

Factores clave en la elección de desenlaces apropiados para estudios clínicos

Carlos Cabrera-Ubilla^a , Vicente Nuñez-Galleguillos^a , Javier Bracchiglione, MSc^{a, b, c*} , Roberto Garnham, MSc^a , Ignacio Pérez^a 

^aEscuela de Medicina, Universidad de Valparaíso, Viña del Mar, Chile; ^bCentro de Estudios Interdisciplinarios en Salud (CIESAL), Universidad de Valparaíso, Viña del Mar, Chile; ^cCentro Cochrane Iberoamericano, Institut de Recerca Sant Pau (IR SANT PAU), CIBERESP, Barcelona, España

RESUMEN La investigación en salud es la base del conocimiento médico y de las recomendaciones en los sistemas de salud. Por ello, la elección de desenlaces apropiados en estudios de intervenciones terapéuticas es un paso fundamental en la producción de evidencia y, posteriormente, para la toma de decisiones. En este artículo proponemos tres factores clave para la elección de desenlaces: la inclusión de desenlaces reportados por pacientes, ya que ponen el foco del efecto de la intervención en la percepción que tienen los propios pacientes de su estado de salud y calidad de vida; la consideración de desenlaces clínicamente relevantes, los cuales son mediciones directas del estado de salud del paciente y, por ende serán determinantes en la toma de decisiones; y la utilización de *Core Outcome Sets*, herramienta que estandariza y permite la homogeneización en la medición e interpretación de desenlaces, facilitando la producción y posterior síntesis de evidencia apropiada para el ecosistema de evidencia. La correcta elección de los desenlaces permitirá que la evidencia generada de estos estudios ayude a los tomadores de decisiones en salud y los profesionales clínicos a entregar cuidados apropiados centrados en el paciente y a optimizar el uso de recursos en salud e investigación clínica.

KEYWORDS Evidence-Based Medicine; Education, Medical; Outcome Assessment, Health Care

INTRODUCCIÓN

La investigación en ciencias de la salud es la base del conocimiento médico y, en gran medida, de las recomendaciones generadas en los sistemas de salud [1]. Desde las ciencias básicas hasta la clínica, existe un largo camino para que una intervención candidata (por ejemplo, una molécula o un fenómeno físico) se convierta en un tratamiento clínicamente aceptado y recomendado. La fundación MAGIC (*Making Grade the Irresistible Choice*) propone un ecosistema de evidencia, conformado por un ciclo desde la producción de evidencia, síntesis, divulgación a clínicos y pacientes, implementación, evaluación y mejora, para nuevamente llegar a la producción de evidencia. Entre cada paso hay un traspaso de información que sirve para sentar las bases de la etapa siguiente. En el centro del modelo, actúan de manera simultánea e integrada cinco elementos: evidencia confiable, entendimiento común de los métodos, datos digitalmente

estructurados, herramientas y plataformas, y la cultura de compartir [2].

Este paso del “saber” al “hacer” es el principal objetivo de la investigación clínica, pues sólo habiendo cumplido con esto, se pueden tomar decisiones en salud informadas por evidencia. Dado que la cantidad de investigación ha ido en aumento constante [3,4], es importante que diferentes estudios relativos a un mismo tópico (enfermedad o condición médica) tributen a la generación de conocimiento de una forma consensuada, con el fin de mejorar la calidad de atención y los resultados en salud de los pacientes. Una interacción descoordinada entre investigadores y tomadores de decisiones puede llevar a generar investigación inapropiada con un uso subóptimo de recursos [5]. Es por ello que se requiere de un plan de trabajo internacional coordinado, a través de comunidades científicas conformadas por equipos multidisciplinares y pacientes o sus representantes, con el fin de priorizar tópicos de investigación, evitar el trabajo redundante y generar recomendaciones clínicas del mayor nivel de evidencia posible [6].

Parte importante de esta acción coordinada consiste en la elección de desenlaces óptimos para ser incluidos en una investigación clínica. La metodología GRADE (*Grading of Recommendations Assessment, Development and Evaluation*) clasifica los desenlaces como “críticos”, “importantes” y “no importantes” según la importancia relativa que le asigna el

* Autor de correspondencia javier.bracchiglione@gmail.com

Citación Cabrera-Ubilla C, Nuñez-Galleguillos V, Bracchiglione J, Garnham R, Pérez I. Factores clave en la elección de desenlaces apropiados para estudios clínicos. *Medwave* 2024;24(4):e2790
DOI 10.5867/medwave.2024.04.2790

Fecha de envío Oct 23, 2023, **Fecha de aceptación** Apr 15, 2024,

Fecha de publicación May 15, 2024

Correspondencia a Angamos 655, Viña del Mar, Chile

IDEAS CLAVE

- Elegir desenlaces apropiados permite generar conocimiento útil para la toma de decisiones en salud.
- Al elegir desenlaces para una investigación, es clave considerar desenlaces clínicamente relevantes, que sean reportados por los pacientes, y que idealmente hayan sido priorizados previamente por un *core outcome set*.
- La significancia clínica debería ser clave al momento de interpretar un desenlace, más allá de la significancia estadística.

panel formulador de recomendaciones. De este modo, un desenlace crítico incide directamente en la toma de decisiones, uno importante probablemente lo hará, y uno no importante no lo hará [7].

