

ARTÍCULO ESPECIAL

Metodología para la investigación y publicación científica en pediatría. Principales diseños metodológicos

Begoña Pérez-Moneo^{a,b,*}, M. Jesús Esparza Olcina^c, Sergio Flores Villar^d, Carolina Blanco Rodríguez^e, M. Victoria Martínez Rubio^f, Manuel Molina Arias^g y Comité de Trabajo de Pediatría Basada en la Evidencia[◇]

^a Departamento de Salud Pública y Materno-Infantil, Facultad de Medicina, Universidad Complutense, Madrid, España

^b Servicio de Pediatría, Hospital Universitario Infanta Leonor, Madrid, España

^c Miembro del Grupo de Pediatría basada en la evidencia de la Asociación Española de Pediatría de Atención Primaria

^d Área de Hospitalización Pediátrica y Cardiología Pediátrica, Hospital Universitario Mutua Terrassa, Terrassa, Barcelona, España

^e Centro de Salud de Sanchinarro, Madrid, España

^f Centro de Salud de Los Fresnos, Torrejón de Ardoz, Madrid, España

^g Servicio de Gastroenterología Infantil, Hospital Universitario La Paz, Madrid, España

Recibido el 14 de marzo de 2025; aceptado el 7 de junio de 2025

PALABRAS CLAVE

Investigación en servicios de salud; Estudios epidemiológicos; Sesgo

Resumen La epidemiología es una disciplina clave en la investigación médica, ya que estudia la distribución y los determinantes de las enfermedades en las poblaciones. Para ello, emplea distintos diseños metodológicos, que se dividen en observacionales y experimentales. Los estudios observacionales incluyen los descriptivos, como series de casos y estudios transversales, y los analíticos, como los de cohortes y casos y controles, que evalúan asociaciones entre exposiciones y salud. Los estudios experimentales, como los ensayos clínicos aleatorizados, establecen relaciones de causalidad. La elección del diseño depende de la pregunta de investigación, la factibilidad y los recursos disponibles.

Es conveniente conocer sus características, las situaciones en que se deben aplicar y los sesgos más frecuentes, para la interpretación correcta de los resultados obtenidos.

© 2025 Asociación Española de Pediatría. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Autor para correspondencia.

Correo electrónico: begona.perez@salud.madrid.org (B. Pérez-Moneo).

◇ El resto de los miembros del Comité de Trabajo de Pediatría Basada en la Evidencia se presentan en el [anexo](#).

<https://doi.org/10.1016/j.anpedi.2025.503933>

1695-4033/© 2025 Asociación Española de Pediatría. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Cómo citar este artículo: B. Pérez-Moneo, M.J. Esparza Olcina, S. Flores Villar et al., Metodología para la investigación y publicación científica en pediatría. Principales diseños metodológicos, Anales de Pediatría, <https://doi.org/10.1016/j.anpedi.2025.503933>

KEYWORDS

Health care research;
Epidemiologic
studies;
Bias

Methodology for research and scientific publication in pediatrics. Main methodological designs

Abstract Epidemiology is a fundamental discipline in medical research, as it examines the distribution and determinants of diseases within populations. To achieve this, it employs various research designs, which are broadly classified into observational and experimental studies. Observational studies include descriptive studies, such as case series and cross-sectional studies, as well as analytical studies, such as cohort and case-control studies, which assess associations between exposures and health outcomes. Experimental studies, particularly randomized controlled trials, establish causal relationships. The choice of study design depends on the research question, feasibility and available resources.

It is essential to understand their characteristics, the contexts in which they should be applied and the most common sources of bias to ensure the accurate interpretation of results.

© 2025 Asociación Española de Pediatría. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

La epidemiología es una disciplina fundamental en la investigación médica, ya que permite comprender la distribución y los determinantes de las enfermedades en las poblaciones. Para lograr este objetivo, se han desarrollado diferentes diseños metodológicos que facilitan la recolección, análisis e interpretación de los datos, cada uno con sus propias ventajas, limitaciones y aplicaciones específicas.

Los diseños epidemiológicos se pueden clasificar en dos grandes grupos: los estudios observacionales y los experimentales. Los primeros incluyen estudios descriptivos, como las series de casos y los estudios transversales, que proporcionan información sobre la frecuencia y distribución de una enfermedad en una población. También comprenden estudios analíticos, como los de cohortes y los de casos y controles, que permiten evaluar posibles asociaciones entre factores de exposición y efectos en la salud. Finalmente, los estudios experimentales, en particular los ensayos clínicos controlados y aleatorizados, constituyen la herramienta más robusta para establecer relaciones de causalidad al intervenir directamente sobre la exposición de los participantes (tabla 1)¹⁻³.

La elección del diseño adecuado dependerá de múltiples factores, como la pregunta de investigación, la factibilidad del estudio y la disponibilidad de recursos. En este artículo, se analizarán las características, aplicaciones y limitaciones de los principales diseños epidemiológicos, con el objetivo de proporcionar una guía para su selección y correcta implementación en la práctica investigadora.

Estudios observacionales

Los estudios observacionales son aquellos en los que los investigadores actúan como meros observadores, sin ejercer ningún tipo de influencia sobre los factores de exposición. Se dividen en descriptivos y analíticos. Los estudios descriptivos exponen los datos encontrados en una muestra de una población, y son las series de casos y los estudios transver-

sales. Los analíticos, que tienen un grupo de comparación, son los estudios de cohortes y de casos y controles⁴.

Series de casos

Se describe un grupo de sujetos que presentan una misma característica o enfermedad.

Pueden servir para formular hipótesis y plantear estudios analíticos o experimentales, pero no para evaluar la presencia de una asociación estadística.

Estudios transversales

Carecen de direccionalidad: todas las mediciones se realizan aproximadamente al mismo tiempo, sin periodo de seguimiento. Sirven para estimar la frecuencia de una enfermedad o característica en una muestra de una población en un momento determinado, por lo que se conocen también como estudios de prevalencia⁵.

