté disponible antes de que se complete el estudio pivotal que incluya una variable clínica.

Pero se ha de tener cuidado en situaciones donde la fisiopatología de la enfermedad, el mecanismo de acción de la intervención o las vías que involucran el biomarcador no están completamente dilucidadas, pues su uso puede llevar a conclusiones falsas.

Un cambio de un biomarcador en el contexto de una enfermedad compleja no se traduce necesariamente en beneficio clínico (Cox, 2018). Cuando un biomarcador está en la vía causal de la enfermedad y se valida para un contexto de uso en una enfermedad, sus cambios posteriores se pueden utilizar para predecir el beneficio clínico después del tratamiento pero nosi no lo está.

Un biomarcador podrá ser utilizado cuando:

- Exista una clara comprensión de la fisiopatología de la enfermedad y el biomarcador esté involucrado en al menos una de las principales vías fisiopatológicas.
- Exista certeza en el mecanismo de acción de la terapia sobre la enfermedad y de que el biomarcador refleje este mecanismo.
- El biomarcador sea sensible y específico a los cambios en las condiciones del paciente.
- La medición del biomarcador tenga una metodología validada.
- Existan estudios preclínicos o modelos in vitro que hayan establecido la relación entre el comportamiento del biomarcador y la respuesta a la dosis.
- El biomarcador tenga un valor predictivo en otra enfermedad similar.

■ 18.3.2. Resultados clínicos centrados en el paciente con enfermedades raras

Las terapias para EE. RR. deben desarrollarse para tratar pacientes, y no solo su enfermedad, y por lo tanto, la capacidad de conseguir resultados que reflejen los beneficios reales desde la perspectiva del paciente es muy importante.

La evaluación de resultados clínicos centrados en el paciente (equivalente a las siglas en inglés PCOM, *Patient-Centered Outcomes Measures*) permite traducir la atención o el beneficio del tratamiento observado en una medida "interpretable" del beneficio del paciente.

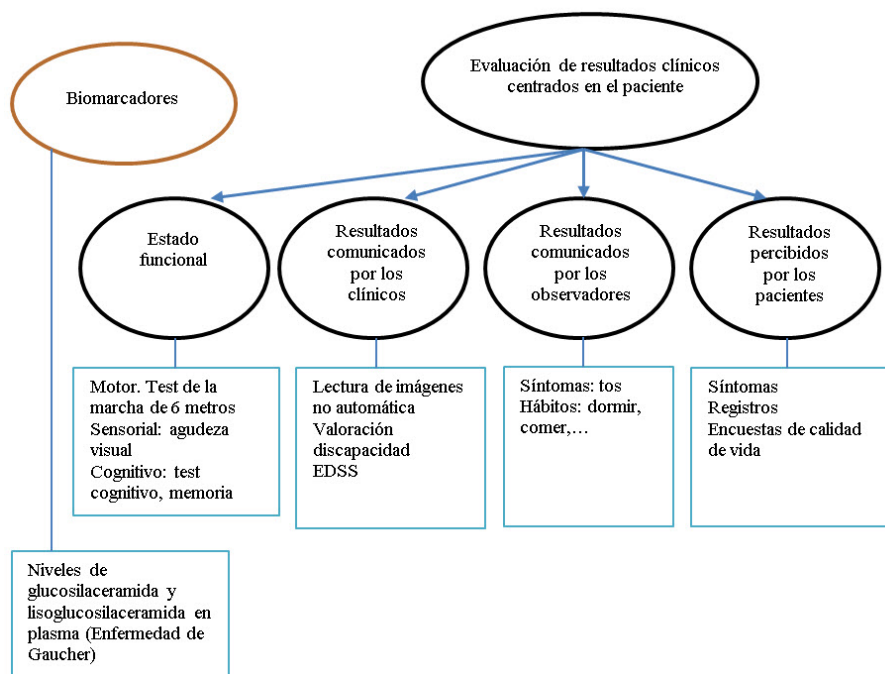
Existen diferentes formas de evaluación de los resultados clínicos:

- resultados percibidos por el paciente (en inglés PRO, *Patient Reported Outcomes*),
- resultados comunicados por el clínico (en inglés ClinRO, *Clinical Reported Outcomes*),
- resultados comunicados por el observador (en inglés ObsRO, *Observer Reported Outcomes*)
- resultados del estado funcional (en inglés PerfO, *Performance Outcomes*).

Todas las medidas de resultados clínicos centradas en los pacientes (MPCO) pueden ser utilizadas para evaluar el beneficio del tratamiento, que en definitiva hace referencia a cómo se siente el paciente, cuál es su estado funcional y su supervivencia. En la Figura 1 se describen ejemplos de los distintos tipos de evaluación de resultados clínicos.

Figura 1.

TIPOS DE EVALUACIÓN DE RESULTADOS CLÍNICOS CON EJEMPLOS DE CADA UNO DE ELLOS



El Consorcio Internacional de Investigación en Enfermedades Raras (IRDiRC) que es una iniciativa creada en 2011 por la Comisión Europea y el Instituto Nacional de Salud de EE. UU. para fomentar la investigación colaborativa y la inversión en campo de las EE. RR., creó un grupo de trabajo en junio de 2015 con la finalidad de profundizar en las medidas de resultado centradas en el paciente con EE. RR. (Morel y Cano, 2017), que puso de manifiesto la necesidad de evaluar los resultados que reflejan los beneficios reales en el paciente y que para avanzar, tanto en investigación como en práctica clínica es necesaria la colaboración de todas las partes interesadas, poniendo en el Centro a la comunidad de pacientes con EE. RR..

■ **18.3.2.1. Dificultades para la selección de medidas de evaluación de resultados centrados en el paciente en enfermedades raras. Estrategias para minimizarlas**

Para desarrollar medidas de evaluación de resultados centrados en el paciente en los ensayos clínicos son necesarias tres fases: comprender la enfermedad, conceptualizar el beneficio del tratamiento y seleccionar y desarrollar instrumentos para medir el resultado.

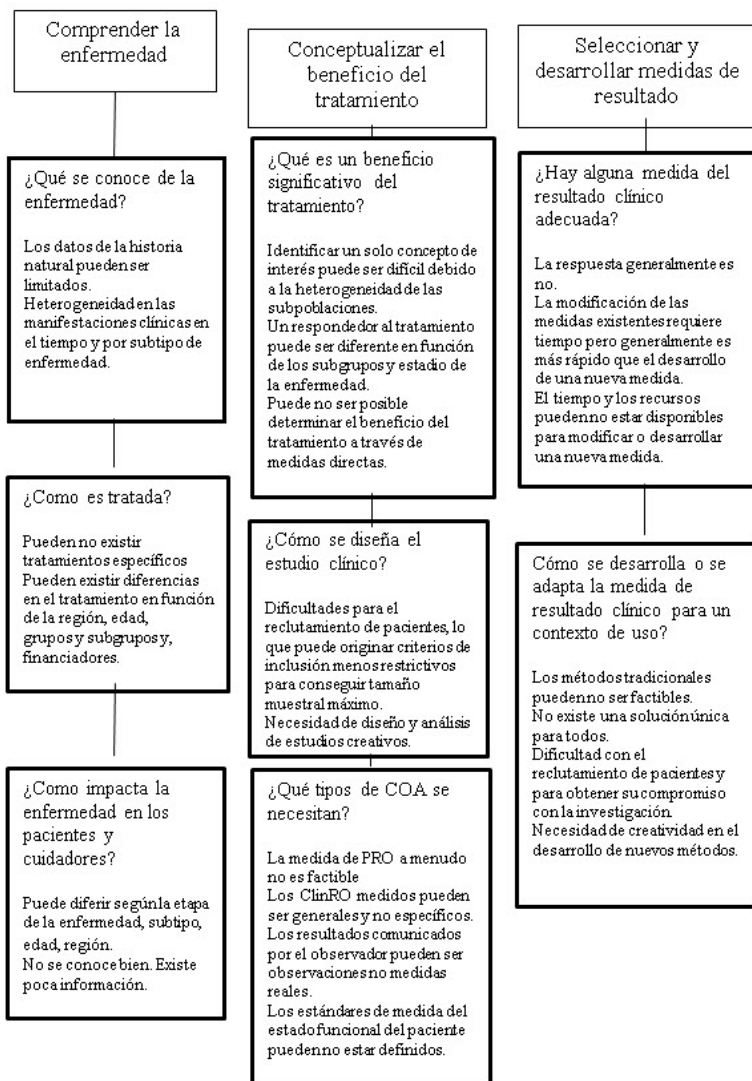
Estas fases presentan características diferenciales en las EE. RR. que se ponen de manifiesto en la Tabla 1.

Algunas de las sugerencias planteadas para salvar las dificultades que existen para evaluar los resultados en EE. RR. son:

- Usar todas las fuentes de datos disponibles, incluidas las de uso no tradicional.
- Tener presente la edad y el estado cognitivo para seleccionar la variable de resultado más adecuada.
- Recopilar la información desde el primer síntoma hasta el diagnóstico para comprender la evolución de la enfermedad.
- Entrevistar a un grupo de pacientes y cuidadores representativos de diferentes edades y etapas de la enfermedad, fenotipos y subgrupos demográficos para comprender las diferencias en la experiencia de la enfermedad, la importancia e impacto de los síntomas y las percepciones de la sociedad.
- Trabajar junto con las asociaciones de pacientes de enfermedades raras para obtener información, diseñar estudios, reclutar pacientes, etcétera.
- Ante la diversidad en la presentación de la enfermedad y en la experiencia de los pacientes se recomienda centrarse en los síntomas más comunes y más importantes para los pacientes, y comprender las consecuencias del retraso diagnóstico en los resultados.

Figura 2.

DIFICULTADES Y CARACTERÍSTICAS DIFERENCIALES, EN LAS 3 FASES DEL PROCESO PARA SELECCIONAR MEDIDAS DE EVALUACIÓN DE RESULTADOS CLÍNICOS CENTRADOS EN EL PACIENTE EN EE. RR.



- Para minimizar las barreras para obtener las aportaciones de los pacientes y de sus cuidadores se aconseja, además de contactar con las asociaciones de pacientes, considerar las redes sociales e Internet para la recopilación

de datos e identificar centros de excelencia o redes de atención clínica que pueden tener un amplio alcance geográfico.

- Ante la ausencia de una variable para medir el resultado clínico adecuada, es posible tratar de seleccionar una nueva variable de resultado pero dadas las dificultades para dicha selección y para el desarrollo de la evidencia que la debe soportar, la adaptación de una ya conocida es una buena opción. Se puede adoptar la misma medida de resultados de una enfermedad similar usando métodos cualitativos y cuantitativos cuando sea posible. También puede usarse un instrumento de evaluación próximo al concepto de interés que pueda desagregarse para incluir sólo las subescalas o ítems más próximos a dicho concepto de interés o bien un instrumento de evaluación de resultados clínicos genérico que si bien puede ser menos sensible a un cambio puede proporcionar una "evaluación estándar" y datos comparativos que pueden ser considerados por los reguladores
- Para solventar las dificultades asociadas para captar pacientes para los estudios, estos pueden realizarse usando también cuidadores, clínicos, incluso otros miembros de la familia o maestros, cualquiera que tenga un conocimiento profundo del paciente.
- El desarrollo de un fármaco centrado en el paciente asegura que se tengan presente la perspectiva del paciente y del cuidador en el diseño de los estudios y en las variables de medida.

■ 18.3.2.2. *Resultados percibidos por los pacientes*

Para monitorizar cambios en la historia natural y la progresión de una ER, así como la respuesta a intervenciones o tratamientos, son importantes los resultados percibidos por los pacientes (RPP), equivalente español de "*PRO: patient reported outcomes*", en las que los pacientes evalúan directamente sus síntomas, estado de salud o calidad de vida sin la interpretación de un médico (Benjamin *et al.*, 2017; U.S. Food and Drug Administration, 2019). Se describen en detalle en el capítulo 1.

Estas evaluaciones subjetivas se realizan sobre cualquier aspecto del estado de salud, incluyendo signos y síntomas, capacidad funcional física, estado psicológico, capacidad funcional social, bienestar general, calidad de vida relacionada con la salud (CVRS), percepciones del tratamiento, satisfacción del paciente (con el tratamiento, con la atención recibida, con la comunicación profesional), utilidades, adherencia al tratamiento, etc. (Rothman *et al.*, 2007). La percepción del paciente sobre la experiencia de la enfermedad puede variar en relación con estándares internos, valores intrínsecos y expectativas (Rothman *et al.*, 2007). Para recoger las percepciones de los pacientes se han desarrollado medidas o instrumentos de resultados percibidos por los pacientes (MR. PP.) (en inglés, *Patient Reported Outcome Measures: PROMs*) que recogen los puntos de vista del paciente sobre su estado

de salud y facilitan la comprensión del impacto de la enfermedad y su tratamiento en la CVRS y los síntomas. Diferentes entidades han desarrollado guías con estándares para el desarrollo, evaluación, implementación y análisis de la MR. PP. Entre ellas se encuentran las agencias reguladoras como la FDA (U.S. Food and Drug Administration, 2019), organizaciones que publican estándares como CONSORT (Calvert *et al.*, 2018), asociaciones profesionales como la International Society for Quality of Life Research (ISOQOL) (Brundage *et al.*, 2013) y la International Society for Pharmacoeconomics and Outcomes Research (ISPOR) (Benajmin *et al.*, 2017).

El impacto de las EE. RR. es muy diverso e incluye pérdida de función física, alteración cognitiva, alteración de la habilidad de comunicación, así como problemas emocionales y aislamiento social (Slade *et al.*, 2018), lo que produce una alteración en las interacciones sociales, el trabajo y la educación y además altera la calidad de vida de los pacientes y los familiares (Johansen *et al.*, 2013).

Con las EE. RR. nos encontramos con la dificultad que entraña diseñar un MR. PP. específico para cada una de ellas y validar los cuestionarios genéricos ya existentes para cada enfermedad para su uso en ensayos clínicos, lo que se traduce en escasez de MR. PP. validadas (Slade *et al.*, 2018). La opinión de los pacientes en el desarrollo de las MR. PP., incluyendo la investigación cualitativa, es esencial para asegurar que medidas de resultado que importan a las personas con EE. RR. sean recogidas de forma apropiada. Dado el gran número de EE. RR., el pequeño número de pacientes que tienen cada enfermedad y el coste del desarrollo de los instrumentos, se necesitan soluciones creativas y pragmáticas para el desarrollo y uso de las MR. PP. (Slade *et al.*, 2018).

Dentro de los RPP, la medición de la calidad de vida respecto a la salud (CVRS) se considera una de las medidas fundamentales, ya que permite una valoración subjetiva acerca de la repercusión de la patología y su tratamiento en la vida diaria. Es la variable de RPP que ha recibido mayor interés y atención a lo largo de las últimas décadas. No obstante, el espectacular desarrollo que ha experimentado su análisis e investigación no ha ido parejo a su implementación en la práctica diaria.

La CVRS es un término que ha ido evolucionando conceptualmente. Se puede definir como la percepción individual del impacto de una enfermedad, discapacidad o síntoma sobre los dominios somáticos, sociales, psicológicos y físicos del funcionamiento y del bienestar (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2013). Es comúnmente aceptada la importancia de estudiar la CVRS para una valoración completa del estado de salud del paciente, de la carga general de la enfermedad y de la evaluación de la respuesta al tratamiento. Su medida puede ayudar en la identificación de problemas de salud (tanto físicos como mentales o sociales) no detectados en la evaluación clínica convencional (Caballero y Prior, 2017).

Para medir la CVRS se utilizan generalmente cuestionarios con varias dimensiones (física, social, emocional, cognitiva, laboral, síntomas, efectos secundarios, ...). Estos cuestionarios de CVRS se pueden clasificar en genéricos (SF-36, EQ-5D,

NHP) o específicos para una enfermedad determinada o síntoma (U.S. Food and Drug Administration, 2019). Los cuestionarios genéricos tienen la ventaja de permitir comparar con otras enfermedades, pero generalmente les falta sensibilidad para evaluar determinados aspectos de una enfermedad particular, por lo que generalmente se prefiere el uso de instrumentos específicos para este objetivo. Para medir la CVRS es aconsejable que los MR. PP. utilizados cumplan los criterios recomendados por los expertos, siendo viabilidad, fiabilidad, evidencias de validez y sensibilidad al cambio las más importantes (Reeve *et al.*, 2013).

Un estudio realizado en EE. UU. ha analizado la CVRS en una muestra amplia (n=1218) de pacientes con diferentes EE. RR., que se categorizaron según las definiciones de Orphanet (Bogart e Irvin, 2017). Para ello utilizaron PROMIS (*Patient-Reported Outcomes Measurement Information System*). Las puntuaciones medias de los 6 dominios del PROMIS de la muestra completa fueron más bajas que los datos normativos de la población Americana y de pacientes con enfermedades crónicas comunes y variaron mucho según la enfermedad. La pobre CVRS se podría atribuir a problemas en el acceso al diagnóstico, a la información médica, el tratamiento, el apoyo psicosocial, la superación del estigma y la incertidumbre. Concluyen que la identificación de maneras de mejorar la CVRS es crucial en el cuidado centrado en el paciente y debería tener una financiación prioritaria.