Este artículo corresponde al decimotercero de una serie metodológica de revisiones narrativas acerca de tópicos generales en bioestadística y epidemiología clínica, las que exploran y resumen en un lenguaje amigable, artículos publicados disponibles en las principales bases de datos y textos de consulta especializados. La serie está orientada a la formación de estudiantes de pre y posgrado. Es realizada por la Cátedra de Medicina Basada en Evidencia de la Escuela de Medicina de la Universidad de Valparaíso, Chile en colaboración con el Departamento de Investigación del Instituto Universitario del Hospital Italiano de Buenos Aires, Argentina y el Centro Evidencia UC, de la Pontificia Universidad Católica de Chile. El objetivo de este manuscrito es discutir los factores clave que deberían considerar, tanto autores como lectores de evidencia científica, al momento de analizar desenlaces de una investigación clínica.

¿POR QUÉ INDICAMOS UN TRATAMIENTO?

En general, las intervenciones terapéuticas buscan cinco objetivos: aumentar la longevidad, prevenir comorbilidades, disminuir los síntomas, mejorar la calidad de vida y optimizar el uso de recursos [8]. Por ende, consideraremos una intervención como efectiva cuando mejore un desenlace que refleje uno o más de estos objetivos de la forma más directa posible.

Por ejemplo, la efectividad de dos abordajes terapéuticos para el cáncer de colon en adultos podría determinarse midiendo la mortalidad, recurrencias, aparición de metástasis, tamaño tumoral, marcadores tumorales, sangrado microscópico, calidad de vida, dolor, reacciones adversas, entre otros. Dado que probablemente no todos los desenlaces tengan la misma importancia relativa para pacientes, familiares, profesionales y tomadores de decisiones, la pregunta fundamental es ¿qué desenlace(s) deberíamos elegir o priorizar para ser medidos en investigación clínica? Vale decir, ¿qué desenlace(s) probablemente serán críticos o importantes para la toma de decisiones en salud? [9]. La importancia de este cuestionamiento radica en que la efectividad de las intervenciones estudiadas se determinará según la medición de los desenlaces seleccionados. Un buen desenlace, por tanto, debiese ser aquel que sea factible de ser medido y que sea considerado crítico o importante para la toma de decisiones en distintos contextos. En este artículo, proponemos tres factores clave que todo investigador debería

considerar para elegir un desenlace apropiado: la inclusión de desenlaces reportados por pacientes, la consideración de desenlaces clínicamente relevantes y la utilización de *Core Outcome Sets*.

PRIMER FACTOR CLAVE: DESENLACES REPORTADOS POR PACIENTES

La adjudicación de un desenlace en investigación clínica se refiere a la determinación de la presencia, ausencia o cambio en la medición de un desenlace. Los desenlaces en salud pueden ser asignados de distintas maneras y por distintos adjudicadores. Por ejemplo, un cambio en la medición plasmática de biomarcadores a través de exámenes sería adjudicado directamente por el laboratorio, un cambio en el tamaño tumoral sería adjudicado por un profesional clínico mediante la interpretación de exámenes imagenológicos, y un cambio en la intensidad del dolor podría ser adjudicado por los mismos pacientes, o en su defecto, por algún observador externo (cuidadores o personal clínico) [10].

Los desenlaces reportados por pacientes (*patient reported outcomes* o PRO por su sigla en inglés), son los “reportes del estado de salud de un paciente que vienen directamente de ellos, sin ninguna interpretación de un profesional clínico u otra persona” [11]. Dado que representan la propia percepción de los pacientes de su estado de salud, calidad de vida y de los desenlaces que son más significativos para ellos [12], los desenlaces reportados por pacientes responden directamente a dos de los objetivos por los cuales indicamos un tratamiento: disminuir los síntomas y mejorar la calidad de vida. De este modo, los desenlaces reportados por pacientes, en general proporcionan información crucial para que los profesionales clínicos puedan tomar decisiones apropiadas [10].

En ocasiones hay una pobre correlación entre los desenlaces reportados por pacientes y los desenlaces reportados por laboratorio, profesionales clínicos u observadores externos. Por ejemplo:

- En pacientes con enfermedad pulmonar obstructiva crónica, existe una baja correlación entre el volumen espiratorio forzado y la calidad de vida reportada por los pacientes [13].
- En pacientes oncológicos, la intensidad de los síntomas puede ser subestimada según la interpretación de personal clínico [14].
- En patologías funcionales como el síndrome de intestino irritable, síndromes dolorosos o síntomas como náuseas

y fatiga, la evaluación del estado de un paciente es limitada y difícilmente objetivable a través de la observación externa (reportada por familiares u otro acompañante), la evaluación clínica o exámenes [10].