Son útiles para generar hipótesis que habrá que comprobar posteriormente con estudios más complejos.

Muestran asociaciones, pero no pueden proporcionar relaciones de causa-efecto.

Estudios de cohortes

Una cohorte designa un grupo de sujetos que tienen una característica en común, generalmente la exposición al factor de estudio.

La cohorte es observada a lo largo del tiempo, comparándola con otra cohorte de sujetos no expuestos hasta que los participantes desarrollan el efecto, hasta que se pierden durante el seguimiento o hasta la finalización del estudio.

Pueden ser prospectivos y retrospectivos. En los prospectivos las dos cohortes, expuesta y no expuesta, se forman antes del desarrollo del resultado o enfermedad de interés. En los retrospectivos, exposición y efecto ya han ocurrido cuando se inicia el estudio. En menos ocasiones, los estudios

Tabla 1 Clasificación de los diferentes estudios epidemiológicos

Según su finalidad			
Clasificación	Características generales	Tipos de estudios	Comentarios
Estudios descriptivos: estudian frecuencia y distribución de los fenómenos de salud y enfermedad	Responden a las preguntas: ¿qué?, ¿cuándo?, ¿cuántos?, ¿cómo?, ¿dónde?	Ecológicos (datos agregados)	Estudios frecuentes para salud pública, dependen de la existencia de registros Aportan datos de frecuencias absolutas y relativas, incidencia, prevalencia, mortalidad, etc.
		Estudio de un caso y series de casos	Pueden ser útiles para establecer nuevas hipótesis En series longitudinales se puede ver la evolución natural de una enfermedad
		Transversales	También denominados de prevalencia Se incluyen los estudios que evalúan pruebas diagnósticas y los de concordancia entre técnicas y observadores
Estudios analíticos: evalúan presuntas relaciones causas-efecto	Responden a la pregunta: ¿por qué?	Observacionales (cohortes y casos y controles) De intervención o experimentales	Su objetivo es evaluar hipótesis. Se clasifican en función de la actitud del investigador respecto al factor evaluado en observaciones o de intervención
Según la unidad de análisis			
Clasificación	Características generales	Tipos de estudios	Comentarios
Individuo	Datos desagregados por cada sujeto	Todos los estudios salvo los indicados en poblaciones	Los datos se presentan por cada participante
Poblaciones	Datos agregados de grupos de individuos (colegios, distritos, familias, etc.)	Estudios ecológicos Estudios comunitarios	Pueden ser de naturaleza descriptiva o de intervención
Según la direccionalidad (orden en que se investiga la causa y el efecto)			
Clasificación	Características generales	Tipos de estudios	Comentarios
Hacia delante (anterógrada)	Secuencia desde la causa al efecto	Estudios de cohortes, estudios de intervención	Solo la direccionalidad anterógrada o hacia delante permite establecer relación temporal entre causa y efecto, que es uno de los criterios de causalidad
Hacia atrás (retrógrada)	Secuencia desde el efecto a la causa	Estudios de casos y controles	
Simultánea o sin direccionalidad	Se evalúa a la vez causa y efecto	Series de casos Estudios transversales	

Tabla 1 (continuación)

Según la selección de la muestra			
Clasificación	Características generales	Tipos de estudio	Comentarios
Muestreo probabilístico	Selección con el objeto de ser representativa de la población diana	Estudios descriptivos de prevalencia	La medida de efecto que dan es la prevalencia
Muestreo a conveniencia	Selección de personas que presentan el efecto o están expuestas a un factor poco frecuente o una combinación de ambas	Estudios analíticos, que analizan hipótesis causales	Será útil en exposiciones o efectos poco frecuentes o que queramos controlar
Según la relación temporal o proximidad (tiempo transcurrido desde que se produjeron los hechos que se analizan respecto al estudio)			
Clasificación	Características generales	Tipos de estudio	Comentarios
Históricos o retrospectivos	La exposición y el efecto ya han sucedido cuando se efectúa el estudio	Cohortes retrospectivas Casos y controles, con casos prevalentes	Información obtenida de registros, por lo que hay que tener en cuenta su calidad Más baratos y rápidos de realizar
Concurrentes o prospectivos	Registra eventos que se producen a partir del inicio del estudio	Ensayos clínicos Cohortes prospectivas Casos y controles, con casos incidentes	Son estudios de mayor coste y dificultad en su ejecución, susceptibles al cambio de circunstancias y pérdida de participantes
Mixtos	En general la exposición es histórica y el efecto sucede durante el estudio	Estudios mixtos	
Según el control de la asignación de los factores de estudio			
Clasificación	Características generales	Tipos de estudio	Comentarios
Estudios no experimentales u observacionales: el factor de estudio se ha producido por un motivo independiente a la investigación	Estudio descriptivo, de grupo único, simultáneo	Estudio de un caso y series de casos	Pueden ser útiles para establecer nuevas hipótesis En series longitudinales se puede ver la evolución natural de una enfermedad
	Estudio descriptivo, de grupo único o con subdivisión, simultáneo, de poblaciones	Estudios ecológicos Estudios descriptivos con datos agregados	En los estudios ecológicos las variables, el análisis y la inferencia se deben hacer todas al mismo nivel
	Descriptivo o analítico, grupo único con subdivisión generalmente, retrospectivos	Estudios transversales	Especialmente indicados en enfermedades crónicas, aportan medidas de prevalencia, rápidos y baratos Permiten analizar varios efectos simultáneamente No establecen relaciones de causalidad No son útiles en enfermedades raras
	Estudio analítico, con grupo de comparación, hacia atrás	Estudios de casos-controles	Útiles en enfermedades poco frecuentes o con periodo de latencia largo Permiten analizar varias exposiciones No se pueden obtener tasas de incidencia

Tabla 1 (continuación)