■ 18.3.2.3. Resultados comunicados por observadores de los pacientes

En las EE. RR. es frecuente que los pacientes no puedan responder a las preguntas de una MR. PP. porque sean demasiado jóvenes o por las alteraciones físicas y cognitivas asociadas a muchas de las enfermedades. Todo ello dificulta la recogida de datos sobre los resultados del tratamiento directamente de los pacientes. En el caso de que los pacientes puedan responder a las preguntas de una MR. PP., pero tengan alteraciones físicas que les dificulten rellenar las encuestas en papel, ordenador o teléfono automatizado, las respuestas a una MR. PP. pueden ser recogidas por un “entrevistador”. En los casos en que no puedan contestarlas es necesario recoger la información de forma indirecta de observadores (clínicos, padres, otras personas) que tengan un conocimiento directo del estado funcional, los síntomas, signos y comportamiento del paciente en relación con la enfermedad. Para ello se utilizan “observer reported outcomes: ObsROs” (Benjamin *et al.*, 2017; Basch y Bennett, 2014).

■ 18.3.2.4. Resultados comunicados por los clínicos

Los resultados comunicados por los clínicos (*Clinician reported outcomes: ClinRO*) no poseen una definición formal hasta la fecha, sin embargo, el grupo de trabajo de ISPOR en su informe sobre buenas prácticas de investigación en la eva-

luación de resultados clínicos (Powers *et al.*, 2017) los define como "evaluación determinada por observadores con capacitación y entrenamiento profesional relevante para la medición". Los observadores pueden ser profesionales sanitarios que hayan recibido entrenamiento especializado para evaluar el estado de salud de los pacientes, no siendo imprescindible que sean médicos.

Están descritos en detalle en el capítulo 1. Los ClinROs pueden estar influenciados por las valoraciones, motivaciones o decisiones de las personas que las realizan.

Aunque la perspectiva del paciente es muy importante para comprender los efectos del tratamiento, hay muchas situaciones en las que esto no es posible y son imprescindibles los resultados comunicados por un clínico capacitado. Por tanto los ClinROs son especialmente prácticos cuando los pacientes no pueden evaluar los síntomas y su calidad de vida y/o cuando se requiere un juicio clínico para realizar la evaluación.

La evaluación de un ClinRO puede tener múltiples subcomponentes que influyen en dicha valoración. Los biomarcadores no son ClinRO, pero pueden incluirse en su evaluación como parte de la información que los clínicos utilizan para formar opiniones o tomar decisiones, pero no serán ClinRO si el biomarcador es el único factor para la decisión clínica.

Los tipos de ClinROs no son diferentes en el ámbito de las EE. RR. versus las comunes, e incluyen lecturas, escalas, y evaluaciones globales clínicas (Powers *et al.*, 2017).

Las lecturas son resultados definidos claramente que pueden observarse e informarse de una manera dicotómica (p. ej.: sí/no, presencia/ausencia). Las lecturas son, a menudo, dicotomizadas en base al juicio clínico, así la variable hospitalización sí o no en un niño con fibrosis quística es contada con facilidad pero es la decisión de hospitalizar al paciente lo que requiere un juicio clínico. En este caso, las razones que llevaron a la hospitalización son lo importante, y no sólo el caso de hospitalización en sí.

Las escalas son categóricas (ya sean ordenadas o no) o medidas continuas, en las cuales la evaluación tiene al menos tres niveles posibles que generan puntuaciones que representan el (los) concepto (s) de interés.

El papel de los expertos clínicos es realmente esencial cuando se usan estas escalas, ya que las mismas presentan, con frecuencia, una forma semiestructurada, siendo el juicio clínico una parte fundamental en la clasificación.

Son muy numerosas las escalas utilizadas en las EE. RR., como ejemplo de ellas señalaremos alguna de las escalas validadas en España para enfermedades neuromusculares:

- *Hammersmith Functional Motor Scale* (HFMS) utilizada en niños con atrofia muscular espinal tipo II.
- Escala EgenKlassifikation (EK): para evaluar la capacidad funcional de niños y adultos con atrofia muscular espinal y distrofia muscular de Duchenne que están en fase de silla de ruedas.
- Escala EgenKlassifikation 2 (EK2): ampliación de la escala EK y sirve para evaluar la capacidad funcional de niños y adultos con atrofia muscular espinal y distrofia muscular de Duchenne que están en fase de silla de ruedas siendo más completa y específica para los afectados de atrofia muscular espinal.

Previamente al uso de las diferentes escalas en una población determinada, estas han debido ser validadas en dicha población, ya que las características ambientales, culturales idiomáticas etc... no son iguales en todos los países.

El entrenamiento de los profesionales en escalas de evaluación es imprescindible y este puede ser especialmente difícil en las EE. RR. por el reducido número de pacientes y la dispersión de los mismos, hechos que pueden originar que un profesional lleve solo a uno ó dos pacientes con una determinada enfermedad, lo que impedirá o dificultará la adquisición de habilidades en el uso de las mismas.

Los valores obtenidos deberán valorarse en un contexto determinado, teniendo presente, entre otros aspectos, la edad, peso y sexo del paciente.

Las evaluaciones globales clínicas (*Clinical Global Assessment CGA*) son evaluaciones basadas en general en el juicio clínico del estado de salud total del paciente o de un aspecto de su estado de salud para el cual no hay una variable bien definida o ni siquiera está definida.

Se ha de tener presente que la evaluación de un ClinRO puede ser adecuada en un contexto de uso y no serlo en un contexto diferente. Por ejemplo, las manifestaciones de la enfermedad pueden diferir entre las etapas tempranas y posteriores de la enfermedad o por otros aspectos distintivos de las poblaciones de la enfermedad.

La mayoría de los ClinROs son evaluaciones indirectas del beneficio de un tratamiento pero a pesar de ello de gran valor siempre que exista una relación, claramente definida, entre los mismos y los resultados en salud de los pacientes.

■ 18.3.3. Variables subrogadas

Las variables subrogadas o intermedias se definen como medidas de laboratorio o signos físicos que se usan en ensayos terapéuticos como sustitutos de una variable clínicamente significativa que es una medida directa sobre lo que siente

un paciente, su estado funcional o su supervivencia y que se espera que prediga el efecto de la terapia. Generalmente incluyen biomarcadores fisiológicos, inmunológicos, histológicos y genéticos, pero también pueden ser de otro tipo.

Las medidas de resultados centradas en los pacientes pueden utilizarse para validar variables subrogadas (Morel y Cano, 2017). Un ejemplo es la mielofibrosis, una ER de la médula ósea caracterizada por la presencia de tejido fibroso en la médula ósea y citopenia, con desplazamiento de células madre de la médula ósea a la sangre, donde colonizan órganos a distancia (fundamentalmente, el bazo y el hígado), lo que puede originar un aumento de estos órganos que causa malestar y dolor abdominal. En esta enfermedad el volumen esplénico es una variable subrogada que puede predecir el beneficio del tratamiento, pero la incorporación en el ensayo clínico con ruxolotinib de una medida directa del beneficio del tratamiento en el paciente a través del cuestionario MSAF (*Myelofibrosis Symptom Assessment Form*), facilitó la aprobación regulatoria rápida en 2011 por la FDA y la EMA y evitó el requisito de ensayos confirmatorios posteriores a la comercialización.

■ 18.3.4. Medidas de resultados en salud económicos

Las distintas variables de medidas de resultados en salud económicos se describen en profundidad en el capítulo 2. Existe una gran carencia de estudios de resultados en salud económicos en EE. RR., ya que la baja prevalencia de cada enfermedad, junto con el gran número de EE. RR. dificulta su realización.

En Europa se planteó el desarrollo de un modelo de estudio de la carga socioeconómica y la calidad de vida en pacientes con EE. RR. y sus cuidadores (BURQOL-RD: *Social economic burden and health-related quality of life in patients with rare diseases in Europe*) (Angelis, Tordrup y Kanavos, 2015; López-Bastida *et al.*, 2016). Fue un proyecto financiado por la Comisión Europea, de tres años de duración (2010-2013), en él participaron inicialmente 23 socios de ocho países miembros de la UE, uno de ellos España. Dos asociaciones de pacientes, FEDER y EURORDIS, participaron como participante asociado y socio colaborador del proyecto, respectivamente. En primer lugar se realizó una selección de las diez EE. RR. para la creación del modelo y realización del estudio piloto. Para esta selección se utilizó un método Delphi de dos vueltas en combinación con un diagrama de Carrol para consensuar la lista. En el diagrama de Carrol se tuvieron en cuenta tres determinantes: prevalencia de la enfermedad $\geq 1/10.000$, disponibilidad de fármacos efectivos y necesidad de cuidador por la afectación físico-mental. Finalmente las enfermedades seleccionadas fueron: fibrosis quística (FQ), síndrome de Prader-Willi (SPW), hemofilia, distrofia muscular de Duchenne (DMD), epidermólisis bullosa (EB), síndrome del X frágil (SXF), esclerodermia (SCL), mucopolisacaridosis (MPS), artritis idiopática juvenil (AIJ) e histiocitosis (HIS). Dentro de este proyecto se realizó una revisión de los estudios existentes sobre los costes directos e indirectos de estas diez enfermedades en Europa (Linertová, García-Pérez y Gorostiza, 2017). La pre-

valencia de las diez enfermedades en Europa era muy diferente y varió desde 1-2 por 100.000 habitantes (SPW, un subtipo de HIS y EB) hasta 42 por 100.000 habitantes (SCL). En general se observó que la evidencia sobre coste de enfermedad en EE. RR. era muy escasa, identificando un total de 77 estudios con distribución desigual entre las 10 enfermedades. La FQ con 29 estudios y la hemofilia con 22 estudios estaban relativamente bien estudiadas en comparación con otras enfermedades en las que los datos de coste de enfermedad eran muy limitados. Los datos de coste durante toda la vida solo estaban disponibles para 4 enfermedades, mientras que los datos de costes anuales totales (incluyendo los costes indirectos) solo se disponían para cinco enfermedades. La disponibilidad de datos se correlacionó con la existencia de un tratamiento farmacológico y los costes indirectos eran una proporción significativa de los costes totales. Las variaciones metodológicas no permitieron una comparación detallada entre enfermedades, pero la mayoría de las EE. RR. estudiadas se asociaban con una carga económica significativa, tanto directa como indirecta (Linertová, García-Pérez y Gorostiza, 2017).

Posteriormente se realizó el estudio piloto del BURQOL-RD en Alemania, Bulgaria, España, Francia, Hungría, Italia, Reino Unido y Suecia (López-Bastida *et al.*, 2016), en el que se utilizaron procedimientos para incrementar el reclutamiento. En el BURQOL-RD se utilizó un método basado en la prevalencia (se estimó el costo en los seis meses previos a la entrevista y se extrapolaron a un año; excepto para los ingresos hospitalarios, en que se recogieron datos del año previo a la entrevista) y con un planteamiento ascendente o de microcosteo (se estimó la carga de enfermedad a partir del consumo de recursos de un grupo de pacientes con la enfermedad y luego la estimación se extrapoló a la población según la prevalencia de cada enfermedad) y se utilizó una perspectiva desde el punto de vista de la sociedad (López-Bastida *et al.*, 2016). Los resultados del estudio piloto se han publicado en un suplemento de *Eur J HealthEcon* en 2016 y han sido resumidos (López-Bastida *et al.*, 2016). Se analizaron 3232 respuestas completas y se calculó la media de los costes anuales totales en 4 categorías: costes directos sanitarios, costes directos no sanitarios formales, costes directos no sanitarios informales, costes indirectos (Angelis, Tordrup y Kanavos, 2015). El cuidado informal (tiempo de cuidador no pagado, generalmente familiares) representó una gran parte de los costes totales en todos los países y para la mayoría de las enfermedades (Angelis, Tordrup y Kanavos, 2015). Este coste no suele tener un presupuesto y es muchas veces invisible, pero desde el punto de vista de la sociedad es muy importante. Una de las fortalezas de este estudio reside en el uso de una metodología común para evaluar los costes y la CdVRS en un amplio espectro de EE. RR. en diferentes países de la UE. Los autores concluyen que a pesar de la relevancia del estudio de los costes sanitarios directos de las EE. RR., los costes sociales pueden ser incluso más elevados, debido a la pérdida de productividad laboral y el cuidado formal e informal necesario (Angelis, Tordrup y Kanavos, 2015).

El BURQOL-RD puede servir como modelo para realizar estudios de carga económica en EE. RR.

■ 18.4. ALTERNATIVAS EN EL DISEÑO DE ENSAYOS CLÍNICOS

Varios factores han obstaculizado el desarrollo de fármacos para EE. RR. (Sun, Zheng y Simeonov, 2017). Por un lado la heterogeneidad en la fisiopatología de una enfermedad puede producir grandes variaciones en la respuesta a los fármacos en pacientes con la misma enfermedad. Por otro lado se desconoce la progresión de muchas EE. RR. por el número limitado de estudios sobre la historia natural, dificultado aún más en enfermedades ultra raras de las que se conocen menos de 100 pacientes en el mundo. Además, el número inadecuado de pacientes reclutados en los ensayos clínicos puede llevar a una carencia de significado estadístico en las medidas de resultados. Finalmente, la ausencia de biomarcadores que puedan medir la enfermedad puede contribuir a la ambigüedad de los estudios clínicos en EE. RR.

Se han desarrollado métodos para fomentar el desarrollo de fármacos en estas enfermedades que incluyen el uso de biomarcadores como medidas de resultados, el diseño innovador de ensayos clínicos, el uso de diferentes análisis estadísticos potentes y la licencia o aprobación progresiva (Gobburu y Pastoor, 2016).

Algunas alternativas al patrón de referencia de los ensayos clínicos, el diseño aleatorizado, paralelo, controlado con placebo, y que se han utilizado en EE. RR. pueden verse en la Tabla 2 (Gobburu y Pastoor, 2016; Cornu *et al.*, 2013; Abrahamyan *et al.*, 2016).

Tabla 2.

DISEÑOS ALTERNATIVOS EN ENSAYOS CLÍNICOS EN EE. RR.

Ensayo cruzado aleatorizado doble ciego
Ensayos con n=1
Diseño aleatorizado de fase-placebo
Ensayos con fase de reclutamiento enriquecido y retirada aleatorizada
Diseño de posicionamiento y selección
Diseños factoriales
Diseños secuenciales
Estudio por pasos
Cuadrado latino
Tres fases
Inicio retrasado o postergado
Diseños combinados

Ensayo cruzado aleatorizado doble ciego: La principal ventaja del ensayo cruzado es eliminar la variabilidad entre-individuos en la respuesta a la intervención, pues estos ensayos valoran el efecto intra-individuo. Sus inconvenientes son:

- Al observar el efecto de un tratamiento se asume la ausencia de efectos residuales del tratamiento anterior. Debe existir un período de blanqueo o lavado suficientemente largo como para asegurar que no permanecen los efectos residuales del primer tratamiento asignado.
- La secuencia con la que se administran los diferentes tratamientos puede asociarse a respuestas psicológicas diferentes (mayor entusiasmo al principio del estudio, que disminuye con el tiempo).
- Este diseño no es posible cuando la nueva terapia es quirúrgica o el primer tratamiento está dirigido a curar definitivamente la enfermedad.
- Resulta difícil adjudicar los efectos secundarios tardíos.
- El análisis estadístico es más complejo que para grupos paralelos y requiere el uso de métodos emparejados o de medidas repetidas: McNemar, t emparejada, Wilcoxon, modelos mixtos de análisis de la varianza.

Los *ensayos con $n=1$* consisten en que en un solo paciente, se usa primero una terapia (p. ej., durante un mes o una semana) y luego otra terapia o placebo. El paciente no sabe qué terapia está recibiendo cada día (enmascaramiento) y sirve como su propio control. Este diseño aborda el concepto de que una determinada intervención raramente funciona en todos los pacientes y ofrece una metodología para abordar los problemas del tamaño muestral y de la heterogeneidad de las respuestas de una forma objetiva. Como todos los datos de la muestra se refieren a una sola persona (valorada en tiempos distintos), las conclusiones sólo se pueden aplicar a esa persona.

Diseño aleatorizado de fase-placebo: Se desarrolló para tratamientos que pueden producir una remisión o respuesta duradera. Los pacientes son aleatorizados a tratamiento experimental o placebo. Tras un periodo corto y fijo los pacientes del grupo control pasan de forma ciega al tratamiento experimental. La variable principal suele ser el tiempo hasta la respuesta.

