Sapiens in Health Science

Sapiens in Health Sciences. 2026;4.1. 006
ISNN: 3072-7960
DOI: https://doi.org/10.71068/hs000106

Artículo: Revisión

Concordance, efficiency and regional application of epidemiological methodologies versus randomized trials: a systematic review and meta-epidemiological meta-analysis

Concordancia, eficiencia y aplicación regional de metodologías epidemiológicas frente al ensayo aleatorizado: revisión sistemática y meta-análisis meta-epidemiológico

[1]Marco Antonio Ramos Espinal.

[1]Universidad Nacional Autónoma de Honduras (UNAH), Ciudad Universitaria, Tegucigalpa, Honduras

***Autor de correspondencia:** Marco Antonio Ramos Espinal

Enviado: 04-08-2026 **Revisado:** 23-08-2026 **Aceptado:** 04-09-2026 **Publicado:** 11-10-2026

Como citar:

Ramos Espinal, M. A. (2026). Concordancia, eficiencia y aplicación regional de metodologías epidemiológicas frente al ensayo aleatorizado: revisión sistemática y meta-análisis meta-epidemiológico. *Sapiens in Health Sciences*, *4*(1), 1-33. https://doi.org/10.71068/hs000106

ABSTRACT

Introduction: the choice of study design conditions the validity, timeliness and usefulness of epidemiological evidence. We aimed to estimate the concordance of non-randomized and emerging designs with the randomized controlled trial (RCT) and to describe their efficiency and regional application.
Method: systematic review and meta-epidemiological meta-analysis following PRISMA 2020. We searched MEDLINE, Embase, Scopus, Web of Science and Cochrane (January 2020 to May 2026). The primary comparison was restricted to five non-randomized versus randomized-trial estimates; a six-estimate mixed-design pool was retained as exploratory. Random-effects analyses included REML, Hartung-Knapp-Sidik-Jonkman and a joint exclusion of potentially overlapping sources. Risk of bias, adapted GRADE certainty and ordinal regional and efficiency matrices were assessed.
Results: Twenty-four studies were included. Five non-randomized versus randomized-trial comparisons yielded a RoR of 1.00 (95% CI 0.92–1.10); the exploratory six-estimate pool yielded 0.98 (0.89–1.06; $I^2 = 77.9\%$). Joint exclusion of Toews and Kimachi yielded approximately 0.95 (0.88–1.02) among the three remaining non-randomized comparisons, reconstructed from published rounded estimates. Subgroup RoRs were 1.05 for observational designs and 0.91 for trial emulation/real-world evidence. Certainty was low and very low, respectively. Of 28 highest regional ratings, 16 relied on indirect evidence.
Conclusions: A pooled RoR close to 1 does not establish equivalence with randomized trials. Between-design comparisons and regional scoring require cautious, family-specific interpretation; the proposed combination of methods remains a hypothesis for future evaluation.

Keywords: epidemiology; research methodology; meta-epidemiology; target trial emulation; real-world evidence; ratio of ratios; health equity

RESUMEN

Introducción: la elección del diseño condiciona la validez, la oportunidad y la utilidad de la evidencia epidemiológica. Estimamos la concordancia de los diseños no aleatorizados y emergentes con el ensayo controlado aleatorizado (ECA) y describimos su eficiencia y aplicación regional.
Método: revisión sistemática y meta-análisis meta-epidemiológico según PRISMA 2020. Consultamos MEDLINE, Embase, Scopus, Web of Science y Cochrane (enero de 2020 a mayo de 2026). Se priorizó la

©2026 por los Autores. Este artículo es de acceso abierto y distribuido según los términos y condiciones bajo una licencia internacional Creative Commons Atribución 4.0. (CC BY 4.0) https://creativecommons.org/licenses/by/4.0/

comparación de cinco estimadores no aleatorizados frente a ECA; el agrupamiento heterogéneo de seis se mantuvo como exploratorio. Se aplicaron efectos aleatorios, REML, Hartung-Knapp-Sidik-Jonkman y exclusión conjunta de posibles fuentes solapadas. Se evaluaron riesgo de sesgo, certeza GRADE adaptada y matrices ordinales de distribución regional y eficiencia.
Resultados: se incluyeron 24 estudios. Los cinco contrastes entre diseños no aleatorizados y ECA produjeron una RoR de 1.00 (IC 95 % 0.92–1.10); el agrupamiento exploratorio de seis estimadores fue 0.98 (0.89–1.06; I^2 = 77.9 %). Al excluir conjuntamente Toews y Kimachi, los tres estimadores no aleatorizados restantes dieron aproximadamente 0.95 (0.88–1.02), reconstruido desde valores publicados redondeados. Las RoR por familias fueron 1.05 para observacionales y 0.91 para emulación/evidencia de mundo real, con certeza baja y muy baja, respectivamente. Dieciséis de 28 puntuaciones regionales máximas se basaron en evidencia indirecta.
Conclusiones: una RoR próxima a 1 no demuestra equivalencia con los ECA. Los contrastes entre diseños y las puntuaciones regionales requieren interpretación prudente y por familias; la combinación de métodos sigue siendo una hipótesis que debe evaluarse.

Palabras clave: epidemiología; metodología de la investigación; meta-epidemiología; emulación de ensayo diana; evidencia de mundo real; razón de razones; equidad en salud

INTRODUCCIÓN

La epidemiología produce evidencia mediante un repertorio amplio de diseños que difieren en la forma de asignar la exposición, controlar la confusión y seleccionar a la población. En los diseños experimentales el investigador asigna la intervención. El ensayo controlado aleatorizado (ECA) individual lo hace al azar y, con ello, equilibra en expectativa los factores de confusión medidos y no medidos, razón por la que ocupa la cúspide de las jerarquías de evidencia (1)(2). Sus variantes por conglomerados, de cuña escalonada, pragmáticas y de plataforma conservan la aleatorización, pero la adaptan a intervenciones poblacionales, a las condiciones de la práctica habitual o a la evaluación simultánea de varios tratamientos (3)(4).

En los diseños observacionales la exposición no depende del investigador. Las cohortes siguen a grupos expuestos y no expuestos, los estudios de casos y controles parten del desenlace y los transversales describen prevalencias en un momento dado. Ganan en factibilidad, tamaño muestral y cercanía a la práctica, y pierden por la confusión residual, el sesgo de selección y el error de medición (5). A ellos se suman diseños que buscan aproximarse a la inferencia causal sin aleatorizar, como la emulación de ensayos diana, el análisis de la evidencia de mundo real (RWE, por sus siglas en inglés), el diseño test-negativo y la aleatorización mendeliana (2)(6-8), y esquemas orientados a la vigilancia poblacional, como la seroepidemiología estandarizada, la vigilancia genómica y la epidemiología de aguas residuales (9-11).

Un tercer grupo responde a preguntas distintas de la eficacia. Los métodos mixtos y los diseños híbridos de efectividad-implementación evalúan a la vez el efecto de una intervención y las condiciones de su adopción (12)(13), y la investigación participativa basada en la comunidad incorpora a la población como coproductora del conocimiento (14). Cada familia tiene, por tanto, un campo de aplicación propio: estimar efectos causales, vigilar y alertar de forma temprana, o transferir intervenciones a contextos concretos.

Desde 2020 varios trabajos han comparado empíricamente estos diseños. La actualización de la revisión metodológica Cochrane encontró diferencias pequeñas, en promedio, entre los efectos estimados por estudios observacionales y por ECA, aunque con amplia variabilidad entre comparaciones (5). Estudios meta-epidemiológicos en nefrología, nutrición e investigación médica general describieron patrones semejantes y mostraron que la concordancia aumenta cuando la población, la exposición, el comparador y el desenlace de ambos cuerpos de evidencia se parecen (15-18). La iniciativa RCT-DUPLICATE emuló ensayos con bases de datos de reclamaciones de seguros y documentó que la fidelidad de la emulación explica buena parte de las discrepancias (19-21). En precisión, los ensayos pequeños sobre tratamientos para la COVID-19 arrojaron efectos más extremos que los grandes (22). En aplicabilidad, los ensayos de plataforma generaron evidencia accionable en semanas (4)(23), mientras que los modelos predictivos de aprendizaje automático mostraron alto riesgo de sesgo y escasa validación externa (24).

Estos antecedentes dejan tres vacíos. El primero es que cada estudio compara un solo par de diseños o un solo dominio clínico, de modo que no se sabe si las desviaciones observadas son propias de cada familia metodológica o comunes a todas. El segundo, que la calidad de un diseño suele juzgarse solo por su concordancia con el ECA, sin considerar la rapidez, la cobertura o la sostenibilidad de la evidencia que produce. El tercero, que las revisiones disponibles se basan casi por completo en datos de países de altos ingresos y no examinan dónde se aplica cada diseño, pese a que la pandemia de COVID-19 mostró asimetrías marcadas de capacidad de secuenciación, vigilancia e investigación clínica entre regiones (9)(10). Comparar las metodologías en un mismo marco es necesario porque agencias financiadoras, reguladores y universidades eligen diseños con recursos limitados, y esa elección condiciona qué preguntas se responden y para qué poblaciones (6)(12).

El objetivo general de este estudio fue estimar la concordancia de los diseños no aleatorizados y emergentes con el ECA y describir su eficiencia y su aplicación por región en el período 2020-2026. Planteamos tres preguntas de investigación. (P1) ¿En qué medida los estimadores de efecto de las metodologías alternativas se apartan de los del ECA de referencia, y difiere esa desviación entre familias metodológicas? (P2) ¿Qué perfil de eficiencia muestra cada esquema en validez interna, validez externa, rapidez, costo y cobertura, concordancia, y equidad y sostenibilidad? (P3) ¿Cómo se distribuye la aplicación documentada de cada esquema entre siete regiones del mundo? La P1 se respondió con un meta-análisis meta-epidemiológico, y la P2 y la P3, con una síntesis semicuantitativa basada en rúbricas explícitas.

REVISIÓN DE LA LITERATURA

Tres líneas de trabajo sostienen la pregunta del estudio: la meta-epidemiología, la inferencia causal con datos observacionales y los diseños orientados a la implementación y la participación comunitaria. La sección cierra con una distinción conceptual entre concordancia, validez y eficiencia.

Meta-epidemiología y jerarquía de la evidencia

La meta-epidemiología estudia la propia investigación. Compara, dentro de preguntas clínicas equivalentes, los efectos estimados por diseños distintos y expresa la diferencia como una razón de

razones (RoR), cociente entre el estimador de un diseño y el del diseño de referencia (25). Una RoR de 1 indica ausencia de diferencia sistemática, y los valores alejados de 1, que un diseño produce en promedio estimadores mayores o menores en la escala de razón. Leer esa dirección como sobrestimación o subestimación del beneficio depende de cómo se codifique el efecto en cada fuente, aspecto que toda síntesis de segundo orden debe armonizar.

La revisión Cochrane de Toews et al. reunió estudios meta-epidemiológicos de múltiples áreas clínicas y encontró, en promedio, diferencias pequeñas entre estudios observacionales y ECA, con certeza baja (5). Los trabajos por especialidad en nefrología (15) y nutrición (16)(17), y el análisis de cuerpos de evidencia procedentes de revisiones Cochrane (18), llegaron a conclusiones parecidas y señalaron que las discrepancias se concentran en los pares con menor similitud de población, exposición o desenlace. Comparten dos limitaciones: se restringen al contraste entre diseño observacional y ECA, y rara vez distinguen la procedencia geográfica de los datos. Ninguno contrastó en una sola síntesis las familias metodológicas emergentes.