Según el control de la asignación de los factores de estudio			
Clasificación	Características generales	Tipos de estudio	Comentarios
Otros estudios observacionales	Estudio analítico, grupo único (cohorte única) o con grupo comparación, hacia delante	Estudios de cohortes	Útiles en exposiciones raras Permiten estudiar una exposición y varios efectos En los estudios de cohortes concurrentes se minimiza el sesgo en la medida de la exposición Se obtienen tasas de incidencia
	Diseños incompletos	Estudios ecológicos de correlación Estudios de mortalidad o morbilidad proporcional Estudios de agregación temporoespacial	Estudian relaciones causa-efecto, pero por las limitaciones de su diseño se consideran exploratorios
	Diseños híbridos	Estudios de cohortes y casos	Muestreo de controles a partir de la cohorte al inicio del seguimiento Su medida de asociación es la <i>odds ratio</i>
		Estudios de casos y controles anidados en una cohorte	Muestreo de densidad de incidencia de los controles Su medida de asociación es la <i>odds ratio</i>
Estudios experimentales: el investigador controla la asignación de la exposición a estudiar	Experimental, con grupo de comparación, prospectivo, individual	Ensayo clínico controlado	Establecen la secuencia temporal entre causa y efecto Mayor control de exposición, de factores de confusión y de algunos sesgos Permiten el estudio de una única intervención
	Experimental, con grupo de comparación, prospectivo, población De intervención no aleatorizados	Ensayo de campo Ensayo comunitario Estudios cuasiexperimentales	Evalúan medidas de prevención primaria en grandes grupos de sujetos Estudios con intervención no controlada o sin asignación aleatoria del factor de estudio Muy susceptibles a incurrir en sesgos Interpretación cautelosa

de cohortes son ambispectivos (recogen datos prospectivos y retrospectivos de una misma cohorte)⁶.

Pueden clasificarse según usen un grupo de comparación interno o externo y también según el momento de inclusión de los sujetos en el estudio hablando de una cohorte fija, o dinámica o abierta cuando los sujetos continúan incluyéndose a lo largo del seguimiento.

Permiten responder a preguntas clínicas sobre pronóstico en las que sería éticamente imposible plantear un ensayo clínico. Además, permiten establecer hipótesis de causalidad como paso previo a la realización de ensayos clínicos.

En general, los estudios de cohortes son menos susceptibles a la existencia de sesgos que otros estudios observacionales, y permiten estudiar el efecto de una misma exposición sobre varias enfermedades diferentes.

Estudios de casos y controles

Se parte de dos grupos de población, diferenciados en función de la presencia (casos) o ausencia (controles) del efecto o enfermedad en estudio evaluando la existencia de un

determinado factor de exposición, con el objetivo de encontrar una asociación entre efecto y exposición^{7,8}.

Los controles deben ser representativos del nivel de exposición al factor de riesgo evaluado y de la población de la que proceden los casos.

Son más vulnerables a la presencia de sesgos que los estudios de cohortes.

Son especialmente útiles para estudiar enfermedades poco frecuentes o con un periodo de latencia muy prolongado. Son menos costosos que los estudios de cohortes, permiten estudiar un mayor número de sujetos con enfermedad y, a diferencia de los de cohortes, pueden evaluar varios factores de riesgo al mismo tiempo.

Existen una serie de sesgos comunes a los estudios observacionales, que pueden ser consultados en la [tabla 2](#).

Ensayos clínicos

Los ensayos clínicos aleatorizados (ECA) son el prototipo de diseño experimental, cuyo objetivo final es evaluar los efectos de una intervención para tratar de establecer una relación causa-efecto, imposible de asegurar con estudios observacionales. El resumen de su diseño se puede observar en la [figura 1](#).

Tienen una direccionalidad anterógrada y una temporalidad concurrente.

Para asegurar la validez interna de los ECA deben tenerse en cuenta 6 puntos fundamentales⁹:

- Pregunta clínica inicial¹⁰. Todo ECA debe buscar la respuesta a una pregunta clínica estructurada definiendo: población, intervención, comparación y resultado.
- Aleatorización. Los participantes deben asignarse a uno de los grupos (intervención o control), de forma que todos los sujetos tengan la misma posibilidad de acceder a los grupos.
- Secuencia de aleatorización. Debe ser oculta para que no se interfiera por la acción consciente o inconsciente de investigadores o participantes.
- Enmascaramiento. Las intervenciones comparadas deben enmascarse (procedimiento ciego) de tal forma que ni los participantes ni los investigadores conozcan qué recibe cada sujeto del estudio. Reduce el sesgo de información y la interferencia de la interpretación de los resultados.
- Seguimiento. Debe ser igual y completo para todos los grupos, evitando pérdidas excesivas.
- Análisis de los resultados¹¹. Los sujetos deberían ser analizados preferiblemente según el grupo al que fueron inicialmente asignados (análisis por intención de tratar) y no según el grupo en el que finalmente participaron (análisis por protocolo) ya que, en caso contrario, se pierden los beneficios de la aleatorización.

En el ECA, el estricto control que se ejerce sobre el diseño del estudio proporciona información con la mayor calidad de la evidencia, pero esto mismo disminuye su validez externa al limitar la generalización de los resultados en un entorno real.

Cuando no es posible la aleatorización, se llamarán cuasiexperimentales. Estos pueden observarse en las siguientes circunstancias: experimentos naturales (la intervención

ocurre de forma natural o circunstancial), estudio con controles históricos (se compara el grupo que recibe la intervención con otro grupo que en el pasado fue tratado con una intervención similar), estudios postintervención (la observación se realiza con posterioridad a la medida de intervención realizada), estudios antes/después (se realiza medición de ciertas variables antes y después de producirse una determinada intervención). Son más sencillos y económicos, pero son susceptibles de presentar sesgos de selección y confusión, comprometiendo su validez externa y la aplicabilidad de los resultados, además del efecto placebo y efecto Hawthorne (mejoran solo por ser observados).

