



Instituto de Evaluación
Tecnológica en Salud

MANUAL

METODOLÓGICO



EVALUACIÓN

DE

TECNOLOGÍAS EN SALUD



Instituto de Evaluación
Tecnológica en Salud

**Manual metodológico para la elaboración
de evaluaciones de efectividad, seguridad
y validez diagnóstica de tecnologías en
salud**

El Instituto de Evaluación Tecnológica en Salud – IETS, es una corporación sin ánimo de lucro, de participación mixta y de carácter privado, con patrimonio propio, creado según lo estipulado en la Ley 1438 de 2011. Su misión es contribuir al desarrollo de mejores políticas públicas y prácticas asistenciales en salud, mediante la producción de información basada en evidencia, a través de la evaluación de tecnologías en salud y guías de práctica clínica, con rigor técnico, independencia y participación. Sus miembros fundadores son el Ministerio de Salud y Protección Social, el Departamento Administrativo de Ciencia, Tecnología e Innovación –Colciencias, el Instituto Nacional de Vigilancia de Medicamentos y Alimentos – INVIMA, el Instituto Nacional de Salud – INS y la Asociación Colombiana de Facultades de Medicina –ASCOFAME.

Autores

Miguel Hernando Díaz Ortega. Bacteriólogo y Laboratorista Clínico, MSc en Epidemiología Clínica. Instituto de Evaluación Tecnológica en Salud – IETS.

Esperanza Peña Torres. Enfermera, MSc en Administración en Salud, MSc en Epidemiología Clínica. Instituto de Evaluación Tecnológica en Salud – IETS.

Aurelio Mejía Mejía. Economista, MSc en Economía de la Salud. Instituto de Evaluación Tecnológica en Salud – IETS.

Iván Darío Flórez Gómez. Médico Pediatra, MSc en Epidemiología Clínica. Instituto de Evaluación Tecnológica en Salud – IETS.

Revisores

Andrés Pichón-Riviere. Médico, MSc en Epidemiología, PhD en Medicina con orientación en Salud Pública. Instituto de Efectividad Clínica y Sanitaria – IECS (Argentina).

Federico Augustovski. Médico, MSc en Epidemiología. Instituto de Efectividad Clínica y Sanitaria – IECS (Argentina).

Sebastián García Martí. Médico MSc en Efectividad Clínica. Instituto de Efectividad Clínica y Sanitaria – IECS (Argentina).

Juan Manuel Lozano León. Médico, Pediatra, MSc en Epidemiología Clínica. South Florida University.

Guillermo Sánchez Vanegas. Médico, Especialista y MSc en Epidemiología Clínica, PhD en Salud Pública. Fundación Universitaria de Ciencias de la Salud – FUCS.

Elizabeth Parody Rúa. Química Farmacéutica, MSc en Economía de la Salud y Gestión Sanitaria, PhD en Salud Pública y Metodología de la Investigación Biomédica. Universidad ICESI.

Ángela Viviana Pérez Gómez. Médica, MSc (c) en Epidemiología Clínica. Instituto de Evaluación Tecnológica en Salud – IETS.

Laura Catalina Prieto Pinto. Médica, Especialista en Epidemiología. Instituto de Evaluación Tecnológica en Salud – IETS.

Lorena Andrea Cañón Betancourt. Médica, Especialista en Epidemiología. Instituto de Evaluación Tecnológica en Salud – IETS.

Mabel Moreno. Economista, especialista en Estadística, MSc (c) en Economía. Instituto de Evaluación Tecnológica en Salud – IETS.

Diana Isabel Osorio. Economista. Instituto de Evaluación Tecnológica en Salud – IETS.

Fuentes de financiación

Instituto de Evaluación Tecnológica en Salud – IETS.

Conflictos de interés

Los autores de este manual declaran que no existe ningún tipo de conflicto financiero, intelectual, de pertenencia o familiar que afecte las recomendaciones incluidas en el presente manual.

Citación

Este documento debe citarse de la siguiente manera:

Instituto de Evaluación Tecnológica en Salud. Manual metodológico para la elaboración de evaluaciones de efectividad, seguridad y validez diagnóstica de tecnologías en salud. Bogotá D.C.: IETS; 2014.

Derechos de autor

El contenido de este documento, sin perjuicio de las citas y referencias bibliográficas enunciadas, es propiedad del Instituto de Evaluación Tecnológica en Salud – IETS. Por tanto, está prohibida su copia, reproducción, fijación, transmisión, divulgación o similares. En consecuencia, cualquier modificación, transmisión, divulgación, publicación, copia parcial o total, o el uso del contenido del mismo sin importar su propósito y sin que medie el consentimiento expreso y por escrito del Instituto de Evaluación Tecnológica en Salud - IETS, constituirá violación a la normativa nacional, convenios y tratados internacionales ratificados por el Estado colombiano aplicables a los derechos de autor, y acarreará las sanciones civiles, comerciales y penales a que haya lugar.

Correspondencia

Instituto de Evaluación Tecnológica en Salud - IETS

Autopista Norte 118 - 30 Of. 201

Bogotá, D.C., Colombia.

www.iets.org.co

subdireccion.etes@iets.org.co

© Instituto de Evaluación Tecnológica en Salud, 2014.

Tabla de Contenido

Autores	2
Revisores	2
Fuentes de financiación	3
Conflictos de interés	3
Citación.....	3
Derechos de autor	3
Correspondencia	3
Presentación.....	8
Resumen ejecutivo	9
Introducción	10
1. Condición de salud y tecnologías de interés.....	11
1.1. Condición de salud de interés	11
1.2. Tecnologías en salud de interés.....	11
2. Preguntas de evaluación.....	12
2.1. Formulación de las preguntas de evaluación	12
Cuadro 1. Pregunta de evaluación en estructura PICOT	12
2.2. Refinamiento de las preguntas de evaluación.....	15
2.3. Clasificación de la importancia de los desenlaces	15
3. Métodos.....	16
3.1. Criterios de elegibilidad.....	17
3.1.1. Criterios de inclusión.....	17
3.1.2. Criterios de exclusión.....	20
3.2. Búsqueda de evidencia.....	20
3.2.1. Búsqueda en bases de datos electrónicas.....	20
3.2.2. Otros métodos de búsqueda	21
3.2.3. Actualización de la búsqueda de estudios primarios para una revisión sistemática	22

3.2.4.	Gestión documental	22
3.3.	Tamización de referencias y selección de estudios.....	23
3.4.	Evaluación de la calidad de la evidencia	23
3.5.	Extracción de datos y síntesis de la evidencia.....	23
3.6.	Análisis estadístico.....	24
3.6.1.	Metanálisis clásico.....	25
3.6.2.	Metanálisis de comparaciones indirectas y de múltiples comparaciones.....	25
3.6.3.	Adición de nuevos datos a las medidas combinadas del efecto	26
4.	Resultados	26
4.1.	Búsqueda de evidencia.....	26
4.2.	Tamización de referencias y selección de estudios.....	26
4.3.	Calidad de la evidencia.....	26
4.4.	Síntesis de la evidencia	27
4.5.	Descripción de los estudios.....	27
4.6.	Efectividad y seguridad / validez y concordancia de pruebas diagnósticas	27
	Cuadro 2. Perfil de evidencia GRADE.....	29
	Cuadro 3. Perfil de evidencia GRADE.....	30
5.	Discusión.....	31
6.	Conclusiones.....	32
	Referencias bibliográficas.....	33
	Anexos.....	35
	Anexo 1. Registros sanitarios vigentes para las tecnologías de interés	35
	Anexo 2. Clasificación de la importancia de los desenlaces	36
	Anexo 3. Reportes de búsqueda de evidencia en bases de datos electrónicas.....	37
	Anexo 4. Diagrama de flujo de la búsqueda, tamización y selección de evidencia (búsqueda <i>de novo</i>).....	44
	Anexo 5. Diagrama de flujo de la búsqueda, tamización y selección de evidencia (actualización de la búsqueda de estudios primarios para una revisión sistemática).....	45
	Anexo 6. Listado de estudios incluidos en la evaluación	46

Anexo 7. Listado de estudios excluidos de la evaluación y razones para su exclusión.....	47
Anexo 8. Calidad de las revisiones sistemáticas incluidas en la evaluación (herramienta AMSTAR)	48
Anexo 9. Calidad de los metanálisis de comparaciones indirectas o de múltiples comparaciones incluidos en la evaluación (herramienta ISPOR)	50
Anexo 10. Calidad de los ensayos controlados aleatorizados incluidos en la evaluación (riesgo de sesgo de la Colaboración Cochrane)	53
Anexo 11. Calidad de los estudios de cohortes analíticas incluidos en la evaluación (riesgo de sesgo de la Colaboración Cochrane).....	54
Anexo 12. Calidad de los estudios primarios de validez diagnóstica incluidos en la evaluación (herramienta QUADAS-2)	55
Anexo 13. Características de las revisiones sistemáticas incluidas en la síntesis de evidencia	57
Anexo 14. Características de los ensayos controlados aleatorizados incluidos en la síntesis de evidencia	57
Anexo 15. Características de los estudios de cohortes incluidos en la síntesis de evidencia	59
Anexo 16. Características de los estudios primarios de validez diagnóstica o de concordancia de pruebas diagnósticas incluidos en la síntesis de evidencia	60
Glosario	61

Lista de abreviaturas y siglas

ATC	Anatomical, Therapeutic, Chemical classification system.
CUPS	Código Único de Procedimientos en Salud.
ECA	Ensayo Controlado Aleatorizado.
GRADE	Grading of Recommendations Assessment, Development and Evaluation.
HR	Hazard Ratio.
IC	Intervalo de Confianza.
IETS	Instituto de Evaluación Tecnológica en Salud.
INVIMA	Instituto Nacional de Vigilancia de Medicamentos y Alimentos.
ITC	Indirect Treatment Comparisons.
MeSH	Medical Subject Headings.
MTC	Multiple Treatment Comparison.
NA	No Aplica.
OMS	Organización Mundial de la Salud.
OR	Odds Ratio.
POS	Plan Obligatorio de Salud.
ROC	Receiver Operating Characteristic curve.
RR	Riesgo Relativo.
RSL	Revisión Sistemática de la Literatura.

Presentación

El reporte de evaluación de efectividad, seguridad y validez diagnóstica de tecnologías en salud es un documento técnico elaborado con base en la mejor evidencia disponible sobre los beneficios, daños o utilidad diagnóstica de las tecnologías, como uno de los criterios empleados para informar la toma de decisiones en políticas relacionadas con el cuidado de la salud.

El objetivo de este manual es apoyar al Instituto de Evaluación Tecnológica en Salud - IETS, centros desarrolladores de evaluación de tecnologías y grupos externos, en la rigurosa ejecución de los requisitos básicos que debe cumplir un reporte de evaluación de efectividad, seguridad y validez diagnóstica de tecnologías en salud para Colombia.

La adopción de pautas metodológicas por los centros y grupos desarrolladores de evaluación de tecnologías en Colombia, facilita el desarrollo sistemático de evaluaciones, disminuye la variabilidad del uso de los métodos y permitiría la incorporación de los resultados de las evaluaciones como un criterio para la formulación de políticas públicas en salud.

El presente documento fue desarrollado con base en una revisión de metodologías de agencias de evaluación de tecnologías como Health Technology Assessment (HTA), Canadian Agency for Drugs and Technologies in Health (CADTH), European network for Health Technology Assessment (EUnetHTA) y el Instituto de Efectividad Clínica y Sanitaria (IECS). Los métodos incluidos tienen como principal referente lo propuesto por The Cochrane Collaboration y GRADE Working Group.

Este manual fue presentado en reuniones con expertos provenientes de centros académicos, industria farmacéutica, agencias de evaluación de tecnologías y se sometió a revisión externa por pares académicos. Se actualizará anualmente en caso de ser necesario.

De acuerdo con la metodología del Instituto de Evaluación Tecnológica en Salud - IETS, el proceso de evaluación de seguridad y efectividad debe ser un proceso participativo y todos los involucrados deben hacer una declaración de conflictos de interés y cuando corresponda deberán firmar acuerdos de confidencialidad. El IETS ha diseñado los formatos para tal fin (1).

Resumen ejecutivo

El reporte debe contener un resumen que no supere las 300 palabras y debe estar construido bajo la siguiente estructura:

- Introducción.
- Objetivo.
- Metodología.
- Resultados.
- Conclusiones.

Introducción

En la elaboración del reporte, la introducción debe contener los siguientes elementos:

- Síntesis del problema a investigar.
- Justificación.
- Alcance.
- Objetivo de la evaluación.

1. Condición de salud y tecnologías de interés

1.1. Condición de salud de interés

La descripción de la condición de salud de interés debe contener los siguientes aspectos: definición, factores de riesgo, historia natural, fisiopatología, sintomatología, diagnóstico, tratamiento, curso clínico, pronóstico y carga de enfermedad.

Para la caracterización epidemiológica de la condición de salud, debe documentarse su prevalencia o incidencia a partir de publicaciones preferiblemente nacionales y fuentes de información local como encuestas nacionales de salud, SISPRO, SIMIGILA, etc. Asimismo, el reporte debe mencionar cuando no sea posible identificar datos específicos de frecuencia nacional, regional o local para Colombia de la condición de salud de interés.