Emulación de ensayos diana y evidencia de mundo real

El marco de emulación de ensayos diana propone especificar el protocolo del ECA hipotético que se querría realizar (elegibilidad, estrategias de tratamiento, asignación, inicio del seguimiento, desenlaces y análisis) y reproducirlo con datos observacionales (2). Su principal aporte es eliminar sesgos de diseño evitables, como el sesgo de tiempo inmortal o la desalineación entre elegibilidad, asignación e inicio del seguimiento. La iniciativa RCT-DUPLICATE lo puso a prueba al emular ensayos publicados con bases de datos de reclamaciones de seguros de salud. Sus primeros resultados mostraron que la concordancia dependía de que el diseño del ensayo pudiera reproducirse con los datos disponibles (21), y el análisis de 32 ensayos lo confirmó (19)(20). Las agencias reguladoras debaten en qué condiciones la RWE puede respaldar decisiones (6).

Dos diseños más amplían este campo. El test-negativo estima la efectividad vacunal comparando la vacunación entre personas con prueba positiva y negativa, lo que permite un monitoreo continuo a bajo costo (7). La aleatorización mendeliana usa variantes genéticas como instrumentos para anticipar el efecto de dianas farmacológicas (8).

Diseños pragmáticos, participativos y de implementación

Los ensayos por conglomerados y de cuña escalonada permiten evaluar intervenciones poblacionales cuando la aleatorización individual no es factible (3). La guía actualizada del Medical Research Council amplió la evaluación de intervenciones complejas hacia preguntas de contexto, implementación y sostenibilidad (12), y los diseños híbridos de efectividad-implementación, cuya tipología se consolidó en la última década, miden en un mismo estudio el efecto y los resultados de implementación (13). El marco consolidado para la investigación en implementación (CFIR) se ha adaptado a países de ingresos bajos y medianos (26).

En paralelo, la investigación participativa basada en la comunidad y la epidemiología crítica latinoamericana sitúan a las comunidades como coproductoras del conocimiento y vinculan la salud con su determinación social (14)(27). En Centroamérica, las metodologías de diálogo intercultural con

pueblos indígenas ilustran esta orientación (28). La evidencia acumulada indica que estos enfoques mejoran la calidad del proceso y la capacidad comunitaria, aunque la evidencia sobre desenlaces de salud es aún limitada (14).

Concordancia, validez y eficiencia: tres criterios distintos

Este estudio distingue tres criterios que la literatura suele usar como equivalentes. La concordancia es una propiedad empírica de los estimadores; indica cuánto se acerca el resultado de un diseño al de un ECA que responde la misma pregunta, y se mide con la RoR. La validez metodológica es una propiedad del diseño y de su ejecución. La validez interna depende del control de la confusión, de la selección y de la medición, y la validez externa, de la representatividad de la población y del contexto (12). La eficiencia, por último, relaciona la evidencia producida con los recursos y el tiempo empleados, e incluye la rapidez, el costo, la cobertura y la sostenibilidad.

Los tres criterios no se implican mutuamente. Un estudio observacional puede concordar con un ECA porque sesgos de signo opuesto se compensan, sin ser por ello válido; un ECA válido puede ser ineficiente si llega tarde o a un costo que impide replicarlo; y un diseño participativo puede ser eficiente en sostenibilidad sin aportar una estimación causal comparable. Por eso, en este artículo la concordancia se analiza cuantitativamente (P1), la validez se valora con herramientas de riesgo de sesgo y de certeza de la evidencia, y la eficiencia se describe en dimensiones separadas (P2), sin agregarlas en un índice único.

MÉTODO

Diseño y reporte

Realizamos una revisión sistemática con síntesis cuantitativa meta-epidemiológica, reportada según PRISMA 2020 (1)(29) y las recomendaciones de reporte de la investigación meta-epidemiológica (25). La unidad de análisis fue el estudio y no el paciente. Comparamos estimadores de efecto producidos por distintos diseños sobre preguntas equivalentes, con un marco PECOS adaptado: cuerpos de evidencia sobre intervenciones o exposiciones en salud (P); metodología alternativa (E) frente al ECA de referencia (C); discrepancia cuantitativa entre estimadores (O); y estudios meta-epidemiológicos, de emulación o de comparación de diseños (S). El protocolo no se registró de forma prospectiva.

Criterios de elegibilidad

Para el meta-análisis incluimos estudios publicados entre enero de 2020 y mayo de 2026 que compararan cuantitativamente los estimadores de efecto de al menos dos esquemas metodológicos, uno de ellos el ECA o su emulación de referencia, y que informaran una medida de discordancia con su incertidumbre. Excluimos los estudios anteriores a 2020, las comparaciones sin estimador cuantificable y las cohortes duplicadas. Para la síntesis semicuantitativa de aplicación regional y eficiencia admitimos además, en la misma ventana temporal, ensayos de referencia, revisiones sistemáticas, iniciativas de vigilancia, estudios de implementación y documentos metodológicos o normativos que aportaran información sobre rapidez, cobertura, costo, equidad o distribución regional de un esquema.

Fuentes de información y estrategia de búsqueda

Se consultaron MEDLINE/PubMed, Embase, Scopus, Web of Science (Core Collection) y Cochrane Library, con revisión manual de referencias (última actualización declarada: mayo de 2026). La Tabla S1 conserva las formulaciones y filtros informados para cada plataforma. La estrategia principal combinó expresiones de meta-epidemiología, ensayo de referencia, diseño alternativo y concordancia, y la complementaria identificó diseños de vigilancia, implementación y participación. Para reconstruir los 1 287 registros originales se necesitan, además de las ecuaciones, la interfaz utilizada, la fecha y hora de cada ejecución, los límites activados, el número recuperado por base y los identificadores exportados. Estos detalles no aparecen en el archivo fuente y no se presentan como si hubiesen sido comprobados retrospectivamente.

Selección de estudios y extracción de datos

Tras depurar duplicados se cribaron títulos y resúmenes y se revisaron a texto completo 111 registros. Se extrajeron autor, año, ámbito, comparación, métrica, intervalo de confianza, dominio clínico y uso en la síntesis; la Tabla 2 muestra los 24 estudios finalmente incluidos y la Tabla 3, los seis estimadores publicados, disponibles también en un conjunto tabular complementario. Los 87 textos excluidos se informaron por motivos generales en PRISMA, pero no se dispone en este expediente de una lista de registros individuales con motivo específico. Esa trazabilidad pendiente se reconoce expresamente, sin reconstruir exclusiones no documentadas.

Selección de estimadores y control de superposición

Para la síntesis cuantitativa se seleccionaron las RoR o medidas equivalentes en escala de razón con incertidumbre, armonizadas como diseño alternativo dividido por el de referencia. Se examinaron coincidencias de fuentes primarias para evitar reutilizar las mismas comparaciones. La independencia, sin embargo, no puede afirmarse de manera absoluta cuando un estudio agregado incorpora revisiones previas sin publicar todos sus identificadores. La Tabla S3 presenta los pares con superposición conocida o potencial y las decisiones adoptadas. Como respuesta adicional a ese riesgo se calculó una sensibilidad que excluye simultáneamente las dos síntesis potencialmente relacionadas, en lugar de depender exclusivamente de exclusiones de una en una.

Los informes de RCT-DUPLICATE de Franklin y Wang (19)(21) se representaron mediante Heyard et al. (20); Schwingshackl (17), mediante el análisis de pares de Stadelmaier (16); y Bröckelmann (18), mediante la revisión Cochrane de Toews (5). Entre las seis fuentes retenidas persiste un solapamiento potencial entre Toews (5) y Kimachi (15), pues una revisión de revisiones puede reutilizar comparaciones de síntesis temáticas. La prueba de sensibilidad adicional omite ambas fuentes simultáneamente, tanto en el conjunto exploratorio de seis como en el subconjunto principal de cinco. La ausencia de identificadores primarios compartidos verificables limita la comprobación definitiva de independencia.

Riesgo de sesgo y certeza de la evidencia

No existe una herramienta validada específica para estudios meta-epidemiológicos, por lo que adaptamos los dominios de ROBIS (30) y las recomendaciones de reporte meta-epidemiológico (25)

en cuatro dominios: D1, representatividad del muestreo de comparaciones (búsqueda sistemática frente a selección por conveniencia o por factibilidad); D2, emparejamiento de las preguntas comparadas (similitud de población, exposición, comparador y desenlace); D3, armonización de la métrica y de la dirección del efecto; y D4, tratamiento de la dependencia entre comparaciones y de la heterogeneidad, y transparencia analítica. Cada dominio se calificó como riesgo bajo, con algunas dudas o alto, y el juicio global correspondió al dominio peor calificado. La herramienta se aplicó a los diez estudios cuantitativos; los documentos de la síntesis semicuantitativa se caracterizaron por su tipo, sin valoración de riesgo de sesgo.

La valoración GRADE se utilizó como adaptación explicativa para la certeza de las comparaciones metodológicas, no como certificación de eficacia de una intervención clínica (31). Para mantener visible la lógica original se fijó un punto de partida convencional alto y se descontaron niveles por riesgo de sesgo, inconsistencia, indirectitud, imprecisión y posible sesgo de publicación. Se añadió una disminución por dependencia potencial en el subgrupo observacional y por indirectitud e imprecisión en emulación/RWE, donde solo hay dos estimadores y uno presenta riesgo alto. La Tabla 6 detalla las decisiones, cuyo juicio sigue siendo de un solo evaluador, sin revisión independiente documentada.

Síntesis cuantitativa

La medida primaria fue la RoR. Para homogeneizar la dirección, todos los estimadores se expresaron como el cociente entre el diseño alternativo y el de referencia (el ECA, su emulación de referencia o, en el estudio sobre tamaño muestral, el ECA grande). Una RoR de 1 indica ausencia de diferencia sistemática, y los valores distintos de 1, estimadores sistemáticamente mayores o menores en la escala de razón. Como la codificación del beneficio no es la misma en todas las fuentes, interpretamos la RoR como desviación relativa y reservamos los términos sobrestimación y subestimación para los casos en que la fuente los define.

Se utilizaron modelos de efectos aleatorios en escala logarítmica, con DerSimonian-Laird, REML, intervalos Hartung-Knapp-Sidik-Jonkman (HKSJ), exclusión individual y análisis por familia (32)(33). El contraste sustantivo principal incluye cinco estimadores de diseños no aleatorizados frente a ECA; el agregado que además incorpora ECA pequeños frente a grandes se conserva únicamente como exploratorio. Se añadió una exclusión conjunta de Toews y Kimachi, primero sobre los seis estimadores y después sobre los cinco del análisis principal, para examinar la influencia de una dependencia posible. Los resultados nuevos se reconstruyeron con RoR e intervalos redondeados de la Tabla 3; por tanto, son aproximaciones reproducidas sobre datos resumidos, no una reestimación de comparaciones individuales.

La heterogeneidad se describió con Q, I^2, τ^2 e intervalo de predicción. La diferencia entre familias se mantuvo como análisis exploratorio, dado que los contrastes tienen objetivos distintos y algunos grupos solo incluyen uno o dos estimadores. Una RoR con intervalo que contenga 1 no demuestra equivalencia, al no haberse fijado un margen clínicamente justificado (34). El análisis conjunto de seis no debe interpretarse como efecto promedio general de diseños alternativos, puesto que incluye una comparación entre tamaños de ensayos aleatorizados. Las pruebas de embudo y Egger con seis

puntos se presentan solo de manera descriptiva por su limitada potencia; el resultado no descarta sesgo de publicación.

