



CENTRO INTERNACIONAL DE ESTUDOS
DE DOUTORAMENTO E AVANZADOS
DA USC (CIEDUS)

TESIS DE DOCTORADO

**LAS FUENTES DE DATOS EN
FARMACOEPIDEMIOLOGÍA COMO
GENERADORAS DE SESGOS Y
HETEROGENEIDAD EN METANÁLISIS**

Guillermo José Prada Ramallal

ESCUELA DE DOCTORADO INTERNACIONAL
PROGRAMA DE DOCTORADO EN EPIDEMIOLOGÍA Y SALUD PÚBLICA

SANTIAGO DE COMPOSTELA

2019





DECLARACIÓN DEL AUTOR DE LA TESIS
**Las fuentes de datos en farmacoepidemiología como
generadoras de sesgos y heterogeneidad en metanálisis**

D. Guillermo José Prada Ramallal

Presento mi tesis, siguiendo el procedimiento adecuado al Reglamento, y declaro que:

- 1) La tesis abarca los resultados de la elaboración de mi trabajo.*
- 2) En su caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.*
- 3) La tesis es la versión definitiva presentada para su defensa y coincide con la versión enviada en formato electrónico.*
- 4) Confirmando que la tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por mí para la obtención de otros títulos.*

En Santiago de Compostela, a 10 de junio de 2019

Fdo. D. Guillermo José Prada Ramallal





AUTORIZACIÓN DE LOS DIRECTORES DE LA TESIS

Las fuentes de datos en farmacoepidemiología como generadoras de sesgos y heterogeneidad en metanálisis

D. Adolfo Figueiras Guzmán
D. El Bahi Takkouche Souilamas

INFORMAN:

*Que la presente tesis, corresponde con el trabajo realizado por D. **Guillermo José Prada Ramallal**, bajo nuestra dirección, y autorizamos su presentación, considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC, y que como directores de ésta no incurre en las causas de abstención establecidas en Ley 40/2015.*

En Santiago de Compostela, a 10 de junio de 2019

Fdo. D. Adolfo Figueiras Guzmán

Fdo. D. El Bahi Takkouche Souilamas





DECLARACIÓN DEL AUTOR DE LA TESIS

Las fuentes de datos en farmacoepidemiología como generadoras de sesgos y heterogeneidad en metanálisis

- No existen conflictos de interés por parte del autor de esta tesis.
- Todas las figuras incluidas en esta tesis son de elaboración propia, y su inclusión dispone de los permisos requeridos (Anexo 5).

En Santiago de Compostela, a 17 de septiembre de 2019

Fdo. D. Guillermo José Prada Ramallal



RESUMEN

Los datos clínicos y terapéuticos procedentes de registros médicos y bases de datos administrativas para facturación farmacéutica (es decir, bases de datos secundarias) presentan un potencial ampliamente reconocido en la investigación farmacoepidemiológica, habiendo aumentado notablemente su utilización durante las últimas décadas. Contribuye a esto la percepción de que los ensayos clínicos no siempre son útiles para la evaluación de las terapias en la práctica del mundo real, en particular los que proporcionan datos de seguridad limitados.

Sin embargo, la disponibilidad de toda esta información puede resultar engañosamente simple. La utilización de bases de datos secundarias implica no solo las limitaciones específicas de la investigación epidemiológica observacional, sino también las inherentes a este tipo específico de fuentes de datos. Para controlar estas situaciones se debe considerar la aplicación de diseños metodológicos apropiados y técnicas de análisis estadístico. En general, antes de iniciar una investigación utilizando bases de datos secundarias, los investigadores deben evaluar en detalle las fuentes disponibles, centrándose en el propósito para el que fueron creadas. En la fase de análisis, el origen de los datos, según sea la fuente de información primaria o secundaria, debe ser tenido en cuenta como variable en el estudio de heterogeneidad de cualquier metanálisis que considere fármacos como factor de exposición.

En este trabajo se revisa la literatura científica de las dos últimas décadas cuyo propósito fuese poner de manifiesto la existencia de algún tipo de sesgo en estudios clínicos observacionales con fuentes de datos secundarias, y se proponen estrategias para la mitigación de dichos sesgos. A continuación se discuten metodológicamente las diferentes técnicas de metanálisis disponibles para resumir la evidencia sobre seguridad de medicamentos, y se ilustra, mediante varios ejemplos, cómo difieren las conclusiones a partir de un mismo metanálisis cuando se considera el tipo de fuente de datos, primaria o secundaria, como una variable en el análisis de heterogeneidad. Por último, se evalúa el grado en que el origen de los datos se explora como una fuente de heterogeneidad en metanálisis de estudios observacionales publicados en las principales revistas médicas.

PALABRAS CLAVE: Fuente de datos; Farmacoepidemiología; Sesgo; Heterogeneidad; Metanálisis.



RESUMO

Os datos clínicos e terapéuticos procedentes de rexistros médicos e bases de datos administrativas para facturación farmacéutica (é dicir, bases de datos secundarias) teñen un potencial amplamente recoñecido na investigación farmacoepidemiolóxica, aumentando notablemente o seu uso nas últimas décadas. Contribúe a isto a percepción de que os ensaios clínicos non sempre son útiles para avaliar terapias na práctica do mundo real, especialmente aqueles que proporcionan datos de seguridade limitados.

Non obstante, a dispoñibilidade de toda esta información pode resultar enganosamente simple. O uso de bases de datos secundarias implica non só as limitacións específicas da investigación epidemiolóxica observacional, senón tamén as inherentes a este tipo específico de fontes de datos. Para controlar estas situacións débese considerar a aplicación de deseños metodolóxicos apropiados e técnicas de análise estatística. En xeral, antes de iniciar unha investigación con bases de datos secundarias, os investigadores deben avaliar en detalle as fontes dispoñibles, centrándose no propósito para o que foron creadas. Na fase de análise, a orixe dos datos, según sexa a fonte de información primaria ou secundaria, debe ser tida en conta como unha variable no estudo de heteroxeneidade de calquera metanálise que considere fármacos como factor de exposición.

Neste traballo revísase a literatura científica das dúas últimas décadas cuxo propósito fose pór de manifesto a existencia dalgún tipo de sesgo en estudos clínicos de observación con fontes de datos secundarias, e proponse estratexias para mitigar eses sesgos. A continuación discútense metodoloxicamente as diferentes técnicas de metanálise dispoñibles para resumir a evidencia sobre seguridade de medicamentos, e ilústrase, mediante varios exemplos, cómo difiren as conclusións a partir dunha mesma metanálise cando se considera o tipo de fonte de datos, primaria ou secundaria, como variable na análise de heteroxeneidade. Por último, avalíase o grao en que a orixe dos datos é explorada como fonte de heteroxeneidade en metanálises de estudos de observación publicadas nas principais revistas médicas.

PALABRAS CHAVE: Fonte de datos; Farmacoepidemioloxía; Sesgo; Heteroxeneidade; Metanálise.



ABSTRACT

Clinical and therapeutic data from medical records and administrative databases for pharmaceutical billing (i.e., secondary databases) have a widely recognised potential in pharmacoepidemiological research. Their use increased dramatically in the last decades. The perception that clinical trials are not always useful for the evaluation of therapies in day-to-day practice, particularly those that provide limited safety data, contributes to the increase in the use.

However, the apparent simplicity of the use of these databases is probably deceptive. The utilization of secondary databases implies the existence of limitations inherent to any observational epidemiological research, but also those inherent to this specific data source. The implementation of appropriate methodological designs and adequate statistical techniques should be considered to control these drawbacks. In general, before initiating a research using secondary databases, researchers should assess in detail the sources of data available, focusing on the purpose for which they were created. At the analysis stage, the origin of the data, either primary or secondary, should be considered as a variable in the assessment of heterogeneity in any meta-analysis involving drugs as an exposure factor.

This dissertation reviews the scientific literature of the last two decades in which the authors highlight the existence of some type of bias in observational clinical studies when using secondary data sources, and proposes strategies for mitigating such biases. Then, the different meta-analysis techniques available to summarize evidence on drug safety are discussed from a methodologic point of view. Several examples are provided to illustrate how conclusions differ from the same meta-analysis when the type of data source, primary or secondary, is introduced as a variable in the heterogeneity analysis. Finally, the extent to which the origin of the data is explored as a source of heterogeneity is assessed in meta-analyses of observational studies published in major medical journals.

KEYWORDS: Source of data; Pharmacoepidemiology; Bias; Heterogeneity; Meta-analysis.



LISTA DE ABREVIATURAS

AAB	Anticuerpo contra agente biológico
AINE	Antiinflamatorio no esteroideo
Ann Intern Med	<i>Annals of Internal Medicine</i>
APACHE	<i>Acute Physiology And Chronic Health Evaluation</i>
ARA	Antagonista del receptor de angiotensina
BMJ	<i>British Medical Journal</i>
CMI	Concentración mínima inhibitoria
DPI	Datos de pacientes individuales
DPP-4	Dipeptidil peptidasa-4
ECA	Ensayo clínico aleatorizado
EE	Error estándar
ENA	Estudio no aleatorizado
FNT	Factor de necrosis tumoral
GLP-1	Péptido-1 similar al glucagón
GP	Grado de predisposición/propensión
HR	Razón de riesgos instantáneos (<i>Hazard Ratio</i>)
IBP	Inhibidor de la bomba de protones
IC	Intervalo de confianza
IECA	Inhibidor de la enzima convertidora de angiotensina
ISRS	Inhibidor selectivo de la recaptación de serotonina
JAMA	<i>Journal of the American Medical Association</i>
JAMA Intern Med	<i>JAMA Internal Medicine</i>
MA	Metanálisis
MeSH	<i>Medical Subject Headings</i>
MOOSE	<i>Meta-analysis Of Observational Studies in Epidemiology</i>
Nat Rev Dis Primers	<i>Nature Reviews Disease Primers</i>
NEJM	<i>New England Journal of Medicine</i>
NND	Número necesario para dañar
NR	No reportado
OR	Razón de odds (<i>Odds Ratio</i>)
PRISMA	<i>Preferred Reporting Items for Systematic reviews and Meta-Analyses</i>
RAM	Reacción adversa a medicamento
RR	Razón de riesgos (<i>Risk Ratio</i>), riesgo relativo
SGLT-2	Cotransportador sodio-glucosa tipo-2
VI	Variable instrumental