Dentro de los ECA existen diferentes tipos de diseños:

- Factorial¹². Evalúan dos o más intervenciones en un mismo estudio (siempre con mecanismos de acción y efectos independientes).
- Cruzado¹³. Cada participante recibe en momentos distintos (con periodo de «blanqueo») las dos intervenciones.
- Secuencial¹⁴. Tamaño de muestra no predeterminada. Se incluyen sujetos hasta la presencia de una diferencia significativa, estableciéndose el periodo de seguimiento y los análisis intermedios.
- Pragmático¹⁵. Condiciones similares a la práctica clínica. Valora el efecto global de un tratamiento en el mundo real.
- No inferioridad¹⁶. Demostrar que un nuevo tratamiento no es menos eficaz que el tratamiento estándar.

En los ECA se ha de tener en cuenta los posibles sesgos que pueden llevar a confundir las interpretaciones finales ([tabla 2](#)).

Revisión sistemática y metaanálisis

Las revisiones sistemáticas (RS), al combinar múltiples estudios, tratan de obtener respuestas que no pueden contestarse con los estudios individuales y así obtener conclusiones firmes sobre un problema de salud específico¹⁷.

Los pasos de una RS son¹⁸: formulación del problema, localización y selección de los estudios, evaluación de la calidad o riesgo de sesgos de los mismos, proceso de extracción de datos y análisis y presentación de los resultados. Este último paso puede ser cualitativo o cuantitativo (metaanálisis [MA]).

Para la formulación del problema es muy importante una correcta construcción de la pregunta de investigación siguiendo la estructura PICO (paciente o problema, intervención, comparación y resultado -outcome-). La estrategia de búsqueda debe asegurar que todas las publicaciones pertinentes e importantes se han revisado y se especificará detalladamente. También se describirá con detalle la selección de los estudios que finalmente se vayan a incluir.

Para evaluar la calidad, deberá valorarse el riesgo de sesgo de cada uno de los estudios, utilizando herramientas validadas que se especificarán. Este análisis es muy importante, porque si la calidad es menor, la evidencia será más débil.

Tabla 2 Sesgos a considerar en los estudios observacionales, ensayos clínicos y revisiones sistemáticas

Tipo de sesgos	Ejemplos de sesgos	Descripción	Medidas de control
Sesgo de selección	Sesgos al seleccionar los participantes (autoselección, muestro no representativo, de incidencia-prevalencia, etc.)	Los participantes no son representativos de la población general Afectan a la validez externa	Aleatorización (cuando sea posible) Asegurar una muestra representativa de la población objetivo y controlar el proceso de selección
	Sesgos al ejecutar el estudio (por pérdida de seguimiento, por no respuesta) Sesgos de publicación Sesgos de recuperación Sesgo de selección de estudios publicados para revisión sistemática	Son típicos de las revisiones sistemáticas Los estudios con resultados positivos o estadísticamente significativos se publican con más frecuencia que los estudios negativos	Mismo proceso de selección para casos y controles en estudios retrospectivos Registro de los estudios y sus resultados en bases de datos públicos Realizar búsquedas sistematizadas
Sesgo de información o de clasificación errónea	No diferenciales o que influyen en todos los grupos (de la enfermedad, de la exposición) Diferenciales (de la enfermedad, de la exposición) Según su origen (de migración diagnóstica, de sospecha diagnóstica, de inaceptabilidad o su contrario de obsequiosidad, sesgo de atención, sesgo de recuerdo, sesgo del observador, de latencia de la enfermedad, de desenmascaramiento)	Hay un error sistemático en la recopilación o medición de datos. Se clasifican según el tipo de clasificación errónea que generan, si es en todos los grupos de estudio o solo en uno y según su origen Pueden afectar a la validez externa e interna	Uso de herramientas de medición validadas y estandarizadas, con definición clara de las categorías de exposición y resultado No cambiar las herramientas a lo largo del tiempo del estudio Entrenamiento adecuado de los investigadores, minimizando el número de personas que participan en la recogida de datos Doble ciego o ciego en la medición de los resultados Reducir el tiempo entre la exposición y la recogida de la información Cuando se pueda, incorporar un grupo control, idealmente con placebo, para disminuir el sesgo de atención
Sesgo de análisis	Se implican variables de confusión y de interacción o modificadoras de efecto	Una tercera variable está relacionada tanto con el factor de exposición como con el resultado, distorsionando la relación entre ambos Se pueden corregir en la fase de diseño y de análisis	En la fase de diseño: asignación aleatoria y restricción En la fase de diseño y análisis: emparejamiento En la fase de análisis: análisis estratificado y multivariante
	Falacia ecológica	Sesgo de los estudios ecológicos en que se extrapolan al nivel individual los resultados obtenidos a nivel grupal	Prudencia en la interpretación de resultados
	Sesgo de causalidad inversa	Una relación se interpreta como causal sin evidencia suficiente de que la secuencia temporal sea la adecuada Sesgo típico de los estudios transversales o series temporales	Prudencia en la interpretación de resultados

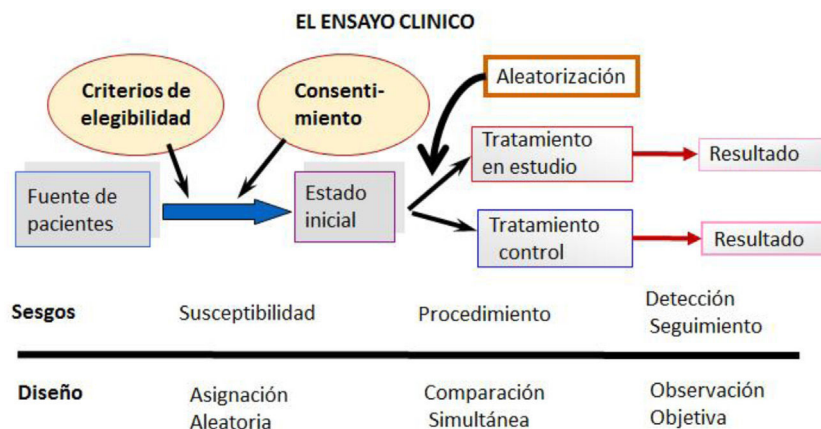


Figura 1 Esquema del diseño de un ensayo clínico básico y potenciales fuentes de los sesgos más comunes.