Ensayos con fase de reclutamiento enriquecido y retirada aleatorizada: Este tipo de diseño consta de dos fases: la fase de reclutamiento en la que el fármaco se administra de forma abierta a todos los pacientes y es utilizada para detectar a los pacientes que responden al tratamiento. Los pacientes respondedores pasan a la segunda fase o de retirada aleatorizada, en la que de forma aleatoria siguen con el tratamiento activo o cambian a placebo. El análisis de eficacia del estudio se realiza únicamente con la segunda fase del mismo.

Diseños factoriales: es una forma eficiente de comparar todas las posibles combinaciones de dos o más intervenciones en un único estudio. El número total de

combinaciones depende del número de factores (p. ej. Tratamientos) a comparar. El diseño más frecuente es aquel en el que se comparan dos productos (2x2) y el placebo: se dividen a los pacientes en cuatro grupos: a) tratamiento A y placebo, b) tratamiento B y placebo, c) ambos tratamientos A y B, d) ningún tratamiento.

Diseños adaptativos: permiten a los investigadores recoger datos durante el propio estudio para adaptar de forma dinámica el diseño del estudio sin menoscabar la validez o integridad del mismo. Se basan principalmente en análisis provisionales sin enmascaramiento (abiertos) que permiten determinar los efectos del tratamiento. Estos diseños tratan de minimizar el número de participantes requeridos para alcanzar resultados, manteniendo la validez estadística.

Diseño de posicionamiento y selección: este tipo de estudio ofrece una alta probabilidad de identificar la mejor opción cuando una opción de tratamiento es superior. Constan generalmente de dos fases, con un estudio estadístico formal realizado únicamente al final de la segunda fase. En la primera fase los pacientes son aleatorizados a todas las opciones de tratamiento. En una segunda fase los tratamientos inferiores son eliminados mediante el uso de reglas preestablecidas, pero generalmente sin un testado formal de hipótesis (retirada de los perdedores), o el tratamiento con mejor resultado es seleccionado para pasar a la siguiente fase (selección del ganador). En la segunda fase el tratamiento o tratamientos seleccionados son comparados con placebo u otras alternativas utilizando un diseño convencional o adaptativo. El análisis final se ajusta para mantener el error tipo I general.

Diseños secuenciales: Los diseños secuenciales son similares al diseño tradicional aleatorizado controlado con placebo, pero permite análisis interinos de datos de forma repetida para guiar la terminación precoz del estudio por razones de seguridad, futilidad, eficacia o unacombinación todos. En diferencia a los ensayos clásicos el número final de participantes necesario se desconoce al inicio del estudio, aunque a veces se limita el número máximo de participantes. El ensayo clínico finaliza cuando el primer análisis interino de datos alcanza criterios preestablecidos de parada (se limita de esta forma el número de pacientes expuestos a un tratamiento inferior, inseguro o fútil, en comparación con un ensayo clínico convencional).

Cuadrado latino: Se comparan más de dos tratamientos. Se trata de un diseño equilibrado en el que cada tratamiento o dosis aparece una sola vez en cada secuencia y en cada periodo de tratamiento.

Inicio retrasado o postergado: Hay dos fases: fase inicial controlada con placebo (los pacientes son aleatorizados a tratamiento o placebo) seguidos de una fase de control activa (todos los pacientes reciben tratamiento). Los pacientes que comienzan con el placebo tienen un inicio retrasado.

Estudio por pasos: Todos los pacientes reciben el tratamiento en estudio al final. La intervención se asigna de forma secuencial a los participantes (bien como individuos o como grupos de individuos). En un diseño de cinco pasos todos los

pacientes comienza con el control y en los siguientes cinco periodos de tiempo los individuos o grupos de individuos son asignados de forma aleatorizada al tratamiento, para finalizar con todos los pacientes recibiendo el tratamiento estudiado.

Tres fases: Consta de una fase inicial aleatorizada controlada con placebo, seguida de una fase de retirada aleatorizada en los respondedores y una tercera fase aleatorizada para los no respondedores a placebo que posteriormente responden al tratamiento.

Diseños combinados: algunos de estos diseños son complementarios y pueden usarse de forma combinada.

La licencia progresiva de un fármaco permitiría que el programa de desarrollo se reestructurara para permitir la aprobación precoz en subgrupos de pacientes con mayor riesgo. La aprobación se reevaluaría en otros momentos adicionales a lo largo del desarrollo del fármaco, según se disponga de datos adicionales de la población en la que se aprobó el fármaco, así como se hayan obtenido nuevos datos para otras poblaciones (Gobburu y Pastoor, 2016).

Los ensayos clínicos con muestras pequeñas son necesarios cuando hay dificultad para reclutar suficientes pacientes para realizar análisis estadísticos frecuentistas convencionales. Estos ensayos son a menudo necesarios para el estudio de EE. RR. así como para poblaciones específicas, tales como niños. La EMA ha publicado una guía sobre los ensayos clínicos en pequeñas poblaciones que se describirá más adelante (European Medicines Agency. Committee for Medicinal Products for Human Use [CHMP], 2006). Una revisión de las publicaciones sobre ensayos clínicos en EE. RR. o subpoblaciones de pacientes mostró 12 posibles métodos alternativos al patrón de referencia del diseño de grupos paralelos, con ventajas y limitaciones. Proponen un algoritmo para escoger el diseño más adecuado (Cornu *et al.*, 2013).

Recientemente se han revisado los métodos de descubrimiento y desarrollo de nuevos fármacos y su aplicación en EE. RR. (Sun, Zheng y Simeonov, 2017). Estos autores se centran en la revisión de las moléculas pequeñas y los tratamientos biológicos. También abordan el replanteamiento de fármacos o reposicionamiento de fármacos ya aprobados para tratar nuevas enfermedades. Un ejemplo del reposicionamiento de fármacos es el sildenafil (Viagra®, Pfizer), que inicialmente se estudió para la hipertensión y la angina de pecho en los 80, se replanteó para el tratamiento de la disfunción eréctil en 1998 y se aprobó en 2005 para el tratamiento de la hipertensión pulmonar (Revatio®, Pfizer). Una forma de acelerar el desarrollo de fármacos para EE. RR. puede consistir en un programa robusto de reposicionamiento. En EE. UU. se ha creado un programa específico denominado *New Therapeutic Uses* (<https://ncats.nih.gov/ntu>).

■ 18.5. DESAFÍOS ESTADÍSTICOS EN ENFERMEDADES RARAS

El paradigma tradicional o frecuentista para el análisis de los ensayos clínicos se centra en el valor P, la probabilidad de que los datos observados o valores más extremos puedan ocurrir si el tratamiento fuera totalmente ineficaz (Ej., hipótesis nula de no efecto del tratamiento). En la situación de contraste de hipótesis, cuando la probabilidad es menor del 0,05 (convencional), el efecto del tratamiento observado se considera estadísticamente significativo y se rechaza la hipótesis nula. La probabilidad de concluir que hay un efecto del tratamiento cuando realmente no lo hay se conoce como error de tipo I (o falso positivo). Aunque la magnitud del valor P está relacionada con el tamaño del efecto, también está influenciada por el tamaño muestral y el diseño del estudio, determinando ambos la precisión del estudio. Por lo tanto, el análisis estadístico tradicional no es ideal para el estudio de las EE. RR., en las que el tamaño de las muestras es reducido (Abrahamyan *et al.*, 2016).

Métodos bayesianos

Hay diversas razones por las que los métodos bayesianos son útiles en el análisis de datos en estudios de EE. RR. (Abrahamyan *et al.*, 2016): permiten la incorporación formal de información apropiada procedente de fuera del estudio; se pueden incluir colecciones de datos cumulativos del estudio y facilitan un diseño flexible del estudio; y ponen su atención en la estimación de efectos y no están lastrados por las expectativas de demostrar significación como en los estudios convencionales. Estas ventajas son aplicables tanto a los diseños convencionales como nuevos. El planteamiento clave de los diseños bayesianos es cómo se modifica el conocimiento previo consistente en una distribución de probabilidad, ante la aparición de una nueva evidencia expresada también como una distribución de probabilidad, de modo que todos los resultados disponibles se tienen en cuenta.

Análisis secuencial por grupos

El análisis de datos acumulados en una serie de momentos predeterminados (métodos secuenciales por grupos, MSG) permite realizar análisis intermedios de los datos con grupos de pacientes de tamaño preestablecido en lugar de analizar la trayectoria paciente a paciente. Permite gestionar la multiplicidad de análisis mediante diversos métodos de ajuste del error de tipo I y finalizar prematuramente un estudio si se rechaza la hipótesis nula a partir de los resultados de los análisis intermedios.

■ 18.6. GUÍAS DE LAS AGENCIAS REGULADORAS SOBRE ENSAYOS CLÍNICOS EN EE. RR.

■ 18.6.1 Guía de la EMA sobre ensayos clínicos en poblaciones pequeñas

Esta guía considera los problemas asociados a los ensayos clínicos cuando hay un número limitado de pacientes disponibles para estudiar, como sucede en las

EE. RR. y en muchos ensayos pediátricos, en los que las condiciones del ensayo pueden verse limitadas por la prevalencia de la enfermedad (European Medicines Agency. Committee for Medicinal Products for Human Use [CHMP], 2006).

La guía insiste en que no existen métodos específicos para diseñar, desarrollar o analizar ensayos clínicos en EE. RR. y las pautas aplicables a los ensayos en enfermedades comunes también les son aplicables, pero sí que existen enfoques para aumentar la eficiencia de los ensayos clínicos en poblaciones pequeñas y en situaciones donde no es posible obtener evidencia de la eficacia y seguridad a través de un ensayo clínico controlado y aleatorizado. En estas situaciones, la autoridad regulatoria puede aceptar diferentes enfoques siempre que se asegure que los intereses de los pacientes están protegidos. Se dará prioridad a la necesidad de resultados clínicos frente a resultados estadísticos.

En enfermedades muy raras, la evaluación combinada de estudios de casos individuales puede ser la única forma de proporcionar evidencia. En tales situaciones, las condiciones del tratamiento y la recopilación de los datos que serán de alta calidad, deberán estar estandarizados y cumplir con las normas de buena práctica clínica. Estos estudios estarán descritos en los protocolos y se planificarán prospectivamente. En ocasiones una revisión sistemática de todos los datos (incluidos datos de otras fuentes) puede añadir peso a la evidencia. También se podrá considerar un análisis combinado de informes de casos individuales y estudios observacionales.

El conocimiento detallado de la fisiopatología de la enfermedad y la farmacología del medicamento pueden ayudar a diseñar los estudios y a identificar las fuentes de heterogeneidad en los pacientes. Los estudios farmacológicos no clínicos pueden ser útiles, especialmente en aquellas situaciones con muy pocos pacientes.

Las variables ideales son las denominadas duras y clínicamente relevantes, como curación o supervivencia, pero se pueden aceptar variables subrogadas, si éstas están plenamente justificadas y su relación con la eficacia clínica está clara para poder evaluar la relación beneficio-riesgo.

Los controles y los grupos de comparación son muy importantes, dado que su ausencia compromete la fiabilidad de los estudios. La situación ideal en el desarrollo de un tratamiento es un ensayo clínico comparativo con comparador activo, pero si esto no es posible se puede considerar placebo como comparador y si este no es adecuado se podrá considerar el mejor estándar de cuidado. Cuando el comparador no ha demostrado evidencia, el fármaco en estudio deberá demostrar superioridad.

Los registros de pacientes pueden proporcionar información importante sobre el curso natural de la enfermedad y pueden ayudar en la evaluación de efectividad y seguridad de los fármacos. Además, pueden usarse como fuente de controles históricos. Los registros, así utilizados deben contener datos de alta calidad.

En la guía también se insiste en la necesidad de evitar los sesgos biológicos, siendo uno de ellos la pérdida de seguimiento de los pacientes y se recogen diver-

sas consideraciones estadísticas relativas al diseño de los estudios y al análisis de los datos.

La EMA recomienda, encarecidamente a los promotores, buscar asesoramiento científico y asistencia para el protocolo durante todas las fases del desarrollo de un fármaco pues recibirán orientación para el enfoque del mismo que, sin duda, facilitará la aceptación posterior del fármaco.

■ 18.6.2. Guía de la FDA para la orientar a los promotores en el desarrollo de medicamentos para enfermedades raras

La FDA, también ha publicado una guía de recomendaciones (U.S. Food and Drug Administration, 2019) para ayudar a las empresas que desarrollan fármacos y productos biológicos, destinados a tratar o prevenir EE. RR., para que los programas de desarrollo de dichos medicamentos sean más eficaces y eficientes.

La guía aborda los siguientes aspectos importantes del desarrollo de medicamentos (si bien por razones de espacio no es posible detallarlos en este documento) :

- Descripción y comprensión de la historia natural de la enfermedad.
- Conocimiento de la fisiopatología de la enfermedad y el mecanismo de acción propuesto para el medicamento.
- Consideraciones farmacotoxicológicas no clínicas para apoyar la investigación clínica propuesta.
- Selección de variables adecuadas para la evaluación de resultados.
- Estándar de evidencia para establecer la eficacia y seguridad.
- Consideraciones para la fabricación de medicamentos durante su desarrollo.
- Participación de pacientes en los programas de desarrollo.

La FDA reconoce que ciertos aspectos del desarrollo de fármacos son factibles para enfermedades comunes, pero pueden no serlo para EE. RR. y admite flexibilidad en la aplicación de las normas regulatorias e insiste en la necesidad de que los promotores se reúnan de manera precoz y frecuente con ella para facilitar y hacer más eficiente este proceso.

La FDA fomenta la participación de los pacientes/ cuidadores en el proceso de desarrollo de los fármacos ya que ellos pueden proporcionar información importante sobre sus experiencias, perspectivas, necesidades y prioridades relacionadas con las variables de resultado, durante la revisión de un medicamento en investigación. Los pacientes pueden participar en comités asesores, servir como representantes

de pacientes específicos de una enfermedad, facilitar iniciativas de desarrollo de fármacos, responder a consultas sobre cuestiones relativas a la medida de resultados clínicamente significativos y participar en estudios de historia natural de la enfermedad.

■ REFERENCIAS

- ABRAHAMIAN, L., FELDMAN, B. M., TOMLINSON, G., FAUGHNAN, M. E., JOHNSON, S. R., DIAMOND, I. R. *et al.* (2016). Alternative for clinical trials in rare diseases. *Am J Med Genet Part C Semin Med Genet*, 9999C, pp. 1-19. doi: 10.1002/ajmg.c.31533.
- ANGELIS, A., TORDRUP, D. y KANAVOS, P. (2015). Socio-economic burden of rare diseases: a systematic review of cost of illness evidence. *Health Policy*, 119, pp. 964–979. doi: 10.1016/j.healthpol.2014.12.016.
- BASCH, E. y BENNETT, A. V. (2014). Patient-Reported Outcomes in Clinical Trials of Rare Diseases. *J Gen Intern Med*, 29(Suppl 3), pp. S801–S803. doi: 10.1007/s11606-014-2892-z.
- BENJAMIN, K., VERNON, M. K., PATRICK, D. L., PERFETT, E., NESTLER-PARR, S. y BURKE, L. (2017). Patient-Reported Outcome and Observer-reported outcome assessment in rare disease clinical trials - emerging good practices: an ISPOR COA emerging good practices report. *Value Health*, 20(7), pp. 838-855. doi: 10.1016/j.jval.2017.05.015.
- BOGART, K. R. e IRVIN, V. L. (2017). Health-related quality of life among adults with diverse rare disorders. *Orphanet J Rare Dis.*, 12(1), p. 177. doi: 10.1186/s13023-017-0730-1.
- BRUNDAGE, M., BLAZEY, J., REVICKI, D., BASS, B., DE VET, H., DUFFY, H. *et al.* (2013). Patient-reported outcomes in randomized clinical trials: development of ISOQOL reporting standards. *Qual Life Res*, 22(6), pp. 1161–1175. doi: 10.1007/s11136-012-0252-1.
- CABALLERO, T. y PRIOR, N. (2017). Burden of Illness and Quality-of-Life Measures in Angioedema Conditions. *Immunol Allergy Clin North Am*, 37(3), pp. 597-616. doi: 10.1016/j.iac.2017.04.005.
- CALVERT, M., KYTE, D., MERCIECA-BEBBER, R., SLADE, A., CHAN, A. W., KING, M.T. *et al.* (2018). Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *JAMA*, 319(5), pp. 483-494. doi: 10.1001/jama.2017.21903.
- CORNU, C., KASSAI B., FISCH, R., CHIRON, C., ALBERTI, C., GUERRINI, R. *et al.* (2013). Experimental designs for small randomised clinical trials: an algorithm for choice. *Orphanet J. Rare Dis.*, 8, p. 48. doi: 10.1186/1750-1172-8-48.
- COX, G. F. (2018). The art and science of choosing efficacy endpoints for rare disease clinical trials. *Am J Med Genet Part A*, 176(4), pp. 759-772. doi: 10.1002/ajmg.a.38629.
- EUROPEAN MEDICINES AGENCY. COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE [CHMP] (2006). *Guideline on Clinical Trials in Small Populations*. London. Recuperado de: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003615.pdf
- EURORDIS. (2009). *The Voice of 12,000 Patients: Experiences and Expectations of Rare Disease Patients on Diagnosis and Care in Europe*. Recuperado de: https://www.eurordis.org/IMG/pdf/voice_12000_patients/EURORDISCARE_FULLBOOKr.pdf
- GOBBURU, J. y PASTOOR, D. (2016). Drugs Against Rare Diseases: Are The Regulatory Standards Higher? *Clinical Pharmacology & Therapeutics*, 100 (4), pp. 322-323. doi: 10.1002/cpt.415.
- HERNÁNDEZ, C. (2016). Marco regulatorio de la investigación con medicamentos huérfanos en la Unión Europea. En: C. AYUSO, R. DAL-RÉ R y F. PALAU. (Ed.), *Ética en la investigación de las Enfermedades raras* (pp. 87-102). Madrid, España: Ergón.