ÍNDICE

1. INTRODUCCIÓN.....	21
1.1. RETOS Y DIFICULTADES EN EL USO DE FUENTES DE DATOS EN FARMACOEPIDEMIOLOGÍA.....	23
1.1.1. Datos primarios y secundarios.....	23
1.1.2. El error sistemático en farmacoepidemiología.....	26
1.1.2.1. Confusión.....	29
1.1.2.1.1. Confusión por indicación de tratamiento.....	29
1.1.2.1.2. Confusión tiempo-dependiente.....	30
1.1.2.1.3. Confusión por variables no medidas.....	30
1.1.2.2. Sesgo de selección.....	31
1.1.2.2.1. Sesgo protopático.....	31
1.1.2.2.2. Pérdidas durante el seguimiento (censuras).....	31
1.1.2.2.3. Depleción de susceptibles (sesgo del usuario prevalente).....	31
1.1.2.2.4. Falta de datos.....	32
1.1.2.3. Sesgo de medición.....	32
1.1.2.3.1. Sesgo de clasificación.....	32
1.1.2.4. Sesgo relacionado con el tiempo.....	34
1.1.2.4.1. Sesgo de tiempo inmortal.....	34
1.1.2.4.2. Otras formas de sesgo relacionado con el tiempo.....	36
1.1.3. Conclusiones.....	36
1.2. ESTRATEGIAS DE CONTROL DE SESGOS EN FARMACOEPIDEMIOLOGÍA.....	37
1.2.1. Diseño del estudio.....	37
1.2.1.1. Restricción.....	37
1.2.1.2. Diseños de usuarios nuevos/incidentes.....	38
1.2.1.3. Diseños de casos.....	39
1.2.1.4. Otros diseños.....	39
1.2.2. Análisis de datos.....	40
1.2.2.1. Estratificación.....	40
1.2.2.2. Regresión.....	40
1.2.2.3. Grado de riesgo de enfermedad.....	41
1.2.2.4. Grado de predisposición a recibir un tratamiento: Análisis con una variable de propensión a la exposición.....	41
1.2.2.5. Análisis de sensibilidad y ajuste externo.....	43
1.2.2.5.1. Análisis simples de sensibilidad de factores de confusión a partir de datos agregados.....	44
1.2.2.5.2. Ajustes externos que abordan la distribución conjunta de varios factores de confusión a partir de datos individuales.....	44
1.2.2.6. Análisis con variables instrumentales.....	44
1.2.2.7. Control de la confusión tiempo-dependiente en el análisis.....	45
1.2.2.7.1. Estimación-g.....	45
1.2.2.7.2. Modelos estructurales marginales.....	46
1.2.2.8. Otras técnicas.....	46
1.2.3. Conclusiones.....	47

2. OBJETIVOS.....	49
3. MÉTODOS.....	51
3.1. EL SESGO EN ESTUDIOS FARMACOEPIDEMIOLÓGICOS QUE UTILIZAN BASES DE DATOS SECUNDARIAS: UNA REVISIÓN EXPLORATORIA.....	51
3.1.1. Fuentes de datos y estrategia de búsqueda.....	51
3.1.2. Selección de artículos y extracción de datos.....	51
3.2. CONCLUSIONES DIVERGENTES A PARTIR DEL MISMO METANÁLISIS EN ESTUDIOS DE SEGURIDAD DE MEDICAMENTOS. RELEVANCIA DEL TIPO DE FUENTE DE DATOS CONSIDERADA.....	53
3.2.1. Criterios de inclusión.....	53
3.2.2. Evaluación de la exposición y el resultado.....	53
3.2.3. Extracción de datos.....	53
3.2.4. Análisis de datos.....	54
3.3. UTILIZACIÓN DE FUENTES DE DATOS Y HETEROGENEIDAD EN METANÁLISIS DE ESTUDIOS OBSERVACIONALES SOBRE EFECTOS DE MEDICAMENTOS: UN ESTUDIO DE LAS PRINCIPALES REVISTAS MÉDICAS.....	55
3.3.1. Selección de metanálisis y proceso de recogida de datos.....	55
3.3.2. Evaluación de la exposición y el resultado.....	55
4. RESULTADOS.....	57
4.1. EL SESGO EN ESTUDIOS FARMACOEPIDEMIOLÓGICOS QUE UTILIZAN BASES DE DATOS SECUNDARIAS: UNA REVISIÓN EXPLORATORIA.....	57
4.1.1. Tendencia de publicación.....	57
4.1.2. Principales sesgos mencionados en los artículos incluidos en la revisión.....	57
4.2. CONCLUSIONES DIVERGENTES A PARTIR DEL MISMO METANÁLISIS EN ESTUDIOS DE SEGURIDAD DE MEDICAMENTOS. RELEVANCIA DEL TIPO DE FUENTE DE DATOS CONSIDERADA.....	64
4.2.1. Inhibidores selectivos de la recaptación de serotonina y sangrado gastrointestinal alto.....	64
4.2.2. Inhibidores de la bomba de protones junto con tratamiento antiplaquetario dual y eventos adversos cardiovasculares mayores....	64
4.2.3. Fármacos antiinflamatorios no esteroideos no aspirina e infarto de miocardio.....	65
4.2.4. Paracetamol en el embarazo y asma infantil.....	66
4.3. UTILIZACIÓN DE FUENTES DE DATOS Y HETEROGENEIDAD EN METANÁLISIS DE ESTUDIOS OBSERVACIONALES SOBRE EFECTOS DE MEDICAMENTOS: UN ESTUDIO DE LAS PRINCIPALES REVISTAS MÉDICAS.....	68
4.3.1. Fuente de datos de la exposición y el resultado.....	68
4.3.2. Fuente de datos en el análisis de la heterogeneidad.....	71

5. DISCUSIÓN.	81
5.1. EL SESGO EN ESTUDIOS FARMACOEPIDEMIOLÓGICOS QUE UTILIZAN BASES DE DATOS SECUNDARIAS: UNA REVISIÓN EXPLORATORIA.	81
5.1.1. Limitaciones del método.	86
5.2. DISCUSIÓN METODOLÓGICA ACERCA DE LAS DIFERENTES TÉCNICAS DE METANÁLISIS PARA RESUMIR LA EVIDENCIA SOBRE SEGURIDAD DE MEDICAMENTOS.	89
5.2.1. Fuentes de datos utilizadas para resumir los resultados de seguridad de medicamentos.	90
5.2.1.1. Ensayos clínicos aleatorizados y estudios no aleatorizados.	90
5.2.1.2. Estudios publicados y no publicados.	92
5.2.1.3. Fuentes de datos primarias y secundarias.	94
5.2.2. Diferentes tipos de síntesis.	95
5.2.2.1. Síntesis cualitativa.	95
5.2.2.1.1. <i>Revisiones narrativas tradicionales.</i>	95
5.2.2.1.2. <i>Revisiones sistemáticas.</i>	95
5.2.2.2. Síntesis cuantitativa.	96
5.2.2.2.1. <i>Metanálisis de datos agregados.</i>	96
5.2.2.2.2. <i>Metanálisis de datos de pacientes individuales.</i>	96
5.2.2.2.3. <i>Metanálisis de datos de pacientes individuales planificados prospectivamente.</i>	96
5.2.2.2.4. <i>Metanálisis en red (metanálisis de tratamiento múltiple o de comparación de tratamiento mixto).</i>	96
5.2.3. ¿Cuándo se pueden realizar metanálisis y revisiones sistemáticas?	97
5.2.4. Metanálisis sobre seguridad de medicamentos: justificación, ventajas y limitaciones.	97
5.2.4.1. Metanálisis de datos agregados.....	97
5.2.4.2. Metanálisis de datos individuales.	99
5.3. CONCLUSIONES DIVERGENTES A PARTIR DEL MISMO METANÁLISIS EN ESTUDIOS DE SEGURIDAD DE MEDICAMENTOS. RELEVANCIA DEL TIPO DE FUENTE DE DATOS CONSIDERADA.	105
5.3.1. Limitaciones del método.	107
5.4. UTILIZACIÓN DE FUENTES DE DATOS Y HETEROGENEIDAD EN METANÁLISIS DE ESTUDIOS OBSERVACIONALES SOBRE EFECTOS DE MEDICAMENTOS: UN ESTUDIO DE LAS PRINCIPALES REVISTAS MÉDICAS.	108
5.4.1. Limitaciones del método.	110
6. CONCLUSIONES GENERALES.	111
7. BIBLIOGRAFÍA.	113
8. APÉNDICE.	135



1. INTRODUCCIÓN.

Tras la segunda guerra mundial, la industrialización progresiva de los medicamentos originó una revolución terapéutica. Paralelamente, el desarrollo de la tecnología permitió acumular grandes cantidades de datos sanitarios procedentes tanto de registros médicos como del uso administrativo para facturación farmacéutica, u otros fines fiscales relacionados con la prestación de salud al paciente. Desde entonces, esta información se ha ido almacenando en bases de datos actualizadas a diario, incrementándose exponencialmente en número y tamaño durante las últimas décadas.

La utilización secundaria de estas bases de datos presenta numerosas ventajas prácticas y metodológicas [1, 2], ofreciendo la posibilidad de un acceso fácil y rápido a grandes cantidades de datos informatizados. Esta disponibilidad de datos ha creado grandes expectativas sobre su potencial aplicación en farmacoepidemiología como fuentes de datos para investigación. Contribuye a esto la percepción de que los ensayos clínicos no siempre son útiles para la evaluación de las terapias en la práctica del mundo real, en particular los que proporcionan datos de seguridad limitados [3].

No obstante, frente a este entusiasmo, no deben descuidarse sus limitaciones [4]. El acceso a toda esta información y su utilización para estudios epidemiológicos observacionales puede resultar engañosamente simple. A las dificultades propias de toda investigación observacional hay que sumarle algunas específicas de las bases de datos secundarias, y también desafíos sociales relacionados con la privacidad y seguridad [5, 6]. Frente a las prometedoras perspectivas de futuro en torno a la utilización de bases de datos secundarias, todas las precauciones son pocas para llevar a cabo una investigación de calidad. La utilización de estas bases de datos en investigación debe abordarse con cautela, siendo numerosos los requisitos de calidad y desafíos metodológicos que deben ser considerados [7].

Si bien las preguntas específicas de investigación se responden idealmente a través de estudios con datos primarios hechos a medida (*ad hoc*) que proporcionan datos más precisos y actualizados, el diseño de un proyecto completamente nuevo puede no representar una estrategia viable. Este hecho se añade al uso creciente de grandes conjuntos de datos en algunos países, publicándose con frecuencia estudios que ponen de relieve las dificultades derivadas de su explotación, y la consiguiente identificación de resultados sesgados [3, 4, 8-12]. Junto con los sesgos de selección y medición, la confusión es una de las más importantes fuentes de error en los estudios observacionales, y cualquier asociación observada en estos estudios está confundida hasta cierto punto.