Debe seguirse una sistemática en la recogida de datos de los estudios que deberá explicitarse, ya que puede tener una gran influencia en la conclusión final.

Cuando la calidad y homogeneidad de los estudios lo permite, los resultados de una RS pueden completarse con un MA. El MA es un conjunto de procedimientos estadísticos cuyo objetivo es hacer una síntesis cuantitativa de los resultados de una RS, obteniendo una medida global del efecto ponderado.

Al agrupar los resultados de varios estudios individuales se obtiene un mayor tamaño muestral, lo que permite aumentar la precisión de la estimación y la potencia del estudio, permitiendo detectar diferencias que pasaron inadvertidas en los estudios individuales, obtener conclusiones cuando los resultados de los estudios individuales son contradictorios y aumentar la generalización de estos.

Para llevar a cabo un MA, el investigador debe obtener toda la información disponible; aunque este es un objetivo fundamental, hay que tener en cuenta que es posible que no se cumpla (tabla 2), por defectos en la búsqueda (sesgo de recuperación) o porque solo se hayan publicado determinados resultados y otros no (sesgo de publicación)¹⁹. También habrá que valorar si los estudios primarios son lo suficientemente homogéneos como para poder «mezclarse» en un análisis de síntesis (estudiar la heterogeneidad), elegir el método matemático más adecuado para combinar los resultados²⁰ y presentar e interpretar los resultados de la síntesis de forma adecuada, habitualmente mediante una representación gráfica conocida como diagrama de efectos o *forest plot* (fig. 2).

Este diagrama permitirá verificar qué variables se comparan y cómo, y localizar la medida de efecto utilizada para medir el resultado. Una figura, en forma de diamante, en el diagrama resume según su posición y amplitud el resultado global del MA²¹. La heterogeneidad de los estudios incluidos puede verse a simple vista mirando si los segmentos que representan los estudios primarios están o no muy dispersos y si se solapan o no, aunque es preferible cuantificarla con los parámetros adecuados, como la Q de Cochran o la I^2 . Habrá que fijarse también en qué lado de la línea de efecto nulo están el efecto global y su intervalo de confianza. Siempre que este no corte la línea de efecto nulo, el resultado será estadísticamente significativo, aunque esto no presupone la importancia clínica del mismo, que deberá

Esquema de la representación de los estudios primarios del metaanálisis mediante undiagrama de efectos o *forest plot*

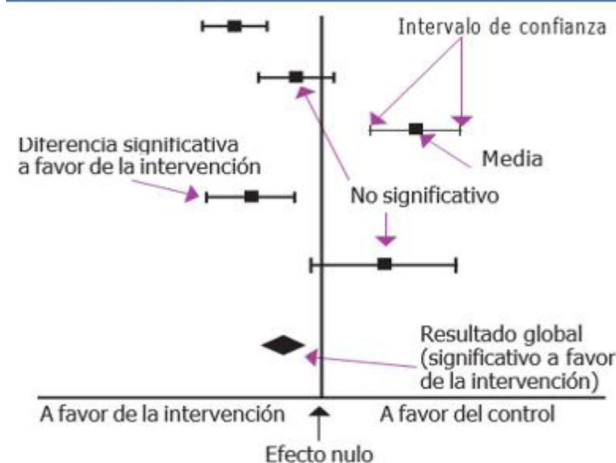


Figura 2 Interpretación del diagrama de efectos (*forest plot*) para la representación gráfica de un metaanálisis.

ser evaluado en función de la propia relevancia de la variable.

Todo lo dicho anteriormente se refiere al MA de ensayos clínicos, el que con más frecuencia encontramos en las publicaciones, no obstante, se debe conocer que hay otros tipos de MA, como los de pruebas diagnósticas o los MA en red, con algunas características diferenciales.

Estudios de pruebas diagnósticas

Una prueba diagnóstica (PD) es cualquier procedimiento que pueda proporcionar información con el fin de tomar una decisión clínica. La capacidad de la prueba de diagnosticar adecuadamente se evalúa estudiando su validez y su fiabilidad^{1,22,23}.

La validez se refiere a la capacidad de una PD para medir lo que debe, y corresponde a la exactitud diagnóstica. Se evalúa comparando sus resultados con los de un patrón de referencia (PR), y dará lugar a los principales indicadores de validez (tabla 3): sensibilidad (S), especificidad (E), cocien-

Tabla 3 Indicadores de validez para pruebas diagnósticas

		Patrón de referencia		
Prueba diagnóstica	+	+	Verdaderos positivos (VP)	Falsos positivos (FP)
		a		b
	-	-	Falsos negativos (FN)	Verdaderos negativos (VN)
		c		d
		a + c		b + d
				a + b
				c + d
Claves				
VP: enfermos con la prueba positiva				
FP: no enfermos con la prueba positiva				
FN: enfermos con la prueba negativa				
VN: no enfermos con la prueba negativa				
a + c: casos con patrón de referencia positivos				
c + d: casos con la prueba diagnóstica negativa				
Sensibilidad (Se) = $a / (a + c)$				
Especificidad (Es) = $d / (b + d)$				
Valor predictivo positivo (VPP) = $a / (a + b)$				
Valor predictivo negativo (VPN) = $d / (c + d)$				
Prevalencia = $a + c / (a + b + c + d)$				
Cociente de probabilidad (CP) positivo = $Se / (1 - Es)$				
Cociente de probabilidad negativo = $(1 - Se) / Es$				
Probabilidad prepueba (Ppre) = $(a + c) / (a + b + c + d)$				
Odds prepueba = $Ppre / (1 - Ppre)$				
Odds postprueba = CP positivo $\times$ odds pre				
Probabilidad postprueba (Ppost) = odds post / $(1 + \text{odds post})$				

tes de probabilidad (CP) o valores predictivos. En PD cuyos resultados se pueden utilizar con distintos puntos de corte, se expresarán como una curva ROC (curva de rendimiento diagnóstico) que refleja la S y la E para distintos puntos de corte, así como el rendimiento global de la prueba. La validez de las PD procede de estudios realizados en muestras de población, por lo que las estimaciones están sujetas a variabilidad aleatoria (deben expresarse en intervalos de confianza), y a sesgos si los estudios no han sido diseñados correctamente (tabla 4)²⁴.