1.2. Tecnologías en salud de interés

Esta sección describe las intervenciones que son objeto de la evaluación. Según corresponda, deben incluirse los siguientes elementos:

- Medicamentos: nombre de la(s) tecnología(s) y la descripción del código ATC (para el caso de los medicamentos en Denominación Común Internacional) mediante una figura titulada "Clasificación ATC para la(s) tecnología(s) de interés", junto con su respectiva fuente (OMS). Con base en los registros sanitarios vigentes (INVIMA) debe especificarse: principio activo, nombre comercial, forma farmacéutica, concentración, presentación, indicaciones y titular del registro. Además debe incluirse información relacionada con el mecanismo de acción, dosificación para la población de interés de la evaluación, forma de administración, contraindicaciones, precauciones, ámbito de administración e identificación según el Código Único de Medicamentos (CUM) correspondiente al laboratorio que produce o comercializa la tecnología.
- Pruebas diagnósticas, dispositivos médicos y reactivos: nombre de las tecnologías y los registros sanitarios vigentes. Se debe especificar el nombre del producto, presentación, usos, observaciones que se consideren relevantes y el titular del registro. En el caso de las pruebas diagnósticas y los dispositivos médicos, debe describirse su proceso de aplicación y contraindicaciones.
- Procedimientos: nombre de las tecnologías, indicación, proceso de aplicación y el CUPS. Para las tecnologías con registro sanitario INVIMA, debe presentarse el Anexo 1, incluyendo la fuente y fecha de corte para la consulta.

Adicionalmente, deben señalarse las fuentes de financiación de las tecnologías (plan de beneficios, fondos o programas especiales, recobros, gasto de bolsillo).

2. Preguntas de evaluación

2.1. Formulación de las preguntas de evaluación

Las preguntas de evaluación deben plantearse mediante la estructura PICOT:

- P: población.
- I: tecnologías de interés.
- C: comparadores.
- O: desenlaces (del inglés *outcomes*).
- T: tiempo (si aplica).

En el caso de evaluaciones sobre pruebas diagnósticas, los términos 'tecnologías de interés' y 'comparadores' se reemplazan por 'pruebas índice y estándar' o 'patrón de referencia', respectivamente.

Debe emplearse uno de los siguientes formatos, según corresponda:

- Efectividad y seguridad: en [población], ¿cuál es la efectividad y seguridad de [tecnologías de interés] comparado con [comparadores], para [desenlaces]?
- Validez diagnóstica: en [población], ¿cuál es la validez diagnóstica de [pruebas índice] comparado con [estándar o patrón de referencia], para [desenlaces]?
- Concordancia de pruebas diagnósticas: en [población], ¿cuál es la concordancia diagnóstica de [prueba A] comparada con [prueba B], para [desenlaces]?

Las preguntas deben presentarse de forma desagregada, tal como lo indica el siguiente cuadro:

Cuadro 1. Pregunta de evaluación en estructura PICOT

P	
I	
C	
O	
T	

Deben tenerse en cuenta las siguientes consideraciones para cada uno de los elementos de la pregunta PICOT:

- Población: la población objeto de la aplicación de la tecnología debe incluir la condición de salud de interés, sospecha clínica, estado, duración de la enfermedad (si aplica), falla a tratamientos previos, comorbilidades, cointervenciones y características sociodemográficas tales como edad, sexo, raza, nivel educativo, estrato socioeconómico, localización geográfica, etc.
- Tecnologías de interés: corresponde a las intervenciones que son objeto de la evaluación. Las tecnologías a evaluar deben presentarse de forma genérica; en este punto no deben incluirse especificidades de las tecnologías como marcas, dosis, concentraciones o formas farmacéuticas (que ya se describieron previamente). Así mismo, debe especificarse el propósito de uso de las tecnologías: prevención, tamización, diagnóstico, tratamiento, seguimiento, rehabilitación, cuidado paliativo u organizacional, y si su ámbito de uso es hospitalario o ambulatorio.

Para el caso de los medicamentos, debe especificarse la modalidad del tratamiento (monoterapia, terapia multimodal, tratamiento coadyuvante, terapia secuencial, etc.) junto con la línea terapéutica de interés. En el caso de tecnologías diagnósticas, debe precisarse si la prueba es única, en paralelo o en serie.

Adicionalmente, debe justificarse el análisis de aquellas tecnologías que son imprescindibles para el uso de las tecnologías de interés, es decir, que son complementarias.

- Comparadores: son las intervenciones que tienen la misma indicación de las tecnologías de interés frente a las cuales se desea conocer, de modo comparativo, los beneficios, daños o utilidad diagnóstica. Idealmente, los comparadores deben ser las tecnologías de referencia en la práctica clínica (estándares de atención).

En el caso de las evaluaciones de validez diagnóstica, deben seleccionarse, siempre que sea posible, patrones de referencia fuertes (por ejemplo, biopsia, autopsia, cirugía, etc.). Si el patrón de referencia seleccionado no se ha validado aún como patrón de oro de la condición de salud de interés, deben presentarse los supuestos para su uso.

Los comparadores deben definirse de forma genérica y sin considerar marcas, dosis, concentraciones o formas farmacéuticas.

Cuando no se identifique información de seguridad para las comparaciones predefinidas, se considera adecuado incluir evidencia de las tecnologías de interés con placebo como comparador.

- Desenlaces: son los resultados en salud atribuibles al uso de las tecnologías.

Para el caso de la efectividad y la seguridad, el grupo desarrollador de la evaluación debe seleccionar desenlaces finales (independientemente del tiempo de seguimiento) que sean relevantes tanto para los pacientes como para los servicios y sistemas de salud, es decir, que midan directamente el beneficio o daño clínico asociado con el uso de las tecnologías para la condición de salud de interés (por ejemplo, morbilidad y mortalidad) sin considerar si han sido o no reportados en estudios previos.

Ante la ausencia de información sobre desenlaces finales, el grupo desarrollador puede considerar relevante incluir evidencia para desenlaces subrogados. El reporte debe ser explícito sobre cuáles desenlaces se consideran en esta categoría y deben estar adecuadamente validados. Su relación con el desenlace final debe haber sido demostrada con base en plausibilidad biológica y evidencia empírica (incluir las referencias bibliográficas de soporte), dando la posibilidad de sustituir de forma confiable a los desenlaces clínicos finales y predecir el efecto (beneficio o daño) de las tecnologías bajo evaluación.

No deben emplearse desenlaces compuestos si existe un desenlace primario sencillo que es relevante.

En cuanto a la evaluación de seguridad, esta debe enfocarse en los eventos adversos serios (2) y frecuentes, considerando los daños durante el uso de la tecnología tras su abandono o suspensión y asociados con interacciones.

En el caso de evaluaciones de validez diagnóstica, debe considerarse por defecto el desenlace de presencia o ausencia de la condición de salud de interés. En este tipo de evaluación debe incluirse evidencia sobre seguridad, cuando se conozca que el uso de la tecnología puede estar asociado a eventos adversos serios y frecuentes.

Todos los desenlaces deben definirse como variables, sin especificar si el desenlace se favorece o no por el efecto de la intervención.

- Tiempo: debe precisarse el tiempo para la medición de los desenlaces cuando se evalúa sobrevida y se considera el tiempo de seguimiento como una variable crítica para los desenlaces de efectividad y seguridad.

Para la formulación preliminar de las preguntas de evaluación, deben seguirse los siguientes pasos:

- a) Consultar los registros sanitarios vigentes de la página del INVIMA para las tecnologías e indicación de interés, registrando la fecha de corte para la consulta. No deben incluirse en la evaluación tecnologías sin el registro sanitario para la indicación de interés. En caso de no identificar registros sanitarios en la base del INVIMA, debe revisarse el listado de medicamentos vitales no disponibles de dicha entidad.

- b)** Identificar (para el caso de evaluaciones de medicamentos) los posibles comparadores mediante la revisión de grupos farmacológicos (clasificación ATC), con base en la información del Centro Colaborador de la OMS para la metodología de estadísticas de medicamentos (http://www.whocc.no/atc_ddd_index/). Este paso debe complementarse con la consulta del listado de medicamentos esenciales de la OMS para adultos o niños, según sea el caso.
- c)** Delimitar la población y los comparadores a través de la revisión de recomendaciones de guías de práctica clínica identificadas en las siguientes fuentes:
- National Guideline Clearinghouse (NGC)
<http://www.guideline.gov/index.aspx>
 - Guidelines International Network (GIN)
<http://www.g-i-n.net/gin>
 - New Zealand Guidelines Group (NZGG)
<http://www.health.govt.nz/>
 - Scottish Intercollegiate Guidelines Network (SIGN)
<http://www.sign.ac.uk/index.html>
 - Ministerio de Salud y Protección Social - IETS
<http://www.iets.org.co/>
 - GuíaSalud
<http://portal.guiasalud.es/web/guest/home>
 - Centro Nacional de Excelencia Tecnológica en Salud (CENETEC)
<http://www.cenetec.salud.gob.mx/>

Esta consulta debe complementarse con un rápido repaso de las revisiones Cochrane, revisiones narrativas, estudios primarios y documentos sobre políticas vigentes de cobertura.

2.2. Refinamiento de las preguntas de evaluación

De acuerdo con la metodología propuesta en el 'Manual de Procesos Participativos del IETS'(1), la pregunta de evaluación debe refinarse mediante consulta con actores relevantes como expertos metodológicos y temáticos, pacientes y cuidadores, profesionales de la salud entre otros. La pregunta con estructura PICOT preliminar debe publicarse en la página web del IETS para comentarios, y tras los ajustes que tengan lugar, debe procederse a su publicación como pregunta definitiva.

2.3. Clasificación de la importancia de los desenlaces

En caso de que la evaluación conlleve a la realización de un perfil de evidencia GRADE, los resultados de la graduación de desenlaces deben documentarse en el Anexo 2. Para las evaluaciones de efectividad y seguridad, los expertos temáticos, pacientes y cuidadores, deben valorar de manera independiente la importancia de las variables de resultado definidas para la evaluación mediante consulta presencial según lo establece el manual de procesos participativos

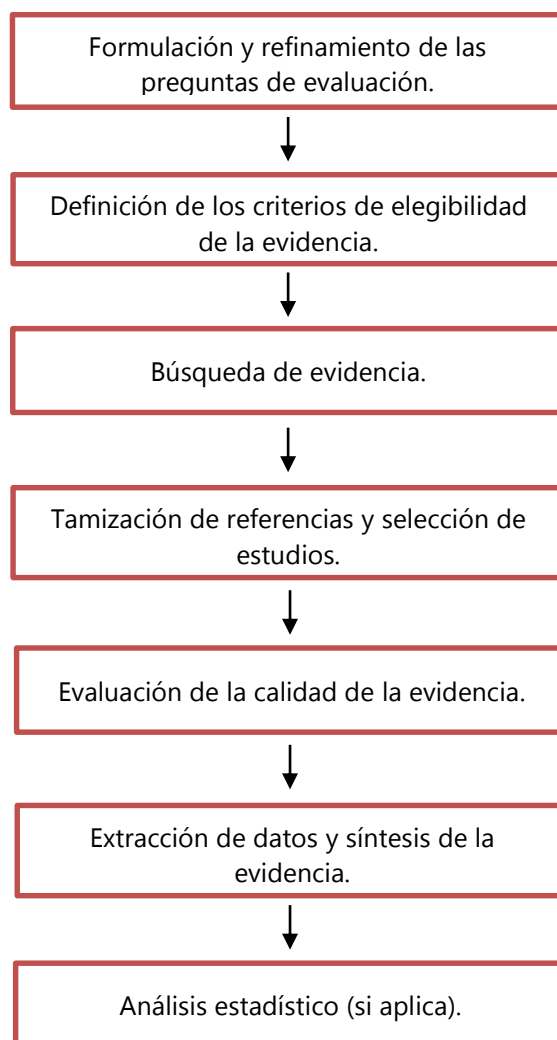
del IEST (1). Para hacer la clasificación se debe aplicar la siguiente escala Likert, según la metodología propuesta por GRADE Working Group (3-4).

1	2	3	4	5	6	7	8	9
Desenlace no importante para la toma de decisiones (no se incluye en el perfil de evidencia).			Desenlace importante, pero no crítico para la toma de decisiones (se incluye en el perfil de evidencia).			Desenlace crítico para la toma de decisiones (se incluye en el perfil de evidencia).		

Con base en la puntuación media del grupo, cada desenlace debe clasificarse en una de las tres categorías descritas.

3. Métodos

Los pasos de la evaluación se presentan en el siguiente diagrama de flujo:



El reporte de la sección de métodos debe ser tan detallado que permita replicar la evaluación, la cual, debe realizarse de acuerdo con un protocolo definido *a priori* por el grupo desarrollador. Este protocolo debe publicarse en la página web del IETS, según la metodología propuesta en el manual de procesos participativos (1), y con previa revisión por pares.

Este protocolo debe contener la siguiente información:

- Preguntas de evaluación en estructura PICOT.
- Métodos

3.1. Criterios de elegibilidad

A partir de las preguntas PICOT se definen los criterios de elegibilidad de la evidencia, en términos de criterios de inclusión y exclusión.

3.1.1. Criterios de inclusión

Los criterios de inclusión definen el perfil de la población, subgrupos, tecnologías de interés, comparadores, desenlaces, tiempo (si aplica) y tipo de estudios que debe contener la evaluación. En caso de considerar relevante analizar la efectividad y seguridad o validez diagnóstica de las tecnologías de interés en subgrupos específicos de la población, estos deben predefinirse como variables.

Criterios de inclusión de estudios relacionados con el formato, idioma y fecha de publicación

Los criterios que se presentan a continuación son orientativos. En todo caso, los criterios adoptados para la evaluación deben ser explícitos.

- El estudio debe estar disponible como publicación completa. Los estudios publicados únicamente en formato de resumen no son considerados porque la información reportada es incompleta para evaluar su calidad metodológica. Adicionalmente, los resultados de los estudios pueden cambiar significativamente entre la presentación inicial en un evento y la publicación final en un medio impreso.
- Estudios en inglés o en español, independientemente de su estado de publicación (publicado, en prensa o literatura gris). Debe especificarse si se considera relevante incluir estudios publicados en un idioma diferente a los mencionados.
- Revisiones sistemáticas publicadas en los últimos 5 años. En este punto debe justificarse si se considera pertinente emplear un rango de fecha de publicación diferente.