Por otra parte, el metanálisis representa una de las herramientas más valiosas para evaluar los efectos de los medicamentos, ya que puede conducir a la mejor evidencia posible en epidemiología [13]. En consecuencia, su uso para tomar decisiones clínicas y regulatorias relevantes sobre la seguridad y eficacia de los medicamentos está aumentando notablemente [14]. La existencia de heterogeneidad en un metanálisis dado es una característica que debe ser cuidadosamente descrita analizando los posibles factores responsables de su generación [15].

No obstante, el uso extensivo del metanálisis ha suscitado cierta controversia [16]. Para obtener conclusiones correctas que puedan derivar en decisiones clínicas y regulatorias importantes, se han sugerido mejoras en el procedimiento de metanálisis, incluyendo un estudio más detallado de la heterogeneidad y el uso de modelos de efectos aleatorios en lugar de modelos de efectos fijos, así como la implementación de técnicas de sensibilidad para comprobar cuán dependientes del modelo seleccionado son los resultados [16, 17]. Mientras que la estratificación por el diseño del estudio se utiliza frecuentemente para evaluar la heterogeneidad en los metanálisis de los estudios observacionales [18, 19], la estratificación por el tipo de fuente de datos (según sea el origen de la información primario o secundario) rara vez se utiliza con este fin [20].

La finalidad de este trabajo fue estudiar las repercusiones de la utilización de bases de datos secundarias en estudios de farmacoepidemiología. A continuación se exponen, con carácter introductorio, los principales mecanismos conocidos generadores de sesgos en estudios epidemiológicos sobre efectos de medicamentos, así como las estrategias disponibles para el control, parcial o total, de dichos sesgos.



1.1. RETOS Y DIFICULTADES EN EL USO DE FUENTES DE DATOS EN FARMACOEPIDEMIOLOGÍA.

1.1.1. Datos primarios y secundarios.

En Salud Pública, al hablar de **datos primarios** (*ad hoc*) nos referimos a la información recogida por el investigador directamente de la realidad a través de sus propios instrumentos, como encuestas, grupos de estudio u observación. Los datos primarios se elaboran durante el tiempo en que se realiza la investigación.

Por otro lado, los **datos secundarios** son los datos recogidos para otros fines. Los investigadores reutilizan la información y le dan un nuevo propósito porque resulta más fácil y menos costoso que recogerla por primera vez. Sin embargo, rara vez suele ser tan útil y precisa como la información primaria. Aunque algunos de ellos son registros que proceden también de un contacto con la práctica, ya han sido recogidos y muchas veces procesados con anterioridad. Dado que fueron generados con un propósito diferente al de la investigación de interés, es posible que surjan problemas de correspondencia entre los datos requeridos y los datos disponibles. Los investigadores no tienen conocimiento acerca de cómo se recogieron los datos, ni tienen control sobre éstos, de modo que pueden no estar actualizados, y en consecuencia no podrán ser idóneamente utilizados en el estudio de una asociación.

Desde el punto de vista del investigador, podemos clasificar los datos secundarios en internos o externos según su procedencia.

- **Datos secundarios internos:** Se considera información interna aquella que se recoge desde la propia empresa, entidad o sistema sanitario donde trabaja el investigador. Los datos internos de los sistemas de información sanitaria, provenientes del Sistema Nacional de Salud, consejerías de salud autonómicas o servicios de atención primaria u hospitalaria, proporcionan información básica sobre los recursos disponibles. Estas bases informatizadas de datos internos recogen información individualizada sobre el uso de recursos sanitarios, como prescripción de medicamentos, visitas ambulatorias, ingresos médicos con diagnósticos al alta, entre otros. Algunos ejemplos son registros como el Conjunto Mínimo de Bases de Datos (CMBD) y la encuesta de morbilidad hospitalaria.

- **Datos secundarios externos:** Los datos secundarios que se publican a nivel externo provienen de fuentes diversas ajenas al ámbito de trabajo del investigador. Algunos ejemplos son los datos generales de publicaciones, como libros, revistas y boletines; fuentes gubernamentales (por ejemplo, el registro de mortalidad del Instituto Nacional de Estadística o el sistema de notificación de accidentes de tráfico de la Dirección General de Tráfico); organizaciones sociales y/o científicas sin ánimo de lucro (por ejemplo, el proyecto BIFAP (Base de Datos para la Investigación Farmacoepidemiológica en Atención Primaria)); o entidades privadas que ofertan bases de datos.

Debe tenerse en cuenta que en esta clasificación la referencia es el propio investigador, lo que le confiere un cierto grado de relatividad (una fuente interna para un investigador puede ser externa para otro, y viceversa).

Desde el punto de vista metodológico se utilizan dos tipos principales de bases de datos secundarias en investigación observacional: las bases de datos administrativas y los registros médicos [21].

- **Bases de datos administrativas:** Registran de forma automatizada consultas, hospitalizaciones y prescripciones farmacológicas, y se recogen de forma rutinaria fundamentalmente para fines de contabilidad. Contienen información general sobre un gran número de pacientes.

La utilidad de estos datos para fines de investigación es un subproducto de la finalidad administrativa primaria de los datos. Los registros de facturación de los sistemas sanitarios se consideran la medida más exacta de la exposición al fármaco, proporcionando una gran cantidad de información, entre la que se incluye la fecha de prescripción, la farmacia de dispensación, el nombre del medicamento, la cantidad, dosis y duración del tratamiento, y por ello suele considerarse el estándar de oro para medir la exposición.

- **Bases de datos de registros médicos:** Se obtienen a partir de las historias clínicas electrónicas actualizadas por los profesionales de la salud. En la historia clínica electrónica, tanto en el ámbito hospitalario como extrahospitalario, se introduce información detallada sobre la medicación prescrita; sin embargo, los registros médicos no reflejan si el paciente obtiene el medicamento de la farmacia ni el grado de exposición al mismo. Estos registros pueden considerarse una fuente de captura de medicamentos de venta libre (*over the counter*), en general no cubiertos por ningún registro de dispensación [22]. A diferencia de las bases administrativas, las bases de datos de registros médicos en general incluyen información adicional de variables, tales como consumo de alcohol, tabaquismo, masa corporal o número de partos/abortos.

Los registros médicos electrónicos se perfilan como una prometedora fuente de datos para la investigación clínica, aunque tienen sus propios retos [23, 24], incluyendo el desarrollo y la armonización de estándares para el registro de datos. La tabla 1 compara las bases de datos de registros médicos con las bases administrativas en base a algunos de los factores que son importantes en los estudios farmacoepidemiológicos.

La creciente tendencia al registro de información está haciendo que el uso de grandes conjuntos de datos sea cada vez más común en la investigación farmacoepidemiológica, especialmente en Canadá y EE.UU., donde se han utilizado bases de datos administrativas como Medicare, Medicaid, entre otras. En Europa esto ha ocurrido con la GPRD (*General Practice Research Database*) en el Reino Unido, y en otras bases de datos de los países nórdicos, los Países Bajos, Alemania e Italia. Esta tendencia es alimentada por el creciente reconocimiento de que los ensayos clínicos aleatorizados, aunque esenciales, no pueden responder a todas las cuestiones importantes sobre eficacia/seguridad de los medicamentos en la práctica clínica diaria [3].

El potencial de investigación que ofrece la utilización de bases de datos secundarias preexistentes en farmacoepidemiología es indudable. Metodológicamente se caracterizan por

Tabla 1. Comparación entre bases de datos de registros médicos y bases administrativas de uso en estudios farmacoepidemiológicos.

Ítem	Registros médicos	Bases administrativas
Origen de los datos	Historias clínicas electrónicas mantenidas por los profesionales de la salud	Subproducto de transacciones económicas relacionadas con la prestación sanitaria
Evento que provocó el registro de prescripciones de pacientes ambulatorios	Prescripción	Dispensación
Prescripciones de múltiples prescriptores	No habitual	Sí
Diagnósticos hospitalarios	Habitual	Sí
Registro de potenciales variables de confusión	Habitual	No
Utilidad	Proporcionan datos más amplios para el ajuste de la comorbilidad en estudios de investigación que pueden ser particularmente susceptibles al sesgo de selección	Adecuadas para cuestiones de investigación que implican resultados poco frecuentes cuando contienen un gran número de pacientes
Ejemplos	<ul style="list-style-type: none"> - German MediPlus - Italian General Practice Research Database - UK MediPlus - United Kingdom General Practice Research Database (GPRD) 	<ul style="list-style-type: none"> - Group Health Cooperative - Health Maintenance Organization Research Network (HMORN) - Medicaid - Medicare - Medicines Monitoring Unit (MEMO) - United Health Group

su elevada representatividad, el bajo riesgo de selección por no participación, un tamaño de muestra grande que se puede seguir durante largos períodos de tiempo con pocas pérdidas y de forma continuada, evitando sesgos como el de causalidad inversa [25]; y una disponibilidad de información precisa de la exposición a fármacos prescritos, que evita los sesgos de memoria y del entrevistador.

El uso de bases de datos secundarias presenta ventajas especialmente de orden práctico, permitiendo el estudio de los efectos de los fármacos de forma rápida y a un coste bajo, en las condiciones de la práctica clínica habitual, y en poblaciones más heterogéneas que las de los ensayos clínicos.

Aunque el potencial de las bases automatizadas para estudios epidemiológicos es ampliamente reconocido, entraña numerosas dificultades. Estos estudios son casi siempre complicados, y por ello requieren una inversión sustancial de tiempo y reflexión [26]. La tabla 2 recoge las principales dificultades y consiguientes sesgos potenciales en estudios de bases de datos secundarias [3].

Tabla 2. Dificultades y sesgos potenciales en estudios con bases de datos secundarias en farmacoepidemiología.