La fiabilidad es la capacidad de una PD de obtener resultados comparables si se utiliza en condiciones idénticas. La variabilidad de las mediciones va a estar influida por múltiples factores que interesa conocer y controlar. Puede ser evaluada, dependiendo del carácter de la variable que se mida, mediante el índice Kappa, el índice Kappa ponderado o el coeficiente de correlación intraclase para pruebas con resultado cualitativo, y mediante el método de Bland-Altman para pruebas con resultado cuantitativo.

Las PD pueden ser evaluadas con cualquier tipo de estudio epidemiológico. Existe una gran variabilidad en el diseño de los estudios sobre evaluación de PD, que dificulta su interpretación y aplicación. Los aspectos que se deben evaluar cuando se analizan estos son los siguientes:

- Diseño en el tiempo: según el momento en que se aplican la PD y el PR, pueden ser de tipo simultáneo (ambas al mismo tiempo), prospectivo (antes la PD) y retrospectivo (en pacientes ya diagnosticados mediante el PR).
- Criterios de elegibilidad únicos o no para todos los participantes.

- Método de muestreo de participantes (consecutivo, aleatorio, conveniencia, disponibilidad, etc.).
- Enmascaramiento en la realización/interpretación de la PD.
- Número de estándares de referencia comparados.
- Diseño de verificación limitada o no.

Según este planteamiento se podrían clasificar los estudios de PD en los siguientes grupos²⁵:

- Diseño completamente emparejado: cada participante recibe todas las pruebas. El de mayor poder estadístico.
- Diseño parcialmente emparejado aleatorizado o no: algunos participantes reciben múltiples PD, pero otros reciben solo una. Mayor riesgo de sesgos. Opción atractiva en PD de coste elevado, invasivas, etc.
- Diseño aleatorizado no pareado: los investigadores asignan a cada participante a una de las PD/PR.
- Diseño no aleatorizado no pareado: diseño con altísimo riesgo de sesgos.

La declaración STARD^{26,27} (*Standard for Reporting of Diagnostic Accuracy*) contiene una lista de verificación de los elementos que deben ser incluidos para la publicación de estudios de precisión diagnóstica. Por otro lado, existen herramientas para analizar la calidad de los estudios de exactitud diagnóstica incluidos en revisiones sistemáticas, entre las que destaca la QUADAS 2 (*Quality Assessment of Diagnostic Accuracy Studies*)^{28,29}.

Tabla 4 Sesgos potenciales de los estudios de pruebas diagnósticas

Tipo de sesgo	Descripción	Ejemplo de sesgo	Medida de control
Sesgo de selección	No se tiene en cuenta el espectro clínico de la enfermedad o comorbilidad		Representar el espectro completo de la muestra, describir el espectro completo en el análisis; análisis de la prueba diagnóstica en los subgrupos
Sesgos de información o de clasificación errónea	No se dispone de una buena prueba de referencia y se usa la disponible (aunque no clasifique bien y discrimine)	Sesgo de la prueba de referencia imperfecta	Seguimiento evolutivo de los pacientes (saber si son enfermos o no); correcciones matemáticas
	Elementos de la prueba diagnóstica forman parte de la prueba de referencia	Sesgo de incorporación	Conceptualización adecuada de la prueba diagnóstica y de la prueba de referencia
	La interpretación de la prueba diagnóstica y la prueba de referencia se realiza conociendo el otro resultado, es decir, de modo no ciego	Sesgo de revisión o valoración ciega	Enmascaramiento de los investigadores que interpretan y realizan la prueba diagnóstica
	El resultado de la prueba diagnóstica condiciona la realización de una prueba de referencia	Sesgo de secuencia o verificación diagnóstica	Realización estándar en todos los pacientes del estudio, y si no es posible hacerlo: seguimiento de pruebas negativas y correcciones matemáticas
	Eventualidad que se produce en cualquier prueba diagnóstica o de referencia	Sesgo por resultados no interpretables	Repetir la prueba si es posible; inclusión en estudio de casos no interpretables
	Diversos observadores que actúan dentro del estudio tienen diferentes sensibilidad (S) y especificidad (E) o porque el mismo observador con el entrenamiento cambia su S y E durante el estudio	Sesgo por variabilidad en la interpretación de resultados	Estudios previos de consistencia interobservadores, correcciones matemáticas

Elección del diseño epidemiológico

Una vez vistos los diferentes diseños epidemiológicos, tendremos que plantearnos cuál será el que mejor se adapte a un determinado proyecto de investigación. Además de tener en consideración la pregunta de investigación y los objetivos del trabajo, la elección dependerá de varios factores, como la posibilidad de controlar la intervención en estudio, la necesidad de comparar diferentes grupos y el criterio de muestreo para la selección de los participantes.

Cuando se busca evaluar el efecto de una exposición en un resultado específico, lo más adecuado es introducir la exposición de manera controlada. En este caso, el estudio de intervención es el más apropiado, destacando el ensayo clínico aleatorizado y controlado como el estándar más alto, ya que permite establecer relaciones de causalidad entre exposición y efecto. No obstante, en situaciones donde no sea

posible implementar un grupo de control o realizar la aleatorización de los participantes, se pueden emplear ensayos no controlados o estudios cuasiexperimentales.