Criterios de inclusión de estudios relacionados con el diseño, la calidad metodológica y de reporte

Los pasos que se describen a continuación orientan de forma secuencial las consideraciones para la inclusión de estudios. En todo caso, los criterios adoptados para la evaluación deben ser explícitos.

Pregunta de efectividad y seguridad:

Paso 1. Revisiones panorámicas (*overview of reviews*) o revisiones sistemáticas de ECA o de estudios de cohortes analíticas.

Se considera actualmente que las RSL aportan información clave para la toma de decisiones en salud. Sus resultados dan cuenta de los efectos (beneficios y daños) de las intervenciones en una población. En comparación con estudios primarios, ofrecen ventajas como la minimización de los sesgos, tienen mayor poder estadístico, reducen el error aleatorio y proporcionan perspectivas complementarias de las opciones bajo comparación (5-7).

Debe verificarse que el estudio:

- a)** Es una RSL. De acuerdo con los criterios propuestos en el manual Cochrane de revisiones sistemáticas de intervenciones, el estudio tiene (6):
 - Un conjunto de objetivos claramente establecidos, con criterios de elegibilidad de estudios previamente definidos.
 - Una metodología explícita y reproducible.
 - Una búsqueda sistemática que identifique todos los estudios que puedan cumplir los criterios de elegibilidad.
 - Una evaluación de la validez de los resultados de los estudios incluidos, por ejemplo, mediante la evaluación del riesgo de sesgos.
 - Una presentación sistemática y una síntesis de las características y resultados de los estudios incluidos.
- b)** Incluye comparaciones directas (cabeza a cabeza) entre las tecnologías de interés.
- c)** Reporta datos cuantitativos para al menos una comparación y desenlace de interés.
- d)** Incluye estudios con un tiempo de seguimiento suficiente para medir los desenlaces de efectividad o seguridad de interés para la evaluación. Mediante consulta virtual, esta información se contrastará con el concepto de los expertos temáticos del grupo desarrollador de la evaluación.

Paso 2. Revisiones sistemáticas con ITC o MTC.

Su inclusión está limitada a la ausencia de RSL con comparaciones directas entre las tecnologías de interés.

Paso 3. Ensayos controlados aleatorizados

Su inclusión está limitada a la ausencia de RSL que cumplan los criterios anteriormente descritos.

Debe verificarse que el estudio:

- a) No esté incluido en al menos una RSL seleccionada para la síntesis de evidencia.
- b) Es un ECA fase IV.
- c) Tiene un registro en al menos una base de datos internacional.
- d) Tiene bajo riesgo de sesgo (remitirse a la sección "Evaluación de la calidad de la evidencia").
- e) Tiene pérdidas al seguimiento menores al 20% (8).
- f) Tiene un tiempo de seguimiento suficiente para medir los desenlaces de efectividad o seguridad de interés para la evaluación. Mediante consulta virtual, esta información se contrastará con el concepto de los expertos temáticos del grupo desarrollador de la evaluación.

Paso 4. Estudios de cohortes analíticos

Su inclusión está limitada a la ausencia de ECA que cumplan los criterios anteriormente descritos.

Debe verificarse que el estudio:

- a) No esté incluido en al menos una RSL seleccionada para la síntesis de evidencia.
- b) Tiene bajo riesgo de sesgo (remitirse a la sección "Evaluación de la calidad de la evidencia").
- c) Tiene pérdidas al seguimiento menores al 20% (9).
- d) Tiene un adecuado control de la confusión.
- e) El tiempo de seguimiento para los desenlaces de efectividad o seguridad es mayor al de los ECA disponibles.

Pregunta de validez y concordancia de pruebas diagnósticas:

Paso 1. Revisiones panorámicas o revisiones sistemáticas de estudios de validez diagnóstica.

Debe verificarse que el estudio:

- a) Es una RSL de acuerdo con los criterios anteriormente señalados.
- b) Reporta datos cuantitativos para al menos una comparación y desenlace de interés.

Paso 2. Estudios primarios de validez diagnóstica

Su inclusión está limitada a la ausencia de RSL que cumplan los criterios anteriormente descritos.

Debe verificarse que el estudio:

- a) No esté incluido en al menos una RSL seleccionada para la síntesis de evidencia.
- b) Tiene buena calidad (remitirse a la sección "Evaluación de la calidad de la evidencia").

Debe considerarse la relevancia de incluir, de forma complementaria, estudios de concordancia entre pruebas diagnósticas; esto es, cuando se tenga como objetivo determinar si una prueba diferente a la habitual tiene una utilidad diagnóstica similar que las haga intercambiables en su uso.

Los estudios primarios siempre se tendrán en cuenta para efectos de la actualización de las RSL seleccionadas para la síntesis de evidencia. Para mayor información, remitirse a la sección "Actualización de la búsqueda de estudios primarios para una revisión sistemática". Los nuevos estudios identificados deben cumplir los criterios de inclusión y no cumplir ningún criterio de exclusión de la revisión que es objeto de la actualización.

3.1.2. Criterios de exclusión

Los criterios de exclusión deben definir las características que obligan a descartar de la evaluación a aquellos sujetos (población), estudios o desenlaces que cumplen los criterios de inclusión, pero que no pueden admitirse por diferentes motivos. Los criterios de exclusión no deben plantearse como una negación de los criterios de inclusión.

Deben precisarse aquellas situaciones médicas, éticas, de conveniencia, adherencia u otras que obliguen a excluir a la población que cumple los criterios de inclusión, por considerar que introducen sesgos en la evaluación.

Asimismo, deben excluirse de la evaluación los desenlaces intermedios (entre la línea de base y el punto final del estudio).

3.2. Búsqueda de evidencia

Con el objetivo de identificar evidencia relevante en relación con las preguntas de evaluación, debe llevarse a cabo una búsqueda sistemática y exhaustiva. Todo el proceso debe acogerse a los estándares de calidad internacional utilizados en RSL (6). Las búsquedas deben llevarse a cabo por personal entrenado.

3.2.1. Búsqueda en bases de datos electrónicas

Para identificar publicaciones indexadas, deben consultarse las siguientes fuentes:

- MEDLINE, de ser posible incluir los repositorios In-Process & Other Non-Indexed Citations y Daily Update (plataforma Ovid o PubMed).

- EMBASE (Elsevier).
- Cochrane Database of Systematic Reviews (plataforma Wiley).
- Database of Abstracts of Reviews of Effects - DARE (plataforma Wiley).
- LILACS (Biblioteca Virtual en Salud - BVS, interfaz iAHx).

Si se requiere la búsqueda de ECA, deben consultarse además de las anteriores, las siguientes bases de datos:

- Cochrane Central Register of Controlled Trials - CENTRAL (plataforma Ovid).
- WHO International Clinical Trials Registry Platform ICTRP portal.
- ClinicalTrials.gov.

En caso de ser necesario, también debe incluirse la identificación de publicaciones a través del motor de búsqueda Google.

Para la realización de las búsquedas, inicialmente se identifican los términos clave en lenguaje natural a partir de las preguntas en formato PICOT. Posteriormente, se diseña una estrategia de búsqueda genérica, compuesta por vocabulario controlado explotado (MeSH, DeCS y Emtree) y lenguaje libre, considerando sinónimos, abreviaturas, acrónimos, variaciones ortográficas y plurales. La estrategia se complementa con identificadores de campo, truncadores, operadores de proximidad y operadores booleanos, y se limita empleando filtros validados por tipo de estudio (10): de alta especificidad para RSL y de alta sensibilidad para estudios primarios. Esta estrategia se valida mediante consulta virtual con expertos temáticos (1) y se adapta para las diferentes bases de datos.

La búsqueda de revisiones sistemáticas debe limitarse a literatura publicada en los últimos cinco años. Cuando se trate de búsquedas *de novo* de estudios primarios, estas deben realizarse sin restricción en la fecha de publicación; deben realizarse sin restricción de idioma.

3.2.2. Otros métodos de búsqueda

Además de buscar en las bases de datos mencionadas, es indispensable incluir los siguientes métodos de búsqueda complementarios. Un listado de los estudios seleccionados junto con los criterios de inclusión y exclusión debe enviarse al grupo de expertos temáticos, a los productores y comercializadores de las tecnologías evaluadas y de sus comparadores, indagando sobre la disponibilidad de estudios adicionales publicados o no publicados que cumplan con los criterios de elegibilidad (de inclusión y de no exclusión) específicos para la evaluación (1).

Debe realizarse una búsqueda manual “en bola de nieve” mediante la revisión de las listas de referencias bibliográficas de los estudios seleccionados. Adicionalmente, debe buscarse información en las bases de datos especializadas sugeridas por los expertos temáticos.

En el caso de medicamentos y dispositivos médicos, debe realizarse una revisión de la información sobre vigilancia poscomercialización para documentar, en caso de ser posible, la

proporción de eventos adversos asociados al uso de las tecnologías de interés. Para esto, deben consultarse la página web del INVIMA y, de forma opcional, otras fuentes como Uppsala Monitoring Centre, EMA (European Medicines Agency), ANVISA (Agencia Nacional de Vigilancia Sanitaria - Brasil), MHRA (Medicines and Healthcare Products Regulatory Agency - Reino Unido) y FDA (Food and Drug Administration - Estados Unidos). En caso de identificar alertas provenientes de otro país, el IETS se encargará junto con el INVIMA del estudio del caso correspondiente.

3.2.3. Actualización de la búsqueda de estudios primarios para una revisión sistemática

Las RSL seleccionadas para la síntesis de evidencia deben ser objeto de una actualización de la búsqueda de estudios primarios de la siguiente forma:

- Revisiones sistemáticas con fecha de búsqueda menor o igual a dos años: deben actualizarse con estudios publicados entre la fecha de búsqueda de la revisión y la fecha actual, consultando las bases de datos WHO International Clinical Trials Registry Platform ICTRP portal y CENTRAL para revisiones sistemáticas de ECA o MEDLINE y EMBASE para revisiones sistemáticas de otro tipo de estudio primario.
- Revisiones sistemáticas con fecha de búsqueda mayor a dos años o RSL con métodos de búsqueda no exhaustivos (esto es cuando la revisión no incluya las bases de datos MEDLINE, EMBASE, CENTRAL y al menos otro método de búsqueda complementario): deben actualizarse sin restringir los estudios por su fecha de publicación, incluyendo las fuentes listadas en la sección “Búsqueda en bases de datos electrónicas”, y emplear al menos otro método de búsqueda complementario.

Para las revisiones sistemáticas con ITC o MTC, debe buscarse ECA con comparaciones directas entre las tecnologías de interés desde la fecha de búsqueda de la revisión hasta la fecha actual, consultando las bases de datos WHO International Clinical Trials Registry Platform ICTRP portal y CENTRAL.

Deben emplearse las estrategias de búsqueda originales de la revisión que es objeto de la actualización, siempre y cuando el reporte de las estrategias de búsqueda sea completo; en caso contrario, deben usarse las estrategias de búsqueda diseñadas para la evaluación, limitándolas por el tipo de estudio correspondiente. Las búsquedas deben realizarse sin restricción de idioma.

3.2.4. Gestión documental

Para cada búsqueda en bases de datos electrónicas debe generarse un reporte, garantizando su reproducibilidad y transparencia. Las estrategias de búsqueda y sus resultados deben almacenarse en formato electrónico.

Los resultados de las búsquedas deben incluirse en una base de datos, preferiblemente en un programa de manejo de referencias bibliográficas. Las publicaciones duplicadas deben eliminarse.

3.3. Tamización de referencias y selección de estudios

Las referencias deben ser tamizadas por dos revisores de forma independiente, revisando los títulos y resúmenes frente a los criterios de elegibilidad predefinidos. Previamente debe realizarse una prueba piloto para asegurar la consistencia en la aplicación de los criterios de elegibilidad.

En caso de duda sobre el cumplimiento de los criterios de elegibilidad, debe revisarse el texto completo del estudio para orientar la decisión. Deben describirse las estrategias empleadas para resolver los desacuerdos entre los pares revisores (consenso, consulta con un tercer revisor, etc.).

A partir del grupo de referencias preseleccionadas, se realiza la selección de estudios; para esto, un revisor, con el apoyo de un experto temático, verifica que el estudio cumpla los criterios de elegibilidad (de inclusión y de no exclusión) específicos para las preguntas de evaluación (población, tecnologías de interés, comparadores y desenlaces) mediante la revisión de cada publicación en texto completo.

Los estudios no recuperados en texto completo al momento de la publicación del reporte de evaluación, deben clasificarse como estudios excluidos.

3.4. Evaluación de la calidad de la evidencia

La calidad de los estudios seleccionados debe ser evaluada por un experto metodólogo, de acuerdo con las herramientas y el tipo de estudio que se señalan a continuación:

- RSL: AMSTAR (11).
- Revisiones sistemáticas con ITC o MTC: ISPOR (12).
- ECA: riesgo de sesgo de la Colaboración Cochrane (6).
- Estudios de cohortes: riesgo de sesgo de la Colaboración Cochrane (herramienta en desarrollo).
- Estudios primarios de validez diagnóstica: QUADAS-2 (13).

3.5. Extracción de datos y síntesis de la evidencia

Los hallazgos deben ser extraídos y resumidos a partir de las RSL valoradas con más alta calidad metodológica. Cuando se identifiquen RSL con calidad metodológica similar para una misma comparación y desenlace, los resultados del reporte deben estar basados en las estimaciones de

mayor precisión. Se considera pertinente incluir más de una revisión si la información contenida en cada publicación es complementaria en términos de comparaciones y desenlaces.

En caso de contar únicamente con una RSL de baja calidad que incluya metanálisis para alguna comparación y desenlace de interés, y que cuente con una búsqueda exhaustiva de evidencia, los hallazgos de la evaluación deben estar basados en los efectos reportados por los estudios primarios con mayor ponderación en el metanálisis (14).