Síntesis semicuantitativa

La distribución regional y la eficiencia se describieron mediante dos matrices de apreciación documental. La primera empleó siete regiones con una puntuación ordinal de 0 a 3 y la segunda seis dominios de eficiencia, en escala de 1 a 5, aplicados mediante los criterios de la Tabla 1. Sus valores no son tasas empíricas de utilización de métodos, ni permiten clasificar de manera concluyente regiones o sistemas de salud. La Tabla S2 distingue puntuaciones basadas directamente en el corpus de aquellas estimadas a partir de fuentes indirectas; estas últimas se conservan como orientación exploratoria y no se interpretan como evidencia de implantación regional comprobada.

Tabla 1.

Criterios de puntuación de las matrices de aplicación regional y de eficiencia

Matriz o dimensión	Regla de asignación
Aplicación regional (0-3)	0: sin aplicación documentada en la región en 2020-2026. 1, emergente: estudios aislados, pilotos o participación minoritaria en iniciativas globales. 2, establecida: aplicación recurrente en varios países de la región o participación sustantiva en redes multinacionales. 3, extensa o líder: programas nacionales, infraestructura de datos o redes lideradas desde la región.
Validez interna (1-5)	5: aleatorización individual con ocultamiento de la asignación. 4: aleatorización por conglomerados o adaptativa, con riesgos específicos controlados. 3: control de la confusión medida por diseño o análisis. 2: sin control formal de la confusión. 1: diseño descriptivo.
Validez externa o traslacional (1-5)	5: evidencia generada en condiciones reales con la población destinataria y orientada a la implementación. 3: población amplia en entornos asistenciales habituales. 1: población muy seleccionada y condiciones controladas.
Eficiencia temporal (1-5)	5: evidencia accionable en días o semanas. 4: semanas a meses. 3: meses a un año. 2: uno a tres años. 1: más de tres años.
Costo y cobertura (1-5)	5: cobertura poblacional con bajo costo marginal. 3: costo moderado por participante y cobertura de cientos o miles de personas. 1: costo alto por participante y cobertura limitada.

Matriz o dimensión	Regla de asignación
Concordancia con el ECA (1-5)	5: referente. 4: concordancia demostrada cuantitativamente (RoR cercana a 1 con heterogeneidad baja o correlación ≥ 0.9). 3: concordancia promedio con heterogeneidad. 2: discordancia frecuente o validación escasa. 1: no evaluable.
Equidad y sostenibilidad (1-5)	5: coproducción comunitaria y transferencia de capacidades documentadas. 3: participación local en la ejecución. 1: dependencia de infraestructura externa sin participación local.

Nota: Las puntuaciones 2 y 4 de las dimensiones de eficiencia corresponden a situaciones intermedias entre los niveles descritos. Elaboración propia.

Para puntuar cada celda se examinaron los países de procedencia de los estudios, datos regionales, redes, programas nacionales, tiempo hasta disponer de resultados, cobertura y participación comunitaria. Después se aplicaron las reglas ordinales de la Tabla 1 y se clasificó el soporte como directo o indirecto. Entre las 28 puntuaciones máximas, doce disponen de respaldo directo y dieciséis se apoyan en indicios indirectos (Tabla S2); este predominio impide atribuir a la matriz precisión censal o diferencias regionales demostradas. La asignación fue realizada por un único evaluador, por lo que se presenta como juicio documental estructurado sin acuerdo interevaluador calculado.

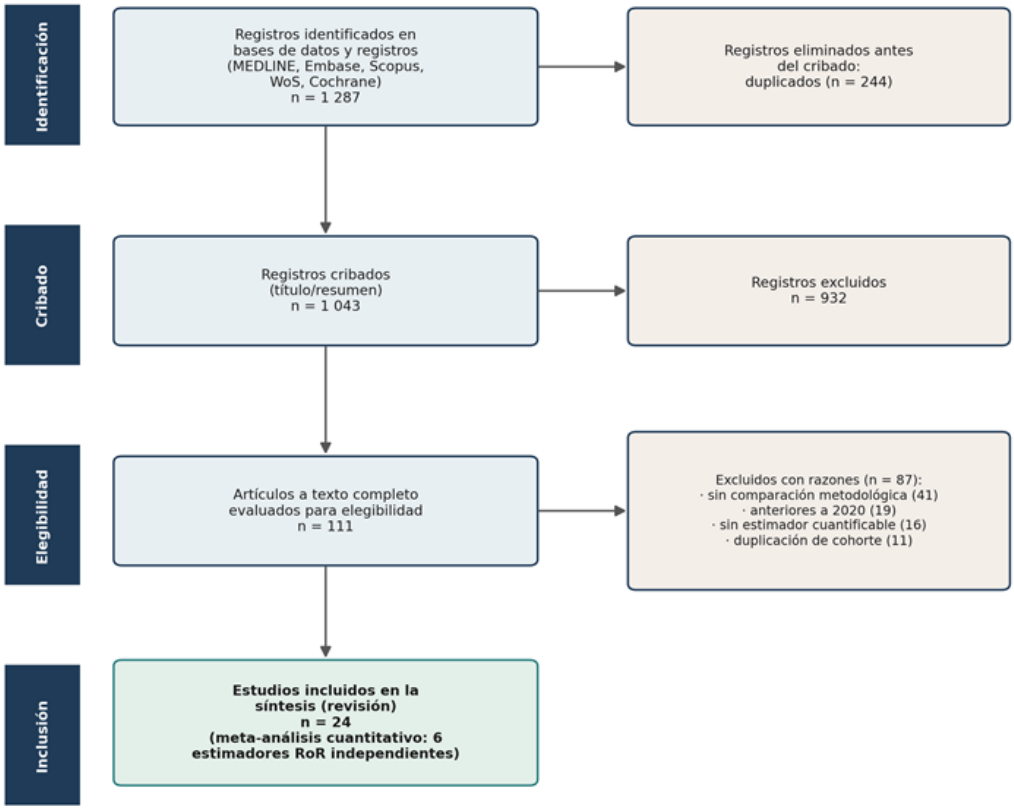
RESULTADOS

Selección de estudios

La búsqueda informó 1 287 registros, 244 duplicados y 1 043 títulos o resúmenes cribados. Se evaluaron 111 textos completos y se excluyeron 87, con inclusión final de 24 estudios (Figura 1). Diez aportaban información cuantitativa; seis proporcionaron estimadores seleccionados y cuatro fueron representados por fuentes con comparaciones superpuestas. Otros catorce se utilizaron en la síntesis semicuantitativa. La Figura 1 comunica recuentos agregados de exclusión, no el motivo específico de cada uno de los 87 textos descartados. La lista individual de decisiones no forma parte de la documentación disponible para esta versión y sería necesaria para una auditoría íntegra del proceso.

Figura 1.

Diagrama de flujo de la selección de estudios (PRISMA 2020).



Nota: Elaboración propia a partir de la búsqueda sistemática.

Características de los estudios

La Tabla 2 presenta 24 estudios: diez de comparación cuantitativa entre diseños y catorce de síntesis semicuantitativa. Entre los cuantitativos figuran cinco estudios meta-epidemiológicos, tres publicaciones de RCT-DUPLICATE, una comparación de RWE y un análisis del tamaño de los ECA. Predominan los equipos de Europa y Norteamérica (9)(10)(26)(28).

Tabla 2.

Matriz de los 24 estudios incluidos en la revisión

N.º	Autor (año) [ref.]	País / ámbito	Diseño	Objetivo	Principales resultados
1	Toews (2024) (5)	Alemania; internacional	Revisión metodológica Cochrane de estudios meta-epidemiológicos	Comparar efectos de salud estimados con diseños observacionales y con ECA	RoR 1.08 (1.01-1.15); diferencias pequeñas en promedio, certeza baja

N.º	Autor (año) [ref.]	País / ámbito	Diseño	Objetivo	Principales resultados
2	Kimachi (2021) (15)	Japón; internacional	Estudio meta-epidemiológico	Comparar estimadores de estudios observacionales y ECA en meta-análisis de nefrología	RoR 1.05 (0.90-1.23); sin diferencia sistemática
3	Stadelmaier (2025) (16)	Alemania; PAI	Meta-epidemiológico de pares individuales ECA-cohorte	Evaluar la concordancia entre ECA y cohortes individuales en nutrición	RRR 1.00 (0.91-1.10)
4	Schwingshackl (2021) (17)	Alemania; PAI	Meta-epidemiológico de cuerpos de evidencia	Comparar cuerpos de evidencia de ECA y de cohortes en nutrición	Concordancia mayor cuando la similitud PI/ECO era alta
5	Bröckelmann (2022) (18)	Alemania; internacional	Meta-epidemiológico de cuerpos de evidencia de revisiones Cochrane	Comparar cuerpos de evidencia de ECA y de cohortes en investigación médica	Concordancia en promedio; discrepancias asociadas a diferencias PI/ECO
6	Franklin (2021) (21)	Estados Unidos	Emulación de ensayos con bases de reclamaciones (RCT-DUPLICATE)	Emular los primeros diez ECA con estudios de RWE	La concordancia dependió de la emulabilidad del diseño
7	Wang (2023) (19)	Estados Unidos	Emulación de 32 ECA (RCT-DUPLICATE)	Evaluar si los análisis de bases de datos reproducen los ECA	Correlación 0.82; 0.93 en emulaciones de alta fidelidad
8	Heyard (2024) (20)	Suiza, Estados Unidos	Meta-análisis de pares ensayo-emulación	Cuantificar la variación entre ECA y emulaciones y sus determinantes de diseño	RoR 0.91 (0.86-0.96); mayor desviación en desenlaces cardiometabólicos

N.º	Autor (año) [ref.]	País / ámbito	Diseño	Objetivo	Principales resultados
9	Huang Bartlett (2020) (35)	Estados Unidos	Comparación de RWE con un ensayo de fase 3	Comparar la SLP en mundo real y en ensayo de terapia endocrina en cáncer de mama metastásico	Cociente de HR 1.04 (0.69-1.56)
10	Kim (2026) (22)	Corea del Sur; internacional	Estudio meta-epidemiológico	Comparar efectos de ECA pequeños y grandes sobre tratamientos para COVID-19	RoR 0.85 (0.76-0.95)
11	RECOVERY (2021) (4)	Reino Unido	ECA de plataforma adaptativo	Evaluar tratamientos para pacientes hospitalizados con COVID-19	Dexametasona redujo hasta en un tercio la mortalidad con ventilación mecánica
12	WHO Solidarity (2021) (23)	30 países	ECA adaptativo multinacional	Evaluar antivirales reutilizados para COVID-19	Más de 11 000 adultos en 405 hospitales; sin efecto relevante en mortalidad
13	Bergeri (2022) (9)	Global (OMS)	Revisión sistemática y meta-análisis	Estimar la seroprevalencia mundial de SARS-CoV-2 con estudios estandarizados	Razón infección:caso cercana a 2:1 en PAI de América y Europa y superior a 100:1 en África