Como regla general (pero por ningún motivo obligatoria) los desenlaces reportados por pacientes deberían ser priorizados en la planificación de investigación clínica por sobre aquellos reportados por terceros, con la notable excepción de los desenlaces relacionados con sobrevida/mortalidad.

SEGUNDO FACTOR CLAVE: DESENLACES CLÍNICAMENTE RELEVANTES

Para la elección de un buen desenlace es crucial saber discernir entre uno clínicamente relevante de uno subrogado. Volviendo al ejemplo de pacientes adultos con cáncer de colon, supongamos que los autores de un ensayo clínico desean saber si el abordaje terapéutico "A" es mejor que "B" midiendo los siguientes desenlaces: mortalidad, calidad de vida y tamaño tumoral.

Mortalidad y calidad de vida probablemente serán desenlaces muy prioritarios para pacientes y clínicos involucrados en la toma de decisiones. Una intervención que logre disminuir la mortalidad con una mejoría de la calidad de vida para pacientes con cáncer de colon, probablemente sea considerado un tratamiento exitoso. Estos desenlaces son relevantes en sí mismos, dado que representan un estado de salud deseable. Por otra parte, la medición del tamaño tumoral no es un desenlace relevante en sí mismo. Su importancia radica en una inferencia o hipótesis causal. Medimos tamaño tumoral porque asumimos, indirectamente, que un menor tamaño tumoral llevará a menor mortalidad y a mejor calidad de vida. Pero este razonamiento causal está sustentado en una lógica fisiopatológica, que muchas veces no se comporta de un modo lineal y directo en la práctica clínica.

Un ejemplo claro de esto ocurre en el tratamiento de la dislipidemia. En pacientes con dislipidemia, podría considerarse que la mejoría del perfil lipídico es un desenlace relevante. Esto se basa en el supuesto teórico de que un paciente con dislipidemia que mejora su perfil lipídico tendrá menor probabilidad de morir o de presentar un evento cardiovascular mayor. Vale decir, estamos asumiendo que hay una relación causal directa y lineal entre mejoría de perfil lipídico con mejoría de sobrevida y disminución de riesgo de eventos cardiovasculares. Sin embargo, esto puede no siempre ser así. A modo de ejemplo, la cerivastatina fue retirada del mercado por aumentar la mortalidad a causa de rhabdomiolisis, pese a la notable mejoría que generaba en el perfil lipídico [15].

El ejemplo hipotético del cáncer de colon y el ejemplo real de la cerivastatina en dislipidemia nos llevan a la conclusión de que hay dos tipos de desenlaces:

1. Desenlaces "clínicamente relevantes": aquellos que son importantes en sí mismos para tomar decisiones, ya que

son mediciones directas del estado de salud de un paciente [16].

2. Desenlaces "subrogados", "sustitutos" o "intermedios": aquellos que no son relevantes en sí mismos, sino que intentan predecir o asociarse a los primeros, representando una medida indirecta de un desenlace clínicamente significativo que se utiliza como predictor en estudios clínicos [17].

Los desenlaces subrogados están relacionados con el mecanismo fisiopatológico, formando parte del proceso biológico que provoca la aparición del desenlace clínicamente relevante, por lo que se espera que puedan predecirse a través de la medición de biomarcadores [16,17]. Sin embargo, como ya hemos visto con la cerivastatina, esto no siempre es así. La Tabla 1 provee ejemplos de estudios en medicamentos como la rosiglitazona [18], el torcetrapib [19] y el flúor [20], que mostraron una disociación entre los hallazgos medidos por desenlaces subrogados y por desenlaces clínicamente relevantes.

Desde una perspectiva mecanicista podrían considerarse más importantes los desenlaces subrogados, dado que facilitan la realización de ensayos clínicos al tender a ser más accesibles, baratos y rápidos [21]. Sin embargo, para determinar si un tratamiento contribuye al bienestar de los pacientes, debemos conocer cómo este afecta los desenlaces que son importantes para ellos, para posteriormente tomar decisiones centradas en los pacientes [22].

Debemos destacar que un desenlace reportado por pacientes no es necesariamente un desenlace clínicamente relevante. Por ejemplo, un paciente podría reportar prurito leve que no incide en su calidad de vida en un tratamiento con quimioterapia. A su vez, un desenlace clínicamente relevante podría no ser reportado por el paciente, como el caso de mortalidad o ciertos exámenes de laboratorios para patologías específicas (por ejemplo, carga viral en pacientes con virus de la inmunodeficiencia humana, VIH).

¿Son malos los desenlaces subrogados?