Si el control de la intervención por parte del investigador no es viable por razones éticas o prácticas, se opta por un estudio observacional, en el cual el investigador se limita a recoger datos sin modificar la exposición de los participantes. Dentro de los estudios observacionales, si no se requiere comparar grupos, se recurre a un estudio descriptivo. Un ejemplo característico es el estudio transversal, que analiza una muestra en un momento específico y que tiene las ventajas de su sencillez y bajo coste. Sin embargo, estos estudios tienen el inconveniente de no permitir establecer relaciones de causalidad. Existen también otros tipos de estudios descriptivos, como las series de casos y los casos clínicos aislados, cuya calidad de evidencia es menor.

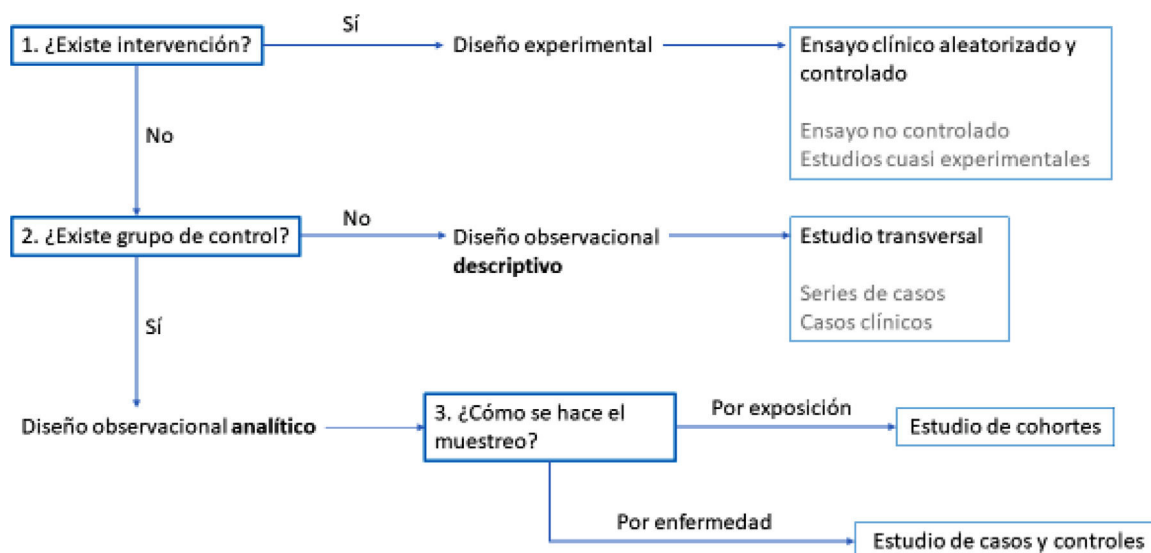


Figura 3 Esquema para la elección del diseño metodológico más apropiado.

Por otra parte, si se necesita comparar grupos, habrá que recurrir a un estudio observacional analítico, que permite evaluar relaciones entre variables. En este caso, la forma en que se seleccionan los participantes determina el tipo de estudio. Cuando los sujetos son elegidos en función de su exposición a un factor, se lleva a cabo un estudio de cohortes, el cual permite evaluar los efectos a lo largo del tiempo y proporciona mayor calidad de evidencia. En cambio, si los participantes se seleccionan según la presencia o ausencia de un efecto, como una enfermedad, se emplea un estudio de casos y controles, que resulta más eficiente cuando el efecto es poco frecuente, aunque es más susceptible a la presencia de sesgos que los estudios de cohortes.

Otro aspecto a tener en cuenta es la unidad de estudio. Generalmente, la investigación se realiza sobre sujetos individuales, pero en ciertos casos puede existir interés en analizar grupos de población. Para estos casos, los estudios ecológicos son una alternativa cuando se quiera hacer un estudio descriptivo, mientras que en los estudios de intervención pueden emplearse ensayos comunitarios.

En conclusión, la elección del diseño metodológico debe fundamentarse en la posibilidad de intervención, la necesidad de comparar grupos y el criterio de muestreo (fig. 3). Los ECA proporcionan la mejor calidad de evidencia, seguidos por los estudios de cohortes y los estudios de casos y controles. Por su parte, los estudios descriptivos, aunque útiles para ciertos propósitos, ofrecen una menor solidez en la evidencia obtenida.

Financiación

Este trabajo no ha recibido financiación.

Conflicto de intereses

Los autores del manuscrito declaran no tener ningún conflicto de intereses.

Anexo. Comité de Trabajo de Pediatría Basada en la Evidencia (resto de miembros)

Pilar Airzpurua Galdeano, M. Salomé Albi Rodríguez, María Aparicio Rodrigo, Nieves Balado Insunza, Albert Balaguer Santamaría, Laura Cabrera Morente, Fernando Carvajal Encina, Jaime Cuervo Valdés, Eduardo Cuestas Montañés, Garazi Fraile Astorga, Álvaro Gimeno Díaz de Atauri, Javier González de Dios, Paz González Rodríguez, Rafael Martín Masot, Carlos Ochoa Sangrador, Eduardo Ortega Páez, Elena Pérez González, M. José Rivero Martín, y Juan Ruiz-Canela Cáceres.

Bibliografía

- Royo Bordonada MÁ, Damián J, García López FJ. *Epidemiología básica*. Madrid: Instituto de Salud Carlos III; 2024.
- Villa Romero AR, Moreno Altamirano L, García de la Torre GS. *Epidemiología y estadística en salud pública*. México DF: McGraw-Hill Interamericana Editores; 2011.
- Comité de Pediatría Basada en la Evidencia de la AEP; Grupo de Trabajo de Pediatría Basada en la Evidencia de la AEPap. *Medicina Basada en la Evidencia. Lo que quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. [Internet]. Madrid: Lúa Ediciones; 2024 [consultado 8 Mar 2025]. Disponible en: <https://evidenciasenpediatria.es/book/>.
- Estudios descriptivos. Argimon Pallás JM, Jiménez Villa J, editores. *Métodos de investigación clínica y epidemiológica*. 5ª ed. Barcelona: Elsevier España; 2019. p. 27–32.
- Rivero Martín MJ, Molina Arias M, Ochoa Sangrador C, Comité de Pediatría Basada en la Evidencia de la AEP; Grupo de Trabajo de Pediatría Basada en la Evidencia de la AEPap. *Estudios observacionales (I)*. En: *Medicina Basada en la Evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. [Internet]. Madrid: Lúa; 2024. p. 103–9 [consultado 8 Mar 2025]. Disponible en: <https://evidenciasenpediatria.es/book/>.
- Martínez Rubio MV, Comité de Pediatría Basada en la Evidencia de la AEP; Grupo de Trabajo de Pediatría Basada en la Evidencia de la AEPap. *Estudios observacionales (II)*. En: *Medicina Basada en la Evidencia. Lo que siempre quiso saber sobre la evidencia*