N.º	Autor (año) [ref.]	País / ámbito	Diseño	Objetivo	Principales resultados
14	Diamond (2022) (11)	Estados Unidos; internacional	Análisis de política sanitaria	Valorar la vigilancia de patógenos en aguas residuales para la salud pública	Alerta temprana y cobertura poblacional a bajo costo
15	Brito (2022) (10)	Global	Análisis de datos de secuenciación	Describir disparidades en la vigilancia genómica de SARS-CoV-2	78 % de PAI frente a 42 % de PIBM secuenciaron más del 0.5 % de los casos
16	Means (2020) (26)	PIBM	Revisión sistemática	Evaluar el uso del CFIR en PIBM	El CFIR requiere adaptar constructos al contexto de PIBM
17	Ortiz (2020) (14)	Estados Unidos; internacional	Meta-revisión de alcance	Sintetizar procesos y resultados de la investigación comprometida con la comunidad	Mejora de la capacidad comunitaria; evidencia limitada sobre desenlaces de salud
18	Andaur Navarro (2021) (24)	Países Bajos; internacional	Revisión sistemática	Evaluar el riesgo de sesgo de modelos predictivos con aprendizaje supervisado	Riesgo de sesgo alto en la mayoría; validación externa escasa
19	MSPAS y OPS (2024) (28)	Guatemala	Guía metodológica	Orientar los diálogos interculturales en salud con pueblos indígenas	Metodología participativa aplicada a servicios de salud

N.º	Autor (año) [ref.]	País / ámbito	Diseño	Objetivo	Principales resultados
20	Dean (2021) (7)	Estados Unidos	Análisis metodológico	Examinar el diseño test-negativo para estimar la efectividad vacunal	Diseño eficiente para vigilancia continua; sensible a sesgos de búsqueda de atención
21	Sanderson (2022) (8)	Reino Unido; internacional	Revisión metodológica	Describir fundamentos y aplicaciones de la aleatorización mendeliana	Útil para anticipar efectos de dianas farmacológicas
22	Concato (2022) (6)	Estados Unidos	Análisis regulatorio	Describir el uso de RWE en decisiones regulatorias	Condiciones para aceptar la RWE como evidencia regulatoria
23	Hemming (2020) (3)	Reino Unido, Canadá	Revisión metodológica	Definir cuándo conviene un ensayo de cuña escalonada	Útil para intervenciones poblacionales de implantación escalonada
24	Skivington (2021) (12)	Reino Unido	Guía metodológica (MRC)	Actualizar el marco para desarrollar y evaluar intervenciones complejas	Amplía la evaluación al contexto, la implementación y la sostenibilidad

Nota: MA: incluido en el meta-análisis; SP: excluido del meta-análisis por superposición de comparaciones primarias (estudio que lo representa entre paréntesis); SC: síntesis semicuantitativa; ECA: ensayo controlado aleatorizado; RWE: evidencia de mundo real; RoR: razón de razones; RRR: razón de razones de riesgo; HR: razón de riesgos instantáneos; SLP: supervivencia libre de progresión; PAI: países de altos ingresos; PIBM: países de ingresos bajos y medianos; PI/ECO: población, intervención o exposición, comparador y desenlace. Elaboración propia.

Los seis estimadores de la Tabla 3 abarcan nefrología, nutrición, oncología, varios dominios clínicos y COVID-19. Tres contrastan estudios observacionales con ECA (5)(15)(16), dos emulación o RWE con ECA (20)(35) y uno compara ECA pequeños y grandes (22). Este último responde a una pregunta sobre efectos del tamaño del ensayo, no a la validez relativa de métodos no aleatorizados. Por ello, el análisis primario se restringe a los cinco primeros contrastes pertinentes y el conjunto de seis se

conserva para ilustrar la heterogeneidad conceptual. La mayoría de las fuentes representan poblaciones de países de altos ingresos.

Tabla 3.

Estimadores RoR seleccionados para la síntesis (con posible dependencia residual)

Fuente (año) [ref.]	Comparación	Métrica	RoR (IC 95 %)	Región	Dominio
Toews 2024 (5)	Observacional vs. ECA	RoR	1.08 (1.01-1.15)	Global	Todas las áreas
Kimachi 2021 (15)	Observacional vs. ECA	RoR	1.05 (0.90-1.23)	Global	Nefrología
Stadelmaier 2025 (16)	Cohorte vs. ECA	RRR	1.00 (0.91-1.10)	Global (PAI)	Nutrición
Heyard 2024 / RCT-DUPLICATE (19)(20)	Emulación / RWE vs. ECA	RoR	0.91 (0.86-0.96)	Global (PAI)	Multidominio
Huang Bartlett 2020 (35)	RWE (emulación) vs. ECA	Cociente de HR	1.04 (0.69-1.56)	Norteamérica	Oncología (mama)
Kim 2026 (22)	ECA pequeño vs. grande	RoR	0.85 (0.76-0.95)	Global	COVID-19

Nota: Todas las RoR se expresan como diseño alternativo / diseño de referencia (en (1), ECA pequeño / ECA grande). PAI: países de altos ingresos; RoR: razón de razones; RRR: razón de razones de riesgo; HR: razón de riesgos instantáneos (hazard ratio). Elaboración propia a partir de los estudios incluidos.

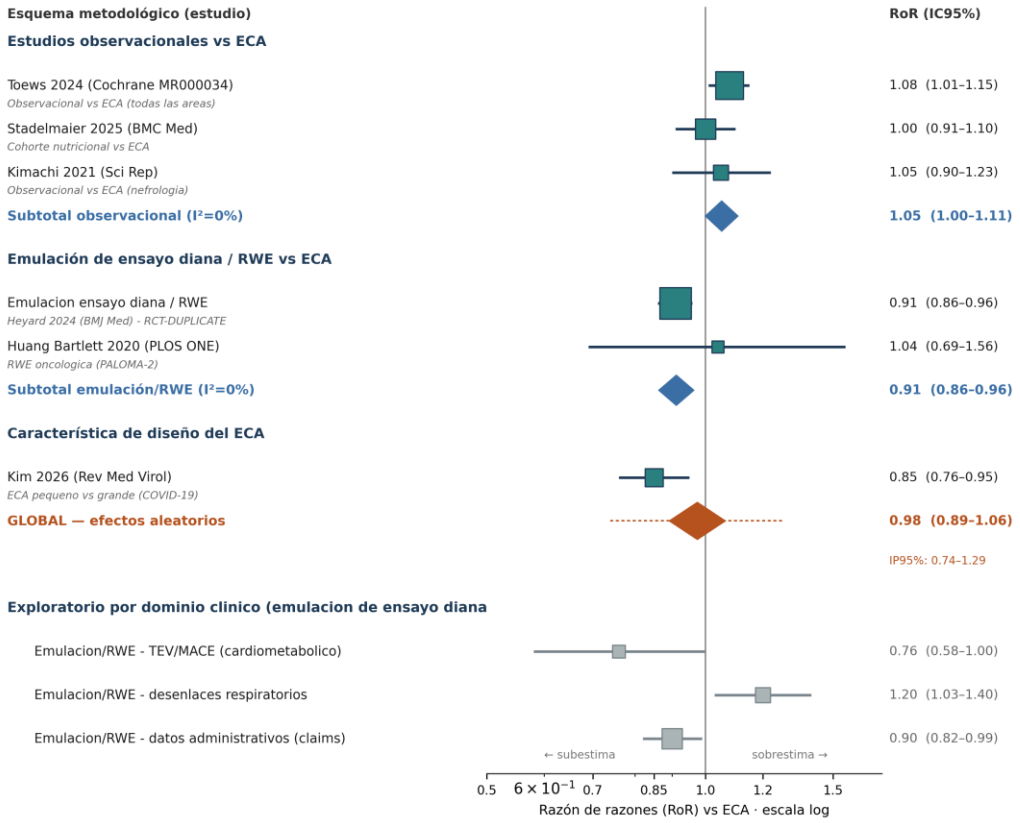
Concordancia con el ECA

La comparación de interés es la discrepancia entre estimadores obtenidos con diseños no aleatorizados y ECA. El agregado restringido a cinco RoR se considera por ello resultado principal. La sexta comparación, ECA pequeños frente a grandes, aborda un problema de tamaño y sesgo de ensayos aleatorizados y se informa separadamente. El conjunto exploratorio de seis ofrece contexto sobre heterogeneidad entre familias, pero no constituye una estimación homogénea de concordancia no aleatorizada. Además, la posible dependencia entre la revisión Cochrane y nefrología obliga a examinar el resultado al retirar ambas fuentes, como se expone en la Tabla 4.

El análisis primario restringido a cinco estimadores no aleatorizados frente a ECA produjo una RoR de 1.00 (IC 95 % 0.92–1.10; I^2 = 75.8 %). La combinación exploratoria de las seis comparaciones, representada en la Figura 2, dio una RoR de 0.98 (IC 95 % 0.89–1.06; p = 0.57), heterogeneidad I^2 = 77.9 % e intervalo de predicción de 0.74–1.29. Ninguno de los dos promedios acredita equivalencia metodológica. La comparación adicional entre tamaños de ECA se mantiene como un estrato separado y no forma parte de la inferencia principal acerca de diseños no aleatorizados.

Figura 2.

Gráfico de bosque del meta-análisis de efectos aleatorios de la razón de razones (RoR) frente al ECA de referencia, con subtotales por familia metodológica.



Nota: el gráfico conserva las seis fuentes para mostrar las tres familias metodológicas. El resultado de seis es exploratorio: RoR 0.98 (IC 95 % 0.89–1.06), I^2 = 77.9 %. La comparación principal excluye ECA pequeños frente a grandes (k = 5; RoR 1.00, IC 95 % 0.92–1.10). Elaboración propia con los estimadores del manuscrito.

La prueba entre subgrupos fue significativa (Q = 20.5; gl = 2; p < 0.001). Los estudios observacionales produjeron estimadores en promedio un 5 % mayores que los del ECA en la escala de razón (RoR 1.05; IC 95 % 1.00-1.11; I^2 = 0 %), en línea con la revisión Cochrane (5) (RoR 1.08) y con la concordancia descrita en nutrición (16) y nefrología (15). La emulación de ensayos diana y la RWE produjeron estimadores un 9 % menores (RoR 0.91; IC 95 % 0.86-0.96; I^2 = 0 %), con mayor desviación en desenlaces cardiometabólicos (RoR 0.76) que en respiratorios (RoR 1.20) (20). Los ensayos pequeños mostraron efectos más extremos que los grandes (RoR 0.85; IC 95 % 0.76-0.95) (22). Las familias se apartan del referente en direcciones opuestas, de modo que el valor global cercano a 1 resulta, en parte, de la compensación entre desviaciones de signo contrario.

El análisis principal restringido se mantuvo próximo a la unidad (k = 5; RoR 1.00; IC 95 % 0.92–1.10), aunque con heterogeneidad considerable. Los resultados de seis estimadores, presentados solo con finalidad exploratoria, variaron según REML, HKSJ y exclusión individual (Tabla 4). En una sensibilidad adicional que retiró conjuntamente Toews y Kimachi, los cuatro estimadores restantes del conjunto exploratorio produjeron aproximadamente RoR 0.92 (IC 95 % 0.86–0.99), mientras que los tres no aleatorizados restantes dieron RoR 0.95 (0.88–1.02). Estas cifras se reconstruyeron desde los RoR e

IC publicados y redondeados, por lo que no equivalen a una comprobación de independencia de las comparaciones primarias. La variación obliga a evitar conclusiones de estabilidad universal.