Si bien en teoría es preferible utilizar desenlaces clínicamente relevantes ya que entregan mejor información a los profesionales clínicos sobre cuándo utilizar un tratamiento [22], hay ocasiones en que pueden ser difícilmente accesibles al ser más costosos, complejos de medir, requerir un seguimiento prolongado o mayor tamaño muestral [16,23]. En casos en que la medición de un desenlace clínicamente relevante supone una excesiva complejidad metodológica y financiera, se puede justificar la utilización de desenlaces subrogados. Por ejemplo, la medición de hemoglobina glicosilada para pacientes con diabetes mellitus y de carga viral para portadores del VIH son desenlaces subrogados ampliamente utilizados debido a que, por la historia natural de estas patologías, los pacientes permanecen asintomáticos por un extenso período previo a la

Tabla 1. Ejemplos de disociación entre desenlaces subrogados y clínicamente relevantes.

Intervención terapéutica	Enfermedad o condición de salud	Desenlace subrogado	Desenlace clínicamente relevante
Cerivastatina	Dislipidemia	Mejora en el perfil lipídico.	Aumento de mortalidad por rhabdomiolisis.
Rosiglitazona	Diabetes mellitus	Disminución de HbA1c.	Aumento de la mortalidad cardiovascular.
Torcetrapib	Dislipidemia	Aumento de niveles plasmáticos de colesterol HDL.	Aumento de la mortalidad por todas las causas.
Flúor	Osteoporosis	Aumento de densitometría mineral ósea.	Aumento del riesgo de fracturas.

HbA1c: hemoglobina glicosilada A1c. HDL: lipoproteínas de alta densidad. Fuente: adaptado de [15][18][19][20].

aparición de desenlaces que podrían ser clínicamente relevantes.

TERCER FACTOR CLAVE: CORE OUTCOME SETS

Debido a la multiplicidad de desenlaces que pueden ser heterogéneamente priorizados entre diferentes grupos de investigadores o tomadores de decisiones, se hace necesario consensuar un conjunto mínimo de desenlaces que sean relevantes de medir en distintos contextos (lo que en terminología GRADE podría asemejarse al concepto de desenlaces críticos e importantes). Los *Core Outcome Sets* (COS) han intentado generar estos consensos [24].

Los *Core Outcome Sets* son conjuntos de desenlaces consensuados y estandarizados, que deberían ser medidos y reportados en todos los ensayos clínicos de una determinada enfermedad o condición de salud [25]. Estos han sido recopilados en recursos especializados, tales como el sitio web de la Iniciativa *Core Outcome Measures in Effectiveness Trials* (COMET) [24]. Para el desarrollo de los *Core Outcome Sets*, se ha elaborado una guía de recomendaciones, el *COS standards for development* (COS-STAD), en la cual se han definido tres pasos. El primer paso consiste en definir el alcance del trabajo, identificando la patología, población, escenario e intervención que se investigará, y cerciorarse de que no existe otro trabajo del mismo tema en la base datos del COMET. El segundo paso es la conformación de un panel multidisciplinario compuesto por investigadores, profesionales de la salud, profesionales de la industria médica, pacientes, familiares y/o cuidadores. El tercer paso comprende la identificación y priorización de los *outcomes*. La identificación se realiza mediante estudios cualitativos primarios o secundarios que evalúen los *outcomes* reportados en los ensayos clínicos. La priorización, es habitualmente realizada a través de la técnica Delphi, la cual consiste en que el equipo investigador envía un listado de los *outcomes* relevantes identificados al panel multidisciplinario, para que estos otorguen un puntaje de acuerdo a la importancia que tienen para ellos. Luego se realiza una segunda ronda, en donde se eliminan los *outcomes* que unánimemente han sido considerados no relevantes y se envía un nuevo listado al panel multidisciplinario con una retroalimentación y el puntaje obtenido en cada *outcome*, de manera que puedan cambiar su puntaje si lo consideran pertinente.

Este proceso se repite hasta lograr un consenso, para finalmente realizar una discusión final de los resultados obtenidos [26,27].

Por ejemplo, el sitio web de la iniciativa COMET incluye un estudio que investigó los *Core Outcome Sets* para el eczema [28], en donde los desenlaces consensuados propuestos para ser priorizados en una investigación en este tópico, fueron:

- Signos clínicos.
- Síntomas.
- Control a largo plazo de las erupciones.
- Calidad de vida.

De esta manera, los *Core Outcome Sets* estandarizan y permiten la homogeneización en la medición de desenlaces, facilitando la producción y posterior síntesis de evidencia apropiada para el ecosistema de evidencia [29].

INTERPRETACIÓN DE DESENLACES: SIGNIFICANCIA ESTADÍSTICA VERSUS SIGNIFICANCIA CLÍNICA

Una vez elegidos o priorizados los desenlaces, la interpretación adecuada de los resultados de una investigación en salud es fundamental para una óptima toma de decisiones clínicas y de salud pública. Por lo tanto, es importante comprender las diferencias entre la significancia estadística y clínica, ya que un desenlace estadísticamente significativo no implica necesariamente que sea clínicamente significativo, y viceversa [30,31].