- JOHANSEN, H., DAMMANN, B., ANDRESEN, I. L. y FAGERLAND, M. W. (2013). Health-related quality of life for children with rare diagnoses, their parents' satisfaction with life and the association between the two. *Health Qual Life Outcomes*, 11, p. 152. doi: 10.1186/1477-7525-11-152.
- LINERTOVÁ, R., GARCÍA-PÉREZ, L. y GOROSTIZA, I. (2017). Cost-of-Illness in Rare Diseases. *Adv Exp Med Biol*, 1031, pp. 283-297. doi: 10.1007/978-3-319-67144-4_17.
- LOGVISS, K., KRIEVINS, D. y PURVINA, S. (2018). Characteristics of clinical trials in rare vs. common diseases: A register-based Latvian study. *PLoS One*, 13(4), p. e0194494. doi: 10.1371/journal.pone.0194494.
- LÓPEZ-BASTIDA, J., OLIVA-MORENO, J., LINERTOVÁ, R. y SERRANO-AGUILAR, P. (2016). Social/economic costs and health-related quality of life in patients with rare diseases in Europe. *Eur J Health Econ*, 17(Suppl 1), pp. 1-5. doi: 10.1007/s10198-016-0780-7.
- MINISTERIO DE SANIDAD, SERVICIOS SOCIALES E IGUALDAD. (2013). *Estrategia en enfermedades raras del Sistema Nacional de Salud (2013). Actualización aprobada por el Consejo Interterritorial del Sistema Nacional de Salud el 11 de junio de 2014*. Madrid. Recuperado de: http://www.msc.es/organizacion/sns/planCalidadSNS/pdf/Estrategia_Enfermedades_Raras_SNS_2014.pdf
- MOREL, T. y CANO, S. J. (2017). Measuring what matters to rare disease patients - reflections on the work by the IRDIRC taskforce on patient-centered outcome measures. *Orphanet J Rare Dis.*, 12(1), p. 171. doi: 10.1186/s13023-017-0718-x.
- NÚÑEZ, M., GIL, B., SACRISTÁN, J. A. y DILLA, T. (2017). Analysis of Orphan Drugs Marketed in Spain during the Period 2010-2015: Epidemiological, Clinical and Economic Characteristics. *Health Econ Outcome Res Open Access*, 3, p. 140. doi: 10.4172/2471-268X/1000140.
- PARLAMENTO EUROPEO. (2000). Regulation (EC) No 141/2000 of the European parliament and of the Council of 16 December 1999 on orphan medicinal products. *Off J Eur Communities*, L18, pp. 1-5.
- PICAVET, E., CASSIMAN, D., HOLLAK, C. E., MAERTENS, J. A. y SIMOENS, S. (2013). Clinical evidence for orphan medicinal products - a cause for concern? *Orphanet J Rare Dis.*, 8(1), p. 164. doi: 10.1186/1750-1172-8-164.
- POWERS, J. H., PATRICK, D. L., WALTON, M. K., MARQUIS, P., CANO, S., HOBART, J. *et al.* (2017). Clinician-Reported Outcome Assessments of Treatment Benefit: Report of the ISPOR Clinical Outcome Assessment Emerging Good Practices Task Force. *Value Health*, 20(1), pp. 2-14. doi: 10.1016/j.jval.2016.11.005.
- REEVE, B. B., WYRWICH, K. W., WU, A. W., VELIKOVA, G., TERWEE, C. B., SNYDER, C. F., *et al.* (2013). ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res*, 22, pp. 1889–1905. doi: 10.1007/s11136-012-0344-y.
- RICHEY, E. A., LYONS, E. A., NEBEKER, J. R., SHANKARAN, V., MCKOYM J. M. *et al.* (2009). Accelerated approval of cancer drugs: improved access to therapeutic breakthroughs or early release of unsafe and ineffective drugs? *J Clin Oncol.*, 27(26), pp. 4398–4405. doi: 10.1200/JCO.2008.21.1961.
- ROTHMAN, M. L., BELTRÁN, P., CAPPELLERI, J. C., LIPSCOMB, J. y TESCHENDORF, B. (2007). Patient-reported outcomes: conceptual issues. *Value Health*, 10(Suppl 2), pp. S66-S75.
- SLADE, A., ISA, F., KYTE, D., PANKHURST, T., KEREČUK, L., FERGUSON, J. *et al.* (2018). Patient reported outcome measures in rare diseases: a narrative review. *Orphanet Journal of Rare Diseases*, 13, p. 61. doi: 10.1186/s13023-018-0810-x.
- SUN, W., ZHENG, W. y SIMEONOV, A. (2017). Drug discovery and development for rare genetic disorders. *Am J Med Genet Part. A*, pp. 1-16.
- U.S. FOOD AND DRUG ADMINISTRATION. (2019). *Rare Diseases: Common Issues in Drug Development Guidance for Industry. Revision 1*. Center for Drug Evaluation and Research (CDER). Recuperado de: <https://www.fda.gov/media/119757/download>



19

INVESTIGACIÓN DE RESULTADOS EN SALUD EN NIÑOS Y ADOLESCENTES

JOSÉ LUIS REVUELTA HERRERO, SILVIA MANRIQUE RODRÍGUEZ,
CRISTINA CALVO REY, CECILIA MARTÍNEZ FERNÁNDEZ-LLAMAZARES

■ 19.1. INTRODUCCIÓN

“Todos los hospitales deberían hacer un seguimiento de los pacientes que atienden hasta determinar si el tratamiento ha sido un éxito, y si no es así, preguntarse por qué, con el objetivo de prevenir fallos semejantes en el futuro”. Esta cita pertenece al cirujano originario de Boston Ernest Amory Codman. A pesar de que se encuentra hoy día de plena actualidad, en realidad data de 1914, lo que pone de manifiesto que la preocupación por los resultados que tienen las intervenciones sanitarias en los pacientes no es algo nuevo y es intrínseca a la actividad asistencial. Recientemente, la medición y análisis de los resultados en salud se ha establecido como una herramienta prioritaria para avanzar hacia el conocido como triple objetivo de los sistemas sanitarios: mejorar la salud de la población, ofrecer una mejor atención sanitaria a los pacientes y reducir los costes de esta atención.

Como ya se ha expuesto en capítulos anteriores de este libro, la investigación de resultados en salud se diferencia de la investigación clínica, principalmente los ensayos clínicos aleatorizados, en que se realiza en condiciones de práctica clínica habitual y pueden tener un horizonte temporal más lejano. Se dirige, por tanto, a determinar la efectividad de las intervenciones sanitarias (por ejemplo, la utilización de un determinado medicamento, la realización de una cirugía con una técnica laparoscópica o la aplicación de un programa de seguimiento del paciente trasplantado pediátrico). La investigación de resultados en salud también se distingue de los estudios puramente fármaco-económicos, los cuales buscan determinar la eficiencia de las intervenciones, aunque también puede integrar costes además de resultados en sus análisis.

Los resultados en salud se han ido incorporando de forma progresiva a la investigación. Esto se debe en parte al desarrollo de una visión más amplia del concepto de salud y que también requiere de la participación del paciente con sus perspectivas. Aunque en áreas como la oncología o las enfermedades autoinmunes es difícil encontrar estudios de investigación en los que no se midan resultados en salud, es cierto que en la investigación en pediatría existe un cierto retraso. En este capítulo se revisarán algunas de las áreas en las que se centra la investigación de resultados en salud en niños y adolescentes:

- Efectividad y costes de las intervenciones sanitarias.
- Preferencias y prioridades de los pacientes.

- Resultados percibidos por los pacientes (*PRO*) y calidad de vida.
- Calidad de la asistencia sanitaria.

La investigación de resultados en salud para que sea verdaderamente útil y nos ayude a incorporar la percepción del paciente sobre el impacto de las intervenciones sanitarias en su salud, nos plantea algunos retos. En el caso de la investigación en niños y adolescentes existen algunas consideraciones previas que son intrínsecas a estas poblaciones y que se han señalado como barreras en algunos casos para el desarrollo de la investigación de resultados en salud en pediatría:

- La relación de dependencia con los progenitores o cuidadores: por ejemplo, a la hora de acceder a los recursos sanitarios, identificar las necesidades y resolverlas, exposición a factores de riesgo, manejo de los estados de salud y enfermedad, y la prevención. No debe olvidarse tampoco el impacto de la familia en la adherencia a las intervenciones sanitarias (tratamiento farmacológico, educación para la salud etc.), así como la capacidad del entorno para dar soporte económico a estas intervenciones.
- Los cambios cognitivos, físicos y sociales que se producen durante la maduración del niño, que ocurren a gran velocidad en comparación con los adultos, y que presentan una gran variabilidad interindividual. La morbilidad generada durante la infancia se manifiesta igualmente durante la adolescencia y, posteriormente, la edad adulta. La evaluación de las intervenciones sanitarias debería tener un horizonte temporal suficiente que permitiera valorar los resultados y los costes a largo plazo.
- Las diferencias epidemiológicas con los adultos. Menos del 5 % de los niños sufren enfermedades como diabetes, enfermedad cardiovascular o artritis, consideradas habitualmente como enfermedades de adultos. Por el contrario, las patologías más comunes en la población pediátrica son agudas y se pueden considerar menores. Las enfermedades graves son poco frecuentes y están relacionadas con patologías congénitas o hereditarias, o con el cáncer (leucemias, linfoma y tumores del sistema nervioso central). Estos pacientes son responsables de una parte importante del gasto puesto que requieren un manejo especializado por parte de equipos multidisciplinarios y que integren distintos cuidados (Forrest, Simpson y Clancy, 1997; Forrest, 2004).
- La importancia del contexto sociocultural en el desarrollo físico, psicológico y social de los niños y adolescentes. Es necesario considerar su entorno: sus padres, su familia, sus compañeros de colegio y el resto de la comunidad. Las relaciones en estos ámbitos tendrán un impacto directo sobre aspectos como el comportamiento, sus capacidades cognitivas, afectivas y su desarrollo social, que son dimensiones que se evalúan en multitud de cuestionarios sobre calidad de vida relacionada con la salud (Matza *et al.*, 2004).

Nuestro sistema sanitario se enfrenta a retos muy importantes, como el envejecimiento de la población o la cronificación de las patologías, que tienen un impacto alto en la sostenibilidad del sistema. Por este motivo, debemos avanzar hacia una asistencia sanitaria basada en resultados y que busque generar valor para los pacientes. Es por esto que necesitamos disponer de medidas, como la calidad de vida relacionada con la salud, que nos permitan incorporar las percepciones de los pacientes. En este sentido, la Convención de Naciones Unidas sobre los Derechos del Niño de 1989, que reconoce el derecho a la salud, refleja en su artículo 12 que la opinión de los niños y los adolescentes debe de ser tenida en cuenta en aquellos aspectos que les afecten.

■ 19.2. EFECTIVIDAD Y COSTES DE LAS INTERVENCIONES SANITARIAS

La principal fuente de que disponemos para conocer la eficacia de las intervenciones sanitarias, es decir los resultados que se obtienen al realizar una determinada intervención sobre una población, son los ensayos clínicos. Sin embargo, sabemos que las condiciones en las que se realizan los ensayos clínicos difieren de la práctica clínica habitual (poblaciones homogéneas de pacientes seleccionados, sin comorbilidades que podrían ser frecuentes en la población general y con niveles de adherencia sobreestimados). Por otro lado, nos encontramos con que la población pediátrica ha sido excluida de los ensayos clínicos de forma sistemática. Los datos sobre dosificación, eficacia y seguridad con frecuencia se extrapolan de los obtenidos en adultos o proceden de estudios de baja calidad. En el campo de los medicamentos, se sabe que los niños presentan características farmacocinéticas y farmacodinámicas que los diferencian de los adultos y que pueden determinar un resultado diferente del binomio beneficio/riesgo estimado para los adultos. Una población especialmente vulnerable son los neonatos, en los que apenas hay investigación. En las últimas décadas las autoridades norteamericana y europea han incorporado diferentes regulaciones con el objetivo de promocionar la investigación de calidad en la población pediátrica. Por su parte, la Unión Europea aprobó los reglamentos 1901/2006 y 1902/2006 para estimular la investigación y el desarrollo de medicamentos pediátricos, promoviendo la disponibilidad de medicamentos autorizados específicamente para esta población. Desde la entrada en vigor de esta legislación, el número de nuevas indicaciones o medicamentos autorizados por el procedimiento centralizado se duplicó con respecto a periodos anteriores (European Medicines Agency y Paediatric Committee, 2016).

A pesar estos esfuerzos, la investigación clínica sigue estando muy por detrás de la realizada en población adulta. Por este motivo, hasta un 79 % de los medicamentos de los medicamentos utilizados en los hospitales se hace fuera de ficha técnica (Shah *et al.*, 2007). Entre las barreras para la realización de ensayos clínicos en niños se encuentran los condicionantes éticos y la baja prevalencia de las enfermedades pediátricas, que encarecen los ensayos y ofrecen finalmente un mercado reducido para la industria farmacéutica. Por otro lado, no debe olvidarse

que como en otros ámbitos, los ensayos clínicos en pediatría muestran un patrón desigual. Reflejan patologías frecuentes en los países desarrollados (como el asma o la migraña, que ofrecen resultados más rentables para la industria) y olvidan otras patologías (por ejemplo, la malaria) que son más prevalentes en países subdesarrollados y responsables en realidad de la mayor carga de enfermedad pediátrica (Bourgeois *et al.*, 2012).

Todos estos aspectos no hacen sino subrayar la importancia que deben tener los estudios de efectividad, es decir, aquellos que buscan determinar el efecto que tienen las intervenciones sanitarias en condiciones de práctica clínica, en la población pediátrica. Sin embargo, es habitual que la evidencia sobre las intervenciones en esta población sea limitada incluso aunque su aplicación esté muy extendida (Moyer y Butler, 2004). Algunas de las barreras para la realización de este tipo de estudios ya se han descrito anteriormente:

- La dificultad para alcanzar tamaños muestrales suficientemente grandes que permitan obtener resultados significativos.
- El alto coste de estos estudios, puesto que frecuentemente se necesitan tiempos de seguimiento altos para conseguir una visión global del resultado de una intervención sanitaria y no solo una fotografía instantánea.
- El problema de que la investigación tiene como objetivo principal prevenir o curar una determinada enfermedad, no a mejorar la calidad de vida en global.
- La falta de referencias bibliográficas y metodológicas para la investigación en resultados en salud en niños y adolescentes, puesto que la mayoría se ha realizado en patologías propias de los adultos.

Metodológicamente, este tipo de estudios presentan una serie de requerimientos que se describen con acierto en el trabajo de Kaplan *et al.*, (2001):

Definición de la intervención sanitaria: puede incluir aspectos específicos como un fármaco, un procedimiento quirúrgico o un test diagnóstico y otros más complejos, como el impacto de un programa de atención al paciente diabético pediátrico que integre diferentes medidas y profesionales (médicos, farmacéuticos, enfermería y trabajadores sociales). Esta definición debe ser lo más concisa y precisa posible.

Elección del resultado en salud: La OMS definió resultado como un “cambio en la salud de un individuo, un grupo de personas o una población que se atribuya a una determinada intervención o serie de intervenciones”. En la práctica se distingue entre tres tipos de resultados: resultados clínicos, económicos y humanísticos. Los resultados clínicos son los más frecuentemente utilizados en los ensayos clínicos y otros tipos de estudios, e incluyen resultados de pruebas diagnósticas o analíticas, sintomatología, evolución de la enfermedad, efectos adversos de las intervenciones etcétera. El objetivo de algunos tipos de intervenciones sanitarias (por ejemplo, un programa de educación nutricional en familias de niños con obesidad) no es siempre

el niño, por lo que si se incluyen los padres u otros familiares, deben incluirse medidas de resultados apropiadas para reflejar los cambios de la intervención en este grupo.