Tabla 4.

Análisis principal restringido y sensibilidades del agrupamiento exploratorio

Solo diseños no aleatorizados vs. ECA (sin Kim 2026)	**5**	**1.00 (0.92-1.10)**	**75.8**	**0.007**	**0.74-1.36**
Análisis	k	RoR (IC 95 %)	I^2 (%)	τ^2	IP 95 %
Análisis	k	RoR (IC 95 %)	I^2 (%)	τ^2	IP 95 %
Sin Toews y Kimachi (solo no aleatorizados vs. ECA)	3	0.95 (0.88–1.02)	35.9	0.002	No calculado
Sin Toews y Kimachi (seis fuentes exploratorias)	4	0.92 (0.86–0.99)	44.5	0.002	No calculado
τ^2 por REML	6	0.97 (0.90-1.06)	77.9	0.007	0.75-1.27
DerSimonian-Laird con HKSJ	6	0.98 (0.88-1.08)	77.9	0.008	0.74-1.29
REML con HKSJ	6	0.97 (0.88-1.08)	77.9	0.007	0.75-1.27
Sin el estimador basado en HR (Huang Bartlett 2020)	5	0.97 (0.89-1.07)	82.2	0.009	0.70-1.35
Sin Toews 2024	5	0.94 (0.88-1.01)	49.7	0.003	0.76-1.16
Sin Kimachi 2021	5	0.96 (0.87-1.06)	81.4	0.009	0.69-1.35
Sin Stadelmaier 2025	5	0.97 (0.87-1.08)	81.9	0.011	0.67-1.41
Sin Heyard 2024	5	1.00 (0.90-1.10)	70.4	0.008	0.72-1.38

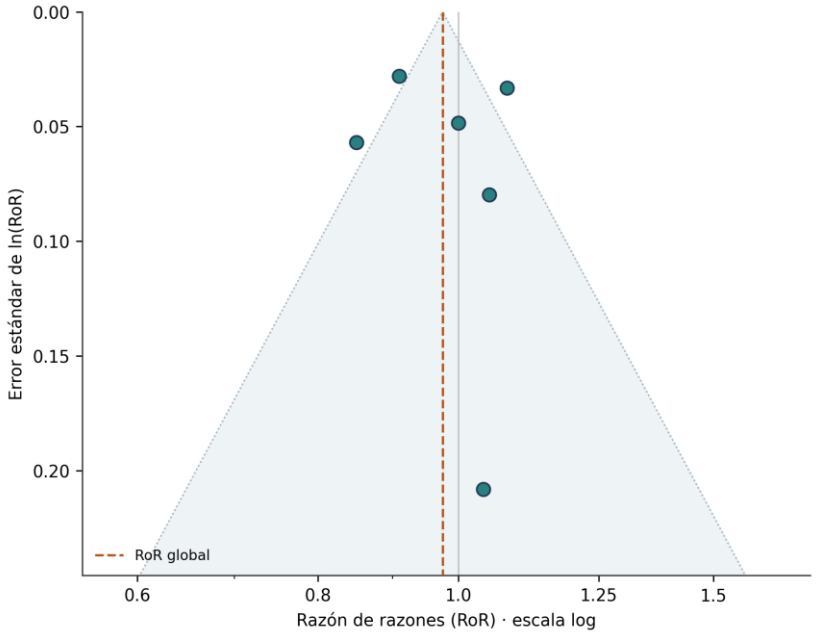
Nota: HKSJ: Hartung-Knapp-Sidik-Jonkman; REML: máxima verosimilitud restringida; IP: intervalo de predicción; k: número de estimadores. Las filas de exclusión conjunta se calcularon a partir de estimaciones y límites redondeados de la Tabla 3 por DerSimonian-Laird; su precisión es aproximada. Los demás valores proceden de la versión anterior del manuscrito. Elaboración propia.

RCT-DUPLICATE muestra qué determina la concordancia. Al emular 32 ECA, la correlación global fue 0.82 y subió a 0.93 cuando el diseño podía emularse con alta fidelidad (19). La aleatorización mendeliana mostró también una concordancia cualitativa notable con ensayos de fármacos (8). El

gráfico de embudo no evidenció asimetría manifiesta y la prueba de Egger no fue significativa (intercepto = 0.39; p = 0.86), aunque con seis estimadores su potencia es limitada (Figura 3).

Figura 3.

Gráfico de embudo de los seis estimadores seleccionados y evaluación exploratoria del sesgo de publicación.



Nota: Prueba de Egger no significativa (intercepto = 0.39; k = 6). Elaboración propia.

Riesgo de sesgo y certeza de la evidencia

De los diez estudios cuantitativos, uno presentó riesgo de sesgo bajo, ocho algunas dudas y uno riesgo alto, este último por basarse en una sola comparación (Tabla 5). Las dudas se concentraron en el dominio D1, por la selección de los ensayos según su factibilidad de emulación en RCT-DUPLICATE, y en el D2, por la similitud imperfecta entre las preguntas comparadas en los estudios de cuerpos de evidencia.

Tabla 5.

Riesgo de sesgo de los estudios cuantitativos incluidos

Estudio	D1	D2	D3	D4	Global
Toews 2024	Bajo	Dudas	Bajo	Bajo	Dudas
Kimachi 2021	Bajo	Dudas	Bajo	Dudas	Dudas
Stadelmaier 2025	Bajo	Bajo	Bajo	Bajo	Bajo
Schwingshackl 2021	Bajo	Dudas	Bajo	Bajo	Dudas
Bröckelmann 2022	Bajo	Dudas	Bajo	Bajo	Dudas
Franklin 2021	Dudas	Bajo	Bajo	Dudas	Dudas
Wang 2023	Dudas	Bajo	Bajo	Bajo	Dudas

Estudio	D1	D2	D3	D4	Global
Heyard 2024	Dudas	Bajo	Bajo	Bajo	Dudas
Huang Bartlett 2020	Alto	Bajo	Bajo	Dudas	Alto
Kim 2026	Bajo	Bajo	Dudas	Dudas	Dudas

Nota: D1: representatividad del muestreo de comparaciones; D2: emparejamiento de las preguntas comparadas; D3: armonización de la métrica y de la dirección del efecto; D4: dependencia, heterogeneidad y transparencia analítica. Bajo: riesgo bajo; Dudas: algunas dudas; Alto: riesgo alto. Adaptado de ROBIS (30) y de Murad y Wang (25). Elaboración propia.

La certeza adaptada de la evidencia se mantuvo muy baja para el promedio exploratorio de seis comparaciones, por heterogeneidad, indirectitud e imprecisión. Para observacional frente a ECA se estimó baja, al considerar el riesgo de sesgo en emparejamiento y la posibilidad de reutilización de comparaciones entre Toews y Kimachi. Para emulación/RWE frente a ECA la certeza se valoró muy baja: solo dos estimadores, selección dependiente de la emulabilidad, uno de riesgo alto e intervalos de estudio muy diferentes. Para ECA pequeños frente a grandes se mantuvo baja por un único estimador y su limitación al contexto de COVID-19. La Tabla 6 transparenta las reducciones aplicadas.

Tabla 6.

Certeza de la evidencia (GRADE) por comparación

Comparación	k	RoR (IC 95 %)	Motivos para disminuir la certeza	Certeza
Global (todas las familias)	6	0.98 (0.89-1.06)	Partida alta; −1 inconsistencia seria (I^2 = 77.9 %), −1 indirectitud (familias y preguntas distintas), −1 imprecisión (intervalo de predicción 0.74–1.29).	Muy baja
Observacional vs. ECA	3	1.05 (1.00-1.11)	Partida alta; −1 riesgo de sesgo (emparejamiento D2), −1 posible dependencia entre Toews y Kimachi no descartada con identificadores primarios. No disminuye por I^2 dentro del subgrupo.	Baja

Comparación	k	RoR (IC 95 %)	Motivos para disminuir la certeza	Certeza
Emulación / RWE vs. ECA	2	0.91 (0.86-0.96)	Partida alta; −1 sesgo (selección por emulabilidad; uno con riesgo alto), −1 indirectitud (dominios distintos), −1 imprecisión (solo k = 2 y un IC individual amplio).	Muy baja
ECA pequeño vs. grande	1	0.85 (0.76-0.95)	Partida alta; −1 indirectitud (solo tratamientos COVID-19), −1 imprecisión (un estimador, sin consistencia entre estudios evaluable).	Baja

Nota: Adaptación de GRADE a comparaciones meta-epidemiológicas, con punto de partida convencional alto y reducciones explícitas. No se redujo por sesgo de publicación exclusivamente sobre la base de Egger, cuya potencia es insuficiente con k = 6; tampoco puede darse ese sesgo por descartado. IP: intervalo de predicción; k: número de estimadores. Evaluación efectuada por un único investigador.

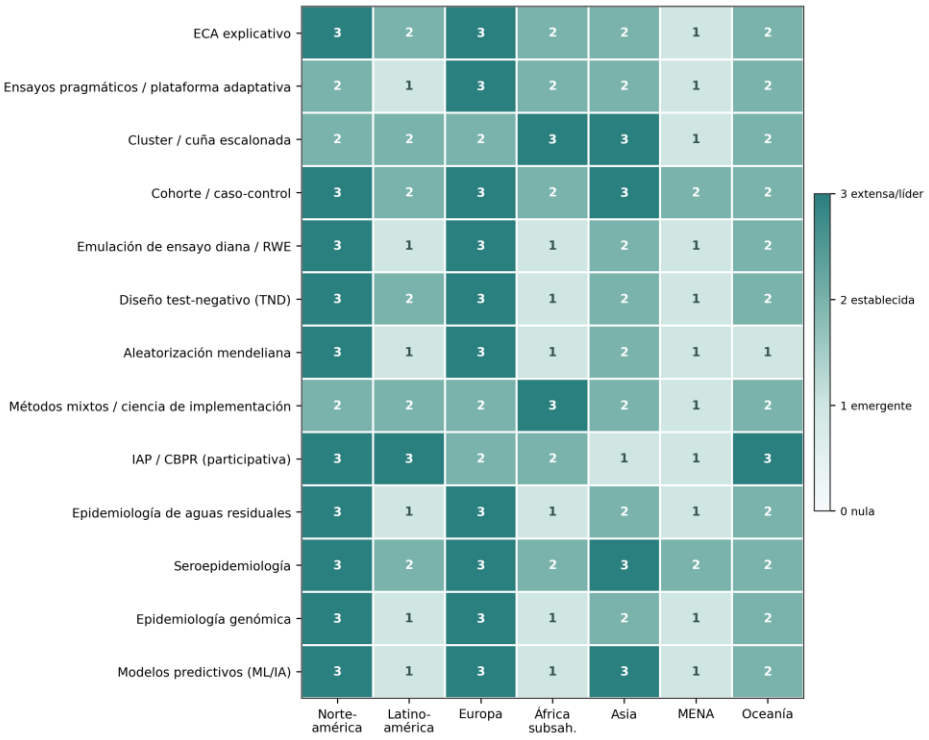
Distribución regional

La Figura 4 representa la clasificación ordinal del material localizado según la rúbrica de la Tabla 1, no una medición censal de utilización regional. De las 28 puntuaciones máximas, doce disponen de soporte directo y dieciséis dependen de documentación indirecta (Tabla S2). Por tanto, la clasificación permite identificar regiones y métodos que requieren investigación adicional, pero no establecer jerarquías empíricas concluyentes de aplicación entre continentes.