La significancia estadística se refiere a que la probabilidad de que un desenlace haya ocurrido por el azar es muy baja. Usualmente se determina por convención, como un valor p menor a 0,05 en un test estadístico determinado, o como un intervalo de confianza al 95% que no cruza la línea de no significancia estadística [32,33]. En cambio, la significancia clínica está delimitada por la diferencia clínica mínimamente importante (*minimally important difference* o MID en inglés), la cual marca el cambio más pequeño en los desenlaces reportados por pacientes que ellos perciben como importante (ya sea beneficioso o perjudicial), y que podría hacer que tanto él como su tratante consideren un cambio en la terapia [34]. La significancia clínica está determinada por la precisión de los resultados [35,36]. Supongamos que para el tratamiento de hipertensión arterial fijamos una diferencia clínica mínimamente importante de 10 milímetros de mercurio. Si el intervalo de confianza (una medida de precisión de los resultados) de

Tabla 2. Posibles desenlaces en un estudio de asma bronquial.

Adjudicador del desenlace	Desenlace medido
Laboratorio	<ul style="list-style-type: none"> • Valor del VEF1 pre-broncodilatador en espirometría. • Niveles de IgE sérica.
Profesional clínico	<ul style="list-style-type: none"> • Sibilancias u otras alteraciones en la auscultación pulmonar. • Hallazgos patológicos en radiografía o tomografía computada de tórax: hiperinsuflación, atelectasias y/o engrosamiento de la pared bronquial. • Remodelación histológica de la vía aérea inferior, a través de lavado o biopsia bronquial.
Familiar o cuidador	<ul style="list-style-type: none"> • Percepción subjetiva de familiares o cuidadores respecto a la calidad de vida del paciente. • Estrés del cuidador.
Paciente	<ul style="list-style-type: none"> • Calidad de vida: evaluación subjetiva o a través de encuestas estandarizadas. • Intensidad y frecuencia de síntomas (disnea, tos, opresión retroesternal). • Tolerancia al ejercicio. • Uso semanal de salbutamol como medicación de rescate. • Efectos adversos del tratamiento farmacológico.

VEF1: volumen espiratorio forzado al primer segundo. IgE: inmunoglobulina E.

Fuente: elaborada por los autores.

un estudio indica que un tratamiento genera un cambio de presión arterial de -20 milímetros de mercurio, con un intervalo de confianza que va de -25 a -15 milímetros de mercurio, podemos decir que se estima que el tratamiento disminuirá la presión arterial en una magnitud tan alta como 25 milímetros de mercurio o tan baja como 15 milímetros de mercurio. En ambos casos (incluso en el peor de los casos) el tratamiento genera un cambio clínicamente significativo de presión arterial (mayor a la diferencia clínica mínimamente importante predefinida). Este resultado es preciso, dado que no cruza la línea de significancia clínica. Por otra parte, si el cambio fuera de -10 milímetros de mercurio con un intervalo de confianza que va de -15 a -5 milímetros de mercurio, tendríamos un escenario más complejo. En el mejor de los casos (-15 milímetros de mercurio) la diferencia sería clínicamente significativa, pero en el peor de los casos (-5 milímetros de mercurio) sería clínicamente irrelevante. En este caso, el intervalo de confianza cruza el umbral de significancia clínica y es, por tanto, impreciso. En este manuscrito no ahondaremos en este tema específico, puesto que ha sido revisado en extenso en otro artículo de esta serie metodológica [36]. Una correcta interpretación de un desenlace incluye la determinación no solo de su significancia estadística, sino también de su significancia clínica, utilizando una diferencia clínica mínimamente importante explícita.

APLICACIÓN A UN PROYECTO DE INVESTIGACIÓN

Para entender mejor los factores clave propuestos para la elección de un desenlace apropiado, analizaremos la aplicación de cada uno al caso hipotético de un grupo de investigadores que debe seleccionar los desenlaces a incluir en un estudio que compara dos tratamientos para el asma bronquial.

Primer factor clave: desenlaces reportados por pacientes

En primer lugar, para tener una visión general podemos clasificar todos los posibles desenlaces a medir de acuerdo a quien lo reporta, los cuales se detallan en la Tabla 2.

También, algunos desenlaces pueden ser reportados por más de un adjudicador. Por ejemplo, el número de hospitalizaciones, consultas a urgencias y exacerbaciones o crisis asmáticas, pueden ser reportadas tanto por el propio paciente, por un familiar o por el profesional clínico.

Como regla general, es recomendable incluir aquellos desenlaces reportados por pacientes (como los mencionados en la Tabla 2) ya que inciden directamente en la percepción de su estado de salud y, por ende, orientarán a tomar decisiones según sus resultados.

Segundo factor clave: desenlaces clínicamente relevantes

En segundo lugar, se deberían priorizar los desenlaces clínicamente relevantes por sobre los subrogados, a menos que sea muy complejo contar con los primeros. Dentro de los desenlaces clínicamente relevantes podemos encontrar la calidad de vida, intensidad y frecuencia de síntomas o tolerancia al ejercicio, ya que son mediciones directas del estado de salud de los pacientes y permitirán tomar decisiones en sus tratamientos.