Elección del grupo control: Desafortunadamente los ensayos clínicos aleatorizados pragmáticos son poco frecuentes y la mayoría de la investigación en efectividad se realiza con diseños no experimentales (mayoritariamente observacionales no controlados). Dado que no se utiliza placebo, el control es el tratamiento considerado estándar y su elección es fundamental para la aplicabilidad de las conclusiones del estudio.

Ajuste de riesgos: Los distintos perfiles de pacientes deben tenerse en cuenta a la hora de realizar comparaciones para que sean significativas. Se deben ajustar los resultados con las variaciones en las características de los pacientes que pueden darse, por ejemplo, entre dos hospitales que atiendan casos con diferente severidad o casuística (*case-mix*).

En definitiva, podemos concluir que la calidad de los estudios de efectividad y el interés de las conclusiones van a depender de los siguientes aspectos:

- El conocimiento de cómo las variables subrogadas se relacionan con los resultados en salud que realmente nos interesan.
- La procedencia de los datos (historia clínica, bases de datos, programas de prescripción farmacológica o registros de dispensación) y las limitaciones que tienen.
- La consideración de las posibles limitaciones (sesgos a la hora de la selección de la población, obtención de datos, etc.).

El interés de los costes para la determinación de los resultados económicos en salud se ha revisado ampliamente en otros capítulos de este libro. Muchas de las dificultades a la hora de definir los costes directos e indirectos tienen que ver con las particularidades de la población pediátrica descritas en la introducción. Por ejemplo, en el caso de una enfermedad crónica en un niño, deberá tenerse en cuenta el impacto económico en términos de productividad a lo largo de toda su vida. Este concepto de productividad no se debe limitarse a la capacidad de trabajo en la edad adulta, sino como una pérdida potencial de educación, desarrollo cognitivo, conductual o social que tendrá impacto cuando sea adulto. Las necesidades (cuidadores, terapeutas, tratamientos farmacológicos etc) y la dependencia respecto a los progenitores se modificarán durante el proceso de crecimiento y maduración del niño, por lo que se deben tener en cuenta las diferentes etapas a la hora de asignar los costes (Ungar, 2007).

■ 19.3. PREFERENCIAS Y PRIORIDADES DE LOS PACIENTES

Los estudios de preferencias y prioridades de los pacientes comparten con la medición de resultados en salud referidos por los pacientes el objetivo de que

la atención sanitaria gire realmente en torno al paciente, haciéndose consciente de sus necesidades y ofreciéndole soluciones que le sean valiosas. Sin embargo, tienen un enfoque diferente que se debe al origen económico de los métodos que aplicados en los estudios de preferencias de los pacientes. Mientras que los estudios de resultados reportados por los pacientes reflejan los resultados obtenidos en un dominio particular, las metodologías de preferencias de los pacientes buscan dilucidar las prioridades y deseos de los pacientes de forma transversal en varios dominios. De esta forma, el valor de un determinado resultado en salud dependería de las prioridades y preferencias de los pacientes.

Fundamentalmente, existen dos metodologías para la realización de estudios de preferencias de los pacientes:

- *Método de las preferencias reveladas*: en las que el investigador estima el valor que para los pacientes tiene un determinado aspecto mediante la observación de su comportamiento en condiciones de práctica clínica habituales.
- *Método de las preferencias expresadas*: en los que el investigador plantea diferentes escenarios hipotéticos sobre los que los sujetos participantes en el estudio ofrecen sus preferencias. Como ya se ha revisado en este libro, existen fundamentalmente dos tipos de estudios: disponibilidad a pagar y análisis de conjunto (conjoint analysis). El análisis de conjunto es una técnica que permite estimar el impacto que un determinado atributo tiene sobre el valor global de una intervención sanitaria.

Los niños y adolescentes y sus padres pueden tener perspectivas distintas sobre el impacto de las intervenciones sanitarias sobre su calidad de vida, u opiniones diferentes sobre cuáles son los resultados más relevantes para ellos respecto de los profesionales sanitarios y los gestores. Este hecho ha sido puesto de manifiesto en diferentes estudios, como el de Kramer *et al.*, que mostró que los padres y los pediatras tenían percepciones diferentes sobre el proceso diagnóstico en niños con fiebre sin foco que acudían a un servicio de urgencias. Los padres concedían más importancia al dolor o la incomodidad de sus hijos, a la vez que relativizaban el riesgo de complicaciones poco frecuentes pero más graves (por ejemplo: pérdida de la capacidad auditiva por meningitis bacteriana) frente a la realización de pruebas diagnósticas molestas (hemograma, radiografía o sonda urinaria) (Kramer *et al.*, 1990).

La toma de decisiones clínicas se basa en la valoración del beneficio frente al riesgo de una determinada intervención sanitaria y también está influenciada por las perspectivas de los niños, los padres y los profesionales sanitarios. Este es el caso de las decisiones terapéuticas en el campo de la oncología pediátrica. La baja incidencia de las enfermedades oncohematológicas pediátricas hace que en muchas ocasiones no exista una evidencia lo suficientemente sólida que permita hacer recomendaciones robustas. Pueden existir grupos de riesgo que aún no estén

definidos y que modifiquen el resultado de las intervenciones sanitarias. Por otro lado, al contrario que en los adultos, el objetivo principal es la curación, lo cual se alcanza en un porcentaje alto de pacientes. Sin embargo, es habitual la aparición de toxicidades que pueden condicionar su supervivencia o calidad de vida a largo plazo. El trabajo de Tomlinson *et al.* (2011) sobre la utilización de quimioterapia paliativa frente a los cuidados de soporte en pacientes oncológicos pediátricos puso de manifiesto las diferencias existentes entre los profesionales sanitarios y los padres. A pesar de la opinión más favorable de los profesionales sanitarios sobre los cuidados paliativos de soporte, los padres preferían la utilización de quimioterapia incluso aunque pudiera reducir la calidad de vida de los niños. El estudio concluye que la esperanza de la curación, a pesar del estadio, pesaría sobre otros factores como la supervivencia o la calidad de vida (Tomlinson *et al.*, 2011). Por el contrario, el estudio de Knowles *et al.* (2001) mostró que existía un acuerdo significativo entre padres y profesionales sanitarios a la hora de valorar diferentes estados de salud en niños con cardiopatías congénitas, lo que pone de manifiesto que es posible que los profesionales, los pacientes y los padres se encuentren alineados (Knowles *et al.*, 2007).

En este sentido, la Organización Mundial de la Salud recomienda que las perspectivas de los pacientes deben incluirse en las guías de práctica clínica (Schünemann, Fretheim y Oxman, 2006). La toma en consideración de estas perspectivas se considera que aumenta la satisfacción de los pacientes con los cuidados, favorece la toma de decisiones por los padres y permite a los gestores sanitarios priorizar intervenciones sanitarias en función de los valores de los pacientes (Sung y Regier, 2013).

Finalmente, reseñar algunas consideraciones que deben tenerse en cuenta a la hora de realizar e interpretar los estudios en este ámbito:

- Los valores pueden no ser estáticos a lo largo de toda la enfermedad, y variar en función de la edad, el pronóstico o la situación clínica.
- Los pacientes pediátricos pueden presentar deficiencias a la hora de expresar información sobre su estado de salud o dificultades para comprender el balance beneficio riesgo de una intervención sanitaria. Además, los aspectos culturales, religiosos o educativos pueden influir también en este ámbito.

■ 19.4. RESULTADOS PERCIBIDOS POR LOS PACIENTES Y CALIDAD DE VIDA

■ 19.4.1. ¿Qué son los resultados percibidos por los pacientes?

Como se ha desarrollado previamente a lo largo de este libro, los resultados se pueden clasificar en clínicos (supervivencia, recidiva, pruebas diagnósticas), eco-

nómicos (costes, ahorro) y humanísticos (percepción sobre el estado de salud). Los resultados pueden ser observados, si la comunicación la hacen los clínicos o los investigadores, o referidos por el paciente. La FDA definió un PRO (del inglés *patient reported outcome, PRO*) como cualquier informe del estado de salud de un paciente que procede directamente del paciente, sin interpretación de la respuesta por parte de un clínico u otro profesional (U.S. Department of Health and Human Services *et al.*, 2009). En pediatría también es habitual hablar de resultados referidos por los cuidadores más cercanos (en inglés: *proxys*), que actúan poniéndose en lugar del paciente. Su uso debería estar limitado a las situaciones en las que los resultados no se puedan obtener de los pacientes directamente.

Los PRO nos pueden aportar información sobre diferentes aspectos relacionados con la salud (Deshpande *et al.*, 2011):

- Sintomatología: frecuencia y severidad de los síntomas, siendo especialmente interesantes aquellos que no se pueden medir, que son subjetivos o psicológicos.
- Estatus funcional: capacidad para realizar actividades específicas para evaluar el funcionamiento cognitivo o físico y evaluar la gravedad de la discapacidad.
- Estilos de vida y conductas para la salud: por ejemplo, hábitos de vida saludables o adherencia al tratamiento.
- Impacto de la enfermedad en la calidad de vida, calidad de vida relacionada con la salud
- Satisfacción del paciente con los cuidados (por ejemplo, los tratamientos farmacológicos) y la experiencia del paciente.

Otro concepto relacionado es la calidad de vida relacionada con la salud (en inglés *health related quality of life, HRQoL*), que incluye todos aquellos aspectos que tienen impacto sobre la salud de una persona, y formaría parte de un concepto más amplio que es la calidad de vida.

■ 19.4.2. ¿Por qué medir resultados percibidos por los pacientes?

Existe evidencia científica que muestra que las percepciones de los profesionales sanitarios sobre la gravedad de los síntomas que sufre un paciente pediátrico pueden no concordar con lo referido por los propios pacientes. Por ejemplo, el estudio de Hockenberry *et al.* (2003) en pacientes pediátricos oncológicos que estaba recibiendo tratamiento con quimioterapia, mostró una escasa concordancia entre lo comunicado por los pacientes, los padres y los profesionales sanitarios respecto a la gravedad de la fatiga (Hinds *et al.*, 1999). Otro estudio mostró diferencias entre adolescentes y sus padres a la hora de valorar la calidad de vida relacionada con la

salud, de forma que estos últimos valoraban como peor el estado de salud de sus hijos adolescentes que los propios pacientes (Huang *et al.*, 2009). Como clínicos, estos ejemplos nos hacen pensar que, obviando la información comunicada por los propios pacientes, es posible que no estemos respondiendo a sus necesidades. Es por esto que la medición de PRO ocuparía un lugar fundamental en el triple objetivo de los sistemas sanitarios que se había comentado al inicio del capítulo.

■ 19.4.3. ¿Cómo medir resultados percibidos por los pacientes?

La medición de resultados percibidos por el paciente se realiza mediante cuestionarios que incluyen una o varias dimensiones (bienestar psicológico, función física, relaciones sociales, función cognitiva...) y deben cumplir una serie de requisitos psicométricos que permitan garantizar la calidad de la información recogida. Entre estos criterios se incluyen la fiabilidad (la capacidad de medir con precisión un determinado concepto), la validez, la sensibilidad y la capacidad de ser utilizado en condiciones de la vida real.

De forma general, los conceptos que se incluyen en los instrumentos para niños y adolescentes son similares a los que se utilizan en los dirigidos a la población adulta. Sin embargo, los intentos de aplicar estos instrumentos para adultos en la población pediátrica han ofrecido resultados deficientes que ponen de manifiesto la necesidad de adaptarlos. Es cierto que la investigación de resultados en salud en niños y adolescentes plantea algunos retos, pero utilizando instrumentos adecuados, existe evidencia que indica que los niños son capaces de comunicar aspectos concretos sobre su salud desde los siete años y, posiblemente, incluso desde los cinco años (Eiser y Morse, 2001). Sin embargo, no debe obviarse que se han identificado varios factores que pueden tener influencia en la validez de estas respuestas. Entre ellos se encuentra una particularidad de la población pediátrica y es que los niños y los adolescentes presentan diferentes niveles de comprensión y expresión de conceptos abstractos en función de su desarrollo cognitivo. La edad, con variaciones interindividuales, e incluso el género pueden ser factores determinantes en este nivel de desarrollo. El documento de buenas prácticas que elaboró ISPOR en 2013 sobre la medición de PRO en pacientes pediátricos revisa la evidencia disponible en cuatro grupos de edad, si bien reconoce que los puntos de corte son artificiales y que debería hacerse una evaluación individualizada de las capacidades cognitivas del niño y su capacidad para expresar conceptos, más o menos abstractos sobre su estado de salud (Matza *et al.*, 2013):

- Niños menores de cinco años: afirma que no existe evidencia que soporte el uso de la información comunicada por el propio paciente, por lo que deberán utilizarse variables clínicas o las observaciones realizadas por los progenitores o profesionales implicados.

- Niños entre cinco y siete años: existen resultados dispares sobre la medición de PRO en este rango de edad. Aunque los niños de esta edad son capaces de expresar de forma sencilla información sobre el dolor o las náuseas, pueden tener problemas de léxico y comprensión conceptos abstractos más complejos. Existen cuestionarios, tanto generales como específicos, que pretenden cubrir esta población. Por ejemplo, el instrumento *Pediatric Quality of Life Inventory* (PedsQL), que se revisará más adelante, incluye este grupo de edad con buenos resultados en cuanto a validez interna. Sin embargo, otros instrumentos no han mostrado esta robustez a la hora de aplicarse en este rango de edad
- Niños entre ocho y once años: a diferencia del punto anterior, existen trabajos que han demostrado que los niños incluidos en este rango de edad presentan una buena capacidad de comprensión para la mayoría de los conceptos relacionados con la salud, que son similares a los alcanzados en edades más avanzadas. Siempre que se pueda asegurar la correcta comprensión de los conceptos del instrumento elegido, se recomienda la utilización de los datos referidos por los propios pacientes en este rango de edades.
- Adolescentes entre doce y dieciocho años: el reto en este grupo es definir qué información es realmente relevante para esta población. Aunque en algunos aspectos se pueden asemejar a los adultos, algunos estudios han puesto de manifiesto que existen diferencias en la información que deben abarcar los dominios. Dada la heterogeneidad de este rango de edad, se recomienda la utilización de cuestionarios apropiados para los adolescentes evitando la aplicación directa de instrumentos dirigidos a adultos sin realizar una validación previa.

Otro aspecto que puede influir en las propiedades psicométricas de los instrumentos es el formato del cuestionario. Respecto a la forma de administración, se distingue entre cuestionarios autoaplicados y los realizados por un entrevistador. En principio, los preferibles son los primeros, puesto que evitan que el entrevistador pueda introducir sesgos. Los cuestionarios administrados por un entrevistador pueden ser una opción razonable cuando la capacidad de comprensión (discapacidad intelectual, analfabetismo o bajo nivel cultural) o de cumplimentación del cuestionario del sujeto pueda estar comprometida (por ejemplo, discapacidad visual o incapacidad de concentración en niños con trastorno por déficit de atención e hiperactividad). Igualmente, el tiempo necesario para completar el cuestionario también es un factor a tener en cuenta. En este sentido, el nivel de comprensión lectora es un punto clave y debe considerarse en el contexto de la población de estudio. A la hora de diseñar un cuestionario puede ser necesario traducir la jerga médica a conceptos que los niños puedan entender más fácilmente (por ejemplo: la sensación de náuseas podría expresarse como “tener el estómago revuelto” y el dolor abdominal como “dolor de tripa”). Si el cuestionario es autoaplicado deben adaptarse igualmente las instrucciones para completarlo y las respuestas, no solo los enunciados. La mayoría de los cuestionarios utilizan escalas ordinales crecientes

(tipo Likert), escalas con expresiones faciales o, menos frecuentemente, escalas visuales analógicas. Existe consenso sobre la utilidad de las escalas Likert en este tipo de estudios, si bien parece que los niños más pequeños (de cuatro a seis años) tienen tendencia a marcar respuestas más polarizadas, por lo que se recomienda utilizar ilustraciones que sirvan como ayuda. El periodo de tiempo sobre el que se hace la evaluación, un parámetro definido por cada cuestionario, es otro factor a tener en cuenta. Mientras que los niños de ocho años son capaces de recordar datos de forma fiable de hasta cuatro semanas, niños más pequeños pueden tener problemas con periodos superiores a una semana (Matza *et al.*, 2013). Por todos estos motivos, la adecuación del formato de los cuestionarios a la población debe comprobarse antes de la realización de un estudio.