Las diferencias mejor respaldadas conciernen a la capacidad de vigilancia genómica y serológica, para la que existen datos por nivel de ingresos y regiones (9)(10). Esos resultados señalan desigualdades en infraestructura y notificación, pero no permiten extrapolar una clasificación completa de todas las metodologías. Las puntuaciones elevadas de diseños participativos e híbridos fuera de Europa o Norteamérica derivan en parte de revisiones y orientaciones que no ofrecen tasas regionales comparables (14)(26)(28). Las celdas indirectas y las correspondientes a Oriente Medio y Norte de África u Oceanía deben interpretarse únicamente como señales documentales, sin adjudicarles mayor frecuencia de aplicación comprobada.

Figura 4.

Intensidad de aplicación documentada de cada esquema metodológico por región del mundo, 2020-2026.



Nota: clasificación exploratoria en escala ordinal 0–3 según la Tabla 1. De las 28 celdas con puntuación 3, dieciséis tienen soporte indirecto, de modo que el mapa no representa tasas regionales observadas ni diferencias confirmadas. Véase la Tabla S2. Fuente. Elaboración propia.

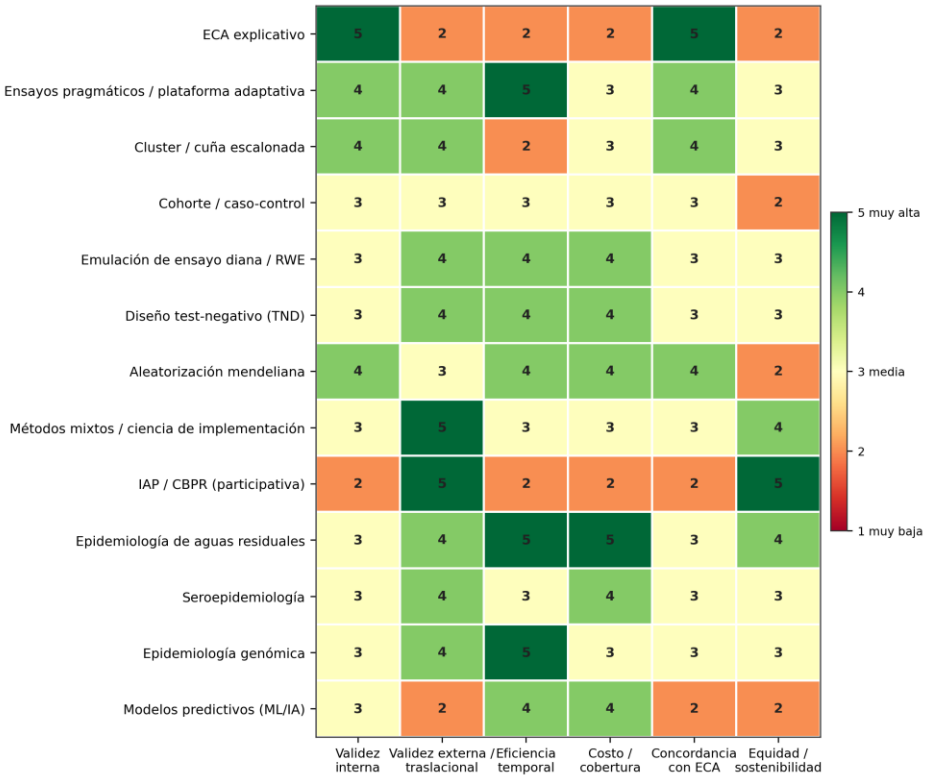
Eficiencia por dimensión

Ningún esquema dominó en las seis dimensiones evaluadas (Figura 5; Tabla 7). El ECA explicativo conserva la máxima validez interna, con una eficiencia temporal modesta. Los ensayos de plataforma adaptativos atenuaron ese cuello de botella. RECOVERY incorporó a su primer paciente unos nueve días después de redactar el protocolo y demostró que la dexametasona reducía hasta en un tercio la mortalidad de los pacientes con ventilación mecánica (4), y SOLIDARITY aleatorizó a más de 11 000 adultos en 405 hospitales de 30 países (23). La vigilancia de aguas residuales destacó por su rapidez y su cobertura a bajo costo (11). La seroepidemiología estandarizada reveló una subdetección masiva, con una razón entre infecciones y casos notificados cercana a 2:1 en los países de altos ingresos de América y Europa y superior a 100:1 en África (9).

Figura 5.

Perfil orientativo de eficiencia por dimensión y esquema metodológico, obtenido mediante rúbrica documental.

https://doi.org/10.71068/hs000106



Nota: Escala ordinal de juicio 1–5 según la Tabla 1. Las puntuaciones no equivalen a mediciones comparativas de costos, rendimiento o efectividad y no se agregan en índice único. Evaluación de un solo autor.

El perfil de la Figura 5 resume apreciaciones documentales, no comparaciones experimentales directas de eficiencia. Los diseños participativos recibieron puntuaciones más altas en equidad y sostenibilidad; los modelos predictivos, en velocidad potencial, con advertencias de riesgo de sesgo y validación externa (14)(24). Los métodos de implementación se consideraron pertinentes para transferencia contextual (3)(12)(26). La Tabla 7 describe estas características y las regiones mencionadas en el corpus, cuyo respaldo directo o indirecto debe consultarse en la Tabla S2. Ninguna de estas puntuaciones demuestra superioridad relativa sin datos comparativos homogéneos.

Tabla 7.

Eficiencia clave y regiones de mayor aplicación por esquema metodológico

Esquema metodológico	**Perfil de eficiencia según evidencia disponible**	**Regiones señaladas en el corpus (orientativo)**	**Ref.**
ECA explicativo	Validez interna máxima; referente de concordancia	Norteamérica, Europa	(2)
Ensayos pragmáticos / plataforma adaptativa	Evidencia accionable en semanas; gran precisión	Europa; 30 países (SOLIDARITY)	(4)(23)

Esquema metodológico	Perfil de eficiencia según evidencia disponible	Regiones señaladas en el corpus (orientativo)	Ref.
Conglomerados / cuña escalonada	Evaluación poblacional en contextos reales	África subsahariana, Asia, América Latina	(3)
Emulación de ensayo diana / RWE	Se aproxima al ECA cuando la emulación es de alta fidelidad; rápida	Norteamérica, Europa	(6)(19)(20)
Diseño test-negativo	Monitoreo continuo de efectividad vacunal	Global (multicontinental)	(7)
Aleatorización mendeliana	Anticipa efectos de dianas farmacológicas	Norteamérica, Europa	(8)
Métodos mixtos / implementación	Cierra la brecha eficacia-práctica; sostenibilidad	África subsahariana, América Latina	(12)(26)
IAP / CBPR (participativa)	Equidad, apropiación y sostenibilidad	América Latina, Oceanía, Norteamérica	(14)(27)(28)
Epidemiología de aguas residuales	Alerta temprana (días a semanas); bajo costo	Norteamérica, Europa, Asia	(11)
Seroepidemiología estandarizada	Mide la infección real; corrige la subnotificación	Global	(9)
Epidemiología genómica	Detección de variantes; rastreo de transmisión	Norteamérica, Europa	(10)
Modelos predictivos (ML/IA)	Rapidez analítica; requiere validación externa	Norteamérica, Europa, Asia	(24)

Nota: RWE: evidencia de mundo real; IAP: investigación-acción participativa; CBPR: investigación participativa basada en la comunidad; ML/IA: aprendizaje automático/inteligencia artificial. Las regiones listadas son orientativas, con evidencia regional de calidad variable. Elaboración propia.

DISCUSIÓN

Hallazgos principales

El resultado sustantivo principal proviene de cinco comparaciones entre diseños no aleatorizados y ECA y arroja RoR 1.00 (IC 95 % 0.92–1.10), con heterogeneidad considerable. Este promedio no equivale a demostrar ausencia de sesgo ni equivalencia, pues las comparaciones individuales y la armonización de las métricas difieren. La inclusión exploratoria del estudio sobre tamaño de ECA produce RoR 0.98 (0.89–1.06), pero mezcla preguntas distintas. La separación entre ambos conjuntos preserva la pregunta de investigación y evita que una desviación relacionada con el tamaño de ensayo se interprete como propiedad de todos los diseños no aleatorizados.

Implicaciones de la heterogeneidad y del número de estimadores

La interpretación está condicionada por pocos estimadores, posible dependencia entre Toews y Kimachi y alta heterogeneidad. Al excluir ambas fuentes conjuntamente, el reanálisis con información redondeada de la Tabla 3 arrojó RoR aproximada de 0.95 (IC 95 % 0.88–1.02) entre los tres contrastes no aleatorizados restantes, y 0.92 (0.86–0.99) cuando se conserva también el estudio sobre tamaños de ECA. La divergencia entre conjuntos evidencia cuánto influye la definición de la pregunta agregada. Ni la exclusión conjunta prueba independencia de las fuentes restantes ni el resultado de una sensibilidad permite corregir todas las diferencias conceptuales o clínicas (32)(33).

Interpretación por familia metodológica

Las RoR por familias conservan interés explicativo, pero la confianza en esas estimaciones se redujo. El grupo observacional aporta tres contrastes, con riesgo de solapamiento entre una revisión general y un estudio temático; la emulación/RWE, dos comparaciones heterogéneas, una de ellas clasificada con riesgo alto. Una RoR cercana a la unidad no identifica ausencia de confusión, y una discrepancia tampoco permite atribuir su origen a un sesgo concreto sin datos del estudio. RCT-DUPLICATE aporta información sobre concordancia condicionada por fidelidad de emulación (19)(20), pero esa observación no garantiza extrapolación a oncología ni a otros dominios. El diseño test-negativo (7) responde además a condiciones distintas de inferencia.

Dimensión regional y equidad

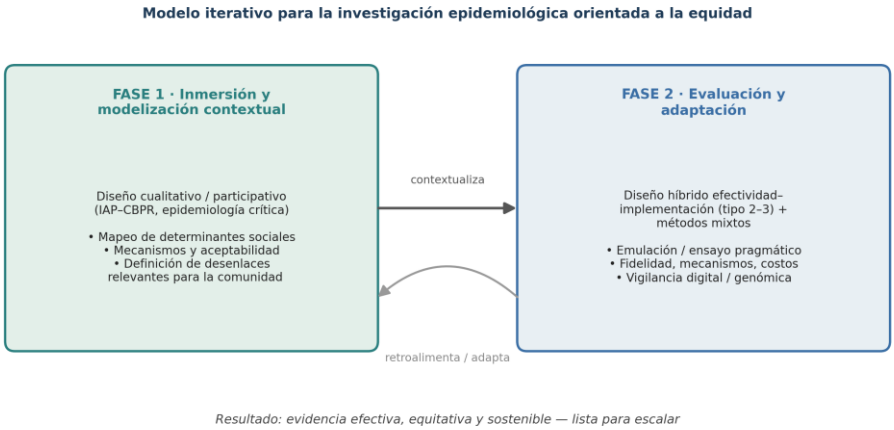
Las matrices regionales deben leerse como cartografía documental provisional. De las 28 puntuaciones máximas, 16 proceden de inferencias indirectas, por lo que no es posible concluir que determinadas familias sean sistemáticamente más utilizadas en América Latina, África o Europa. Sí existe evidencia directa sobre desigualdades de secuenciación genómica y vigilancia serológica (9)(10), así como ejemplos contextualizados de participación comunitaria (14)(28). Sin denominadores regionales comparables, conteos de estudios por diseño y auditoría independiente de las rúbricas, las demás diferencias de la Figura 4 son hipótesis de distribución y no hechos empíricos consolidados. La equidad constituye una consideración normativa relevante, distinta del puntaje descriptivo.