- aplicada a la práctica clínica sin morir en el intento. [Internet]. Madrid: Lúa Eds; 2024. p. 111–9 [consultado 8 Mar 2025]. Disponible en: <https://evidenciasenpediatria.es/book/>.
7. Estudios de casos, controles. Argimon Pallás JM, Jiménez Villa J, editores. Métodos de investigación clínica y epidemiológica, 5ª ed. Barcelona: Elsevier España; 2019. p. 77–88.
 8. Molina Arias M, Ochoa Sangrador C. Estudios observacionales (III). Estudios de casos y controles. Evid Pediatr. 2014;10:33.
 9. Akobeng AK. Assessing the validity of clinical trials. J Pediatr Gastroenterol Nutr. 2008;47:277–82.
 10. Esparza Olcina M, Molina Arias M, Ochoa Sangrador C, Flores Villar S, Comité de Pediatría Basada en la Evidencia de la AEP; Grupo de Trabajo de Pediatría Basada en la Evidencia de la AEPap. Ensayo clínico (I). En: Medicina Basada en la Evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento. [Internet]. Madrid: Lúa Eds; 2024. p. 137–69 [consultado 8 Mar 2025]. Disponible en: <https://evidenciasenpediatria.es/book/>.
 11. Molina Arias M, Ochoa Sangrador C. Ensayo clínico (IV). Resultados. Variables y medidas de efecto e impacto. Evid Pediatr. 2015;11:33.
 12. Montgomery AA, Astin MP, Peters TJ. Reporting of factorial trials of complex interventions in community settings: a systematic review. Trials. 2011;12:179.
 13. Lathyris DN, Trikalinos TA, Ioannidis JP. Evidence from crossover trials: empirical evaluation and comparison against parallel arm trials. Int J Epidemiol. 2007;36, 422–30.
 14. Monleón-Gentino T, Barnadas-Mollins A, Roset-Gamisans M. Diseños secuenciales y análisis intermedio en una investigación clínica: tamaño frente a dificultad. Med Clin (Barc). 2009;132:437–42.
 15. Little P, Moore M, Kelly J, Williamson I, Leydon G, McDermott L, et al. Ibuprofen, paracetamol and steam for patients with respiratory tract infections in primary care: pragmatic randomised factorial trial. BMJ. 2013;347, f6041.
 16. Molina Nadal A. Ensayos clínicos de no inferioridad. FMC. Form Méd Contin Aten Primaria. 2020;27:345–8.
 17. González de Dios J, Balaguer Santamaría A. Revisión sistemática y metaanálisis (1): conceptos básicos. Evid Pediatr. 2007;3:107.
 18. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA(eds). Cochrane Handbook for Systematic Reviews of Interventions version 6.5 [actualizado Ago 2024]. Cochrane. 2024 [consultado 12 Mar 2025]. Disponible en www.training.cochrane.org/handbook.
 19. Molina Arias M. Aspectos metodológicos del metaanálisis (1). Rev Pediatr Aten Primaria. 2018;20:297–302.
 20. Molina Arias M. Aspectos metodológicos del metaanálisis (2). Rev Pediatr Aten Primaria. 2018;20:401–5.
 21. Ried K. Interpreting and understanding meta-analysis graphs-a practical guide. Aust Fam Physician. 2006;35:635–8.
 22. Ochoa Sangrador C, González de Dios J, Buñuel Álvarez JC. Evaluación de artículos científicos sobre pruebas diagnósticas. Evid Pediatr. 2007;3:24.
 23. Torregroza Diaz Granados EJ. Pruebas diagnósticas: razones de probabilidad. Rev Colomb Cir. 2021;36:403–10.
 24. Ochoa Sangrador C, González de Dios J, Buñuel Álvarez JC. Evaluación de artículos científicos sobre pruebas diagnósticas. Evid Pediatr. 2007;3:24.
 25. Yang B, Olsen M, Vali Y, Langendam MW, Takwoingi Y, Hyde CJ, et al. Study designs for comparative diagnostic test accuracy: A methodological review and classification scheme. J Clin Epidemiol. 2021;138:128–38.
 26. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open. 2016;6:e012799.
 27. Aizpurua Galdeano P, González Rodríguez P, Comité de Pediatría Basada en la Evidencia de la AEP; Grupo de Trabajo de Pediatría Basada en la Evidencia de la AEPap. STARD 2015. Guía para la elaboración de manuscritos de estudios de precisión diagnóstica o pronósticos. En: Medicina Basada en la Evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento. [Internet]. Madrid: Lúa Eds; 2024. p. 137–69 [consultado 8 Mar 2025]. Disponible en: <https://evidenciasenpediatria.es/book/>.
 28. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155:529–36.
 29. Aizpurua Galdeano P, González Rodríguez P, Comité de Pediatría Basada en la Evidencia de la AEP; Grupo de Trabajo de Pediatría Basada en la Evidencia de la AEPap. QUADAS-2. Herramienta para la valoración de la calidad de los estudios de precisión diagnóstica incluidos en revisiones sistemáticas. En: Medicina Basada en la Evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento. [Internet]. Madrid: Lúa Eds; 2024. p. 137–69 [consultado 8 Mar 2025]. Disponible en: <https://evidenciasenpediatria.es/book/>.