Etapa de la investigación	Condición que debe cumplir	Dificultades	Sesgos potenciales
Obtención de los datos	Datos completos y precisos para la población definida	Falta de familiaridad de los investigadores con la base de datos Calidad inadecuada de los datos Errores de vinculación con otras bases de procedencia distinta Datos anonimizados	-
Reconocimiento de la población diana	Grupos de exposición debidamente identificados	Asegurar la elegibilidad de los individuos Cambios en la asignación a los grupos de exposición Cuestiones de muestreo	Sesgo del usuario prevalente Sesgo de tiempo inmortal
Medición de la exposición	Definición consistente con el mecanismo que lo relaciona con el resultado	Disponibilidad insuficiente de datos para la definición Elaboración insuficiente de los datos existentes	Sesgo de mala clasificación de la exposición
Medición del resultado	Definición con valor predictivo positivo y sensibilidad adecuados	Condición clínica no diagnosticada de forma fiable	Sesgo de mala clasificación del resultado
Registro de factores de confusión	Inclusión de factores de confusión importantes	Variables no medidas Variables infrautilizadas Dificultad para medir las variables relacionadas con el consumo de fármacos y la salud Tamaño de efecto pequeño	Confusión Efecto del usuario (cumplidor del tratamiento) sano
Garantía de confidencialidad	Protección de datos	Aplicación de estándares de privacidad	-

1.1.2. El error sistemático en farmacoepidemiología.

La farmacoepidemiología presenta una serie de dificultades metodológicas específicas condicionadas por el tipo de exposición y la disponibilidad de datos. Junto con los sesgos de selección y medición, la confusión es una de las más importantes fuentes de error en los estudios observacionales, y cualquier asociación observada en este tipo de estudios está confundida hasta cierto punto [27]. La tabla 3 sintetiza la amplia variedad de sesgos potenciales a los que puede estar sometido cualquier estudio observacional de medicamentos [28], y la tabla 4 ejemplifica los más habituales.

Tabla 3. Clasificación y nomenclatura de los potenciales errores sistemáticos en farmacoepidemiología.

Confusión	<ul style="list-style-type: none"> • Confusión por indicación de tratamiento <ul style="list-style-type: none"> ◦ Confusión por gravedad de la enfermedad • Confusión por contraindicación de tratamiento • Confusión tiempo-dependiente • Confusión por variables no medidas <ul style="list-style-type: none"> ◦ Confusión por fragilidad ◦ Efecto del usuario (cumplidor de tratamiento) sano. Sesgo de adhesión • Confusión residual • Elección inadecuada de las variables: Sobreajuste <ul style="list-style-type: none"> ◦ Sesgo debido al control de una variable intermedia ◦ Sesgo debido al control de una variable instrumental (<i>Z-bias</i>) ◦ Sesgo debido al control de un factor de colisión (<i>M-bias</i>)
Sesgo de selección	<ul style="list-style-type: none"> • Definición inapropiada de la población elegible <ul style="list-style-type: none"> ◦ Riesgos competitivos ◦ Sesgo por acceso diferencial a los servicios de salud ◦ Muestreo con sesgo de duración ◦ Depleción de susceptibles. Efecto de la cohorte superviviente. Sesgo de los riesgos cruzados. Sesgo del usuario prevalente/superviviente ◦ Sesgo de Neyman. Sesgo de incidencia-prevalencia. Sesgo de supervivencia selectiva ◦ Sesgo de selección del tratamiento superviviente. Sesgo de tiempo inmortal ◦ Sesgo de Berkson ◦ Sesgo de exclusión/inclusión ◦ Sobrepajamiento • Falta de precisión en el muestreo <ul style="list-style-type: none"> ◦ Sesgo de muestreo no aleatorio • Procedimientos de diagnóstico desiguales en la población diana <ul style="list-style-type: none"> ◦ Sesgo de detección • Durante el curso del estudio <ul style="list-style-type: none"> ◦ Sesgo por pérdida/retirada diferencial durante el seguimiento. Censura informativa ◦ Falta de datos en el análisis multivariable ◦ Sesgo de no respuesta <ul style="list-style-type: none"> ▪ Sesgo del voluntario sano • Sesgo protopático
Sesgo de medición (sesgo de información)	<ul style="list-style-type: none"> • Mala clasificación (diferencial frente a no diferencial) <ul style="list-style-type: none"> ◦ Sesgo de mala clasificación de la exposición <ul style="list-style-type: none"> ▪ Sesgo de tiempo inmortal. Sesgo de tiempo de espera ▪ Sesgo de tiempo no medible ◦ Sesgo de mala clasificación del resultado ◦ Sesgo de detección ◦ Sesgo del observador ◦ Sesgo de memoria ◦ Sesgo de notificación/infranotificación • Regresión a la media <ul style="list-style-type: none"> ◦ Sesgo por dilución de la regresión • Ambigüedad temporal

Puede consultarse una breve descripción de alguno de estos sesgos, y el diseño de estudio afectado en cada caso, en [28].

Tabla 4. Ejemplos prácticos de los principales sesgos en farmacoepidemiología.

Sesgo	Ejemplo
Confusión por indicación	Si se estudia el riesgo de mortalidad asociado a la vacunación antigripal, puede encontrarse que los vacunados tengan un mayor riesgo de muerte, debido a que la vacuna está indicada para personas con factores de riesgo de mortalidad, como la edad avanzada o el cáncer, entre otras.
Confusión tiempo-dependiente	Si se estudia la relación entre el uso de aspirina con el riesgo de infarto de miocardio (IAM) y muerte cardíaca, el IAM previo es un factor de confusión para el efecto del uso de aspirina sobre el riesgo de muerte cardíaca, porque el IAM previo es una causa del (posterior) uso de aspirina, y es también un factor de riesgo causal para la (posterior) muerte cardíaca. Sin embargo, el uso de aspirina (anterior) también impide causalmente el IAM previo. Por lo tanto, el IAM previo actúa al mismo tiempo como factor de confusión (causa el uso de aspirina) y etapa intermedia (estando afectado por el uso de aspirina), y por lo tanto es un factor de confusión tiempo-dependiente afectado por el tratamiento previo.
Confusión por variables no medidas	Variables habitualmente no disponibles incluyen información sobre comportamiento: tabaquismo, consumo de alcohol, nutrición y hábitos alimentarios, ejercicio, abuso de sustancias, medidas de cumplimiento terapéutico y automedicación; información de laboratorio: análisis de sangre, pruebas de función pulmonar, y otras pruebas de funcionalidad; información de la exposición: productos de venta libre, historia de inmunización, uso de otros medicamentos; y marcadores de gravedad de la enfermedad, entre otros muchos.
Censura informativa	En un país con un sistema sanitario no universal y más de un seguro de salud donde poder elegir asistencia de pago, se realiza un estudio en una población homogénea para evaluar el consumo de un medicamento. Es más probable que los pacientes más jóvenes, o bien individuos sanos con problemas económicos, se den de baja por precisar menos asistencia médica y consumo de medicamentos, permaneciendo los pacientes crónicos. De esta manera pueden asociarse al consumo de ese fármaco condiciones propias de diabéticos, polimedcados, o edad avanzada, entre otros.
Depleción de susceptibles	Si se estudia la relación entre el riesgo de hemorragias gastrointestinales y el consumo de antiinflamatorios no esteroideos (AINEs), puede observarse un riesgo mayor entre los que consumen un AINE por vez primera, debido a un proceso de autoselección de los expuestos a los antiinflamatorios: los pacientes que tomaron un AINE y lo toleraron mantendrán el fármaco, mientras que los que padecieron algún síntoma gastrointestinal lo habrán cambiado por otro.
Mala clasificación	<i>No diferencial:</i> La clasificación errónea de la exposición para un fármaco utilizando los datos de un registro de prescripción es igualmente probable con independencia de la variable de resultado. <i>Diferencial:</i> Si para recibir la medicación un paciente ha de estar incluido en el sistema sanitario, aumenta la probabilidad del registro de un diagnóstico, de modo que los expuestos a un fármaco tienen una menor probabilidad de error de clasificación del resultado. Los no expuestos son mucho más propensos a ser mal clasificados como no enfermos, lo cual es un artefacto de no incluirse en el sistema sanitario.
Sesgo protopático	Si se estudia la relación entre el consumo de AINEs y el riesgo de cáncer puede encontrarse una asociación positiva debido a un efecto de causalidad inversa. El uso de analgésicos en respuesta al dolor causado por un tumor no diagnosticado podría llevar a la conclusión errónea de que el analgésico causó el tumor.
Sesgo de tiempo inmortal	Si se estudia la efectividad de los B-bloqueantes en la prevención de un segundo infarto comparando dos cohortes, y se establece como criterio de inclusión haber retirado el fármaco de la farmacia durante un periodo de tiempo específico después del alta, ese periodo de espera se considera inmortal: cada individuo del grupo expuesto tiene necesariamente que sobrevivir hasta que se cumpla el criterio de inclusión (el expuesto que fallece durante el periodo después del alta se excluye del estudio); mientras que si muere antes de iniciar el tratamiento, entonces se encuadra en el grupo de comparación no expuesto (el que fallece durante ese periodo sí es contabilizado como evento).

A continuación se explican los sesgos más influyentes, bien por su frecuencia de presentación o por su impacto en los resultados, de los estudios farmacoepidemiológicos que se basan, parcial o totalmente, en datos secundarios.

1.1.2.1. Confusión.

Se define clásicamente como un fenómeno que distorsiona la relación entre la exposición y el resultado [29]: un factor de confusión C se asocia con la exposición E , contribuye a parte o la totalidad del resultado (efecto) R observado, y no interviene en la vía causal. Cuando el riesgo de R es diferente entre los individuos con y sin C ; el sesgo se produce si la distribución de C es diferente entre los grupos de exposición. La distorsión puede sobrestimar, subestimar, o revertir el efecto de la exposición.

1.1.2.1.1. Confusión por indicación de tratamiento.

Es una distorsión en la asociación entre la exposición y el resultado que se produce cuando la elección de un fármaco se ve afectada por las características basales específicas de los pacientes, como un peor estado de la enfermedad al inicio del estudio (*confusión por gravedad de la enfermedad*); en consecuencia, un fármaco se prescribe preferentemente a un determinado tipo de pacientes. Suponiendo que la gravedad de la enfermedad sea un factor de riesgo para el resultado de interés, la confusión se produce si los médicos prescriben el tratamiento en función de la eficacia percibida entre diferentes opciones de fármacos a pacientes con diferentes niveles de gravedad [30, 31]. Así, la estimación del efecto real del tratamiento podría confundirse con el efecto de un estado de enfermedad más severo, que llevó a prescribir a los pacientes un tratamiento dado [32]. Esto es, la prescripción de un fármaco (E) se relaciona con una indicación (C : conjunto de factores por los que un paciente recibe un tratamiento concreto), que a su vez es un factor de riesgo de la enfermedad (R) (Fig. 1A).