Propuesta de un marco de dos fases

A partir de esta complementariedad proponemos, como hipótesis de trabajo y no como resultado del análisis cuantitativo, un marco operativo de dos fases (Figura 6). La primera, de inmersión y modelización contextual, emplea un diseño cualitativo o participativo para mapear determinantes y definir los desenlaces relevantes. La segunda, de evaluación y adaptación, recurre a un diseño híbrido de efectividad-implementación con emulación de alta fidelidad o con un ensayo pragmático (12)(13)(26). Que esta combinación sea más eficiente en equidad y sostenibilidad que los diseños aislados es una proposición interpretativa, derivada de la síntesis semicuantitativa, que deberá contrastarse con estudios que comparen de forma prospectiva estrategias metodológicas.

Figura 6.

Marco metodológico sintético de dos fases para la investigación epidemiológica orientada a la equidad.



Nota: Propuesta interpretativa. Elaboración propia.

Fortalezas y limitaciones

La contribución de la revisión consiste en definir comparaciones por familia, separar un resultado principal restringido de otro exploratorio y presentar los factores que condicionan su interpretación. La sensibilidad con exclusión simultánea de Toews y Kimachi amplía la evaluación del posible solapamiento. También se documenta de manera diferenciada el carácter directo o indirecto del respaldo regional y se explicitan las decisiones de disminución de certeza GRADE. La Tabla 2 y el conjunto tabular complementario permiten identificar los estudios y las cifras agregadas de las seis comparaciones usadas. Estas mejoras favorecen la transparencia, aunque no suplen el acceso a los registros primarios de cribado y extracción.

Persisten limitaciones relevantes: no se cuenta con la lista individual de los 87 textos completos excluidos ni con bitácoras de ejecución que reproduzcan el conteo por base de datos. La posible dependencia entre algunas síntesis no puede resolverse definitivamente sin cotejo de ensayos y cohortes de origen. Los seis estimadores combinan comparaciones distintas y la métrica primaria restringida descansa solo en cinco. La certeza y las matrices regionales fueron evaluadas por una sola persona, sin segunda evaluación independiente; las comprobaciones aritméticas de agregados no sustituyen esa validación. Más de la mitad de las puntuaciones máximas regionales procede de evidencia indirecta. El protocolo no fue inscrito prospectivamente y la evidencia disponible limita la generalización de las conclusiones.

Implicaciones para la política científica

Las agencias financiadoras, las revistas y las universidades podrían valorar cada esquema por su adecuación a la pregunta y al contexto, en lugar de tratar el ECA como estándar único de calidad, y exigir transparencia sobre la fidelidad de las emulaciones y el reporte de la implementación (1)(2)(12). Reducir la brecha de capacidades del Sur global favorecería que la evidencia más oportuna represente la carga real de enfermedad.

https://doi.org/10.71068/hs000106

CONCLUSIONES

Ningún diseño mostró superioridad universal. El resultado principal, restringido a cinco comparaciones entre métodos no aleatorizados y ECA, fue RoR 1.00 (IC 95 % 0.92–1.10), sin demostrar equivalencia. La sensibilidad que excluyó conjuntamente dos fuentes potencialmente dependientes no eliminó toda la incertidumbre. El promedio exploratorio de seis estimadores, que incluye el contraste entre ECA pequeños y grandes, debe interpretarse como heterogéneo y conceptualmente mixto. La certeza GRADE adaptada de los subgrupos fue baja o muy baja, por lo que las diferencias agregadas no permiten proponer un factor universal de corrección por diseño.

A partir de la síntesis semicuantitativa proponemos, como hipótesis interpretativa, que la combinación de métodos puede mejorar la eficiencia de la investigación epidemiológica: emulación de alta fidelidad y ensayos pragmáticos o de plataforma para estimar efectos; vigilancia genómica, serológica y ambiental para ganar cobertura y oportunidad; y diseños participativos e híbridos de implementación para asegurar pertinencia y sostenibilidad. Esta proposición no se deriva del meta-análisis y requiere contraste empírico.

La evidencia regional más sólida describe asimetrías de capacidad de vigilancia, especialmente genómica y serológica. Las restantes clasificaciones regionales se basan en una matriz documental con importante componente indirecto y no permiten afirmar diferencias observadas de adopción para todos los diseños. Estas puntuaciones orientan prioridades para estudios comparativos futuros y auditorías con participación de evaluadores independientes.

REFERENCIAS BIBLIOGRÁFICAS

1. Kim DH, et al. The impact of study size on COVID-19 treatment outcomes: a meta-epidemiological study comparing large and small randomized controlled trials. Rev Med Virol. 2026;36(1):e70125. https://doi.org/10.1002/rmv.70125.
2. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71.
3. Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. JAMA. 2022;328(24):2446-7. https://doi.org/10.1001/jama.2022.21383.
4. Hemming K, Taljaard M. Reflection on modern methods: when is a stepped-wedge cluster randomized trial a good study design choice? Int J Epidemiol. 2020;49(3):1043-52. https://doi.org/10.1093/ije/dyaa077
5. RECOVERY Collaborative Group; Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, et al. Dexamethasone in hospitalized patients with Covid-19. N Engl J Med. 2021;384(8):693-704. https://doi.org/10.1056/NEJMoa2021436.
6. Toews I, Anglemyer A, Nyirenda JLZ, Alsaid D, Balduzzi S, Grummich K, et al. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials: a meta-epidemiological study. Cochrane Database Syst Rev. 2024;1(1):MR000034. https://doi.org/10.1002/14651858.MR000034.pub3.
7. Concato J, Corrigan-Curay J. Real-world evidence: where are we now? N Engl J Med. 2022;386(18):1680-2. https://doi.org/10.1056/NEJMp2200089

8. Dean NE, Hogan JW, Schnitzer ME. Covid-19 vaccine effectiveness and the test-negative design. N Engl J Med. 2021;385(15):1431-3. https://doi.org/10.1056/NEJMe2113151.
9. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, et al. Mendelian randomization. Nat Rev Methods Primers. 2022;2:6. https://doi.org/10.1038/s43586-021-00092-5.
10. Bergeri I, Whelan MG, Ware H, Subissi L, Nardone A, Lewis HC, et al. Global SARS-CoV-2 seroprevalence from January 2020 to April 2022: a systematic review and meta-analysis of standardized population-based studies. PLoS Med. 2022;19(11):e1004107. https://doi.org/10.1371/journal.pmed.1004107.
11. Brito AF, Semenova E, Dudas G, Hassler GW, Kalinich CC, Kraemer MUG, et al. Global disparities in SARS-CoV-2 genomic surveillance. Nat Commun. 2022;13(1):7003. https://doi.org/10.1038/s41467-022-33713-y.
12. Diamond MB, Keshaviah A, Bento AI, Conroy-Ben O, Driver EM, Halden RU, et al. Wastewater surveillance of pathogens can inform public health responses. Nat Med. 2022;28(10):1992-5. https://doi.org/10.1038/s41591-022-01940-x.
13. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ. 2021;374:n2061. https://doi.org/10.1136/bmj.n2061.
14. Curran GM, Landes SJ, McBain SA, Pyne JM, Smith JD, Fernandez ME, et al. Reflections on 10 years of effectiveness-implementation hybrid studies. Front Health Serv. 2022;2:1053496. https://doi.org/10.3389/frhs.2022.1053496.
15. Ortiz K, Nash J, Shea L, Oetzel J, Garoutte J, Sanchez-Youngman S, et al. Partnerships, processes, and outcomes: a health equity-focused scoping meta-review of community-engaged scholarship. Annu Rev Public Health. 2020;41:177-99. https://doi.org/10.1146/annurev-publhealth-040119-094220.
16. Kimachi M, Onishi A, Tajika A, Kimachi K, Furukawa TA. Systematic differences in effect estimates between observational studies and randomized control trials in meta-analyses in nephrology. Sci Rep. 2021;11(1):6088. https://doi.org/10.1038/s41598-021-85519-5.
17. Stadelmaier J, Bantle G, Gorenflo L, Kiesswetter E, Nikolakopoulou A, Schwingshackl L. Evaluating agreement between individual nutrition randomised controlled trials and cohort studies: a meta-epidemiological study. BMC Med. 2025;23. https://doi.org/10.1186/s12916-025-03860-2
18. Schwingshackl L, Balduzzi S, Beyerbach J, Bröckelmann N, Werner SS, Zähringer J, et al. Evaluating agreement between bodies of evidence from randomised controlled trials and cohort studies in nutrition research: meta-epidemiological study. BMJ. 2021;374:n1864. https://doi.org/10.1136/bmj.n1864.
19. Bröckelmann N, Balduzzi S, Harms L, Beyerbach J, Petropoulou M, Kubiak C, et al. Evaluating agreement between bodies of evidence from randomized controlled trials and cohort studies in medical research: a meta-epidemiological study. BMC Med. 2022;20(1):174. https://doi.org/10.1186/s12916-022-02369-2.
20. Wang SV, Schneeweiss S; RCT-DUPLICATE Initiative. Emulation of randomized clinical trials with nonrandomized database analyses: results of 32 clinical trials. JAMA. 2023;329(16):1376-85. https://doi.org/10.1001/jama.2023.4221.

21. Heyard R, Held L, Schneeweiss S, Wang SV. Design differences and variation in results between randomised trials and non-randomised emulations: meta-analysis of RCT-DUPLICATE data. BMJ Med. 2024;3(1):e000709. https://doi.org/10.1136/bmjmed-2023-000709.
22. Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, et al. Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results from the RCT DUPLICATE initiative. Circulation. 2021;143(10):1002-13. https://doi.org/10.1161/CIRCULATIONAHA.120.051718.
23. WHO Solidarity Trial Consortium; Pan H, Peto R, Henao-Restrepo AM, Preziosi MP, Sathiyamoorthy V, et al. Repurposed antiviral drugs for Covid-19: interim WHO Solidarity trial results. N Engl J Med. 2021;384(6):497-511. https://doi.org/10.1056/NEJMoa2023184.
24. Andaur Navarro CL, Damen JAA, Takada T, Nijman SWJ, Dhiman P, Ma J, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. BMJ. 2021;375:n2281. https://doi.org/10.1136/bmj.n2281.
25. Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. Evid Based Med. 2017;22(4):139-42. https://doi.org/10.1136/ebmed-2017-110713.
26. Means AR, Kemp CG, Gwayi-Chore MC, Gimbel S, Soi C, Sherr K, et al. Evaluating and optimizing the consolidated framework for implementation research (CFIR) for use in low- and middle-income countries: a systematic review. Implement Sci. 2020;15(1):17. https://doi.org/10.1186/s13012-020-0977-0.
27. Breilh J. Critical epidemiology and the people's health. New York: Oxford University Press; 2021. ISBN: 9780190492793.
28. Ministerio de Salud Pública y Asistencia Social de Guatemala, Organización Panamericana de la Salud. Guía de diálogos interculturales en salud. Guatemala: MSPAS, OPS; 2024. Disponible en: https://www.paho.org/es/documentos/guia-dialogos-interculturales-salud
29. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021;372:n160. https://doi.org/10.1136/bmj.n160.
30. Whiting P, Savović J, Higgins JPT, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol. 2016;69:225-34. https://doi.org/10.1016/j.jclinepi.2015.06.005.
31. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383-94. https://doi.org/10.1016/j.jclinepi.2010.04.026.
32. IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol. 2014;14:25. https://doi.org/10.1186/1471-2288-14-25.
33. Langan D, Higgins JPT, Jackson D, Bowden J, Veroniki AA, Kontopantelis E, et al. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. Res Synth Methods. 2019;10(1):83-98. https://doi.org/10.1002/jrsm.1316.
34. Altman DG, Bland JM. Absence of evidence is not evidence of absence. BMJ. 1995;311(7003):485. https://doi.org/10.1136/bmj.311.7003.485.
35. Huang Bartlett C, Mardekian J, Cotter MJ, Huang X, Zhang Z, Parrinello CM, et al. Concordance of real-world versus conventional progression-free survival from a phase 3 trial of endocrine

therapy as first-line treatment for metastatic breast cancer. PLoS One. 2020;15(4):e0227256. https://doi.org/10.1371/journal.pone.0227256.