Para la extracción de resultados de revisiones sistemáticas de ITC o MTC, no deben considerarse válidas las probabilidades de rankings de tratamientos si el efecto ajustado de la tecnología de interés frente a placebo no es significativo. Para las revisiones sistemáticas de estudios de cohortes, solamente deben incluirse las medidas del efecto ajustadas por factores de confusión.

En caso de no contar con una medida resumen del efecto y sí con estudios primarios no combinables, debe presentarse un rango de los efectos con base en lo reportado por dichos estudios, resaltando el resultado de mayor precisión.

Los resultados deben extraerse tal y como son provistos por los autores en los estudios seleccionados y no deben realizarse modificaciones o cálculos adicionales. Cuando sea necesario, debe contactarse a los autores de los estudios para obtener información sobre hallazgos no reportados o incompletos. En caso de identificar múltiples publicaciones de un mismo estudio, independientemente del diseño de la investigación, debe tomarse en cuenta el reporte con fecha de publicación más reciente.

Debe realizarse un control de calidad confrontando la información incluida en el reporte de evaluación con la presentada en los estudios seleccionados para la síntesis de evidencia.

El reporte de evaluación debe incluir un perfil de evidencia GRADE para efectividad y seguridad o validez diagnóstica cuando:

- La RSL seleccionada para la síntesis de evidencia incluya el perfil de evidencia GRADE con las comparaciones y desenlaces de interés.
- Se generan estimaciones combinadas del efecto, bien sea, metanálisis *de novo*, clásico o indirecto.

Para mayor información sobre la metodología, se sugiere revisar los manuales de GRADE Working Group (3) y Singh y cols (15).

3.6. Análisis estadístico

A continuación se presentan las metodologías para el desarrollo de los metanálisis *de novo*. Deben presentarse las razones por las cuales se decide realizar o no un metanálisis y justificar la selección del modelo (efectos fijos, efectos aleatorios, enfoque frecuentista o bayesiano, etc.).

3.6.1. Metanálisis clásico

Debe contemplarse la opción de un metanálisis en caso de no identificar una medida combinada del efecto en las revisiones seleccionadas y de disponer de al menos dos estudios primarios similares. De forma previa al análisis, debe explorarse la presencia de heterogeneidad clínica, metodológica y estadística, y si existe reporte insuficiente de desenlaces o variación en su definición operativa.

En caso de tomar la decisión de agrupar los estudios en un metanálisis, la estimación promedio del efecto debe obtenerse a través de un metanálisis, ponderando por el tamaño muestral (prueba estadística de Mantel-Haenszel). La significancia estadística del efecto debe evaluarse mediante la prueba estadística Z.

El análisis de la heterogeneidad estadística debe considerar el valor de la prueba de inconsistencia (I^2), el valor p de la prueba de ji-cuadrado y una inspección visual de la consistencia del efecto a través de los estudios con base en un gráfico de bosque (6). En caso de identificar heterogeneidad importante, debe explorarse la fuente de la misma mediante alguna de estas herramientas: estimaciones de modelos de efectos fijos *versus* efectos aleatorios, análisis de subgrupos o meta-regresión.

En todos los casos, el análisis de subgrupos debe realizarse sobre las variables predefinidas en los criterios de inclusión.

Debe realizarse un análisis de sensibilidad para estudiar la influencia del riesgo de sesgo en los resultados del metanálisis. La presencia de sesgo de publicación debe evaluarse mediante un gráfico de embudo y la prueba estadística de Egger (6).

Debe reportarse el nivel de confiabilidad, de significancia estadística y el programa empleado para el análisis estadístico.

Para mayor información sobre los aspectos anteriormente señalados, se sugiere revisar el manual Cochrane de revisiones sistemáticas de intervenciones (6). Para información específica sobre el desarrollo de metanálisis de estudios de validez diagnóstica, se sugiere revisar el Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy (16).

3.6.2. Metanálisis de comparaciones indirectas y de múltiples comparaciones

Debe contemplarse la opción de un metanálisis de ITC o MTC ante la ausencia de evidencia directa (estudios cabeza a cabeza) entre las tecnologías de interés y sus comparadores, la disponibilidad de estudios primarios con baja heterogeneidad y un comparador común entre sí.

Para información específica sobre el desarrollo de este tipo de metanálisis, se sugiere revisar las referencias de Glenny y cols. (17) y Wells y cols. (18).

3.6.3. Adición de nuevos datos a las medidas combinadas del efecto

Cuando se identifiquen estudios primarios adicionales a los incluidos en una revisión seleccionada para la síntesis de evidencia que cumplan los criterios de elegibilidad (de inclusión y de no exclusión) de la revisión y se disponga de un metanálisis clásico (con comparaciones directas entre las tecnologías de interés), debe evaluarse la posibilidad de incluir los datos de estos nuevos estudios primarios en la estimación combinada del efecto, independientemente de la magnitud, dirección y precisión de los nuevos datos. Para esto, debe evaluarse si los nuevos estudios son combinables con el grupo de estudios previamente agrupados, es decir, si cumplen los criterios de elegibilidad del estudio original y no generan importante heterogeneidad (clínica, metodológica o estadística).

4. Resultados

La sección de resultados debe presentarse de acuerdo con la siguiente estructura y consideraciones:

4.1. Búsqueda de evidencia

Los resultados de las búsquedas deben presentarse según aplique en los formatos del Anexo 3 y mediante el diagrama de flujo PRISMA (19) (Anexos 4 o 5, según corresponda).

4.2. Tamización de referencias y selección de estudios

Los resultados de la tamización de referencias y selección de estudios deben presentarse en los Anexos 4 o 5, según corresponda. El listado de los estudios incluidos y excluidos debe presentarse en los Anexos 6 y 7, respectivamente.

En esta sección debe señalarse cuando no se identifiquen estudios que cumplan los criterios de elegibilidad (de inclusión y de no exclusión) o cuando no se excluyan estudios de la evaluación.

4.3. Calidad de la evidencia

La calidad de los estudios incluidos debe presentarse en los Anexos 8-12, según corresponda. En el caso particular del reporte de calidad de los ECA, debe incluirse como anexo una figura resumen del riesgo de sesgo a través de los estudios.

4.4. Síntesis de la evidencia

Debe señalarse el número de estudios en los que se basan los resultados del reporte, incluyendo su calidad y la referencia bibliográfica de cada estudio.

4.5. Descripción de los estudios

Las características de los estudios incluidos en la síntesis de evidencia deben presentarse en los Anexos 13-16, según corresponda. Cuando aplique, debe incluirse en el reporte las figuras de la red de intervenciones o riesgo de sesgo presentadas por los estudios originales, referenciando la fuente bibliográfica correspondiente.

4.6. Efectividad y seguridad / validez y concordancia de pruebas diagnósticas

El reporte de los hallazgos debe contener los siguientes elementos:

- Comparación. En caso de que la evidencia para la tecnología evaluada de cuenta de las dosis, concentraciones o formas farmacéuticas, deben documentarse estas precisiones en el reporte según lo presentando por el estudio original.
- Desenlace. Debe especificarse el método, técnica o herramienta de evaluación y el tiempo de seguimiento.
- Medidas del efecto o asociación (RR, OR, HR, diferencia de medias o diferencia estandarizada de medias), indicadores de validez diagnóstica (sensibilidad, especificidad y valores predictivos junto con su respectivo punto de corte, razones de verosimilitud, OR diagnóstico) o índice de concordancia entre pruebas diagnósticas (Kappa, Kappa ponderado), según corresponda con el propósito de la evaluación y lo reportado por los estudios seleccionados.
- IC 95% o error estándar, según corresponda con el estimador reportado. Si no está disponible el IC y sí el valor p, debe reportarse este último.
- I^2 (sólo aplica para metanálisis).
- Número de estudios (sólo aplica para metanálisis).
- Tamaño de muestra.
- Referencia bibliográfica del estudio.

Las estimaciones del efecto deben reportarse en términos de su significancia estadística.

Debe precisarse si la evidencia presentada es indirecta o si no se identificó evidencia para alguna comparación o desenlace de interés.

Los resultados pueden ser presentados en texto, cuadros o figuras, sin repetirlos; se prefiere incluir gráficos de bosque, rankings de tratamientos o curvas ROC en lugar de texto, referenciando la fuente bibliográfica correspondiente. Los cuadros y figuras deben ser auto-explicativos, incluyendo las respectivas leyendas al pie. Los números decimales deben expresarse mediante el punto decimal.

En el reporte del perfil de evidencia GRADE debe emplearse el Cuadro 2 o 3 para efectividad y seguridad o validez diagnóstica respectivamente, aplicando la metodología propuesta por GRADE Working Group (3).

Cuadro 2. Perfil de evidencia GRADE

Comparación: [XXX] versus [XXX]												
Núm. de estudios/n	Evaluación de la calidad						Resumen de los hallazgos				Calidad de la evidencia	Importancia del desenlace
							Núm. de eventos / Núm. de pacientes (%)		Efecto			
	Diseño	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	*	[Intervención]	[Comparador]	Relativo (%) [IC 95%]	Absoluto [IC 95%]		
Desenlace: [XXX] (método / técnica de evaluación y tiempo de seguimiento)												
Autor año (Ref.) (#/#)		No / Serio / Muy serio	No / Serio / Muy serio	No / Serio / Muy serio	No / Serio / Muy serio	No detectado / Fuertemente sospechado					⊕⊕⊕⊕ ⊕⊕⊕ ⊕⊕ ⊕	Crítico / Importante / No importante
Desenlace: [XXX] (método / técnica de evaluación y tiempo de seguimiento)												
Autor año (Ref.) (#/#)		No / Serio / Muy serio	No / Serio / Muy serio	No / Serio / Muy serio	No / Serio / Muy serio	No detectado / Fuertemente sospechado					⊕⊕⊕⊕ ⊕⊕⊕ ⊕⊕ ⊕	Crítico / Importante / No importante
Calidad global	Alta / Moderada / Baja / Muy baja											

* Sesgo de publicación / Tamaño del efecto / Posibles confusores / Gradiente dosis-respuesta.

- ⊕⊕⊕⊕ Calidad alta: es poco probable que otras investigaciones cambien nuestra confianza en el cálculo aproximado del efecto.
- ⊕⊕⊕ Calidad moderada: es posible que otras investigaciones tengan un impacto importante en nuestra confianza en el cálculo aproximado del efecto y que puedan cambiar el cálculo.
- ⊕⊕ Calidad baja: es muy probable que otras investigaciones tengan un impacto importante en nuestra confianza en el cálculo aproximado del efecto y es probable que cambien el cálculo.
- ⊕ Calidad muy baja: no estamos muy seguros sobre el cálculo aproximado del efecto.

Tomado de (3).

5. Discusión

Debe incluir los siguientes elementos:

- Interpretación de los hallazgos: deben interpretarse los hallazgos más relevantes del estudio (su significado), y las relaciones y generalizaciones que estos indican. Debe considerarse la relevancia clínica de los efectos de las tecnologías, independientemente de su significancia estadística.
- Establecer similitudes y diferencias de los hallazgos y sus interpretaciones respecto a estudios publicados.
- Limitaciones: considerar la validez interna (calidad) de la evidencia, las fortalezas y los potenciales sesgos de la evaluación, las consecuencias de la inclusión de evidencia indirecta, el uso de desenlaces subrogados, etc.
- Validez externa: considerar la generalizabilidad de los hallazgos de la evaluación a otras poblaciones.
- Transferibilidad: considerar que los resultados de la evaluación pueden ser usados por los tomadores de decisiones en un contexto específico.
- Aspectos organizacionales: debe responderse a la pregunta ¿la introducción y uso potencial de una nueva tecnología implicaría cambios organizacionales? Si la respuesta es afirmativa, debe reflexionarse sobre los posibles cambios organizacionales que aplicarían en términos de infraestructura y talento humano.
- Identificar vacíos del conocimiento (puntos no resueltos) y concretar aspectos que deben abordar los futuros trabajos de investigación en el tema tratado.

6. Conclusiones

Las conclusiones deben formularse con base en los hallazgos más relevantes de la evaluación y su significado. Deben ser explícitas respecto a si las tecnologías de interés frente a sus comparadores tienen menor, similar o mayor efectividad, seguridad o rendimiento diagnóstico y si las diferencias son o no estadísticamente significativas. También deben hacer referencia a si no se identifica evidencia o si la evidencia no es concluyente (por ejemplo, cuando sólo hay evidencia de efectividad para el grupo terapéutico y no para la tecnología de interés).

En todos los casos, debe especificarse la comparación, el tipo de comparación (directa o indirecta) y el desenlace.

La calidad de la evidencia debe considerarse de forma explícita en la formulación de las conclusiones, bien sea la calidad de las RSL, del estudio primario o del grupo de estudios primarios con base en los cuales se presentan los resultados del reporte. Cuando este contiene un perfil de evidencia GRADE, debe incluirse la calidad de la evidencia por desenlace, la calidad global para el cuerpo de la evidencia y la importancia de los desenlaces.

Esta sección debe presentarse en el siguiente formato:

- Efectividad / validez y concordancia de pruebas diagnósticas.
- Seguridad.

Una vez finalizado el reporte de evaluación de efectividad y seguridad, este debe publicarse en página web para recepción de comentarios, según lo estipula el Manual de Procesos Participativos del IETS (1). Después de realizar los ajustes a que haya lugar, se procede a la publicación definitiva del reporte.

Finalmente, debe elaborarse el informe del proceso participativo, de acuerdo con el formato establecido en el Manual de Procesos Participativos del IETS (1).