Existen discrepancias sobre la utilización de la información proporcionada por los representantes (progenitores o cuidadores cercanos) frente a la comunicada por los propios pacientes en niños y adolescentes. Esta concordancia puede depender de factores como el dominio que se está estudiando (es mayor para los dominios que implican funcionalidad y menor para los más subjetivos), la edad del niño o la presencia o ausencia de enfermedad. Además, parece razonable pensar que la información procedente de estos cuidadores cercanos estará sesgada por el impacto que les produce a ellos mismos la circunstancia de su hijo (Matza *et al.*, 2013; Janssens *et al.*, 2015). Tanto la EMA como la FDA indican en sus documentos de recomendaciones que los *PRO* deberían recogerse directamente del paciente, siendo conscientes de las dificultades que esto tiene en algunos grupos de edad (U.S. Department of Health and Human Services *et al.*, 2009; Committee for Medicinal Products for Human Use (CHMP) y European Medicines Agency (EMA), 2016 y European Medicines Agency, Paediatric Committee, 2016). Como se ha comentado, la calidad de vida relacionada con la salud es multidimensional, y aunque algunas dimensiones son las mismas que en adultos (función física, relaciones sociales, bienestar psicológico etc.), los aspectos concretos que abarcan son diferentes. De la propia definición de calidad de vida relacionada con la salud se desprende que es subjetiva, por lo que debería medirse desde la perspectiva del paciente.

■ 19.4.4. Tipos de instrumentos:

Cuestionarios genéricos: abarcan aspectos variados del estado de salud, por lo que incluyen la mayoría de las dimensiones. Se pueden utilizar tanto en pacientes como en la población general, por lo que permiten realizar comparaciones. Habitualmente son menos sensibles a los cambios en el estado de salud, por lo que pueden no ser los más útiles para evaluar el impacto de una intervención sanitaria. Algunos ejemplos de instrumentos genéricos son KIDSCREEN o PedsQL. Existen repositorios como biblioPRO¹ donde se pueden encontrar cuestionarios tanto genéricos como específicos.

¹ www.bibliopro.org

KIDSCREEN es un instrumento multidimensional desarrollado en trece países europeos para la evaluación del estado de salud y el bienestar subjetivo en niños y adolescentes (de 8 a 18 años) sanos o con enfermedades crónicas. Consta de tres cuestionarios:

- El cuestionario KIDSCREEN-52 abarca 10 dimensiones (bienestar físico, bienestar psicológico, estados de ánimo, autopercepción, autonomía, relación con los padres y vida familiar, recursos económicos, amigos y apoyo social, entorno escolar y aceptación social) en 52 preguntas.
- El cuestionario KIDSCREEN-27 es una versión resumida del anterior, que abarca 5 dimensiones en 27 preguntas, por lo que el tiempo necesario para completarlo es menor.
- El cuestionario KIDSCREEN-10 se deriva a su vez del cuestionario anterior y está formado por 10 preguntas. Su utilidad principal es como herramienta de cribado poblacional.

Entre las ventajas de KIDSCREEN están el ser cuestionarios transculturales, por lo que pueden servir para comparar países diferentes, y que ha demostrado una validez y fiabilidad aceptables. Todos los cuestionarios están disponibles en una versión para los padres y otra para los niños.

El instrumento PedsQL, por el contrario, fue desarrollado en Estados Unidos y se dirige a niños y adolescentes de 5 a 12 años tanto sanos como con patologías agudas y crónicas. El cuestionario genérico (*23-item PedsQL Generic Core Scales*) dispone de dos versiones, una para la autoadministración por los niños (de 5 a 18 años) y otra para los padres (2-18 años). Es multidimensional, abarcando el funcionamiento físico, emocional, social y escolar. También se han desarrollado múltiples módulos específicos de cada patología (artritis, asma, diabetes, cáncer y enfermedades cardíacas) que complementan el cuestionario genérico. Cabe destacar que ha demostrado tener una fiabilidad y validez adecuadas, y se encuentra disponible en multitud de idiomas, entre ellos el castellano.

Los cuestionarios de utilidad son también genéricos, de forma que muestran las preferencias de los pacientes por distintos estados de salud que oscilan entre el 0 (la muerte) y el 1 (el mejor estado de salud). Estos instrumentos se han utilizado para el cálculo del año de vida ajustado por calidad (AVAC). Algunos ejemplos de cuestionarios adaptados para su uso en población pediátrica son el EQ-5D-Y y el Health Utility Index Mark 2 (HUI2).

El cuestionario EQ-5D-Y es una adaptación del instrumento EQ-5D para niños y adolescentes entre ocho y quince años. Consta de una parte descriptiva en la que se analizan cinco dimensiones (movilidad, autocuidado, actividades de la vida cotidiana, dolor, y sentimientos de tristeza, ansiedad o depresión) en una escala de tres niveles (sin problemas, algunos problemas o muchos problemas). La puntuación obtenida se complementa con una escala visual analógica en la que el encuestado

debe valorar su estado de salud entre el peor estado de salud que se pueda imaginar (cero puntos) y el mejor (cien puntos) en el día de realización del cuestionario. También existen versiones del cuestionario para cuidadores cercanos.

El instrumento HUI2 fue desarrollado originalmente para evaluar resultados y secuelas en individuos que hubiesen sufrido enfermedades oncohematológicas en la infancia. Evalúa siete dimensiones (sensaciones, movilidad, emociones, capacidad cognitiva, autocuidado, dolor y fertilidad) con escalas de entre tres y cinco niveles (desde “funcionamiento normal para su edad” a “completamente incapaz”). Se diferencia del EQ-5D-Y en que busca determinar la capacidad para funcionar en estos dominios, en lugar de valorar el desempeño en estas áreas.

Cuestionarios específicos: se centran en aspectos de un problema de salud en particular (por ejemplo: epilepsia) o una intervención sanitaria (por ejemplo: pacientes en tratamiento con quimioterapia). Se utilizan frecuentemente en el contexto de los ensayos clínicos porque permiten demostrar los beneficios de una determinada intervención (son sensibles) y sirven para comparar poblaciones con el mismo problema de salud. Pueden ser instrumentos tanto multidimensionales como unidimensionales. Se han desarrollado multitud de cuestionarios específicos de patologías, algunos de los cuales se encuentran recogidos en la tabla 1.

Tabla 1.

EJEMPLOS DE CUESTIONARIOS ESPECÍFICOS DE PATOLOGÍAS

Nombre cuestionario	Edad (años)	Informantes	Dominios
<i>Alergia:</i>			
Pediatric Allergic Disease Quality of Life Questionnaire (PADQLQ)	6-16	Niños	Actividades, emociones, síntomas.
<i>Asma:</i>			
Childhood Asthma Control Test for children (C-ACT)	4-11	Niños y padres	Control del asma.
	12-18	Niños	
About my Asthma (AMA)	6-12	Niños	Calidad de vida relacionada con la salud.
Childhood Asthma Questionnaires (CAQs)	CAQ-A: 4-7	Niños	A: calidad de vida, angustia.
	CAQ-B: 8-11		B: calidad de vida (actividades), calidad de vida (actividades pasivas), angustia, gravedad de los síntomas.
	CAQ-C: 12-16		C: calidad de vida (actividades), calidad de vida (actividades sociales), angustia, gravedad de los síntomas, reactividad.

Tabla 1. (continuación)

EJEMPLOS DE CUESTIONARIOS ESPECÍFICOS DE PATOLOGÍAS

Nombre cuestionario	Edad (años)	Informantes	Dominios
Asma:			
Pediatric Asthma Quality of Life Questionnaire (PAQLQ)	7-17	Niños	Sintomatología, función emocional, limitaciones en la actividad.
Cáncer:			
Escala de Calidad de Vida para Niños Oncológicos (ECVNO)	6-18	Niños	Aislamiento social, carencias emocionales, sufrimiento emocional, obstáculos para relacionarse.
Pediatric Oncology Quality of Life Scale (POQOLS)	3-18	Padres	Función física y restricciones, angustia, respuesta al tratamiento activo.
Enfermedad inflamatoria intestinal:			
Impact Questionnaire (IMPACT)	9-18	Niños	Función intestinal, imagen corporal, deterioro social/funcional, daño emocional, tratamientos y pruebas, deterioro global.
Epilepsia:			
Impact of Childhood Neurologic Disability Score (ICND)	2-18	Padres	Comportamiento, funciones físicas y neurológicas, funciones, aprendizaje cognitivo, epilepsia.
Obesidad:			
Sizing me up	5-18	Niños	Función física, función emocional, burla y marginación, atributos positivos, rechazo social.
Sizing them up		Padres	Función física, función emocional, burla y marginación, atributos positivos, rechazo social, tiempo de la comida, colegio, adaptación y desarrollo del adolescente.
Impact of weight on quality of life-Kids (IWQOL-Kids)	11-19	Niños	Bienestar físico, autoestima corporal, vida social, relaciones familiares.

Fuentes: Adaptado de Matza *et al.*, (2004) y Solans *et al.*, (2008).

DISABKIDS es un instrumento elaborado en siete países europeos dirigido a niños con enfermedades crónicas. Consta de varios cuestionarios: DISABKIDS-37, DISABKIDS-12 y DISABKIDS-Smilely. El cuestionario DISABKIDS-37 consta de 37 puntos que evalúan seis dimensiones (independencia, emociones, inclusión social, exclusión social, limitaciones físicas y tratamientos). Existen versiones tanto para la aplicación a cuidadores cercanos como para los pacientes. El cuestionario DISABKIDS-12 es una versión abreviada del anterior, evaluando en 12 puntos las seis dimensiones anteriores. Mientras que los cuestionarios anteriores están dirigidos a

niños y adolescentes de ocho a dieciséis años, el cuestionario DISABKIDS-Smiley se puede utilizar en niños de entre cuatro y siete años. Se trata de una versión del cuestionario original adaptada al nivel cognitivo de esta población, en la que las respuestas se ofrecen en forma de caras representando emociones. Por último, este instrumento también consta de cuestionarios específicos para algunas patologías como la artritis reumatoide, el asma, la dermatitis atópica, la parálisis cerebral, la fibrosis quística, la diabetes infantil y la epilepsia.

Finalmente, reseñar que existen cuestionarios desarrollados específicamente para subpoblaciones (por ejemplo: adolescentes) que por su espectro se pueden ser igualmente genéricos o específicos. Este es el caso del instrumento *Youth Quality of Life* (YQoL), que es un cuestionario genérico multidimensional dirigido a adolescentes entre once y dieciocho años sanos o con enfermedades crónicas. Se trata de un cuestionario autoaplicado, del que también se encuentra disponible una versión reducida (YQoL-SF), y que abarca cuatro dominios: relaciones sociales, identidad personal, entorno (“oportunidades y obstáculos”) y calidad de vida global.

■ 19.4.5. ¿Cómo elegir el instrumento adecuado?

En las últimas décadas se ha producido un crecimiento exponencial del número de instrumentos para la medida de calidad de vida relacionada con la salud en niños y adolescentes, especialmente de cuestionarios específicos, lo que dificulta su elección. Hay varios puntos que deben tenerse en cuenta.

En primer lugar, la selección del instrumento debe hacerse teniendo en cuenta el objetivo del estudio y la población elegida. Como ya se ha comentado, los cuestionarios específicos son más sensibles que los genéricos, por lo que son los preferidos a la hora de evaluar los cambios ocurridos en el estado de salud o el impacto de las intervenciones sanitarias. En la práctica, es habitual que en un mismo estudio se utilicen tanto cuestionarios genéricos como específicos de la patología con el objetivo disponer de información complementaria. La realización de estudios multicéntricos internacionales plantea el reto de la adaptación de los cuestionarios a los distintos idiomas y culturas, que como ya se ha comentado es un aspecto importante en la validez de los resultados. La mejor opción en estos casos es utilizar cuestionarios que hayan sido desarrollados a la vez en varios países. Como se ha comentado, deberá priorizarse la recogida de datos comunicados por los propios pacientes frente a los comunicados por los cuidadores siempre que sea posible. Muchos instrumentos tienen cuestionarios dirigidos tanto a cuidadores como pacientes. La recogida de ambos tipos de cuestionarios podría aportar información complementaria.

En segundo lugar, los cuestionarios deberían elegirse teniendo en cuenta las dimensiones que abarcan y que estas tengan relación con el objetivo del estudio. No hay que perder la perspectiva de lo que se está evaluando, principalmente por-

que los instrumentos habitualmente incluyen aspectos subjetivos (salud percibida) y objetivos (capacidad funcional) que no deben intercambiarse. Un ejemplo de esto es la llamada “paradoja del discapacitado”, que han puesto de manifiesto diversos estudios que muestran cómo individuos con una capacidad funcional muy disminuida por enfermedades crónicas refieren un nivel alto de calidad de vida (Rosenbaum *et al.*, 2007). Muchos de los aspectos de la población que condicionan la elección del instrumento ya han sido comentados anteriormente, y entre los que se encuentran su comprensión lingüística, desarrollo cognitivo y discapacidades. No debe olvidarse el importante impacto que tiene el contexto social, económico y cultural en el que se desarrollan las medidas, y que tienen influencia sobre la participación o las propias respuestas.

Finalmente es necesario tener en cuenta las características psicométricas del instrumento. La mayoría de ellos aportan información sobre validez y consistencia interna. Sin embargo, la existencia de evidencia publicada sobre la sensibilidad o la correlación test-retest de los cuestionarios es menos frecuente (Solans *et al.*, 2008).

■ 19.4.6. *Patient-reported outcomes measurement information system (PROMIS)*

Es una iniciativa desarrollada por los National Institutes of Health (NIH) norteamericanos que tiene como objetivo proporcionar una base de datos con preguntas (*items*) sobre PRO en diferentes dominios dirigidas a pacientes crónicos adultos, adolescentes y niños. Esta base de datos, accesible de forma gratuita², contiene medidas destinadas a niños entre ocho y diecisiete años, y también para niños desde cinco años en adelante para ser contestadas por cuidadores cercanos. Inicialmente ofrecía cuestiones genéricas que evaluaban el bienestar físico, mental y social, que podían ser utilizados para la realización de medidas que abarcaban varios dominios y subdominios (fatiga, movilidad, intensidad del dolor, ansiedad, relaciones sociales...) y que permitían su uso en diferentes patologías al mismo tiempo. Esta base de datos se puede utilizar para como escalas o test breves (si solo se utilizan cuestiones pertenecientes a un único dominio), para la realización de perfiles de salud (si se incorporan combinaciones de cuestiones procedentes de diferentes dominios en un mismo instrumento) o como banco de preguntas para elaborar un test adaptativo informatizado (*computer adaptive test, CAT*). Para la población pediátrica existen tres instrumentos de perfiles de salud, PROMIS-25, PROMIS-37 y PROMIS-49, que se diferencian en el número de preguntas de cada dominio que incluyen. También existen los perfiles globales de salud PROMIS centrados en la función física y cognitiva que constan de siete o nueve preguntas únicamente. Los tests adaptativos informatizados utilizan un banco de cuestiones de la librería PROMIS para construir una prueba personalizada, de forma que cada pregunta se selecciona en base a lo respondido anteriormente por el encuestado según un algoritmo, de forma que se alcanzan altos grados de fiabilidad y eficiencia.

² www.healthmeasures.net

■ 19.4.7. Experiencias midiendo resultados percibidos por el paciente en la práctica clínica diaria

Experiencia KLIK ePROfile: se trata de un proyecto llevado a cabo en cuatro hospitales pediátricos de Ámsterdam en el que participaron niños y adolescentes con artritis juvenil idiopática. Los pacientes (8-18 años) o sus padres (0-7 años) recibían un enlace a la página web KLIK³ donde podían completar un cuestionario electrónico de recogida de PRO. Los instrumentos utilizados fueron TAPQOL (niños entre 0-5 años) y PedsQL (resto de edades). Los clínicos podían utilizar estas respuestas en la consulta para hacer un seguimiento de los resultados. La implantación de la medición de PRO en estos pacientes pediátricos mejoró la comunicación entre los profesionales, los niños y sus padres (Haverman *et al.*, 2013).

Experiencia en el Cincinnati Children's Hospital Medical Center: desde 2011 este hospital pediátrico tiene un programa institucional de medición de PRO como parte de la actividad clínica. La recogida de resultados tiene lugar en las consultas, de forma que los pacientes pueden rellenar los cuestionarios en los dispositivos electrónicos disponibles. Tanto los instrumentos (genéricos o específicos de una patología) como la frecuencia se han consensuado con el equipo clínico, de forma que los pacientes son notificados de la necesidad de rellenar los cuestionarios cuando acuden a la consulta. Los resultados se integran en la historia clínica electrónica, por lo que están disponibles para que el profesional sanitario y el paciente puedan discutir sobre ellos (Gerhardt *et al.*, 2018).