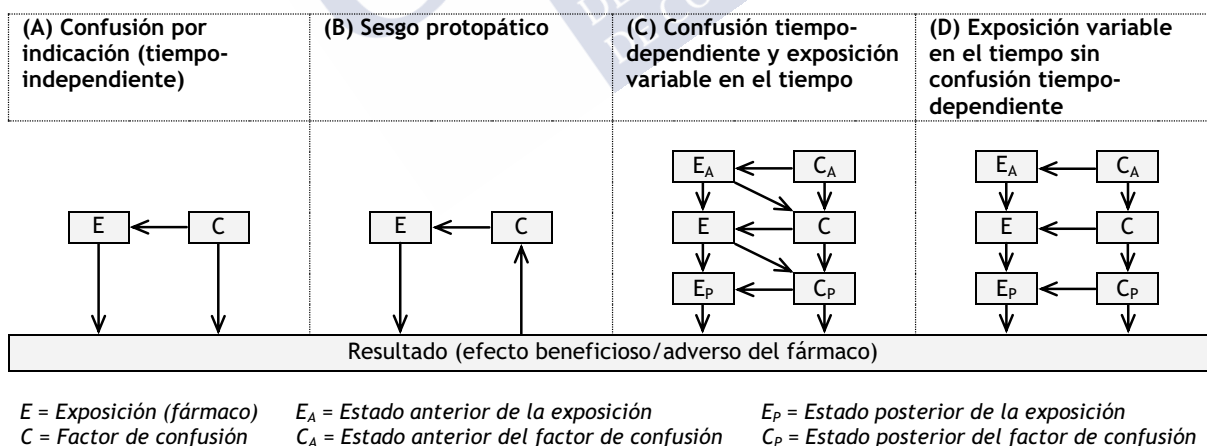


Figura 1.

Confusión y sesgo protopático. (A) La indicación (condición clínica) es un factor de confusión porque se correlaciona con el tratamiento, es un factor de riesgo para el resultado, y no constituye un paso intermedio en la vía causal. (B) El resultado condiciona la condición clínica (mientras que en la confusión por indicación la condición clínica condiciona el resultado). (C) El efecto del fármaco se ve afectado por un factor de confusión tiempo-dependiente, el cual se ve afectado por la exposición anterior al fármaco y afecta a la exposición posterior y el resultado. (D) El estado de exposición anterior no afecta al valor presente del factor de confusión.

En los estudios observacionales, los factores pronósticos y/o de riesgo, conocidos y desconocidos, influyen en la selección del tratamiento, y en consecuencia pueden repartirse desigualmente entre los grupos. A diferencia de los estudios que utilizan fuentes primarias de datos con recogida prospectiva, no siempre es posible ajustar por todos los determinantes conocidos de la selección del tratamiento al emplear datos secundarios, debido a que no recogen factores importantes de la decisión clínica [33].

1.1.2.1.2. Confusión tiempo-dependiente.

Se define como una alteración en la asociación entre E y R debido a una variable C que puede variar con el tiempo y actúa como: (i) factor de confusión entre la exposición actual y el resultado, e (ii) intermediaria entre la exposición anterior y la exposición actual. Es decir, un factor de confusión tiempo-dependiente es una variable predecible por la exposición anterior, y un predictor de la exposición posterior [27]. En farmacoepidemiología, los efectos de los medicamentos dependen del tiempo, viéndose afectados por factores de confusión tiempo-dependientes (Fig. 1C). Las conclusiones extraídas de los estudios que no tienen en cuenta los factores de confusión tiempo-dependientes pueden resultar espurias. La confusión por indicación puede adoptar la forma de confusión tiempo-dependiente [22].

1.1.2.1.3. Confusión por variables no medidas.

Incluso las fuentes de datos más completas pueden dejar de incluir información sobre posibles factores de confusión, que permanecerán sin medir, y por lo tanto sin controlar en un estudio. Mucha información se recoge por causas diferentes a la hipótesis de la investigación, por lo que rara vez se dispone de información suficiente sobre todos los factores de confusión relevantes [34]. Los factores de confusión no medidos pueden ser medibles o no medibles (desconocidos o difíciles de medir).

La *confusión residual* se refiere a la confusión que no se ha controlado completamente, de manera que todavía permanece el efecto de confusión de algunos factores en el efecto final observado. No tener en cuenta todos los factores de confusión puede deberse a variables no medidas o errores de medición en las variables registradas [35]. Por otro lado, la categorización de un factor de confusión medido en una escala continua también puede resultar en un control inadecuado de la confusión, sobre todo cuando algunas categorías presentan un escaso número de observaciones [36].

Confusión por fragilidad: La fragilidad es una condición cercana a la muerte caracterizada por un mal estado de salud general que no se relaciona con una patología específica [25]. A diferencia de los ensayos clínicos, las poblaciones de estudio construidas a partir de grandes bases de datos sanitarios incluyen pacientes frágiles. La fragilidad puede reducir el efecto de un tratamiento si se trata la condición clínica principal del paciente y no se inician terapias útiles para condiciones secundarias. Debido a que la fragilidad es difícil de medir, y un fuerte factor de riesgo para resultados desfavorables, dará lugar a confusión no medida y residual [37]. Al comparar un grupo tratado con otro no tratado, paradójicamente, este fenómeno puede hacer que el fármaco parezca mejor [38].

Efecto del usuario sano: Un mayor nivel de educación y un comportamiento de búsqueda de salud son características que se asocian con el acceso a los recursos sanitarios y al inicio de la terapia, particularmente si el medicamento es un tratamiento preventivo o para

una enfermedad crónica asintomática. Además, los pacientes que cumplen el tratamiento durante períodos prolongados de tiempo tienden a ser más saludables que los que interrumpen la medicación (*efecto del cumplidor del tratamiento sano*) [25, 39, 40]. El ajuste inadecuado de este tipo de comportamientos (habitualmente no registrados en bases de datos secundarias) puede infraestimar la asociación entre un tratamiento y el riesgo de acontecimientos adversos. De forma parecida a la confusión por fragilidad, el sesgo de selección de usuarios sanos tiende a ser más acusado para la mortalidad por cualquier causa [41].

1.1.2.2. Sesgo de selección.

Es una distorsión en la asociación entre E y R que surge de los errores sistemáticos debidos a la forma en que se ha elegido la muestra del estudio, de modo que ésta no será representativa de la población diana a la que se extrapolarán los resultados.

En investigación observacional, la elección del tratamiento se basa en las características de los pacientes y la enfermedad, o la preferencia personal del médico. El sesgo de selección se relaciona con el fenómeno de confusión a través del sesgo de prescripción/indicación, o sesgo de canalización (*channeling bias*), que se produce cuando el fármaco se prescribe preferentemente a pacientes con alto riesgo de padecer el resultado de interés [42, 43].

1.1.2.2.1. Sesgo protopático.

Se produce cuando el fármaco se asocia, bien por indicación o por contraindicación, a los estados subclínicos de la enfermedad, de modo que el síntoma inicial de la enfermedad de estudio es la causa del consumo del medicamento [44]. A diferencia de la confusión por indicación, donde la condición clínica/indicación (C) que determina la selección del medicamento es un factor de riesgo de la enfermedad, en el sesgo protopático esa condición clínica es una consecuencia de la enfermedad (R) (Fig. 1B). El sesgo protopático es un tipo de causalidad inversa. La hipótesis habitual es que el fármaco aumenta el riesgo de enfermedad, pero en presencia de sesgo protopático ocurre lo contrario: los primeros síntomas de la enfermedad son los que aumentan la prescripción del fármaco [45].

1.1.2.2.2. Pérdidas durante el seguimiento (censuras).

En el contexto del análisis de supervivencia, la censura se produce cuando las observaciones no se siguen el tiempo suficiente para observar el resultado de interés. El investigador suele definir las condiciones para establecer el final de la observación. Cuando la censura se debe a razones que no están bajo el control del investigador, se habla de censura aleatoria. El mecanismo de censura puede relacionarse con la exposición o el resultado (la pérdida es diferencial según el grupo de exposición o el riesgo de desarrollar el resultado), se dice entonces que la censura es de tipo informativo. Si los individuos censurados constituyen una submuestra sesgada (por ejemplo, tienen riesgos sistemáticamente mayores o menores de observar el resultado que los individuos no censurados), la censura informativa puede dar lugar a estimaciones sesgadas de la asociación entre la exposición y el resultado [46]. En ocasiones la censura se debe a riesgos competitivos; cuando la tasa del evento competitivo es mayor en uno de los grupos de exposición, la estimación del efecto final estará sesgada.

1.1.2.2.3. Depleción de susceptibles (sesgo del usuario prevalente).

Este fenómeno es común en los estudios que incluyen usuarios prevalentes en lugar de incidentes, favoreciendo la aparición de sesgos relacionados con la insuficiente

verificación de los efectos adversos que ocurren al inicio del tratamiento. Los usuarios prevalentes (“supervivientes” del primer período de tratamiento) pueden no tener el mismo riesgo de un evento adverso que los incidentes (nuevos usuarios): los que toleran el fármaco continúan utilizándolo, y los que no (susceptibles al evento adverso), han dejado de usarlo [47]. Dado que la mayoría del tiempo-persona proviene de los prevalentes, el efecto puede ser infraestimado si se consideran conjuntamente expuestos incidentes y prevalentes.

La inclusión de individuos prevalentes conlleva un sesgo de selección por dos mecanismos: (i) la pérdida de individuos susceptibles, y (ii) el ajuste por factores de confusión que están influidos por exposiciones pasadas. La inclusión de usuarios prevalentes en un estudio complica el control de posibles factores de confusión debido a que estos factores a menudo están afectados por el propio tratamiento. En este caso, los investigadores pueden ajustar por las covariables, cometiendo así el error de ajustar por factores de la vía causal, o en caso de no hacerlo, introducir confusión [48-50].

1.1.2.2.4. Falta de datos.

Una de las grandes dificultades en el análisis de los datos de observación es la ausencia de datos [23]. En su forma más extrema, las observaciones pueden carecer por completo de una variable importante. En el análisis multivariante, como los modelos de regresión, la mayoría de los paquetes de software estadístico eliminan las observaciones cuando se ha perdido alguno de los valores de una variable incluida en el modelo. En consecuencia, las observaciones con valores perdidos pueden conducir a una pérdida sustancial del tamaño de la muestra, aunque el grado de ausencia de datos sea bajo para cualquiera de las variables [51]. Si esta falta de información está asociada con una característica importante (por ejemplo, gravedad, fragilidad) se produce un sesgo de selección.

1.1.2.3. Sesgo de medición.

Es una distorsión en la asociación entre la exposición y el resultado que surge de los errores sistemáticos por el modo en que se miden las variables de interés para los grupos de comparación [27]. En un análisis retrospectivo de una base de datos, la información con respecto a las verdaderas exposiciones, resultados y otras variables (constructo no observado, E, R, C), se registra en forma de indicadores (medidas observadas, E_o, R_o, C_o). Cuando estas medidas observadas no reflejan con exactitud la realidad, se producen los errores de medición (eE, eR, eC) [52].