Referencias

- 1.** Instituto de Evaluación Tecnológica en Salud. Manual de Procesos Participativos del IETS. Bogotá DC.,2014.
- 2.** Asociación Médica Mundial. Declaración de Helsinki. Principios éticos para la investigación médica en seres humanos. Adoptada por la 18ª Asamblea Médica Mundial, Helsinki, Finlandia. 1964.
- 3.** Lavis JN, Wilson MG, Oxman AD, Grimshaw J, Lewin S, Fretheim A. SUPPORT Tools for evidence-informed health Policymaking (STP) 5: Using research evidence to frame options to address a problem. Health Research Policy and Systems. 2009;7(Suppl 1):S5.
- 4.** GRADE Working Group. Publications [acceso 5 de junio de 2014] Disponible en:<http://www.gradeworkinggroup.org/publications/>
- 5.** de Helsinki D. Principios éticos para las investigaciones médicas en seres humanos. Adoptada por la. 2004;18:20-3.
- 6.** Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions 5.1.0 [updated March 2011]2011.
- 7.** Lavis JN, Wilson MG, Oxman AD, Grimshaw J, Lewin S, Fretheim A. SUPPORT Tools for evidence-informed health Policymaking (STP) 5: Using research evidence to frame options to address a problem. Health research policy and systems / BioMed Central. 2009;7 Suppl 1:S5. PubMed PMID: 20018112. Pubmed Central PMCID: PMC3271832. Epub 2009/12/19. eng.
- 8.** Fewtrell MS, Kennedy K, Singhal A, Martin RM, Ness A, Hadders-Algra M, et al. How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? Archives of disease in childhood. 2008;93(6):458-61.
- 9.** De La Asociación DDH. Médica Mundial. Principios éticos para las investigaciones médicas en seres humanos. 2000;59.
- 10.** Glasziou PP, Shepperd S, Brassey J. Can we rely on the best trial? A comparison of individual trials and systematic reviews. BMC medical research methodology. 2010;10(1):23.
- 11.** Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. Journal of clinical epidemiology. 2009 Oct;62(10):1013-20. PubMed PMID: 19230606. Epub 2009/02/24. eng.
- 12.** Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2011 Jun;14(4):417-28. PubMed PMID: 21669366. Epub 2011/06/15. eng.
- 13.** Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine. 2011 Oct 18;155(8):529-36. PubMed PMID: 22007046. Epub 2011/10/19. eng.

- 14.** Glasziou PP, Shepperd S, Brassey J. Can we rely on the best trial? A comparison of individual trials and systematic reviews. *BMC medical research methodology*. 2010;10:23. PubMed PMID: 20298582. Pubmed Central PMCID: PMC2851704. Epub 2010/03/20. eng.
- 15.** Singh S, Chang SM, Matchar DB, Bass EB. Chapter 7: grading a body of evidence on diagnostic tests. *Journal of general internal medicine*. 2012 Jun;27 Suppl 1:S47-55. PubMed PMID: 22648675. Pubmed Central PMCID: PMC3364356. Epub 2012/06/08. eng.
- 16.** Pathela P, Zahid Hasan K, Roy E, Huq F, Kasem Siddique A, Bradley Sack R. Diarrheal illness in a cohort of children 0–2 years of age in rural Bangladesh: I. Incidence and risk factors. *Acta Paediatrica*. 2006;95(4):430-7.
- 17.** Ministerio de Salud y Protección Social, Colciencias, Universidad de Antioquia. Guía de práctica clínica para prevención, diagnóstico y tratamiento de la enfermedad diarreica aguda en niños menores de 5 años. Bogotá D.C,2013.
- 18.** Cifuentes E, Suárez L, Solano M, Santos R. Diarrheal diseases in children from a water reclamation site in Mexico city. *Environmental Health Perspectives*. 2002;110(10):A619.
- 19.** Liberati A, Altman D, Tetzlaff J, Mulrow C, Gotzsche P, Ioannidis J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of clinical epidemiology*. 2009 Oct;62(10):e1-34. PubMed PMID: 19631507. Epub 2009/07/28. eng.

Anexos

Anexo 1. Registros sanitarios vigentes para las tecnologías de interés.

Núm.	Registro sanitario	Nombre comercial	Consideraciones farmacéuticas	Indicación	Titular registro

Anexo 2. Clasificación de la importancia de los desenlaces.

Desenlace	Puntuación media del grupo

Anexo 3. Reportes de búsqueda de evidencia en bases de datos electrónicas.

Reporte de búsqueda electrónica Núm. #	
Tipo de búsqueda	Nueva / Actualización
Bases de datos	<ul style="list-style-type: none"> ▪ MEDLINE ▪ MEDLINE In-Process & Other Non-Indexed Citations ▪ MEDLINE Daily Update
Plataforma	Ovid
Fecha de búsqueda	dd/mm/aaaa
Rango de fecha de búsqueda	aaaa-aaaa / Sin restricción
Restricciones de lenguaje	Ninguna
Otros límites	Filtro de tipo de estudio / Ninguno
Estrategia de búsqueda (resultados)	XXX (#)
Referencias identificadas	#
Referencias sin duplicados	#

Reporte de búsqueda electrónica Núm. #	
Tipo de búsqueda	Nueva / Actualización
Base de datos	EMBASE
Plataforma	Elsevier
Fecha de búsqueda	dd/mm/aaaa
Rango de fecha de búsqueda	aaaa-aaaa / Sin restricción
Restricciones de lenguaje	Ninguna
Otros límites	Filtro de tipo de estudio / Ninguno
Estrategia de búsqueda (resultados)	XXX (#)
Referencias identificadas	#
Referencias sin duplicados	#

Reporte de búsqueda electrónica Núm. #	
Tipo de búsqueda	Nueva / Actualización
Bases de datos	<ul style="list-style-type: none"> ▪ Cochrane Database of Systematic Reviews ▪ Database of Abstracts of Reviews of Effects - DARE
Plataforma	Wiley
Fecha de búsqueda	dd/mm/aaaa
Rango de fecha de búsqueda	aaaa-aaaa / Sin restricción
Restricciones de lenguaje	Ninguna
Otros límites	Filtro de tipo de estudio / Ninguno
Estrategia de búsqueda (resultados)	[XXX] (#)
Referencias identificadas	#
Referencias sin duplicados	#

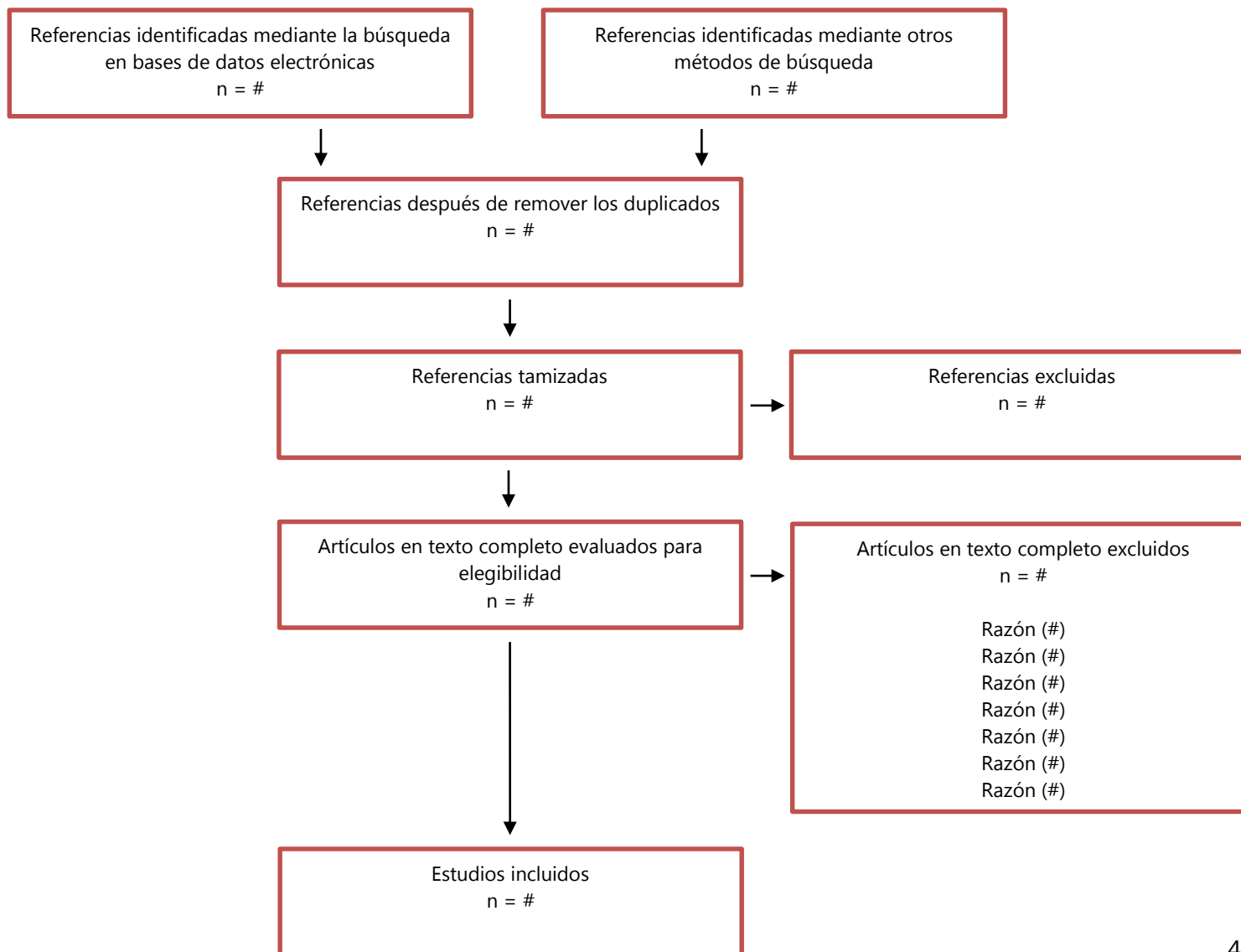
Reporte de búsqueda electrónica Núm. #	
Tipo de búsqueda	Nueva / Actualización
Base de datos	LILACS
Plataforma	Biblioteca Virtual en Salud - BVS, interfaz iAHx
Fecha de búsqueda	dd/mm/aaaa
Rango de fecha de búsqueda	aaaa-aaaa / Sin restricción
Restricciones de lenguaje	Ninguna
Otros límites	Filtro de tipo de estudio / Ninguno
Estrategia de búsqueda (resultados)	[XXX] (#)
Referencias identificadas	#
Referencias sin duplicados	#

Reporte de búsqueda electrónica Núm. #	
Tipo de búsqueda	Nueva / Actualización
Base de datos	Cochrane Central Register of Controlled Trials -CENTRAL
Plataforma	Ovid
Fecha de búsqueda	dd/mm/aaaa
Rango de fecha de búsqueda	aaaa-aaaa / Sin restricción
Restricciones de lenguaje	Ninguna
Otros límites	Filtro de tipo de estudio / Ninguno
Estrategia de búsqueda (resultados)	[XXX] (#)
Referencias identificadas	#
Referencias sin duplicados	#

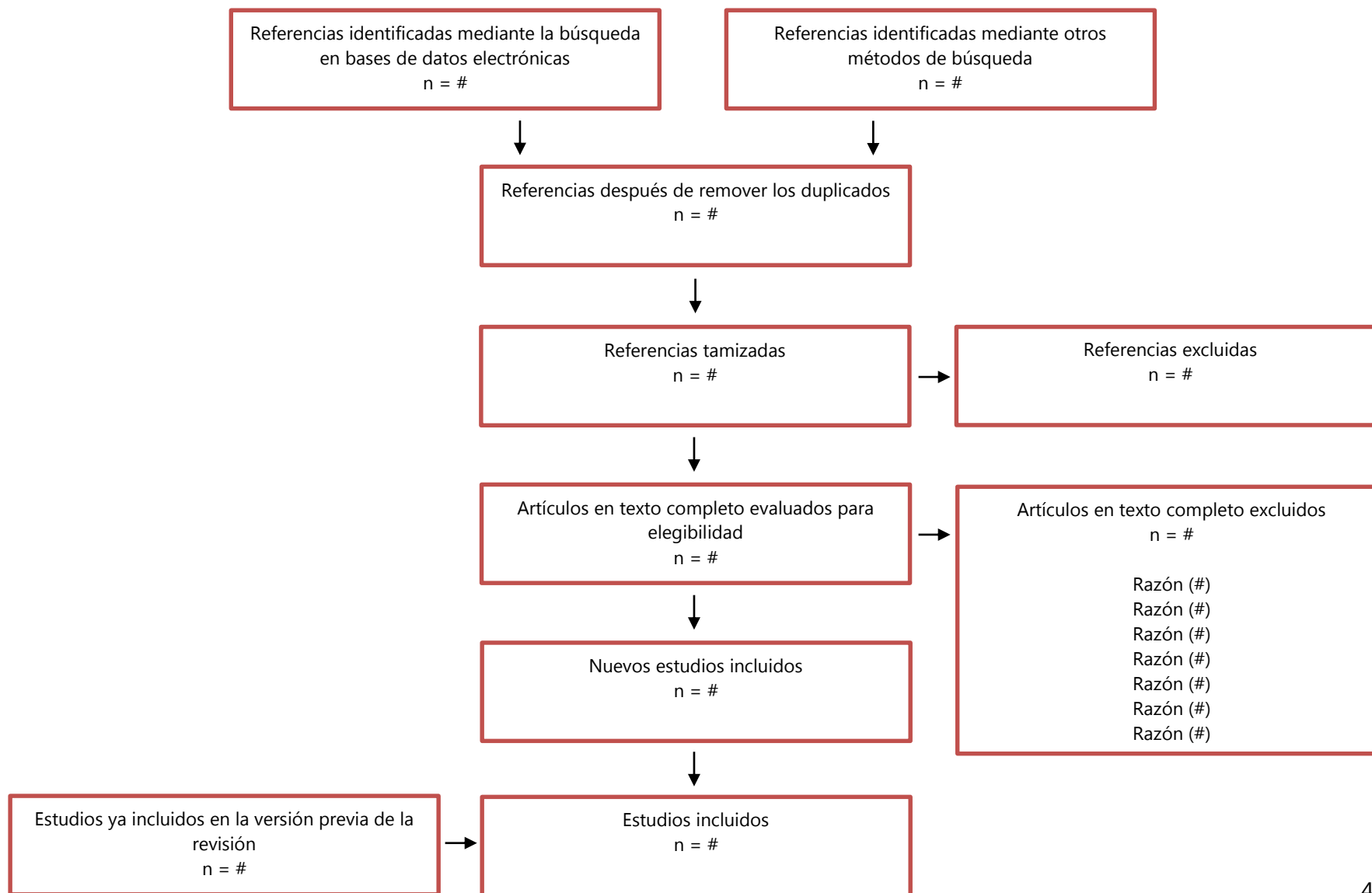
Reporte de búsqueda electrónica Núm. #	
Tipo de búsqueda	Nueva / Actualización
Base de datos	WHO International Clinical Trials Registry Platform ICTRP portal
Plataforma	-
Fecha de búsqueda	dd/mm/aaaa
Rango de fecha de búsqueda	aaaa-aaaa / Sin restricción
Restricciones de lenguaje	Ninguna
Otros límites	Filtro de tipo de estudio / Ninguno
Estrategia de búsqueda (resultados)	[XXX] (#)
Referencias identificadas	#
Referencias sin duplicados	#

Reporte de búsqueda electrónica Núm. #	
Tipo de búsqueda	Nueva / Actualización
Base de datos	ClinicalTrials.gov
Plataforma	-
Fecha de búsqueda	dd/mm/aaaa
Rango de fecha de búsqueda	aaaa-aaaa / Sin restricción
Restricciones de lenguaje	Ninguna
Otros límites	Filtro de tipo de estudio / Ninguno
Estrategia de búsqueda (resultados)	[XXX] (#)
Referencias identificadas	#
Referencias sin duplicados	#

Anexo 4. Diagrama de flujo de la búsqueda, tamización y selección de evidencia (búsqueda *de novo*).