■ 19.4.8. Resultados percibidos por el paciente en el contexto de los ensayos clínicos

Si bien es cierto que las decisiones regulatorias y de financiación de los medicamentos en el pasado se han tomado teniendo como referencia los resultados de eficacia y seguridad demostrados en los ensayos clínicos, la tendencia actual es a incorporar otras medidas para la toma de estas decisiones. En este contexto surge el interés por demostrar el valor de los medicamentos, especialmente en los innovadores, de forma que se perciban como una inversión y no como un gasto. La FDA presentó en 2009 una guía con recomendaciones para la utilización de resultados en salud referidos por los pacientes en los procesos de desarrollo y autorización de los medicamentos, que incluye un apartado específico para niños y adolescentes (U.S. Department of Health and Human Services *et al.*, 2009). Igualmente, la EMA ha elaborado una guía que ofrece recomendaciones sobre la utilización de PRO en el desarrollo de fármacos oncológicos (Committee for Medicinal Products for Human Use (CHMP) y European Medicines Agency (EMA), 2016; European Medicines Agency y Paediatric Committee, 2016).

Los beneficios para la industria farmacéutica de incluir resultados en salud durante el proceso de desarrollo del medicamento son claros: poner en valor el

³ www.hetklikt.nu

producto y favorecer su acceso al mercado en un contexto de reducción de costes y disponibilidad de alternativas. Sin embargo, la adopción de estas estrategias en los ensayos clínicos pediátrica es lenta. Para los profesionales sanitarios, gestores e incluso para los propios pacientes supone avanzar en el conocimiento del impacto de las intervenciones sanitarias en la calidad de vida de los pacientes pediátricos, es decir, determinar la relación las variables subrogadas (por ejemplo, parámetros analíticos) y su significancia clínica (impacto en la calidad de vida del paciente).

■ 19.5. LOS RESULTADOS PERCIBIDOS POR LOS PACIENTES Y LA CALIDAD EN PEDIATRÍA

Este capítulo finaliza recordando nuevamente al cirujano Ernest Codman que, después de graduarse en Harvard y ejercer como cirujano durante varias décadas, fundó el Comité para la Estandarización Hospitalaria asociado al Colegio Americano de Cirujanos en 1910. El objetivo de este comité era recoger y estudiar resultados en salud de centros hospitalarios y evaluar cómo se podían mejorar.

Desde entonces, los clínicos nos hemos ido familiarizando con los indicadores de calidad centrados en los procesos o en la actividad (reingresos, mortalidad, porcentaje de niños vacunados), que responden a la visión de estas organizaciones tan complejas que son los sistemas sanitarios. Sin embargo, nuestro compromiso debe ser con la mejora de la vida de los pacientes. Esto significa que deberíamos conocer el impacto de nuestras intervenciones sobre, por ejemplo, el dolor, la función física o el desarrollo social del niño. En el momento actual, y teniendo presente el triple objetivo de los sistemas sanitarios, numerosas instituciones han reconocido la necesidad de incorporar los resultados referidos por los pacientes a los procesos de mejora continua. Esta información podría servir para ofrecer una información más transparente a los pacientes, de forma que puedan efectuar decisiones sobre su salud de una forma más fundamentada. Por otro lado, estos datos serían una fuente de información de gran utilidad para los clínicos que permitiría reducir la variabilidad clínica no justificada. Finalmente, estas medidas podrían ser utilizadas por los gestores como herramienta para planificar y asignar los recursos del sistema sanitario eficientemente.

La recogida de resultados referidos por los pacientes es un proceso laborioso y existen pocas experiencias de recogida rutinaria y sistemática de *PRO*. El National Health Service inglés (NHS), a través de su iniciativa NHS *PROM* Programme, promueve la recogida de mediciones de resultados referidos por los pacientes en artroplastia de rodilla y cadera desde 2009, con el objetivo de determinar el impacto de estas intervenciones sanitarias incluyendo la perspectiva de los propios pacientes. Igualmente están trabajando en recoger de forma sistemática resultados referidos en enfermedades crónicas como la EPOC. Algunos ejemplos de experiencias en la recogida de *PRO* sistemática son los registros nacionales suecos de calidad en reumatología y nefrología pediátricas (Nilsson, Orwelius y Kristenson, 2016). En el caso

norteamericano, existe la Children's Health Insurance Program Reauthorization Act (abreviada como CHIPRA), que desde 2009 promovió la elaboración por parte del Department of Health and Human Services de indicadores de calidad de la asistencia sanitaria a niños y adolescentes. A pesar de que esta ley requería que estos indicadores abarcaran los distintos ámbitos de la calidad asistencial, la versión de 2018 incluye veintiséis indicadores de los que veinticinco son indicadores de actividad o cobertura sanitaria (por ejemplo: porcentaje de niños inmunizados o visitas a urgencias) y solo uno recoge la experiencia del paciente. En España, comunidades autónomas como la de Madrid o Cataluña disponen de observatorios de resultados en salud. Sin embargo, incluyen pocos indicadores específicos para la población pediátrica (porcentaje de niños inmunizados o satisfacción con el pediatra asignado).

Algunas de las barreras para la incorporación de este tipo de medidas al ámbito de la gestión sanitaria y los programas de mejora de la calidad ya se han expuesto en este capítulo: dudas sobre la calidad de la información recogida en la población pediátrica, abundancia de instrumentos o el coste de realizar estas medidas. Otras limitaciones que se han descrito son:

- Privacidad y uso de los datos recogidos: por un lado, la recogida y uso posterior de los datos debería cumplir tanto lo recogido en la Ley Orgánica de Protección de Datos y la Ley 41/2002 de Autonomía del Paciente. Respecto al uso de la información referida por los pacientes, no puede limitarse al cálculo de indicadores. En nuestra opinión, no sería éticamente admisible recoger toda esta información y no incorporarla al proceso clínico para ofrecer al paciente una mejor atención sanitaria.
- Falta de consenso sobre la estandarización de estas medidas de resultados: por el momento siguen existiendo discrepancias sobre la terminología en resultados en salud, las metodologías más adecuadas para la obtención de datos y su interpretación para obtener conclusiones significativas (Shaller, 2004).

Esta falta de consenso es el punto de partida del International Consortium for Health Outcomes Measurement (ICHOM), una organización que trabaja en la elaboración de conjuntos estandarizados de medidas de resultados en salud enfocados en las patologías responsables de la mayor carga de la enfermedad. Según su visión, la utilización de indicadores basados en resultados referidos por los pacientes permitiría identificar buenas prácticas y disminuir la variabilidad clínica, de forma que se ofrezca una atención sanitaria en el valor. En la preparación de estos estándares de medidas se tienen en cuenta las perspectivas de los gestores sanitarios, de los clínicos y de los pacientes para seleccionar los resultados que son más relevantes. En el ámbito pediátrico, actualmente están disponibles dos conjuntos de medidas de resultados para patologías congénitas como son la queilosquisis (labio leporino) y el paladar hendido, y la microsomía hemifacial, de los que ya existen algunas experiencias de implantación y uso (Bittar *et al.*, 2018). También existe un

set de medidas dirigido a las mujeres que comprende el proceso del embarazo y el parto. Incluye tanto medidas de mortalidad y morbilidad materno-infantiles, como de satisfacción de la madre con la experiencia del parto y los cuidados recibidos, lactancia o depresión posparto. Todos estos conjuntos de medidas se encuentran disponibles en la web del ICHOM⁴.

La recogida de PRO integrada en la asistencia sanitaria es un proceso laborioso que precisa del compromiso de los profesionales sanitarios y de las organizaciones con la calidad. Se han identificado numerosas barreras, entre las que se pueden destacar la falta de integración de las medidas de PRO en los sistemas de historia clínica, la carga de trabajo de los profesionales sanitarios y la falta de estímulos económicos a la hora de implantar estas mejoras. Sin embargo, existe evidencia de que la recogida rutinaria de este tipo de medidas mejora el control de los síntomas, la calidad de la atención sanitaria y la comunicación entre los profesionales sanitarios y el paciente (Chen, Ou y Hollis, 2013; Valderas *et al.*, 2008).

■ 19.6. CONCLUSIONES

- La investigación de resultados en salud en la población pediátrica busca conocer el impacto de las intervenciones sanitarias en condiciones de práctica clínica en niños y adolescentes. Los resultados pueden tener carácter clínico, económico o humanístico.
- Las percepciones de los profesionales sanitarios sobre la gravedad de los síntomas, el funcionamiento físico o cognitivo de los pacientes, o el impacto de una enfermedad en la calidad de vida son diferentes a las que tienen los propios pacientes.
- Los resultados percibidos por los pacientes pediátricos se pueden medir con cuestionarios validados y con formatos adecuados a los diferentes niveles de comprensión y expresión que tienen los niños y adolescentes.
- Aunque existen experiencias de medición de resultados en salud percibidos por los pacientes en condiciones de práctica clínica, existen algunas barreras para su implantación generalizada.
- La investigación de resultados en salud es una fuente de información muy valiosa para llevar a cabo la transformación que precisa el sistema sanitario, involucrando aún más a los pacientes en las decisiones sobre su salud y ofreciendo servicios orientados a una atención sanitaria basada en el valor.

■ REFERENCIAS

BITTAR, P. G., CARLSON, A. R., MABIE-DE RUYTER, A., MARCUS, J. R. y ALLORI, A. C. (2018). Implementation of a standardized data-collection system for comprehensive appraisal of cleft care. *Cleft Palate Craniofacial Journal*, 55(10), pp. 1382-1390.

⁴ www.ichom.org

- BOURGEAIS, F. T., MURTHY, S., PINTO, C., OLSON, K. L., IOANNIDIS, J. L. y MANDL, K. D. (2012). Pediatric versus adult drug trials for conditions with high pediatric disease burden. *Pediatrics*, 130(2), pp. 285–292.
- CHEN, J., OU, L. y HOLLIS, S. J. (2013). A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Services Research*, 13, p. 211.
- COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE (CHMP) y EUROPEAN MEDICINES AGENCY (EMA) (2016). Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. Disponible en: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2016/04/WC500205159.pdf
- DESHPANDE, P. R., RAJAN, S., SUDEEPTHI, B. L. y ABDUL NAZIR, C. P. (2011). Patient-reported outcomes: A new era in clinical research. *Perspectives in Clinical Research*, 2(4), pp. 137–144.
- EISER, C. y MORSE, R. (2001). Quality-of-life measures in chronic diseases of childhood. *Health Technology Assessment*, 5(4), pp. 1–157.
- EUROPEAN MEDICINES AGENCY y PAEDIATRIC COMMITTEE (2016). 10-year Report to the European Commission. General report on the experience acquired as a result of the application of the Paediatric Regulation. Descargado de https://ec.europa.eu/health/sites/health/files/files/paediatrics/2016_pc_report_2017/ema_10_year_report_for_consultation.pdf.
- FORREST, C. B. (2004). Outcomes research on children, adolescents, and their families: directions for future inquiry. *Medical Care*, 42(4 Suppl), pp. III19-III23.
- FORREST, C. B., SIMPSON, L. y CLANCY, C. (1997). Child health services research. Challenges and opportunities. *Journal of the American Medical Association*, 277(22), pp. 1787–1793.
- GERHARDT, W. E., MARA, C. A., KUDEL, I., MORGAN, E. M., SCHOETTKER, P. J., NAPORA, J., BRITTO, M. T. y ALESSANDRINI, E. A. (2018). Systemwide Implementation of Patient-Reported Outcomes in Routine Clinical Care at a Children's Hospital. *Joint Commission Journal of Quality and Patient Safety*, 44(8), pp. 441-451.
- HAVERMAN, L., VAN ROSSUM, M. A. J., VAN VEENENDAAL, M., VAN DEN BERG, J. M., DOLMAN, K. M., SWART, J., KUIJPERS, T. W. y GROOTENHUIS, M. A. (2013). Effectiveness of a web-based application to monitor health-related quality of life. *Pediatrics*, 131(2), pp. e533-543.
- HINDS, P. S., HOCKENBERRY-EATON, M., GILGER, E., KLINE, N., BURLESON, C., BOTTOMLEY, S. y QUARGNENTI, A. (1999). Comparing patient, parent, and staff descriptions of fatigue in pediatric oncology patients. *Cancer Nursing*, 22(4), pp. 277-288.
- HOCKENBERRY, M. J., HINDS, P. S., BARRERA, P., BRYANT, R., ADAMS-McNEILL, J., HOOKE, C., RASCO-BAGGOTT, C., PATTERSON-KELLY, K., GATTUSO, J. S. y MANTEUFFEL, B. (2003). Three instruments to assess fatigue in children with cancer: the child, parent and staff perspectives. *Journal of Pain and Symptom Management*, 25(4), pp. 319-328.
- HUANG, I. C., SHENKMAN, E. A., LEITE, W., KNAPP, C. A., THOMPSON, L. A. y REVICKI, D. A. (2009). Agreement was not found in adolescents' quality of life rated by parents and adolescents. *Journal of Clinical Epidemiology*, 1;62(3), pp. 337–346.
- JANSENSSENS, A., ROGERS, M., THOMPSON COON, J., ALLEN, K., GREEN, C., JENKINSON, C., TENNANT, A., LOGAN, S. y MORRIS, C. (2015). A systematic review of generic multidimensional patient-reported outcome measures for children, part II: evaluation of psychometric performance of English-language versions in a general population. *Value in Health*, 18(2), pp. 334–345.
- KAPLAN, S. H., GREENFIELD, S., CONNOLLY, G. A., BARLOW, S. E. y GRAND, R. (2001). Methodologic issues in the conduct and interpretation of pediatric effectiveness research. *Ambulatory Pediatrics*, 1(1), pp. 63–70.
- KNOWLES, R. L., GRIEBSCHE, I., BULL, C., BROWN, J., WREN, C. y DEZATEUX, C. (2007). Quality of life and congenital heart defects: comparing parent and professional values. *Archives of Disease in Childhood*, 92(5), pp. 388–393.

- KRAMER, M. S., MACLELLAN, A., CIAMPI, A., ETEZADI-AMOLI, J. y LEDUC, D. G. (1990). Parents'vs physicians' utilities (values) for clinical outcomes in potentially bacteriemic children. *Journal of Clinical Epidemiology*, 43(12), pp. 1319–1325.
- MATZA, L. S., PATRICK, D. L., RILEY, A. W., ALEXANDER, J. J., RAJMIL, L., PLEIL, A. M. y BULLINGER, M. (2013). Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. *Value in Health*, 16(4), pp. 461–479.
- MATZA, L. S., SWENSEN, A. R., FLOOD, E. M., SECNIK, K. y LEIDY, N. K. (2004). Assessment of health-related quality of life in children: a review of conceptual, methodological, and regulatory issues. *Value in Health*, 7(1), pp. 79–92.
- MOYER, V. A. y BUTLER, M. (2004). Gaps in the evidence for well-child care: a challenge to our profession. *Pediatrics*, 114(6), pp. 1511–1521.
- NILSSON, E., ORWELIUS, L. y KRISTENSON, M. (2016). Patient-reported outcomes in the Swedish National Quality Registers. *Journal of Intern Medicine*, 279(2), pp. 141–153.
- ROSENBAUM, P. L., LIVINGSTON, M. H., PALISANO, R. J., GALUPPI, B. E. y RUSSELL, D. J. (2007). Quality of life and health-related quality of life of adolescents with cerebral palsy. *Developmental Medicine and Child Neurology*, 49(7), pp. 516–521.
- SCHÜNEMANN, H. J., FRETHEIM, A. y OXMAN, A. D. (2006). Improving the use of research evidence in guideline development: 10. Integrating values and consumer involvement. *Health Research Policy and Systems*, 4, p. 22.
- SHAH, S. S., HALL, M., GOODMAN, D. M., FEUER, P., SHARMA, V., FARGASON, C., HYMAN, D., JENKINS, K., WHITE, M. L., LEVY, F. H., LEVIN, J. E., BERTOCH, D. y SLONIM, A. D. (2007). Off-label drug use in hospitalized children. *Archives of Pediatrics and Adolescent Medicine*, 161(3), pp. 282–290.
- SHALLER, D. (2004). Implementing and using quality measures for children's health care: perspectives on the state of the practice. *Pediatrics*, 113(1 Pt 2), pp. 217–227.
- SOLANS, M., PANE, S., ESTRADA, M. D., SERRA-SUTTON, V., BERRA, S., HERDMAN, M., ALONSO, J. y RAJMIL, L. (2008). Health-related quality of life measurement in children and adolescents: a systematic review of generic and disease-specific instruments. *Value in Health*, 11(4), pp. 742–764.
- SUNG, L. y REGIER, D. A. (2013). Decision making in pediatric oncology: evaluation and incorporation of patient and parent preferences. *Pediatric Blood Cancer*, 60(4), pp. 558–563.
- TOMLINSON, D., BARTELS, U., GAMMON, J., HINDS, P. S., VOLPE, J., BOUFFET, E., REGIER, D. A., BARUCHEL, S., GREENBERG, M., BARRERA, M., LLEWELLYN-THOMAS, H. y SUNG, L. (2011). Chemotherapy versus supportive care alone in pediatric palliative care for cancer: comparing the preferences of parents and health care professionals. *Canadian Medical Association Journal*, 183(17), pp. E1252-1258.
- UNGAR, W. J. (2007). Paediatric health economic evaluations: a world view. *Healthcare Quality*, 10(1), pp. 134–146.
- U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, FOOD AND DRUG ADMINISTRATION, CENTER FOR DRUG EVALUATION AND RESEARCH (CDER), CENTER FOR BIOLOGICS EVALUATION AND RESEARCH (CBER) y CENTER FOR DEVICES AND RADIOLOGICAL HEALTH (CDRH) (2009). *Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*. Recuperado de: <https://www.fda.gov/downloads/drugs/guidances/UCM193282.pdf>
- VALDERAS, J. M., KOTZEVA, A., ESPALLARGUES, M., GUYATT, G., FERRANS, C. E., HALYARD, M. Y., REVICKI, D.A., SYMONDS, T., PARADA, A. y ALONSO, J. (2008). The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Quality of Life Research*, 17(2), pp. 179–193.