1.1.2.3.1. Sesgo de clasificación.

El sesgo de mala clasificación en farmacoepidemiología se clasifica atendiendo a: (i) la relación de dependencia entre los errores de medición en la exposición y el resultado (independiente/dependiente), y (ii) la influencia de la clasificación errónea de la exposición –o el resultado– por el verdadero valor del resultado –o la exposición– (diferencial/no diferencial) (Fig. 2) [51, 53].

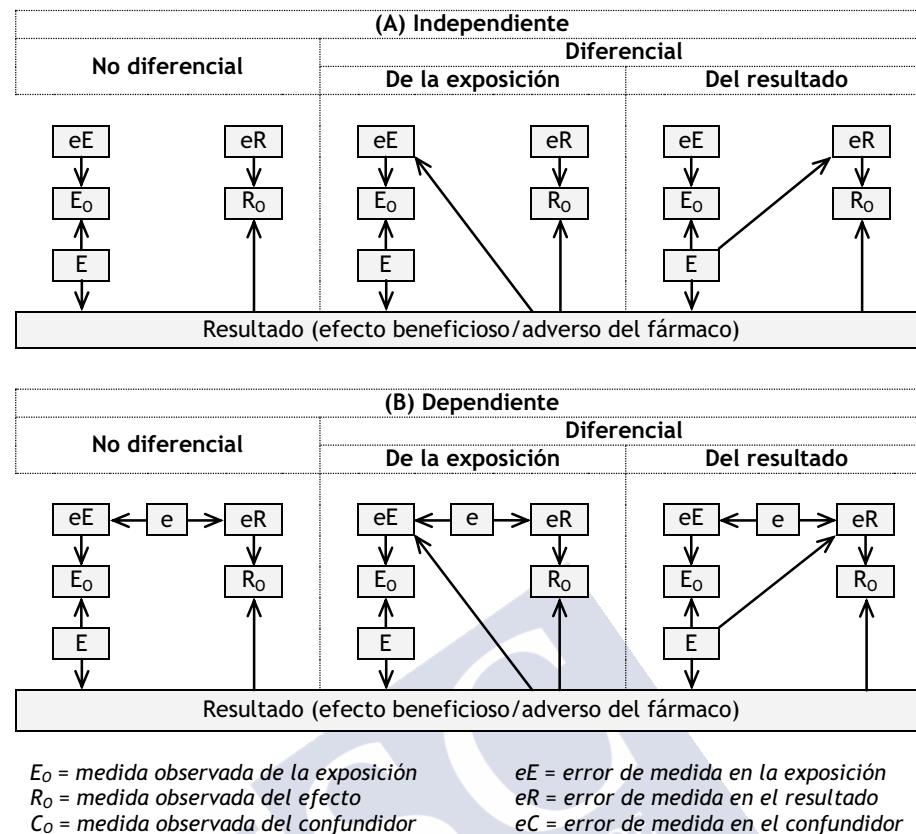


Figura 2.

Sesgo de medición por mala clasificación. (A) *Independiente no diferencial*: se refiere a la independencia entre los errores de medición en la exposición y en el resultado, no viéndose afectada la mala clasificación de la exposición por el verdadero valor de resultado (y del mismo modo en la mala clasificación del resultado). *Independiente diferencial*: se refiere a la independencia entre los errores de medición en la exposición y en el resultado; sin embargo, el verdadero valor del resultado afecta a la clasificación de la exposición (y de manera similar en la mala clasificación del resultado). (B) *Dependiente*: ocurre cuando los errores de medición en la exposición y en el resultado son dependientes entre sí (por ejemplo, comparten el mismo mecanismo de obtención de información), puede ser *no diferencial* y *diferencial* (de la exposición o del resultado).

La mala clasificación diferencial puede exagerar o infraestimar las estimaciones de la asociación entre la exposición y el resultado. La mala clasificación no diferencial conduce a estimaciones hacia la hipótesis nula en variables dicotómicas; sin embargo, en exposiciones de múltiples categorías, si los errores de medición de la exposición y el resultado son dependientes, o cuando la mala clasificación depende de la clasificación errónea de otras variables (por ejemplo, factores de confusión), las estimaciones pueden alejarse del nulo. Además, la categorización de variables continuas o la agrupación en un menor número de categorías puede revertir la mala clasificación no diferencial en diferencial [54].

El sesgo de clasificación también se puede categorizar, según la dirección de la clasificación errónea, como unidireccional o bidireccional. Por ejemplo, los pacientes se pueden identificar como expuestos al fármaco cuando realmente no lo están, y/o no expuestos cuando sí lo están. Igualmente, puede existir una probabilidad de que los casos aparezcan como controles, y los controles aparecen como casos. Es importante considerar la dirección de las posibles fuentes de error de clasificación y cómo podría influir en el rechazo de la hipótesis nula antes de realizar un estudio [22].

1.1.2.3.1.1. Sesgo de mala clasificación de la exposición.

En estudios farmacoepidemiológicos a partir de bases de datos retrospectivas, las medidas de la exposición al fármaco no son necesariamente un reflejo exacto de su consumo real [55]. Las fuentes de sesgo más habituales, además de una medición deficiente, son el incumplimiento del tratamiento, la falta de control sobre el consumo de medicamentos de venta libre, y el uso inadecuado de las ventanas de tiempo para la clasificación de la exposición. La exposición intermitente es más habitual en la práctica médica real de lo que generalmente se reconoce en los estudios observacionales [49]. El sesgo se puede clasificar según la mala clasificación afecte a los expuestos o a los no expuestos.

- *Dirección de los no expuestos:* En los registros ambulatorios existe una mayor probabilidad para la mala clasificación en la dirección de los no expuestos, dadas las diversas vías a través de las cuales los usuarios pueden recibir sus medicamentos al margen de la prescripción dentro del Sistema Nacional de Salud. Algunas de estas fuentes incluyen las hospitalizaciones o la conducta fraudulenta. Además, en los datos administrativos de recetas, el sesgo de clasificación está presente en la medición de la exposición a fármacos de venta libre [22].

- *Dirección de los expuestos:* En las bases de datos la prescripción de un medicamento no significa necesariamente que el paciente haya adquirido el producto y lo haya consumido según las instrucciones. El paciente puede no haber estado realmente expuesto a una cantidad variable del fármaco, pero clasificarse como expuesto [56].

1.1.2.3.1.2. Sesgo de mala clasificación del resultado.

Suele observarse en estudios retrospectivos, generalmente como resultado de la utilización de criterios de diagnóstico no uniformes, o datos de diagnóstico con valores desconocidos. Algunos factores como la ambigüedad clínica en el diagnóstico o la codificación utilizada pueden producir errores de clasificación [57]. Se puede producir un error sistemático en la clasificación de los resultados si el investigador no tiene en cuenta las actualizaciones en los códigos; por ejemplo, los derivados de la transición de la Clasificación Internacional de Enfermedades (CIE-9) a los sistemas de codificación CIE-10 [22].

1.1.2.4. Sesgo relacionado con el tiempo.

Es una forma de sesgo de mala clasificación diferencial debida a una consideración inadecuada del tiempo de seguimiento y el estado de la exposición en las fases de diseño o análisis del estudio. Puede producir resultados sesgados en estudios observacionales a partir de base de datos, tanto de cohortes como de casos y controles [58].

En los estudios sobre efectos adversos, la ventana de riesgo de exposición constituye el número de días de exposición asignados a cada prescripción. El diseño de esta ventana de tiempo influye en la validez de las estimaciones de riesgo; el diseño ideal contempla ventanas de prescripción que cubren solo el período de exceso de riesgo potencial (durante el cual el fármaco puede producir el efecto) [59].

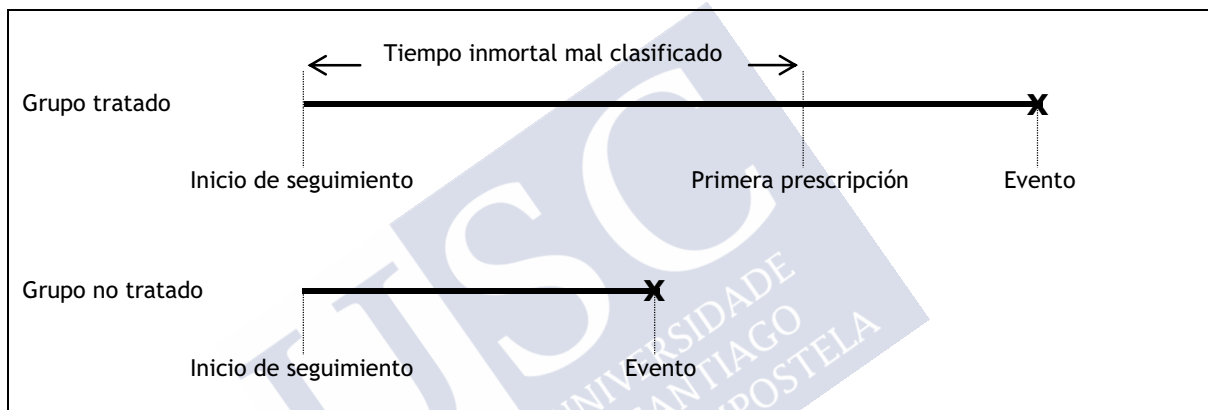
1.1.2.4.1. Sesgo de tiempo inmortal.

El tiempo inmortal es un lapso en el seguimiento de una cohorte en el que, por el tipo de diseño, el resultado del estudio no puede ocurrir. Se produce cuando la determinación del

estado de exposición de un sujeto implica un retraso, o período de espera, respecto la fecha de comienzo del estudio, durante el cual el tiempo de seguimiento se acumula; por ejemplo, si se espera a que una prescripción se dispense tras el alta hospitalaria cuando esta fecha representa el inicio del seguimiento. Este período se considera inmortal porque cada individuo del grupo expuesto tiene necesariamente que sobrevivir hasta que se cumpla la definición de tratamiento, mientras que si muere (o presenta otro evento de censura) antes de iniciar el tratamiento, entonces se encuadra en el grupo no expuesto [60, 61].

El sesgo se introduce en el estudio cuando este período de inmortalidad: (i) se clasifica erróneamente en relación al estado de tratamiento, y se atribuye al grupo tratado en un análisis tiempo-independiente; o bien (ii) se excluye del análisis porque el inicio del seguimiento en el grupo tratado se define por el comienzo del tratamiento, que por diseño es posterior al inicio de seguimiento del grupo no tratado [62] (Fig. 3).