CONSIDERACIONES ÉTICAS

El estudio se desarrolló mediante una revisión de literatura basada exclusivamente en fuentes documentales previamente publicadas. No involucró la participación directa de seres humanos o animales, la realización de intervenciones, el uso de muestras biológicas ni el acceso a datos personales identificables. En consecuencia, no requirió aprobación de un comité de ética ni la obtención de consentimiento informado. Durante el proceso investigativo se observaron los principios de integridad científica, atribución adecuada de la autoría, citación rigurosa y uso responsable de las fuentes consultadas.

DECLARACIÓN DE USO DE LA INTELIGENCIA ARTIFICIAL (IA)

Los autores afirman que la herramienta de inteligencia artificial se utilizó exclusivamente para apoyar la organización del artículo y la traducción del texto. Su uso no determinó el contenido científico, el diseño del estudio, la selección de fuentes, el análisis de datos, la interpretación de resultados ni las conclusiones. La responsabilidad total del contenido sigue siendo de los autores.

FINANCIACIÓN

Los autores declaran que este estudio no recibió ningún tipo de financiación externa por parte de agencias públicas, privadas, ni de organizaciones sin ánimo de lucro. Todas las actividades de investigación, análisis y desarrollo fueron realizadas con recursos propios.

CONFLICTO DE INTERESES

Los autores declaran que no tienen conflictos de intereses relacionados con este estudio y que todos los procedimientos seguidos cumplen con los estándares éticos establecidos por la revista. Asimismo, confirman que este trabajo es inédito y no ha sido publicado, ni parcial ni totalmente, en ninguna otra publicación.

CONTRIBUCIÓN DEL AUTOR

Nombres e iniciales de los autores: Marco Antonio Ramos Espinal (MARE)
1. Conceptualización: MARE
2. Curación de datos: MARE
3. Análisis formal: MARE
4. Adquisición de financiación: no aplicable
5. Investigación: MARE
6. Metodología: MARE
7. Administración de proyectos: MARE
8. Recursos: no aplicable
9. Software: no aplicable
10. Supervisión: MARE
11. Validación: MARE
12. Vista: MARE
13. Borrador original: MARE

14. Redacción, revisión y edición: MARE

MATERIAL SUPLEMENTARIO

Tabla S1.

Estrategias de búsqueda por base de datos

Base de datos	Estrategia (operadores booleanos, descriptores y términos libres)	Filtros aplicados
MEDLINE/PubMed	("meta-epidemiolog*"[tiab] OR "meta-research"[tiab] OR "ratio of ratios"[tiab] OR "ratio of odds ratios"[tiab] OR "ratio of risk ratios"[tiab] OR "target trial emulation"[tiab] OR "trial emulation"[tiab] OR "real-world evidence"[tiab]) AND ("Randomized Controlled Trials as Topic"[Mesh] OR "randomized controlled trial*"[tiab] OR "randomised controlled trial*"[tiab] OR "randomized trial*"[tiab]) AND ("Observational Studies as Topic"[Mesh] OR observational[tiab] OR "non-randomized"[tiab] OR nonrandomized[tiab] OR cohort[tiab] OR "study design*"[tiab]) AND (agreement[tiab] OR concordance[tiab] OR discrepanc*[tiab] OR "effect estimate*"[tiab] OR replicat*[tiab])	Fecha de publicación: 2020/01/01 a 2026/05/31
Embase (Elsevier)	('meta epidemiology'/exp OR 'meta-epidemiolog*':ti,ab OR 'ratio of ratios':ti,ab OR 'ratio of odds ratios':ti,ab OR 'target trial emulation':ti,ab OR 'real world evidence'/exp OR 'real-world evidence':ti,ab) AND ('randomized controlled trial (topic)'/exp OR 'randomi?ed controlled trial*':ti,ab OR 'randomi?ed trial*':ti,ab) AND ('observational study'/exp OR observational:ti,ab OR 'non-randomi?ed':ti,ab OR cohort:ti,ab) AND (agreement:ti,ab OR concordance:ti,ab OR discrepanc*:ti,ab OR 'effect estimate*':ti,ab)	[2020-2026]/py; fecha de publicación hasta mayo de 2026
Scopus	TITLE-ABS-KEY(("meta-epidemiolog*" OR "ratio of ratios" OR "ratio of odds ratios" OR "target trial emulation" OR "real-world evidence") AND ("randomized controlled trial*" OR "randomised controlled trial*" OR "randomized trial*") AND (observational OR "non-randomized" OR nonrandomized OR cohort OR "study design*") AND (agreement OR concordance OR discrepanc* OR "effect estimate*"))	PUBYEAR > 2019 AND PUBYEAR < 2027; registros hasta mayo de 2026
Web of Science (Core Collection)	TS=(("meta-epidemiolog*" OR "ratio of ratios" OR "ratio of odds ratios" OR "target trial emulation" OR "real-world evidence") AND ("randomi?ed controlled trial*" OR "randomi?ed trial*") AND (observational OR "non-randomi?ed" OR cohort OR "study design*") AND (agreement OR concordance OR discrepanc* OR "effect estimate*"))	PY=2020-2026; registros hasta mayo de 2026

Cochrane Library	("meta-epidemiolog*" OR "ratio of ratios" OR "trial emulation" OR "real-world evidence"):ti,ab,kw AND (observational OR "non-randomized" OR cohort):ti,ab,kw AND (randomized OR randomised):ti,ab,kw	Publicación entre enero de 2020 y mayo de 2026; Cochrane Database of Systematic Reviews y Methodology Review Group
Búsqueda complementaria (todas las bases)	("platform trial*" OR "adaptive trial*" OR "stepped wedge" OR "test-negative" OR "mendelian randomi?ation" OR "wastewater surveillance" OR "wastewater-based epidemiology" OR seroprevalence OR "genomic surveillance" OR "implementation science" OR "hybrid effectiveness-implementation" OR "community-based participatory research" OR "participatory action research" OR ("machine learning" AND "prediction model*")) AND ("systematic review" OR "meta-analysis" OR evaluation OR framework OR guidance)	Mismo período; sintaxis de campo adaptada a cada base

Nota: estas son las cadenas e interfaces declaradas en el manuscrito; no se adjuntó bitácora de fecha exacta, número de resultados por base, filtros de interfaz ni exportaciones originales, necesarios para repetir exactamente el conjunto de 1 287 registros. [Mesh]: MeSH; /exp: Emtree; [tiab], :ti,ab y TITLE-ABS-KEY: campos de búsqueda; TS: tema; * y ?: comodines según plataforma. La sintaxis debe contrastarse en la interfaz correspondiente. **Fuente.** Elaboración propia.

Tabla S2.

Base documental de las celdas con puntuación 3 en la matriz de aplicación regional (Figura 4)

Esquema metodológico	**Regiones con puntuación 3**	**Evidencia del corpus**	**Sustento**
ECA explicativo	Norteamérica, Europa	Procedencia de los datos de los estudios meta-epidemiológicos (PAI) (5)(16)	Indirecto
Ensayos pragmáticos / plataforma	Europa	RECOVERY (Reino Unido) (4)	Directo
Conglomerados / cuña escalonada	África subsahariana, Asia	Revisión metodológica sin desagregación regional (3)	Indirecto
Cohorte / caso-control	Norteamérica, Europa, Asia	Fuentes de las comparaciones observacionales, sin desagregación regional (5)(15)	Indirecto

Emulación de ensayo diana / RWE	Norteamérica; Europa	Bases de reclamaciones de Estados Unidos (19)(21); Europa sin estudio propio en el corpus	Directo (Norteamérica); indirecto (Europa)
Diseño test-negativo	Norteamérica; Europa	Análisis metodológico desde Estados Unidos (7); Europa sin estudio propio en el corpus	Directo (Norteamérica); indirecto (Europa)
Aleatorización mendeliana	Norteamérica, Europa	Revisión metodológica sin desagregación regional (8)	Indirecto
Métodos mixtos / implementación	África subsahariana	Revisión sistemática del CFIR en PIBM (26)	Directo
IAP / CBPR (participativa)	Norteamérica, América Latina; Oceanía	Meta-revisión de Estados Unidos (14); guía de Guatemala y epidemiología crítica (27)(28); Oceanía sin estudio propio en el corpus	Directo (Norteamérica, América Latina); indirecto (Oceanía)
Epidemiología de aguas residuales	Norteamérica; Europa	Análisis desde Estados Unidos (11); Europa sin estudio propio en el corpus	Directo (Norteamérica); indirecto (Europa)
Seroepidemiología	Norteamérica, Europa, Asia	Meta-análisis de la OMS con estimaciones por región (9)	Directo
Epidemiología genómica	Norteamérica, Europa	Disparidades de secuenciación por nivel de ingresos (10)	Directo
Modelos predictivos (ML/IA)	Norteamérica, Europa, Asia	Revisión sistemática sin desagregación regional (24)	Indirecto

Nota: Directo = estudio de la región o con resultados desagregados por región; indirecto = procedencia general, iniciativas globales o antecedentes sin medición regional particular. Solo 12 de las 28 celdas con puntuación 3 se sustentan directamente; las otras 16 no demuestran despliegue regional efectivo. Los puntajes corresponden al juicio de un solo evaluador. Elaboración propia.

Tabla S3.

Matriz de posible solapamiento de comparaciones y decisiones de inclusión

Fuentes comparadas	**Relación documental**	**Decisión en la síntesis**	**Alcance del control**
Franklin (21) / Wang (19) y Heyard (20)	Publicaciones de RCT-DUPLICATE con	Retener Heyard (20); omitir los otros de la agregación	Solapamiento conocido; regla de exclusión aplicada

	pares integrados		
Schwingshackl (17) y Stadelmaier (16)	Comparaciones de nutrición relacionadas	Retener Stadelmaier (16)	Solapamiento tratado conforme a la selección declarada
Bröckelmann (18) y Toews (5)	Comparaciones procedentes de evidencia Cochrane	Retener Toews (5)	Solapamiento tratado conforme a la selección declarada
Toews (5) y Kimachi (15)	Posible reutilización de comparaciones en revisión general y nefrología	Ambos en análisis publicado; excluir ambos simultáneamente en sensibilidad	Dependencia residual no descartada; no se dispone de lista de ensayos para matriz estudio a estudio