Últimos números publicados

- N.º 40. DOS ENSAYOS SOBRE FINANCIACIÓN AUTONÓMICA**
(*Serie ECONOMÍA Y SOCIEDAD*),
por Carlos Monasterio Escudero e Ignacio Zubiri Oria.
- N.º 41. EFICIENCIA Y CONCENTRACIÓN DEL SISTEMA BANCARIO ESPAÑOL**
(*Serie ANÁLISIS*),
por Fernando Maravall, Silviu Glavan y Analistas Financieros Internacionales.
- N.º 42. ANÁLISIS DE REFORMAS DEL IMPUESTO SOBRE LA RENTA PERSONAL A PARTIR DE MICRODATOS TRIBUTARIOS** (*Serie ANÁLISIS*),
por José Félix Sanz Sanz, Juan Manuel Castañer Carrasco y Desiderio Romero Jordán.
- N.º 43. COMPORTAMIENTO ESTRATÉGICO DE LA BANCA AL POR MENOR EN ESPAÑA: FUSIONES Y ESPECIALIZACIÓN GEOGRÁFICA** (*Serie TESIS*),
por Cristina Bernad Morcate.
- N.º 44. LA VERTIENTE CUALITATIVA DE LA MATERIALIDAD EN AUDITORÍA: MARCO TEÓRICO Y ESTUDIO EMPÍRICO PARA EL CASO ESPAÑOL** (*Serie TESIS*),
por Javier Montoya del Corte.
- N.º 45. LA DECISIÓN DE INTERNACIONALIZACIÓN DE LAS EMPRESAS: UN MODELO TEÓRICO CON INVERSIÓN HORIZONTAL Y VERTICAL** (*Serie TESIS*),
por Jaime Turrión Sánchez.
- N.º 46. FINANCIACIÓN DE LA ENSEÑANZA OBLIGATORIA: LOS BONOS ESCOLARES EN LA TEORÍA Y EN LA PRÁCTICA** (*Serie ECONOMÍA Y SOCIEDAD*),
por Javier Díaz Malledo (coordinador), Clive R. Belfield, Henry M. Levin, Alejandra Mizala, Anders Böhlmark, Mikael Lindahl, Rafael Granell Pérez y María Jesús San Segundo.
- N.º 47. SERVICIOS Y REGIONES EN ESPAÑA** (*Serie ECONOMÍA Y SOCIEDAD*),
por Juan R. Cuadrado Roura y Andrés Maroto Sánchez.
- N.º 48. LAS EMPRESAS DEL SECTOR DE LA CONSTRUCCIÓN E INMOBILIARIO EN ESPAÑA: DEL BOOM A LA RECESIÓN ECONÓMICA** (*Serie ECONOMÍA Y SOCIEDAD*),
por Belén Gill de Albornoz (Dir.), Juan Fernández de Guevara, Begoña Giner y Luis Martínez.
- N.º 49. INSTRUMENTOS PARA MEJORAR LA EQUIDAD, TRANSPARENCIA Y SOSTENIBILIDAD DE LOS SISTEMAS DE PENSIONES DE REPARTO** (*Serie TESIS*),
por M.ª del Carmen Boado-Penas.
- N.º 50. EL IMPUESTO DE FLUJOS DE CAJA EMPRESARIAL: UNA ALTERNATIVA AL IMPUESTO SOBRE LA RENTA DE SOCIEDADES** (*Serie TESIS*),
por Lourdes Jerez Barroso.
- N.º 51. LA SUBCONTRATACIÓN DE SERVICIOS DE I+D: EVIDENCIA DE EMPRESAS EUROPEAS Y DE EE.UU.** (*Serie TESIS*),
por Andrea Martínez Noya.
- N.º 52. IMPOSICIÓN EFECTIVA SOBRE LAS RENTAS DEL CAPITAL CORPORATIVO: MEDICIÓN E INTERPRETACIÓN. EL IMPUESTO SOBRE SOCIEDADES EN ESPAÑA Y EN LOS PAÍSES DE LA UNIÓN EUROPEA EN EL CAMBIO DE MILENIO** (*Serie ANÁLISIS*),
por José Félix Sanz Sanz, Desiderio Romero Jordán y Begoña Barruso Castillo.
- N.º 53. ¿ES RENTABLE EDUCARSE? MARCO CONCEPTUAL Y PRINCIPALES EXPERIENCIAS EN LOS CONTEXTOS ESPAÑOL, EUROPEO Y EN PAÍSES EMERGENTES** (*Serie ECONOMÍA Y SOCIEDAD*),
por José Luis Raymond (coordinador).
- N.º 54. LA DINÁMICA EXTERIOR DE LAS REGIONES ESPAÑOLAS** (*Serie ECONOMÍA Y SOCIEDAD*),
por José Villaverde Castro y Adolfo Maza Fernández.
- N.º 55. EFECTOS DEL STOCK DE CAPITAL EN LA PRODUCCIÓN Y EL EMPLEO DE LA ECONOMÍA** (*Serie TESIS*),
por Carolina Cosculluela Martínez.

- N.º 56. LA PROCICLICIDAD Y LA REGULACIÓN PRUDENCIAL DEL SISTEMA BANCARIO**
(Serie TESIS),
por Mario José Deprés Polo.
- N.º 57. ENSAYO SOBRE ACTIVOS INTANGIBLES Y PODER DE MERCADO DE LAS EMPRESAS. APLICACIÓN A LA BANCA ESPAÑOLA** *(Serie TESIS)*,
por Alfredo Martín Oliver.
- N.º 58. LOS ATRACTIVOS DE LOCALIZACIÓN PARA LAS EMPRESAS ESPAÑOLAS. EXPLOTACIÓN DE LA ENCUESTA SOBRE ATRACTIVOS DE LOCALIZACIÓN** *(Serie ECONOMÍA Y SOCIEDAD)*,
por Encarnación Cerejijo, David Martín, Juan Andrés Núñez, Jaime Turrión y Francisco J. Velázquez.
- N.º 59. ESTUDIO ECONÓMICO DE LOS COSTES DE LA ENFERMEDAD: APLICACIÓN EMPÍRICA AL CASO DEL ALZHEIMER Y LOS CONSUMOS DE DROGAS ILEGALES** *(Serie TESIS)*,
por Bruno Casal Rodríguez.
- N.º 60. BUBBLES, CURRENCY SPECULATION, AND TECHNOLOGY ADOPTION** *(Serie TESIS)*,
por Carlos J. Pérez.
- N.º 61. DISCAPACIDAD Y MERCADO DE TRABAJO: TRES ANÁLISIS EMPÍRICOS CON LA MUESTRA CONTINUA DE VIDAS LABORALES** *(Serie TESIS)*,
por Vanesa Rodríguez Álvarez.
- N.º 62. EL ANÁLISIS DE LOS IMPUESTOS INDIRECTOS A PARTIR DE LA ENCUESTA DE PRESUPUESTOS FAMILIARES** *(SERIE ANÁLISIS)*,
por José Félix Sanz Sanz, Desiderio Romero Jordán y Juan Manuel Castañer Carrasco.
- N.º 63. EUROPA, ALEMANIA Y ESPAÑA: IMÁGENES Y DEBATES EN TORNO A LA CRISIS** *(Serie ECONOMÍA Y SOCIEDAD)*,
por Víctor Pérez-Díaz, Juan Carlos Rodríguez y Elisa Chulía.
- N.º 64. INTEGRACIÓN, INMIGRANTES E INTERCULTURALIDAD: MODELOS FAMILIARES Y PATRONES CULTURALES A TRAVÉS DE LA PRENSA EN ESPAÑA (2010-11)** *(Serie ECONOMÍA Y SOCIEDAD)*,
por Enrique Uldemolins, Alfonso Corral, Cayetano Fernández, Miguel Ángel Motis, Antonio Prieto y María Luisa Sierra.
- N.º 65. SOSTENIBILIDAD DEL SISTEMA DE PENSIONES DE REPARTO EN ESPAÑA Y MODELIZACIÓN DE LOS RENDIMIENTOS FINANCIEROS** *(Serie TESIS)*,
por Clara Isabel González Martínez.
- N.º 66. EVOLUCIÓN DE LAS FUNDACIONES BANCARIAS ITALIANAS: DE HOLDING DE SOCIEDADES BANCARIAS A UN MODELO INNOVADOR DE "BENEFICIENCIA PRIVADA"** *(Serie ECONOMÍA Y SOCIEDAD)*,
por Paolo Baroli, Claudia Imperatore, Rosella Locatelli y Marco Trombetta.
- N.º 67. LAS CLAVES DEL CRÉDITO BANCARIO TRAS LA CRISIS** *(Serie ECONOMÍA Y SOCIEDAD)*,
por Santiago Carbó Valverde, José García Montalvo, Joaquín Maudos y Francisco Rodríguez Fernández.
- N.º 68. ENTRE DESEQUILIBRIOS Y REFORMAS. ECONOMÍA POLÍTICA, SOCIEDAD Y CULTURA ENTRE DOS SIGLOS** *(Serie ECONOMÍA Y SOCIEDAD)*,
por Víctor Pérez-Díaz y Juan Carlos Rodríguez.
- N.º 69. REFORMA DEL MERCADO DE SERVICIOS PROFESIONALES EN ESPAÑA** *(Serie ECONOMÍA Y SOCIEDAD)*,
por María Paz Espinosa, Aitor Ciarreta y Aitor Zurimendi.
- N.º 71. BUILDING A EUROPEAN ENERGY MARKET: LEGISLATION, IMPLEMENTATION AND CHALLENGES** *(Serie ECONOMÍA Y SOCIEDAD)*,
por Tomás Gómez y Rodrigo Escobar.

- N.º 72. ESSAYS IN TRADE, INNOVATION AND PRODUCTIVITY**
(Serie TESIS),
por Aránzazu Crespo Rodríguez.
- N.º 73. ENDEUDAMIENTO DE ESPAÑA: ¿QUIÉN DEBE A QUIÉN?**
(SERIE ECONOMÍA Y SOCIEDAD),
por Analistas Financieros Internacionales (AFI).
- N.º 74. AGENTES SOCIALES, CULTURA Y TEJIDO PRODUCTIVO EN LA ESPAÑA ACTUAL**
(SERIE ECONOMÍA Y SOCIEDAD),
por Víctor Pérez-Díaz, Juan Carlos Rodríguez, Joaquín Pedro López-Novo y Elisa Chuliá.
- N.º 75. EVOLUCIÓN RECIENTE DEL CRÉDITO Y LAS CONDICIONES DE FINANCIACIÓN: ESPAÑA EN EL CONTEXTO EUROPEO**
(SERIE ECONOMÍA Y SOCIEDAD),
por Joaquín Maudos.
- N.º 76. EFICIENCIA DE LOS SISTEMAS REGIONALES DE INNOVACIÓN EN ESPAÑA**
(SERIE ANÁLISIS),
por Mikel Buesa, Joost Heijs, Thomas Baumert y Cristian Gutiérrez.
- N.º 77. ENCOURAGING BLOOD AND LIVING ORGAN DONATIONS**
(Serie TESIS),
por María Errea y Juan M. Cabasés (director).
- N.º 78. EMPLEO Y MATERNIDAD: OBSTÁCULOS Y DESAFÍOS A LA CONCILIACIÓN DE LA VIDA LABORAL Y FAMILIAR** *(Serie ECONOMÍA Y SOCIEDAD),*
por Margarita León Borja (coordinadora).
- N.º 79. PEOPLE MANAGEMENT IN MICRO AND SMALL COMPANIES - A COMPARATIVE ANALYSIS. EMPLOYEE VOICE PRACTICES AND EMPLOYMENT RELATIONS,**
(Serie ANÁLISIS),
por Sylvia Rohlfel, con la colaboración de Carlos Salvador Muñoz y Alesia Slocum.
- N.º 80. LA CRISIS, ¿UNA OPORTUNIDAD PARA LA ECONOMÍA SOCIAL ESPAÑOLA**
(Serie ECONOMÍA Y SOCIEDAD),
por Pierre Perard.
- N.º 81. UN TRIÁNGULO EUROPEO: ELITES POLÍTICAS, BANCOS CENTRALES Y POPULISMOS**
(Serie ECONOMÍA Y SOCIEDAD),
por Víctor Pérez Díaz, Juan Carlos Rodríguez y Elisa Chuliá.
- N.º 82. EL MERCADO ESPAÑOL DE ELECTRICIDAD**
(Serie ECONOMÍA Y SOCIEDAD),
por Aitor Ciarreta, María Paz Espinosa y Aitor Zurimendi.
- N.º 83. THREE ESSAYS IN LONG-TERM ECONOMIC PERSISTENCE**
(Serie TESIS),
por Felipe Valencia Caicedo.
- N.º 84. ROLE OF MICROPARTICLES IN ATHEROTHROMBOSIS**
(Serie TESIS),
por Rosa Suades Soler.
- N.º 85. IBERISMOS. EXPECTATIVAS PENINSULARES EN EL SIGLO XIX**
(Serie TESIS),
por César Rina Simón.
- N.º 86. MINING STRUCTURAL AND BEHAVIORAL PATTERNS IN SMART MALWARE**
(Serie TESIS),
por Guillermo Suárez-Tangil.
- N.º 87. LA VOZ DE LA SOCIEDAD ANTE LA CISIS**
(Serie ECONOMÍA Y SOCIEDAD),
por Víctor Pérez-Díaz.

- N.º 88. ECONOMÍA SUMERGIDA Y FRAUDE FISCAL EN ESPAÑA: ¿QUÉ SABEMOS? ¿QUÉ PODEMOS HACER?**
(Serie ECONOMÍA Y SOCIEDAD),
por Santiago Lago Peñas.
- N.º 89. CONSTRUCCIÓN EUROPEA, IDENTIDADES Y MEDIOS DE COMUNICACIÓN**
(Serie ECONOMÍA Y SOCIEDAD),
por Víctor Pérez-Díaz, Juan Carlos Rodríguez y Josu Mezo.
- N.º 90. LA INTEGRACIÓN DE LOS INMIGRANTES EN EUROPA Y EN ESPAÑA: MODELOS E INDICADORES PARA LAS POLÍTICAS PÚBLICAS**
(Serie ECONOMÍA Y SOCIEDAD),
por Cayetano Fernández, Alfonso Corral, Antonio Prieto María Luisa Sierra y Enrique Uldemolins
- N.º 91. SOLEDAD, DISCAPACIDAD Y MERCADO DE TRABAJO**
(Serie ECONOMÍA Y SOCIEDAD),
por Miguel Ángel Malo y Ricardo Pagán
- N.º 92. CRISIS ECONÓMICA Y DESIGUALDAD DE LA RENTA EN ESPAÑA. EFECTOS DISTRIBUTIVOS DE LAS POLÍTICAS PÚBLICAS**
(Serie ECONOMÍA Y SOCIEDAD),
por Samuel Calonge Ramírez y Antonio Manresa Sánchez
- N.º 93. LAS DESIGUALDADES ECONÓMICAS EN ESPAÑA: REALIDADES Y PERCEPCIONES**
(Serie ECONOMÍA Y SOCIEDAD),
por Víctor Pérez-Díaz y Juan Carlos Rodríguez
- N.º 94. INNOVACIÓN, CRECIMIENTO Y COMPETITIVIDAD: EL PAPEL DE LA POLÍTICA TECNOLÓGICA EN ESPAÑA**
(Serie ECONOMÍA Y SOCIEDAD),
por Joost Heijs, Mikel Buesa, Delia Margarita Vergara, Cristian Gutiérrez, Guillermo Arenas y Alex Javier Guerrero
- N.º 95. 40 AÑOS DE DESCENTRALIZACIÓN EN ESPAÑA (1978-2018): BALANCE Y PERSPECTIVAS**
(Serie ECONOMÍA Y SOCIEDAD),
por Santiago Lago Peñas

ESTUDIOS
DE LA FUNDACIÓN

SERIE ECONOMÍA Y SOCIEDAD

Información:

Funcas

Caballero de Gracia, 28

28013 Madrid

Teléfono: 91 596 54 81

Fax: 91 596 57 96

publica@funcas.es

www.funcas.es

ISBN 978-84-17609-49-8