Anexo 5. Diagrama de flujo de la búsqueda, tamización y selección de evidencia (actualización de la búsqueda de estudios primarios para una revisión sistemática).



Anexo 6. Listado de estudios incluidos en la evaluación.

Incluir las referencias en estilo Vancouver (<http://www.icmje.org>).

Anexo 7. Listado de estudios excluidos de la evaluación y razones para su exclusión.

Incluir las referencias en estilo Vancouver (<http://www.icmje.org>) y enseguida, la razón para su exclusión.

Anexo 8. Calidad de las revisiones sistemáticas incluidas en la evaluación (herramienta AMSTAR).

Criterio	Autor año (Ref.)	Autor año (Ref.)
1. ¿Fue provisto un diseño <i>a priori</i> ? Se establecieron antes de conducir la revisión: <ul style="list-style-type: none"> ▪ La pregunta de investigación. ▪ Los criterios de inclusión. 	Si / No / NA	Si / No / NA
2. ¿La selección de los estudios y la extracción de datos fue duplicada?: <ul style="list-style-type: none"> ▪ Existieron por lo menos dos revisores independientes para la extracción de datos. ▪ Existió un procedimiento de consenso para resolver los desacuerdos. 	Si / No / NA	Si / No / NA
3. ¿Se realizó una búsqueda exhaustiva de literatura?: <ul style="list-style-type: none"> ▪ La búsqueda incluyó al menos dos fuentes electrónicas. ▪ El reporte incluyó los años y las bases de datos usadas (por ejemplo, CENTRAL, EMBASE y MEDLINE). ▪ Las palabras clave, los términos MeSH y la estrategia de búsqueda fueron reportados. ▪ Todas las búsquedas fueron complementadas por consulta de revisiones, libros de texto, registros especializados, expertos en el campo de estudio específico y por revisión de las referencias en los estudios identificados. 	Si / No / NA	Si / No / NA
4. ¿El estado de la publicación (por ejemplo, literatura gris) fue utilizado como criterio de inclusión?: <ul style="list-style-type: none"> ▪ Los autores declararon que se buscó literatura independientemente del estado de la publicación. ▪ Se indicó si se excluyó o no algún reporte de la revisión sistemática, con base en su estado de publicación, lenguaje, etc. 	Si / No / NA	Si / No / NA
5. ¿Se presentó un listado de estudios incluidos y excluidos?	Si / No / NA	Si / No / NA
6. ¿Se presentaron las características de los estudios incluidos?: <ul style="list-style-type: none"> ▪ Se presentaron los datos de los estudios originales: participantes, intervenciones y desenlaces. ▪ Se reportaron las características de todos los estudios analizados: edad, raza, sexo, nivel socioeconómico, estado, duración y severidad de la enfermedad o comorbilidades. 	Si / No / NA	Si / No / NA
7. ¿La calidad de los estudios incluidos fue evaluada y documentada?: <ul style="list-style-type: none"> ▪ Se definieron <i>a priori</i> los criterios de evaluación de calidad (por ejemplo, para estudios de efectividad, si los autores decidieron incluir solamente estudios aleatorizados, doble ciego, placebo controlados, o si el ocultamiento de la asignación se empleó como un criterio de inclusión). ▪ Se consideraron ítems relevantes para otro tipo de estudios. 	Si / No / NA	Si / No / NA

Criterio	Autor año (Ref.)	Autor año (Ref.)
8. ¿Se utilizó adecuadamente la calidad de los estudios para formular las conclusiones?: <ul style="list-style-type: none"> ▪ Los resultados del rigor metodológico y la calidad científica se consideraron en el análisis, las conclusiones y explícitamente en la formulación de recomendaciones. 	Si / No / NA	Si / No / NA
9. ¿Fueron apropiados los métodos para combinar los hallazgos de los estudios?: <ul style="list-style-type: none"> ▪ Se realizó una prueba para asegurar que los estudios fueran combinables, para evaluar su homogeneidad (por ejemplo, prueba ji cuadrado, I²). ▪ Si la heterogeneidad existió, se debió usar un modelo de efectos aleatorios o se debió tomar en consideración la pertinencia clínica de combinar los estudios (¿es factible combinarlos?). 	Si / No / NA	Si / No / NA
10. ¿Se evaluó la probabilidad de sesgo de publicación?: <ul style="list-style-type: none"> ▪ La evaluación del sesgo de publicación debió incluir una combinación de ayudas gráficas (por ejemplo, gráfico de embudo u otras pruebas disponibles) o pruebas estadísticas (por ejemplo, prueba de Egger). 	Si / No / NA	Si / No / NA
11. ¿Se declaró la existencia de conflictos de interés?: <ul style="list-style-type: none"> ▪ Las fuentes de financiación o apoyo fueron claramente reconocidas tanto en la revisión, como en los estudios incluidos. 	Si / No / NA	Si / No / NA
Calidad global †	#/# Alta / Media / Baja	#/# Alta / Media / Baja

NA: no aplica.

† Calidad global: baja (puntaje entre 0-3), media (puntaje entre 4-7) y alta (puntaje entre 8-11).

Nota: traducción libre realizada por funcionarios de la Subdirección de Evaluación de Tecnologías en Salud, Instituto de Evaluación Tecnológica en Salud - IETS

Anexo 9. Calidad de los metanálisis de comparaciones indirectas o de múltiples comparaciones incluidos en la evaluación (herramienta ISPOR).

Criterio		Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)
Introducción	¿La justificación y los objetivos del estudio se presentan con claridad?	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA
Métodos	¿La sección de métodos incluye lo siguiente?: <ul style="list-style-type: none"> ▪ Descripción de los criterios de elegibilidad. ▪ Fuentes de información. ▪ Estrategias de búsqueda. ▪ Proceso de selección de los estudios. ▪ Extracción de datos. ▪ Evaluación de la calidad de los estudios incluidos. 	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA
	¿Las medidas de resultado son descritas?	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA
	¿Existe una descripción de los métodos para el análisis y síntesis de la evidencia? ¿Los métodos descritos incluyen lo siguiente?:	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA

	<ul style="list-style-type: none"> ▪ Descripción de los métodos de análisis/modelos. ▪ Manejo de sesgos potenciales /inconsistencia. ▪ Marco de análisis. 						
	¿Los análisis de sensibilidad son presentados?	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA

Criterio		Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)
Resultados	¿Los resultados incluyen un resumen de los estudios incluidos en la red de evidencia? ¿Se incluyen datos de estudios individuales? ¿Se incluye la red de estudios?	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA
	¿El estudio describe una evaluación del ajuste del modelo? ¿Los modelos competidores son comparables?	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA
	¿Los resultados de la síntesis de evidencia (ITC/MTC) son presentados claramente?	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA
	¿Los análisis de sensibilidad/de escenarios son presentados?	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA
Discusión	¿La discusión incluye lo siguiente?: <ul style="list-style-type: none"> ▪ Descripción/resumen de los hallazgos principales ▪ Validez interna del análisis ▪ Validez externa ▪ Implicaciones de los resultados para la población blanco 	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA	Si / No / NA
Calidad global		## Alta / Baja	## Alta / Baja	## Alta / Baja	## Alta / Baja	## Alta / Baja	## Alta / Baja

NA: no aplica.

ITC: metanálisis de comparaciones indirectas.

MTC: metanálisis de comparaciones múltiples.

Nota: traducción libre realizada por funcionarios de la Subdirección de Evaluación de Tecnologías en Salud, Instituto de Evaluación Tecnológica en Salud - IETS.

Anexo 10. Calidad de los ensayos controlados aleatorizados incluidos en la evaluación (riesgo de sesgo de la Colaboración Cochrane).

Dominio	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)
Generación de la secuencia de aleatorización (sesgo de selección).	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Ocultamiento de la asignación (sesgo de selección).	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Cegamiento de los participantes y del personal (sesgo de realización).	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Cegamiento de los evaluadores del resultado (sesgo de detección).	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Datos de resultado incompletos (sesgo de desgaste).	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Reporte selectivo de los resultados (sesgo de notificación).	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Otras fuentes de sesgo †	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Resumen del riesgo de sesgo ††	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro

† [XXX]

†† Resumen del riesgo de sesgo: bajo (bajo riesgo de sesgo para todos los dominios), alto (alto riesgo de sesgo para uno o más dominios) y no claro (riesgo de sesgo no claro para uno o más dominios).

Anexo 11. Calidad de los estudios de cohortes analíticas incluidos en la evaluación (riesgo de sesgo de la Colaboración Cochrane).

Dominio	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)
Sesgo debido a factores de confusión en la línea de base.	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Sesgo en la selección de participantes del estudio.	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Sesgo debido a desviación de las intervenciones de interés.	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Sesgo debido a datos perdidos.	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Sesgo en la toma de mediciones.	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Sesgo en la selección de resultados reportados.	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro

Nota: adaptado de Cochrane Risk of Bias tool - draft versión. Traducción libre realizada por funcionarios de la Subdirección de Evaluación de Tecnologías en Salud, Instituto de Evaluación Tecnológica en Salud - IETS.

Anexo 12. Calidad de los estudios primarios de validez diagnóstica incluidos en la evaluación (herramienta QUADAS-2).

Dominio		Criterio	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)
Selección de pacientes	A. Riesgo de sesgo	¿Se realizó una muestra consecutiva o aleatoria de los pacientes reclutados?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		¿Se evitó un diseño de casos y controles?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		¿El estudio evitó exclusiones inadecuadas?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		¿Podría la selección de los pacientes haber introducido un sesgo?	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
B. Aplicabilidad	¿Existe preocupación que los pacientes incluidos no correspondan a la pregunta de investigación?	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	
Prueba índice	A. Riesgo de sesgo	¿Fueron los resultados de la prueba índice interpretados sin conocimiento de los resultados del estándar de referencia?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		Si se usó un umbral, ¿éste fue especificado previamente?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		¿La conducción o interpretación de la prueba índice podría haber introducido un sesgo?	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
	B. Aplicabilidad	¿Existe preocupación acerca que la prueba índice, su conducción o interpretación no correspondan con la pregunta de investigación?	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
Estándar de referencia	A. Riesgo de sesgo	¿Es probable que el estándar de referencia clasifique correctamente la condición de interés?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		¿Fueron los resultados del estándar de referencia interpretados sin conocimiento de los resultados de la prueba índice?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		¿La conducción o interpretación del estándar de referencia podría haber introducido un sesgo?	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro
	B. Aplicabilidad	¿Existe preocupación acerca que la condición de interés definida por el estándar de referencia no corresponda a la pregunta de investigación?	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro

Dominio		Criterio	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)
Flujo de pacientes y tiempos	A. Riesgo de sesgos	¿Hubo un intervalo de tiempo adecuado entre la prueba índice y el estándar de referencia?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		¿Todos los pacientes recibieron el estándar de referencia?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		¿Los pacientes recibieron el mismo estándar de referencia?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		¿Fueron incluidos todos los pacientes en el análisis?	Si / No / No claro	Si / No / No claro	Si / No / No claro	Si / No / No claro
		¿Podría el flujo de pacientes haber introducido un sesgo?	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro	Bajo / Alto / No claro

Nota: traducción libre realizada por funcionarios de la Subdirección de Evaluación de Tecnologías en Salud, Instituto de Evaluación Tecnológica en Salud - IETS.

Anexo 13. Características de las revisiones sistemáticas incluidas en la síntesis de evidencia.

	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)
Tipo de revisión						
Población						
Subgrupos						
Comparaciones						
Desenlaces						
Número de estudios (diseño) y pacientes incluidos						
Bases de datos consultadas						
Fuentes de literatura gris						
Fecha de búsqueda						
Rango de fecha de búsqueda						
Restricciones de lenguaje						
Otros límites empleados						
Calidad de los estudios primarios						
Fuentes de financiación						
Conclusiones						

Anexo 14. Características de los ensayos controlados aleatorizados incluidos en la síntesis de evidencia.

	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)
Estado de publicación						
Diseño						
Población						
Lugar						
Comparaciones						
Desenlaces						
Hipótesis de investigación y tipo de análisis (por protocolo o intención de tratar)						
Tamaño de muestra (# de sujetos aleatorizados)						
Tiempo de seguimiento						
Pérdidas (%)						
Fuentes de financiación						
Conclusiones						

Anexo 15. Características de los estudios de cohortes incluidos en la síntesis de evidencia.