(A) Tiempo inmortal mal clasificado



(B) Tiempo inmortal excluido

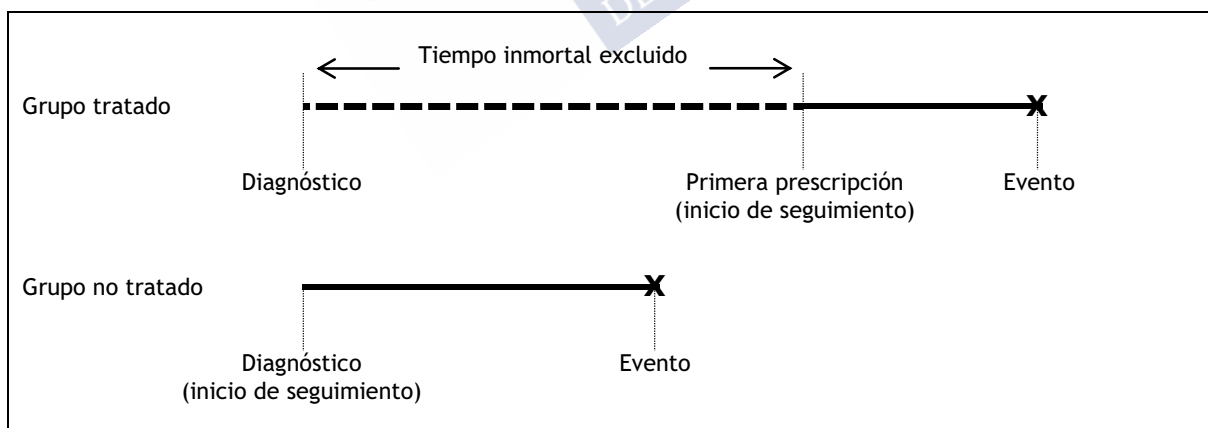


Figura 3.

Sesgo de tiempo inmortal (estudio de cohortes). **(A) Sesgo de mala clasificación:** La exposición al fármaco se define después de la entrada a la cohorte, por lo que el tiempo entre el inicio de seguimiento y la primera prescripción es inmortal para los expuestos (el sujeto debe sobrevivir para recibir el fármaco), y además mal clasificado como expuesto. **(B) Sesgo de selección:** La entrada a la cohorte se define por la primera exposición al fármaco para los expuestos, y por un evento (por ejemplo, un diagnóstico) para los no expuestos; el tiempo entre el evento y la primera prescripción es inmortal, y no se contabiliza en el grupo no tratado.

Este fenómeno es particularmente problemático porque confiere una ventaja de supervivencia espuria al grupo tratado. No tener en cuenta el tiempo-persona inmortal deriva en estimaciones sesgadas de la asociación exposición-resultado, siendo la magnitud del sesgo proporcional a la duración del tiempo-persona inmortal [63].

1.1.2.4.2. Otras formas de sesgo relacionado con el tiempo.

En muchos estudios realizados a partir de bases de datos secundarias se desconoce el estado de exposición durante la hospitalización. Puede producirse un sesgo de mala clasificación de la exposición según se continúe o interrumpa la exposición a los fármacos prescritos antes de la hospitalización, y si los días de hospitalización se consideran o no lagunas de exposición [57]. El sesgo diferencial derivado de no considerar o disponer de información sobre las hospitalizaciones producidas durante el período de observación (sesgo de tiempo no medible/inconmesurable) [64] puede ser problemático al estudiar enfermedades crónicas graves que requieren una amplia utilización de medicamentos y hospitalizaciones múltiples [65].

1.1.3. Conclusiones.

Antes de iniciar un proyecto farmacoepidemiológico con una base de datos secundaria, los investigadores deben estudiar detenidamente la propia fuente de información, considerando si la pregunta de investigación en particular puede ser abordada válidamente con cualquier base de datos. Existen guías orientativas para la selección y uso de los recursos de información [66]. Los datos de facturación de recetas médicas siguen siendo considerados la referencia para medir la exposición a fármacos, pero carecen de la medición de los resultados de estudio. Por otro lado, los registros médicos se consideran normalmente la referencia para la captura de resultados intermedios y finales, pero no registran todos los medicamentos realmente consumidos, y en general no se consideran adecuados para la identificación de la exposición. Mediante la vinculación de bases de datos administrativas y registros médicos, los investigadores pueden llegar a desarrollar medidas de resultados intermedios útiles para la vigilancia poscomercialización debido al uso sinérgico de las ventajas de cada tipo de fuente de datos.

Posiblemente, la mayor limitación de los estudios realizados con bases de datos secundarias es la dificultad de disponer de información sobre potenciales factores de confusión para un correcto análisis e interpretación de los resultados. En este sentido, resulta conveniente combinar la utilización de datos primarios y secundarios para un estudio más riguroso que minimice la aparición de sesgos.

Por último, el impacto en la política sanitaria de la investigación observacional sobre el riesgo/beneficio de los medicamentos es de gran importancia tanto a nivel sanitario como económico, por lo que resulta imprescindible una mayor vigilancia para reconocer y evitar la aparición de sesgos en estos estudios, cada vez más frecuentes debido al aumento del número de bases de datos secundarias.

1.2. ESTRATEGIAS DE CONTROL DE SESGOS EN FARMACOEPIDEMIOLÓGÍA.

Aunque son diversos los errores sistemáticos que pueden sesgar los resultados de la investigación no experimental, al estudiar los efectos de los fármacos resulta especialmente preocupante el fenómeno de confusión. En este apartado se describen las principales estrategias de control de la confusión, así como del resto de sesgos, que se pueden utilizar en la fase de diseño del estudio y en el análisis de los datos.

1.2.1. Diseño del estudio.

En la fase de diseño, la confusión por variables medidas y otros sesgos se pueden abordar mediante estrategias como la restricción (a través de criterios de inclusión y exclusión), el emparejamiento, o la implementación de un diseño de usuarios nuevos. Para controlar la confusión no medida, los diseños de casos cruzados son especialmente útiles (Fig. 4).

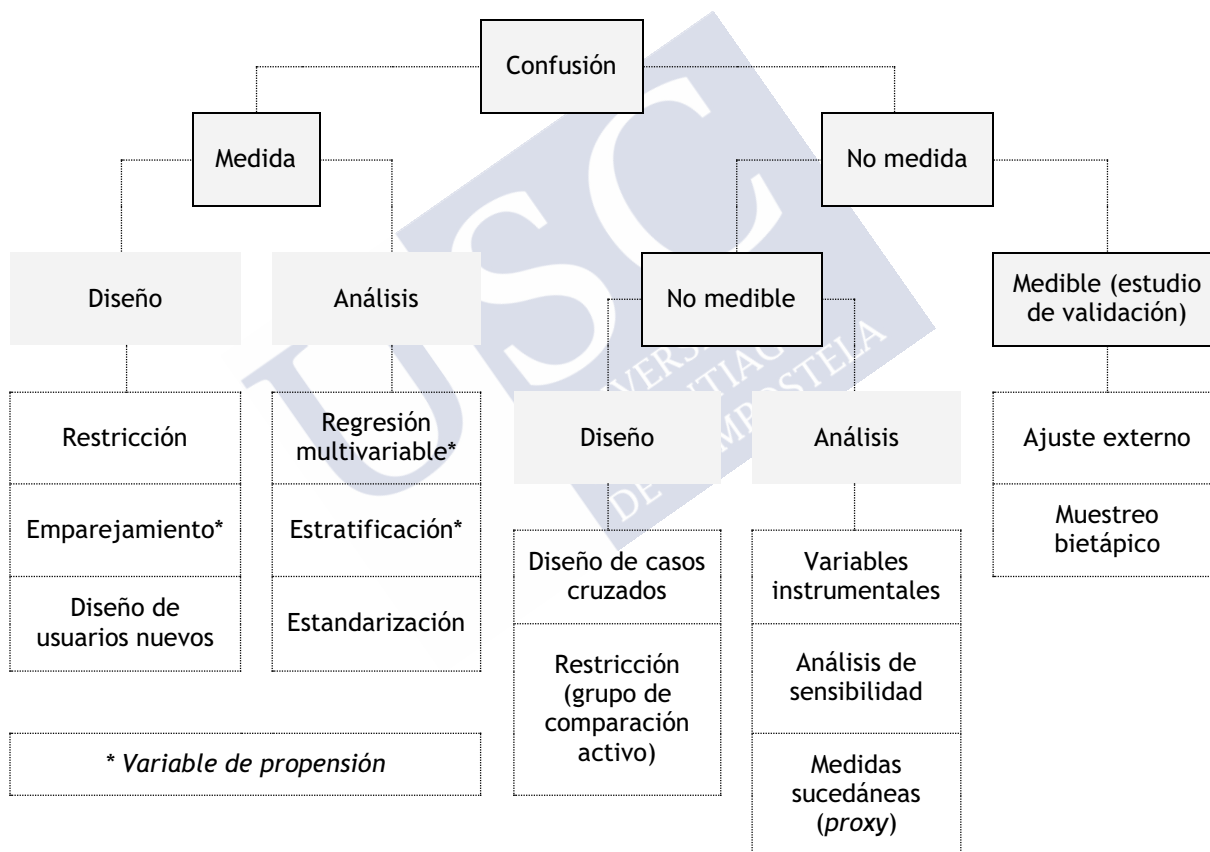


Figura 4.
Métodos farmacoepidemiológicos para controlar la confusión.

1.2.1.1. Restricción.

La restricción de pacientes homogéneos en cuanto a su indicación para el fármaco estudiado conducirá a un mayor equilibrio de las variables predictoras del resultado de interés en los grupos de exposición, reduciendo la confusión, especialmente si existen variables no disponibles en los datos que influyen en las decisiones de prescripción. También puede

reducir la posible modificación del efecto porque aumenta la probabilidad de que todos los sujetos incluidos tengan una respuesta similar a la terapia [67].

El grupo de comparación debe incluir pacientes con distribuciones similares de los factores de riesgo medidos y no medidos del resultado de interés. Al seleccionar fármacos para la misma indicación médica se reduce la confusión. Por ejemplo, los nuevos fármacos competidores dentro de una misma clase suelen comercializarse por su mayor eficacia/seguridad, o indicaciones ligeramente ampliadas, lo que influye en las decisiones de prescripción de los médicos produciendo confusión por indicación [51].