Tercer factor clave: Core Outcome Sets

Para evaluar este factor, buscamos en el sitio web COMET-initiative.org y seleccionamos un estudio que desarrolló un Core Outcome Sets para evaluación de asma bronquial [37]. Los desenlaces consensuados incluyeron:

a. Exacerbación asmática.

- b. Cambios en el control del asma (percepción subjetiva, presencia de sintomatología y valor del volumen espiratorio forzado pre-broncodilatador).
- c. Calidad de vida medida por encuestas estandarizadas de asma.
- d. Número de hospitalizaciones relacionadas con el asma.
- e. Consultas en unidad de urgencias debido al asma.

Para concluir, un set de desenlaces apropiados para esta pregunta sería: calidad de vida, intensidad y frecuencia de síntomas, tolerancia al ejercicio, número de hospitalizaciones o consultas al servicio de urgencias, número de crisis o exacerbaciones asmáticas, ya que son reportados por el paciente o afectan directamente en cómo se siente, son clínicamente relevantes y están priorizados en un *Core Outcome Sets*.

RECOMENDACIONES FINALES

Un desenlace que cumple con los tres aspectos propuestos durante este artículo, y que es interpretado considerando diferencias clínicamente significativas, será un desenlace probablemente relevante para la toma de decisiones en diversos contextos. Asimismo, permitirá una mejor producción y utilización de la evidencia, lo que debe ser considerado por los investigadores a modo de optimizar recursos. La consideración de estos factores clave para priorizar e interpretar desenlaces contribuye a las primeras etapas del ecosistema de evidencia (producción y síntesis de evidencia). En consecuencia, también aporta a un mejor desarrollo de las siguientes etapas del ecosistema (divulgación, implementación, evaluación y mejora), de manera que los tomadores de decisiones en salud y los profesionales clínicos puedan entregar cuidados apropiados en salud a la vez que optimizan el uso de recursos en salud e investigación clínica [38].

Proponemos las siguientes recomendaciones para una apropiada elección de desenlaces a evaluar al momento de diseñar protocolos de investigación clínica:

1. Priorizar desenlaces reportados directamente por los pacientes (con la excepción de desenlaces relacionados con mortalidad/sobrevivencia), como manera de evaluar desenlaces que inciden directamente en la percepción, calidad de vida y/o bienestar que tienen los pacientes sobre su condición de salud.
2. Siempre que sea posible, preferir desenlaces clínicamente relevantes, ya sean reportados por pacientes o no, puesto que estos favorecen una adecuada toma de decisiones clínicas.
3. Una vez identificados los potenciales desenlaces a incluir en un estudio, revisar el *Core Outcome Sets* de la condición de salud en el sitio COMET u otro afín, con el objetivo de estar seguros de que se incluirán los desenlaces mínimos que por un consenso multidisciplinario han sido definidos como imprescindibles para evaluar una condición de salud.

4. Se sugiere considerar la inclusión de pacientes en el diseño del protocolo y discusión de los resultados obtenidos en estudios clínicos, de manera de generar un acercamiento a la investigación de su propia patología desde la confección de estudios primarios experimentales. Creemos que podría ser beneficioso involucrarlos desde etapas tempranas en la generación de conocimiento clínico.

Autoría CCU: investigación, metodología, redacción del borrador original, revisión y edición. VNG: investigación, metodología, redacción del borrador original, revisión y edición. JB: conceptualización, metodología, investigación, revisión y edición, visualización, supervisión. RG: metodología, investigación, revisión y edición, visualización, supervisión. IP: conceptualización, metodología, investigación, revisión y edición, visualización, supervisión.

Agradecimientos Agradecemos a la Cátedra de Medicina Basada en Evidencia de la Escuela de Medicina de la Universidad de Valparaíso, Chile por impulsar esta serie.

Conflictos de intereses Los autores completaron la declaración de conflictos de interés de ICMJE y declararon que no recibieron fondos por la realización de este artículo; no tienen relaciones financieras con organizaciones que puedan tener interés en el artículo publicado en los últimos tres años y no tienen otras relaciones o actividades que puedan influenciar en la publicación del artículo.

Financiamiento Los autores declaran no poseer fuentes de financiamiento externas asociadas a la realización de este artículo.

Idioma del envío Este artículo fue enviado a revisión por pares en su versión en español.

Origen y revisión por pares Este artículo es parte de la colección de "Notas metodológicas" que es fruto de un convenio de colaboración entre Medwave y la Cátedra de Metodología de la Investigación Científica y la cátedra de Medicina Basada en Evidencia de la Escuela de Medicina de la Universidad de Valparaíso. Con revisión por pares externa por tres pares revisores, a doble ciego.