	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)
Diseño						
Población						
Lugar						
Comparaciones						
Desenlaces						
Tamaño de muestra						
Tiempo de seguimiento						
Pérdidas (%)						
Técnicas para el control de la confusión						
Fuentes de financiación						
Conclusiones						

Anexo 16. Características de los estudios primarios de validez diagnóstica o de concordancia de pruebas diagnósticas incluidos en la síntesis de evidencia.

	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)	Autor año (Ref.)
Diseño						
Población						
Lugar						
Comparaciones						
Interpretación de las pruebas						
Desenlaces						
Tamaño de muestra						
Fuentes de financiación						
Conclusiones						

Glosario

Para dar claridad a los lectores del reporte, se sugiere hacer un listado de aquellos conceptos considerados clave dentro de la evaluación.

Los términos presentados a continuación son tomados, literalmente, del glosario de la Red Internacional de Agencias de Evaluación de Tecnologías (INAHTA, por sus siglas en inglés), desarrollado con el propósito de proveer un lenguaje común para la evaluación de tecnologías en salud (5).

- Base de datos: cualquiera de una amplia variedad de bancos de datos (a menudo informatizados) que recogen observaciones e información sobre un grupo de pacientes (p. ej. varones residentes en Gotemburgo), una enfermedad (p. ej. hipertensión), una intervención (p. ej. farmacoterapia antihipertensiva) u otros eventos o características. En función de los criterios de inclusión de la base de datos, las observaciones pueden tener controles. Aunque pueden resultar útiles, diversos factores de confusión (p. ej. no hay aleatorización y el proceso de registro de pacientes o eventos presenta un sesgo de selección potencial) los convierte en métodos relativamente débiles a la hora de determinar relaciones causales entre una intervención y un resultado.
- Biblioteca Cochrane: conjunto de bases de datos publicado en CD-ROM e Internet y actualizado trimestralmente, diseñado para proporcionar información y evidencia para respaldar la toma de decisiones en la atención sanitaria. Las bases de datos son las siguientes: Cochrane Database of Systematic Reviews, Cochrane Database of Methodology Reviews, Cochrane Central Register of Controlled Trials, DARE (Database of Abstracts of Reviews of Effects), Cochrane Methodology Register, Health Technology Assessment y NHS Economic Evaluation.
- Búsqueda manual: búsqueda planeada de una publicación página por página (es decir, manual), e incluye editoriales, cartas, etc., para identificar todos los estudios pertinentes. Normalmente, una persona empezará manualmente por el año en curso de una publicación y buscará hacia atrás, hasta que la búsqueda casi no dé resultados de ensayos o se haya llegado al volumen 1.
- Calidad de la evidencia: el grado en el que el diseño y el desarrollo de un estudio han evitado los errores sistemáticos (sesgos).
- Calidad metodológica: grado en que el diseño y la realización de un estudio probablemente hayan prevenido errores sistemáticos (sesgo). La variación de calidad puede explicar la variación en los resultados de estudios incluidos en una revisión sistemática. Los ensayos (de mejor 'calidad') diseñados con rigor presentan una mayor probabilidad de obtener resultados más próximos a la 'verdad'. (Véase también validez externa y validez).

- Cegamiento: desconocimiento por parte de pacientes y/o investigadores de si los sujetos están recibiendo las intervenciones en investigación o las de control (o estándar) en un ensayo clínico. El objetivo del cegamiento es eliminar la posibilidad de que el conocimiento de la intervención recibida influya los desenlaces sobre el paciente o sobre los comportamientos del investigador, y afecte de ese modo los resultados. El cegamiento no siempre resulta práctico (p. ej. cuando se compara la cirugía con un tratamiento farmacológico), pero debe utilizarse siempre que sea posible y compatible con un cuidado óptimo del paciente. (Véase también ocultamiento de la asignación, simple ciego, doble ciego y triple ciego). Sinónimo: enmascaramiento.
- CINAHL (Cumulative Index to Nursing and Allied Health Literature): base de datos electrónica que contiene bibliografía sobre enfermería y ciencias afines. Años de cobertura: 1981 hasta la actualidad.
- Cochrane Central Register of Controlled Trials (CENTRAL): base de datos electrónica que contiene bibliografía sobre enfermería y ciencias afines. Años de cobertura: 1981 hasta la actualidad.
- Coeficiente de correlación: medida numérica entre -1 y 1 que expresa la fuerza de una asociación lineal observada entre dos variables. Expresada como r , el valor $r = 0$ indica, o bien la ausencia de una relación lineal o una posible relación no lineal entre las dos variables.
- Coeficiente de probabilidad (Likelihood ratio): 1. Cociente entre la probabilidad de obtener un resultado positivo (o negativo) en aquellas personas que tienen una determinada enfermedad y la citada probabilidad en personas sin la mencionada enfermedad. La razón de verosimilitud para un test con resultado positivo es sensibilidad/ (1 menos especificidad). La razón de verosimilitud de un test con resultado negativo es (1 menos sensibilidad)/especificidad. 2. Índice estadístico que compara la bondad del ajuste de dos modelos relacionados, lo que permite contrastar hipótesis en un gran número de situaciones.
- Comorbilidad: presencia de enfermedades coexistentes o adicionales a la que se está estudiando.
- Contraindicación: síntoma clínico o circunstancia que indica que el uso de una intervención normalmente aconsejable resultaría inapropiado.
- Cribado: servicio de salud pública en el cual se formula una pregunta u ofrece una prueba a miembros de una población definida, que no necesariamente perciben que están en peligro de sufrir una enfermedad o sus complicaciones, para identificar a aquellas personas que presentan más probabilidades de beneficiarse que de perjudicarse por la realización de otras pruebas o tratamientos. Sinónimos: tamización.
- Curva (ROC) de características operativas del receptor: gráfico que representa la relación entre la tasa de verdaderos positivos (sensibilidad) y la tasa de falsos positivos ("1 -

especificidad"), como una función del nivel de corte del marcador de una enfermedad (o condición clínica). Las curvas ROC ayudan a demostrar cómo el incremento o disminución del punto de corte que define una prueba como positiva afecta al equilibrio entre la identificación correcta de los enfermos (verdaderos positivos) y la denominación incorrecta como positivos de los sanos (falsos positivos). El área bajo la curva ROC es una expresión del potencial diagnóstico de un marcador independiente de la población de pacientes y puede utilizarse para comparar uno o más marcadores.

- Curva SROC: gráfico bidimensional que en un metanálisis de pruebas diagnósticas, representa en un eje la sensibilidad, y en el otro (1-especificidad) para cada estudio individual y la curva que mejor se ajusta a ese conjunto de puntos.
- DARE (Database of Abstracts of Reviews of Effects): base de datos de revisiones sistemáticas cuya calidad ha sido evaluada sobre los efectos de las intervenciones de salud. (Véase también biblioteca Cochrane).
- Declaración de conflicto de intereses: manifestación explícita realizada por las personas que participan en un informe o una revisión de la evidencia, sobre los intereses personales, económicos o de otra índole que podrían haber influido en los resultados del informe o en su interpretación.
- Desarrollo de consenso: diversas formas de opinión de grupos en las que un grupo (o panel) de expertos interactúa para la evaluación de una intervención y la formulación de propuestas por votación u otro proceso, para llegar a un acuerdo general. Este proceso puede ser informal o formal y conlleva técnicas como la de grupo nominal y la Delphi. Sinónimo: métodos de consenso.
- Doble ciego: ni los participantes en un ensayo ni los investigadores (evaluadores de los resultados) conocen la intervención que cada participante recibe. El objeto de cegar a los investigadores (evaluadores de los resultados, que también podrían ser los proveedores del cuidado) es proteger contra el sesgo de detección. (Véase también cegamiento, simple ciego, y triple ciego).
- Efectividad: beneficio (p. ej. en resultados en salud) que supone utilizar una tecnología para un determinado problema en condiciones generales o habituales; por ejemplo, para el médico de un hospital comunitario o para un paciente en su casa. (Véase también efectividad clínica).
- Efectividad clínica: grado en que una determinada intervención, procedimiento, régimen o servicio tiene el efecto previsto en circunstancias normales en comparación con condiciones controladas. Específicamente, es una evaluación de la razón riesgo/beneficio de una intervención en un entorno clínico estándar, utilizando resultados que midan cuestiones

importantes para los pacientes (p. ej. capacidad para realizar las actividades diarias, prolongación de la vida, etc.).

- Efecto adverso: efecto no deseado o no intencionado de una intervención. (Véase también evento adverso y efecto secundario).
- Efecto del tratamiento: efecto de un tratamiento (intervención) sobre los resultados, es decir, sólo atribuible al efecto de la intervención. Los investigadores tratan de estimar el verdadero efecto del tratamiento utilizando la diferencia entre los resultados observados en el grupo de tratamiento y el grupo control. (Véase también tamaño del efecto).
- Efecto placebo: efecto sobre los resultados (mejora o empeoramiento) de un paciente que puede producirse debido a la expectativa del mismo (o proveedor) de que una determinada intervención tenga un efecto. El efecto placebo (también conocido como Hawthorne) es independiente del efecto verdadero (farmacológico, quirúrgico, etc.) de una intervención dada. Para controlar este factor, puede administrarse un placebo al grupo de control de un ensayo.
- EMBASE: base de datos biomédica y farmacológica (Excerpta Medical). Contiene un gran número de publicaciones europeas. Años de cobertura: 1947 hasta la actualidad.
- Emtree: tesoro de vocabulario controlado de la base de datos EMBASE.
- Enmascaramiento: (véase cegamiento).
- Ensayo controlado aleatorio: experimento de dos o más intervenciones en el que se asigna a personas elegibles una intervención mediante aleatorización. El uso de la aleatorización permite utilizar de forma válida una variedad de métodos estadísticos para comparar los resultados de las intervenciones.
- Ensayo de equivalencia: ensayo cuyo objetivo principal es reflejar que la respuesta a dos o más intervenciones difiere en una cantidad que no es clínicamente importante. Por lo general, se demuestra reflejando que la verdadera diferencia de tratamiento probablemente se encuentre entre un margen de equivalencia inferior y otro superior de diferencias clínicamente aceptables.
- Estimación del efecto: en estudios de los efectos de la atención sanitaria, es la relación observada entre una intervención y un resultado, expresada, por ejemplo, como el número necesario a tratar, el *Odds ratio*, la diferencia de riesgos, el riesgo relativo, la diferencia de media estandarizada o la diferencia de media ponderada. (Véase también efecto del tratamiento).
- Estimación puntual: resultados (p. ej. media, diferencia de media ponderada, *Odds ratio*, riesgo relativo, diferencia de riesgos) obtenidos de una muestra (estudio o metanálisis) que

se utilizan como la mejor estimación de lo que ocurre en la realidad en la población de la cual se ha obtenido la muestra. Un intervalo de confianza es una medición de la incertidumbre (debida al azar) asociada a esa estimación. (Véase también intervalo de confianza).

- Estrategia de búsqueda: combinación de fuentes, términos y límites utilizados en la búsqueda bibliográfica para identificar información para la revisión sistemática o la evaluación de tecnología sanitaria.
- Estudios de fase I, II, III, IV: fase de los ensayos clínicos de nuevas tecnologías (por lo general fármacos) en el proceso de desarrollo y aprobación exigido por los organismos fiscalizadores. Los ensayos de fase I normalmente requieren la participación de aproximadamente entre 20 y 80 voluntarios sanos para determinar la seguridad de un fármaco, el rango de dosificación seguro, la absorción, la actividad metabólica, la excreción y la duración de la actividad. Los ensayos de fase II son ensayos controlados en los que participan aproximadamente entre 100 y 300 pacientes voluntarios (con la enfermedad) para determinar la eficacia y las reacciones adversas del fármaco (a veces se dividen en ensayos piloto de fase IIa y ensayos controlados de fase IIb). Los ensayos de fase III son ensayos controlados más grandes que determinan la eficacia y observan los eventos adversos durante el uso prolongado (a veces se dividen en ensayos de fase IIIa realizados antes de la presentación al organismo regulador y ensayos de fase IIIb realizados después de la presentación a este, pero antes de su aprobación). Los ensayos de fase IV son estudios posteriores a la comercialización cuyo fin es observar los efectos a largo plazo y aportar más información sobre seguridad y eficacia, incluido en los grupos de pacientes con regímenes diferentes.
- Evaluación de calidad: función de medición y observación del aseguramiento de la calidad que sirve para determinar en qué medida se ofrece una atención sanitaria adecuada, según los estándares aplicables o las directrices de cuidado aceptables. (Véase también garantía de la calidad y calidad de la atención).
- Evaluación de Tecnologías Sanitarias (ETS): proceso sistemático de valoración de las propiedades, los efectos y/o los impactos de la tecnología sanitaria. Puede abordar tanto las consecuencias directas y deseadas de las tecnologías como las indirectas y no deseadas. Su objetivo principal es informar la toma de decisiones en atención sanitaria. La evaluación de tecnologías sanitarias es realizada por grupos multidisciplinares utilizando marcos analíticos explícitos basados en diversos métodos.
- Evento adverso: cualquier cambio nocivo, patológico o no intencionado en las funciones anatómicas, físicas o metabólicas evidente por signos físicos, síntomas y/o cambios microbiológicos producidos durante cualquier fase de un estudio clínico, estén o no relacionados con el tratamiento. Incluye la exacerbación de dolencias o eventos preexistentes, enfermedades intercurrentes, accidentes, la interacción con otros fármacos o el empeoramiento significativo de la enfermedad.

- Exactitud diagnóstica: grado en que una medición (p. ej., la estimación de la media del efecto de un tratamiento) es verdadera o correcta. Una estimación puede ser exacta pero no precisa si se basa en un método no sesgado que proporcione observaciones con una gran variación (es decir, de magnitudes muy diferentes). (Véase también precisión).
- Gráfico de embudo: gráfico bidimensional utilizado en los metanálisis, en la que se representa en un eje el tamaño de la muestra de cada estudio, y en el otro, el tamaño del efecto. Puede utilizarse para investigar el sesgo de la publicación. Sinónimo: Gráfica de Chimenea.
- Heterogeneidad: en el metanálisis, la heterogeneidad se refiere a la variabilidad o a las diferencias en las estimaciones de los efectos entre los estudios. En ocasiones, se distingue entre la "heterogeneidad estadística" (diferencias en los efectos declarados), la "heterogeneidad metodológica" (diferencias en el diseño de estudios) y la "heterogeneidad clínica" (diferencias entre estudios en las características fundamentales de participantes, intervenciones o mediciones de resultados). Las pruebas estadísticas de heterogeneidad se utilizan para evaluar si la variabilidad observada en los resultados de los estudios tamaño del efecto, es mayor a la que cabría esperar por azar. Sin embargo, estas pruebas tienen escasa potencia estadística. (Véase también homogeneidad).
- Indicación: síntoma clínico, factor de riesgo o circunstancia en la cual es apropiada la realización de una intervención sanitaria, de acuerdo a lo determinado o especificado, por ejemplo, en una guía de práctica clínica, en el estándar asistencial, por un órgano regulador o por otra fuente autorizada.
- Intervalo de confianza: representa el rango de incertidumbre sobre una estimación del efecto de un tratamiento. Se calcula a partir de las diferencias observadas en los resultados de los grupos de tratamiento y control, y del tamaño de la muestra de un estudio. El intervalo de confianza (IC) es el rango de valores por encima y por debajo de la estimación puntual que, con determinada probabilidad, incluye el verdadero valor del efecto del tratamiento. El uso de los IC presupone que un estudio proporciona una entre las numerosas muestras posibles que se podrían extraer si el estudio se repitiera muchas veces. Los investigadores normalmente utilizan IC del 90%, 95% o 99%. Por ejemplo, un IC del 95% indica que existe una probabilidad del 95% de que el IC calculado a partir de un estudio dado incluya el verdadero valor del efecto de un tratamiento. Si el intervalo incluye un efecto de tratamiento nulo (por lo general es 0,0, pero es 1,0 si el efecto del tratamiento se calcula como *Odds ratio* o riesgo relativo, no se puede rechazar la hipótesis nula de que el tratamiento no tiene efecto.
- Literatura gris: informes de investigación y otras publicaciones en formatos impreso y electrónico que no se encuentran en las publicaciones habituales revisadas por expertos o controladas por editoriales comerciales. Ejemplos: las monografías de agencias gubernamentales, las actas de simposios e informes del sector.
- Magnitud del efecto del tratamiento: se refiere al tamaño (o a la distancia desde el valor nulo que indica que el tratamiento no tiene efecto) de la medida resumen (o estimación puntual)

del efecto del tratamiento y los valores incluidos en el intervalo de confianza del 95% correspondiente.

- MEDLINE: base de datos electrónica generada por la Biblioteca Nacional de Medicina de los Estados Unidos (United States National Library of Medicine). Años de cobertura: 1948 hasta la actualidad. (Véase también PubMed).
- MeSH: el Medical Subject Headings (Descriptor de Ciencias de la Salud) es un vocabulario controlado de términos biomédicos que identifican el contenido de cada artículo en la base de datos MEDLINE.
- Niveles de calidad de la evidencia: los estudios frecuentemente se agrupan en niveles según su validez o el grado en que están libres de sesgos. La jerarquía indica cuáles de los estudios deberían considerarse como los más importantes en una evaluación. En ETS se utilizan diversas jerarquías.
- Número Necesario a Tratar: es una medida del efecto de un tratamiento que proporciona el número de personas que es preciso tratar para prevenir un resultado adverso adicional. Es el valor inverso de la reducción de riesgo absoluto ($1 \div \text{reducción de riesgo absoluto}$); es decir, $1,0 \div (P_c - P_t)$. Por ejemplo, si los resultados de un ensayo indican que la probabilidad de muerte en el grupo control fue del 25% y la probabilidad de muerte en el grupo de tratamiento es del 10%, el número necesario a tratar sería $1,0 \div (0,25 - 0,10) = 6,7$ pacientes. (Véase también reducción absoluta del riesgo, reducción relativa del riesgo, y *Odds ratio*).
- Número necesario de pacientes a tratar para producir un efecto perjudicial (NNH): con respecto a los efectos adversos, el número necesario para perjudicar (mejor dicho, el número necesario a tratar para producir un efecto perjudicial) es la cantidad de pacientes que hay que tratar para provocar un efecto adverso adicional y se calcula como el valor inverso de la diferencia del riesgo absoluto. El término se creó como una variante número necesario a tratar.
- *Odds Ratio* (OR): es una medida del efecto del tratamiento que compara la probabilidad de padecer un evento en el grupo de tratamiento con la probabilidad de padecerlo en el grupo control, es decir, $P_t \div (1 - P_t) P_c \div (1 - P_c)$. Por ejemplo, si los resultados de un ensayo indican que la probabilidad de muerte en el grupo control es del 25% y la probabilidad de muerte en el grupo de tratamiento del 10%, la *Odds ratio* sería $0,10 \div (1,0 - 0,10) \div (0,25 \div (1,0 - 0,25)) = 0,33$. (Véase también reducción absoluta del riesgo, número necesario a tratar y riesgo relativo).
- Patrón oro: método, procedimiento o medición generalmente aceptado como el mejor de su clase, con respecto al cual deben compararse las intervenciones nuevas. Es especialmente importante en los estudios de exactitud (o validez) de las pruebas diagnósticas. Sinónimo: estándar de referencia.

- Prueba de Mantel-Haenszel: prueba Ji-cuadrado para datos estratificados que contrasta la asociación controlando la posible confusión. En los metanálisis, la prueba Mantel-Haenszel se utiliza para analizar los datos estratificados (agrupados) por estudio.
- Prueba estadística: procedimiento o función matemática que se utiliza, entre otras cosas, para determinar si las diferencias en los resultados entre el grupo de tratamiento y el grupo control son lo suficientemente grandes como para concluir que son estadísticamente significativas. Este tipo de prueba estadística genera un valor asociado a un valor P concreto. Entre las pruebas estadísticas más comunes se encuentran las siguientes: Pruebas F, t, Z y Ji-cuadrado. La elección de la prueba depende de las condiciones de un estudio, por ejemplo el tipo de variable de resultado utilizada, si los pacientes fueron seleccionados aleatoriamente de una población más grande, y si puede asumirse que los valores de los resultados de la población tienen una distribución normal u otro tipo de distribución.
- Pubmed: sistema de búsqueda de la Biblioteca Nacional de Medicina de Estados Unidos que contiene de citas de artículos biomédicos que datan desde 1947. Estas citas son de MEDLINE y otras publicaciones de ciencias de la vida. Pubmed contiene enlaces a muchos sitios web que ofrecen artículos en texto completo y otros recursos relacionados. Además de MEDLINE, Pubmed incluye otras bases de datos, como por ejemplo OLDMEDLINE, registros facilitados por la editorial y registros en curso, y se puede acceder a ella de forma gratuita en Internet. (Véase también MEDLINE).
- Revisión de la literatura: resumen e interpretación de los hallazgos de investigación publicados en la literatura. Puede incluir revisiones cualitativas no estructuradas por autores individuales, así como diversos procedimientos sistemáticos y cuantitativos como el metanálisis. Sinónimo: descripción general.
- Revisión narrativa: descripción general de los estudios primarios que no se han identificado o analizado de forma sistemática (estandarizada y objetiva).
- Revisión por pares: proceso por el cual los manuscritos enviados a revistas de salud, biomédicas y otras publicaciones de tipo científico se someten a una revisión por otros expertos con experiencia en el área (normalmente anónimos para el autor) a fin de determinar si ofrecen la calidad necesaria para ser publicados.
- Revisión sistemática: forma de revisión estructurada de la literatura que aborda una cuestión formulada para ser respondida mediante un análisis de la evidencia, y que requiere medios objetivos de búsqueda en la literatura, la aplicación de criterios de inclusión y exclusión predeterminados a esta literatura, la evaluación crítica de la literatura pertinente, y la extracción y síntesis de los datos extraídos basados en la evidencia para formular hallazgos. Para analizar y resumir los resultados de los estudios incluidos, pueden utilizarse opativamente métodos estadísticos (metanálisis). (Véase también revisión Cochrane).
- Revisor: persona responsable de preparar y, en el caso de las revisiones Cochrane, de mantener actualizada una revisión sistemática. El término "revisor" también se utiliza para

referirse a la persona que participa en la revisión por pares o a la que dirime las discrepancias entre evaluadores.

- Riesgo: es la razón de personas que presentan el evento en un grupo respecto al total del grupo.
- Riesgo Relativo: razón de riesgo (estadístico) en el grupo de intervención con respecto al riesgo en el grupo de control. Un riesgo relativo de 1 indica que no hay diferencia entre los grupos comparados. Para efectos no deseados, un RR inferior a 1uno indica que la intervención resultó eficaz para reducir el riesgo de ese resultado.
- Seguridad: juicio sobre la aceptabilidad del riesgo (medida de la probabilidad de un resultado adverso y su gravedad) asociado al uso de una tecnología en una situación concreta, p.ej. en el caso de un paciente con un problema de salud determinado, atendido por un clínico con una determinada experiencia, o en un lugar de tratamiento específico.
- Sesgo: en general, cualquier factor que distorsione la verdadera naturaleza de un evento u observación. En las investigaciones clínicas, un sesgo es cualquier factor sistemático que no sea la intervención de interés que afecta la magnitud (es decir, tiende a aumentar o disminuir) de una diferencia observada en los resultados de un grupo de tratamiento y un grupo de control. El sesgo disminuye la exactitud (aunque no necesariamente la precisión) de una observación. La aleatorización es una técnica utilizada para disminuir esta forma de sesgo. El sesgo también se refiere a un prejuicio o punto de vista parcial que podría afectar la interpretación de un problema. Para disminuir este tipo de sesgo se utiliza una técnica denominada doble ciego.
- Significación estadística: a partir de los resultados de un test estadístico se valora que una intervención tiene un efecto estadísticamente significativo cuando las diferencias observadas entre el grupo de tratamiento y el grupo control son lo suficientemente grandes como para que no sea probable que esas diferencias se hayan debido al azar. En ese caso se rechaza la hipótesis nula. El valor de significación estadística indica la probabilidad de que las diferencias observadas se hayan producido por azar, no aportando información sobre la magnitud del efecto. Por ejemplo, un valor p de 0,05 para una diferencia de riesgo del 10% indicaría que la probabilidad de que esa diferencia de riesgo se haya producido por azar es menos que una entre veinte, resultando estadísticamente significativa. Aunque los valores de corte son siempre arbitrarios, es habitual establecer un nivel p de 0,05; 0,01 en metanálisis, ó 0,10 para la evaluación de interacciones. (Véase valor P).
- Simple ciego: el investigador conoce el tratamiento o la intervención administrada al participante, pero éste no. (Véase también cegamiento, doble ciego y triple ciego). Sinónimo: enmascaramiento simple.
- Solidez de la evidencia: magnitud, precisión y reproducibilidad del efecto de la intervención (incluye la magnitud del tamaño del efecto, la amplitud del intervalo de confianza, el valor p

y la exclusión de efectos sin importancia clínica). En el caso de estudios no aleatorios, pueden tenerse en cuenta factores adicionales como la verosimilitud biológica, el gradiente biológico y la temporalidad de las asociaciones.

- Tecnología sanitaria emergente: tecnología que aún no ha sido adoptada por el sistema de salud (los fármacos suelen encontrarse en ensayos clínicos fase II o fase III, tal vez en pre-lanzamiento; los dispositivos están a 6 meses o menos de introducirse al mercado, ya se comercializan pero tienen una difusión menor al 10% o se comercializan en unos cuantos centros), o tecnología existente a la que se le han hecho cambios de indicación o uso.
- Validez del estudio: grado en que se garantizan las inferencias extraídas del estudio cuando se tienen en cuenta los métodos del estudio, la representatividad de la muestra del mismo y la naturaleza de la población de la cual se extrae validez interna, validez externa, aplicabilidad y generalizabilidad.
- Validez externa: el grado en que los hallazgos obtenidos de una investigación realizada en determinadas circunstancias son generalizables a otras poblaciones distintas. En la medida en que las circunstancias de una determinada investigación (p. ej. características de pacientes o la manera de administrar el tratamiento) difieran de las circunstancias de interés, la validez externa de los resultados de esa investigación puede ser cuestionada. (Véase también aplicabilidad y generalizabilidad).
- Validez interna: grado en que los resultados de un estudio representan con exactitud la relación causal entre una intervención y un resultado en las circunstancias particulares de ese estudio. La validez interna de un ensayo puede ser dudosa cuando determinados tipos de sesgo en el diseño o la realización de un ensayo podrían haber afectado los resultados, ocultando así la verdadera dirección, magnitud o certeza del efecto del tratamiento.
- Vocabulario controlado: sistema de términos en el que intervienen, por ejemplo, definiciones, estructura jerárquica y referencias cruzadas. Su objetivo es eliminar la ambigüedad del lenguaje natural al indexar los documentos. Ejemplos de vocabularios controlados son el MeSH, utilizado en MEDLINE, y Emtree, utilizado en EMBASE.



Instituto de Evaluación
Tecnológica en Salud

-
-  Autopista Norte #118-30, oficina 201
Bogotá D.C.
 -  contacto@iets.org.co
 -  www.iets.org.co
 -  [ietscolombia](#)
 -  ietscolombia.blogspot.com
 -  [@ietscolombia](#)
-