REFERENCIAS

1. Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ*. 2016;353: i2016. <https://doi.org/10.1136/bmj.i2016>
2. MAGIC Foundation. In: MAGIC – Trustworthy guidelines, evidence summaries and decision aids that we can all use and share [Internet]. <https://magicvidence.org/>
3. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med*. 2010;7. <https://doi.org/10.1371/journal.pmed.1000326>
4. Cruzat B, Reveco-Guzmán K, Encina-Meneses M, Ortiz-Muñoz L, Bracchiglione J. Approaching the body of evidence: Key concepts of Overviews. *Medwave*. 23: e2704–e2704. <http://dx.doi.org/10.5867/medwave.2023.05.2704> <https://doi.org/10.5867/medwave.2023.05.2704>

5. Schünemann HJ, Reinap M, Piggott T, Laidmäe E, Köhler K, Pöld M, et al. The ecosystem of health decision making: from fragmentation to synergy. *Lancet Public Health*. 2022;7: e378–e390. [https://doi.org/10.1016/S2468-2667\(22\)00057-3](https://doi.org/10.1016/S2468-2667(22)00057-3)
6. Ravaud P, Créquit P, Williams HC, Meerpohl J, Craig JC, Boutron I. Future of evidence ecosystem series: 3. From an evidence synthesis ecosystem to an evidence ecosystem. *J Clin Epidemiol*. 2020;123: 153–161. <https://doi.org/10.1016/j.jclinepi.2020.01.027>
7. Sanabria AJ, Rigau D, Rotaecche R, Selva A, Marzo-Castillejo M, Alonso-Coello P. Sistema GRADE: metodología para la realización de recomendaciones para la práctica clínica. *Atención Primaria*. 2015;47: 48–55. <https://doi.org/10.1016/j.aprim.2013.12.013>
8. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol*. 2013;66: 719–25. <https://doi.org/10.1016/j.jclinepi.2012.03.013>
9. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *Journal of Clinical Epidemiology*. 2011;64: 395–400. <https://doi.org/10.1016/j.jclinepi.2010.09.012>
10. Johnston BC, Patrick DL, Devji T, Maxwell LJ, Bingham III CO, Beaton D, et al. Chapter 18: Patient-reported outcomes. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, editors. *Cochrane Handbook for Systematic Reviews of Interventions* version. 2019;6: 0. www.training.cochrane.org/handbook
11. Food and Drug Administration. In: Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims [Internet]. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims>
12. Value and use of patient-reported outcomes (PROs) in assessing effects of medical devices. US Food and Drug Administration. 2022; Available from. <https://www.fda.gov/files/about%20fda/published/Value-and-Use-of-Patient-Reported-Outcomes-%28PROs%29-in-Assessing-Effects-of-Medical-Devices.pdf>
13. Westwood M, Bourbeau J, Jones PW, Cerulli A, Capkun-Niggli G, Worthy G. Relationship between FEV1 change and patient-reported outcomes in randomised trials of inhaled bronchodilators for stable COPD: a systematic review. *Respir Res*. 2011;12: 40. <https://doi.org/10.1186/1465-9921-12-40>
14. Laugsand EA, Sprangers MAG, Bjordal K, Skorpen F, Kaasa S, Klepstad P. Health care providers underestimate symptom intensities of cancer patients: a multicenter European study. *Health Qual Life Outcomes*. 2010;8. <https://doi.org/10.1186/1477-7525-8-104>
15. Furberg CD, Pitt B. Withdrawal of cerivastatin from the world market. *Curr Control Trials Cardiovasc Med*. 2001;2: 205–207. <https://doi.org/10.1186/cvm-2-5-205>
16. Hincapié Tabares D, Pérez Carrillo V, Donado Gómez JH. Clasificación de los desenlaces en los ensayos clínicos. *MedUPB*. 2019;38: 147–157. <https://revistas.upb.edu.co/index.php/Medicina/issue/view/815/showToc> <https://doi.org/10.18566/medupb.v38n2.a07>
17. Institute of Medicine (US). Micheel CM, Ball JR, editors. Committee on Qualification of Biomarkers and Surrogate Endpoints in Chronic Disease. Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease. Washington (DC): National Academies Press (US); 2014.
18. Nissen SE, Wolski K. Rosiglitazone revisited: an updated meta-analysis of risk for myocardial infarction and cardiovascular mortality. *Arch Intern Med*. 2010;170: 1191–1201. <https://doi.org/10.1001/archinternmed.2010.207>
19. Bishop BM. Systematic review of CETP inhibitors for increasing high-density lipoprotein cholesterol: where do these agents stand in the approval process? *Am J Ther*. 2015;22: 147–58. <https://doi.org/10.1097/MJT.0b013e31828b8463>
20. Lin SY, Hung MC, Chang SF, Tsuang FY, Chang JZC, Sun JS. Efficacy and Safety of Postmenopausal Osteoporosis Treatments: A Systematic Review and Network Meta-Analysis of Randomized Controlled Trials. *JCM*. 2021;10: 3043. <http://dx.doi.org/10.3390/jcm10143043> <https://doi.org/10.3390/jcm10143043>
21. Knopf K, Baum M, Shimp WS, Bennett CL, Faith D, Fishman ML, et al. Interpretation of surrogate endpoints in the era of the 21st Century Cures Act. *BMJ*. 2016;355: i6286. <https://doi.org/10.1136/bmj.i6286>
22. Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. *BMJ*. 2011;343: d7995. <https://doi.org/10.1136/bmj.d7995>
23. Molenberghs G, Alonso A, Van derW, Buyse M. In: Statistical Evaluation of Surrogate Endpoints in Clinical Studies [Internet]. 14 Nov 2014. https://www.researchgate.net/publication/278739754_Statistical_Evaluation_of_Surrogate_Endpoints_in_Clinical_Studies
24. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. *Trials*. 2017;18. <https://doi.org/10.1186/s13063-017-1978-4>
25. COMET Initiative. <https://www.comet-initiative.org>
26. Webbe J, Sinha I, Gale C. Core Outcome Sets. *Arch Dis Child Educ Pract Ed*. 2018;103: 163–166. <https://pubmed.ncbi.nlm.nih.gov/28667046/> <https://doi.org/10.1136/archdischild-2016-312117>
27. Kirkham JJ, Williamson P. Core outcome sets in medical research. *bmjmed*. 2022;1: e000284. <https://bmjmedicine.bmj.com/content/1/1/e000284> <https://doi.org/10.1136/bmjmed-2022-000284>
28. Schmitt J, Spuls P, Boers M, Thomas K, Chalmers J, Roekevisch E, et al. Towards global consensus on outcome measures for

- atopic eczema research: results of the HOME II meeting. *Allergy*. 2012;67: 1111–7. <https://doi.org/10.1111/j.1398-9995.2012.02874.x>
29. Musgrove E, Gasparini L, McBain K, Clifford SA, Carter SA, Teede H, et al. Synthesizing Core Outcome Sets for outcomes research in cohort studies: a systematic review. *Pediatr Res*. 2022;92: 936–945. <https://doi.org/10.1038/s41390-021-01801-2>
30. Molina Arias M. Aspectos metodológicos del metaanálisis (2). *Rev Pediatr Aten Primaria*. 2018;20(80):401–5.
31. Sgarbossa N, Ibáñez Cobaisse M, González Cianciulli G, Bracchiglione J, Franco JVA. Systematic reviews: Key concepts for health professionals. *Medwave*. 2022;22. <http://dx.doi.org/10.5867/medwave.2022.09.2622> <https://doi.org/10.5867/medwave.2022.09.2622>
32. Madrid Aris E, Martínez Lomakin F. Statistics for the faint of heart – how to interpret confidence intervals and p values. *Medwave*. 2014;14: e5892. <https://doi.org/10.5867/medwave.2014.01.5892>
33. Jiménez-Paneque R. El valor de p en entredicho: significación estadística, clínica y práctica. *Medwave*. 2016;16: e6534–e6534. <https://doi.org/10.5867/medwave.2016.08.6534>
34. Johnston BC, Ebrahim S, Carrasco-Labra A, Furukawa TA, Patrick DL, Crawford MW, et al. Minimally important difference estimates and methods: a protocol. *BMJ Open*. 2015;5. <https://doi.org/10.1136/bmjopen-2015-007953>
35. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol*. 2011;64: 1283–93. <https://doi.org/10.1016/j.jclinepi.2011.01.012>
36. Salas Apaza JA, Franco JVA, Meza N, Madrid E, Loézar C, Garegnani L. Minimal clinically important difference: The basics. *Medwave*. 2021;21: e8149. <https://doi.org/10.5867/medwave.2021.03.8149>
37. Tejwani V, Chang H-Y, Tran AP, Naber JA, Gutzwiller FS, Winders TA, et al. A multistakeholder Delphi consensus core outcome set for clinical trials in moderate-to-severe asthma (coreASTHMA). *Ann Allergy Asthma Immunol*. 2021;127: 116–122. <https://doi.org/10.1016/j.anai.2021.03.022>
38. Bonfill Cosp X. Conceptos básicos para entender y mejorar la adecuación clínica. *Educación Médica*. 2023;24: 100813. <https://doi.org/10.1016/j.edumed.2023.100813>

Key factors in the choice of appropriate outcomes for clinical trials

ABSTRACT

Health research is the foundation of medical knowledge and healthcare system recommendations. Therefore, choosing appropriate outcomes in studies of therapeutic interventions is a fundamental step in producing evidence and, subsequently, for decision-making. In this article, we propose three key factors for the choice of outcomes: the inclusion of patient-reported outcomes, since they focus on the patient's perception of their health status and quality of life; the consideration of clinically relevant outcomes, which are direct measurements of the patient's health status and, therefore, will be decisive in decision-making; and the use of core outcome sets, a tool that standardizes the measurement and interpretation of outcomes, facilitating the production and synthesis of appropriate evidence for the evidence ecosystem. The correct choice of outcomes will help health decision-makers and clinicians deliver appropriate patient-centered care and optimize the use of resources in healthcare and clinical research.



This work is licensed under a Creative Commons Attribution 4.0 International License.