El cumplimiento del tratamiento es más habitual en pacientes que toleran bien el fármaco y que perciben algún beneficio terapéutico, es decir, la prescripción crónica tiende a asociarse con la no susceptibilidad a cualquier efecto adverso del medicamento. Esto significa que el riesgo de un evento será mayor en los usuarios agudos, y un estudio con pacientes mixtos prevalentes e incidentes (según la cronicidad de utilización del medicamento) conducirá a la subdetección de los primeros eventos [68].

1.2.1.2. Diseños de usuarios nuevos/incidentes.

En primer lugar, se identifican todos los pacientes en una población definida que inician el tratamiento a estudio. El seguimiento comienza al mismo tiempo que el inicio de la terapia (t_0) y se restringe a los pacientes con un período mínimo de ausencia de utilización antes de t_0 , definido previamente, excluyéndose así los usuarios prevalentes [49].

Los datos de los pacientes se obtienen justo antes de t_0 , incluyendo los potenciales factores de confusión, de modo que éstos no podrán estar influenciados por el propio fármaco de estudio. Los valores basales de dichas variables pueden utilizarse para ajustar las diferencias entre grupos [69].

Este diseño presenta algunas limitaciones [49], como la dificultad logística que supone identificar el momento en que se inició la medicación y la recogida de información sobre posibles factores de confusión. El uso de bases de datos informatizadas y su vinculación puede contribuir a identificar usuarios nuevos, pues a menudo incluyen información detallada sobre la prescripción y dispensación de medicamentos, además de historiales médicos.

Por otro lado, la restricción para nuevos usuarios puede reducir el tamaño de la muestra, y por ello la potencia del estudio. Los usuarios prevalentes incluyen muchos usuarios crónicos, especialmente importantes para evaluar el riesgo de los efectos relacionados con exposiciones crónicas. En este sentido, algunas bases de datos longitudinales incluyen información suficiente para el estudio de un fármaco desde su introducción. También se puede valorar *a priori* la magnitud de los potenciales sesgos relacionados con los usuarios prevalentes incluidos (por ejemplo, utilizando métodos de análisis de supervivencia para determinar la forma en que varía la función de riesgo según el tiempo desde el inicio del tratamiento, o realizando análisis comparativos entre usuarios nuevos y prevalentes).

Igualmente, los pacientes que son malos cumplidores o consumen inadecuadamente pueden estar sobrerrepresentados entre los que inician la terapia; por ello, se deben recoger y evaluar datos sobre indicación y factores de comportamiento como posibles modificadores del efecto.

1.2.1.3. Diseños de casos.

Diseño de casos cruzados (case-crossover) [70]: Consiste en seleccionar solo casos, evitando de ese modo un posible sesgo de selección, y comparar la prevalencia de exposición durante el período de tiempo previo al evento (ventana de exposición) con la exposición del mismo sujeto en uno o varios períodos control más distantes para obtener una razón de riesgos. Así, las posibles variables de confusión (conocidas y desconocidas) estables en el tiempo quedan automáticamente controladas [71].

Los requisitos para poder aplicar esta metodología son [72]: (i) la aparición del efecto debe ser aguda, y (ii) la exposición al tratamiento debe ser transitoria. Puede producirse un sesgo como consecuencia de cambios temporales en la prescripción o por confusión dentro del propio individuo, debido a una indicación transitoria o cambios en la gravedad de la enfermedad. Una limitación de este diseño es la confusión temporal cuando la prevalencia de la exposición es mayor en la ventana más próxima a la fecha del evento con respecto a la ventana más distante, debido al aumento natural del consumo a lo largo del tiempo (la prescripción puede aumentar al disponer progresivamente de una mayor evidencia sobre los efectos del fármaco).

Para situaciones en que las exposiciones sean prolongadas se ha desarrollado una extensión denominada *diseño caso-tiempo-control (case-time-control)*. Se utilizan datos del historial de la exposición de un grupo que no experimenta el evento de interés, y se analizan de manera similar a los casos para estimar y ajustar el sesgo de los cambios temporales en la prevalencia de exposición [73].

El análisis caso-tiempo-control no requiere que la exposición sea estable en el tiempo, pero supone que la tendencia de exposición no está confundida. Requiere también un número suficiente de pacientes que presenten el evento y la exposición al tratamiento de interés en cualquiera de las ventanas de exposición predefinidas, y todavía puede ser problemático el problema del sesgo de selección al comparar casos y controles [74]. Aunque elimina el efecto de la confusión tiempo-independiente no medida, este diseño es sensible al supuesto de intermitencia de utilización del medicamento y la amplitud de la ventana de tiempo de exposición [75]. La principal desventaja en estudios farmacoepidemiológicos es la necesidad de inferir los períodos de riesgo, cuya infraestimación o sobrestimación puede diluir el efecto del factor de exposición.

Diseño de serie de casos autocontrolados [76]: El período de observación después de cada exposición para cada caso se divide en uno o varios períodos de riesgo (período(s) de tiempo justo después de cada exposición) y un período de control (el tiempo de observación restante), a continuación se comparan las tasas de incidencia dentro de ambos períodos. El análisis de una serie de casos autocontrolados presenta la ventaja de controlar implícitamente cualquier factor de confusión conocido o desconocido que sea estable en el tiempo [58, 77]. Deben definirse también intervalos de tiempo relevantes para los períodos de riesgo y control [78].

1.2.1.4. Otros diseños.

Ensayos pragmáticos: Los ensayos aleatorizados simples a gran escala se caracterizan por un tamaño muestral grande y unos criterios de inclusión flexibles. Las

variables dependientes de estudio son objetivas, procedentes de bases de datos secundarias, y las variables de seguimiento se recogen en la práctica clínica habitual. El análisis se realiza por intención de tratar, evitando el sesgo de selección por pérdidas diferenciales entre los grupos de comparación.

Metanálisis de ensayos aleatorizados: Los ensayos clínicos carecen del poder estadístico suficiente para estudiar la seguridad de un medicamento cuando el efecto es de baja incidencia, lo que puede resolverse combinando varios ensayos individuales mediante la técnica de metanálisis.

1.2.2. Análisis de datos.

Algunos de los siguientes métodos están bien establecidos en la investigación farmacoepidemiológica, como la estratificación y la regresión; otros se utilizan con relativa frecuencia, como el grado de propensión y el análisis de variables instrumentales. Los modelos estructurales marginales y los modelos de ecuaciones estructurales se utilizan menos frecuentemente; requieren una amplia formación e interpretación adecuada de los resultados.

1.2.2.1. Estratificación.

Consiste en clasificar datos en subcategorías que se pueden analizar por separado. Sus aplicaciones incluyen la estandarización, el control de la confusión, el análisis de subgrupos en presencia de modificación del efecto, y el abordaje de un tipo de sesgo de selección que se produce en los estudios de casos y controles emparejados. Cuando una cohorte se clasifica por el tiempo de seguimiento, también se puede evitar el sesgo de los riesgos competitivos y las pérdidas en el seguimiento [51].

Una ventaja de la estratificación es que no requiere supuestos restrictivos. Algunas limitaciones incluyen la posibilidad de estratos de pequeño tamaño que reducen la precisión, la pérdida de información cuando las variables continuas se dividen en categorías elegidas arbitrariamente, y la dificultad del análisis cuando el número de estratos es grande.

La heterogeneidad significativa entre estratos sugiere la presencia de modificación del efecto. Se trata de una característica del efecto que se estudia, y no una fuente de sesgo que deba ser eliminada. Pueden calcularse también estimaciones combinadas del efecto para obtener una estimación global; en este caso, si difiere sustancialmente de las estimaciones específicas por estrato sugiere la presencia de confusión. Existen diversos métodos para la estimación de los efectos combinados, la principal diferencia es cómo asignan los pesos a los estratos.

1.2.2.2. Regresión.

Cuando se trabaja con numerosas categorías en un análisis estratificado, o se necesita ajustar por varios potenciales factores de confusión, se puede utilizar la regresión múltiple, al tiempo que se controla por los efectos de todos los demás factores independientes incluidos en la ecuación de regresión. Otro uso importante es utilizar la ecuación para predecir los resultados en otros pacientes, como ocurre con la regresión logística múltiple en el uso del grado de propensión [51].

Deben identificarse todas las posibles variables de confusión que pueden influir en la selección del tratamiento o la medición del resultado, y cualquier modificación o interacción

de efecto sospechada. Al utilizar fuentes de datos administrativas es habitual no disponer de medidas válidas para algunas de estas variables. Aun cuando todos estos factores se incluyan en la ecuación, todavía puede haber sesgo debido a factores de confusión desconocidos.

El tipo de variable dependiente condiciona la elección del modelo de regresión. Las variables continuas suelen analizarse con una regresión lineal de mínimos cuadrados ordinarios, mientras que los resultados dicotómicos se modelan mediante regresión logística. Ambas pertenecen a una categoría más amplia, los modelos lineales generalizados, que permiten incorporar diferentes especificaciones de estructuras de covarianza cuando se viola el supuesto de independencia entre las observaciones. Son habituales en estudios farmacoepidemiológicos los análisis longitudinales donde se realizan varias mediciones en el tiempo del mismo sujeto; una situación análoga produce el mismo tipo de correlación cuando los sujetos se muestrean repetidamente de una misma ubicación (por ejemplo, un hospital).

Debido a que la exposición al tratamiento puede cambiar con el tiempo (cambio de fármaco o de dosis), el modelo de regresión de riesgos proporcionales de Cox modeliza de un modo más realista la asociación fármaco-efecto, permitiendo la inclusión de covariables tiempo-dependientes. Mientras que el análisis de regresión logística supone que no hay pérdida de seguimiento ni riesgos competitivos en la población estudiada, estos supuestos no son necesarios en el modelo de Cox (donde la variable dependiente es el tiempo para el evento, en vez de la probabilidad de que ocurra el evento durante un período determinado).

Datos perdidos: El método más sencillo consiste en sustituir cada observación ausente por un valor medio utilizando los valores observados de la variable. Otra opción consiste en sustituir el valor predicho a partir de un modelo de regresión; sin embargo, ambos métodos sustituyen el mismo valor para todos los pacientes (o pacientes con características similares). En consecuencia, se reduce la variabilidad en los datos. Esto puede no suponer un problema si el patrón de ausencia de datos es aleatorio y no afecta a un gran número de observaciones para una variable dada. Existen métodos más sofisticados que preservan la variabilidad en los datos, como la imputación múltiple [79].

1.2.2.3. Grado de riesgo de enfermedad.

Un método para controlar un gran número de variables de confusión es la construcción de una puntuación de confusión multivariable que resuma los posibles factores de confusión en un solo índice, como el grado de riesgo de enfermedad, que estima la probabilidad o frecuencia de aparición de la enfermedad condicionado a no estar expuesto. La asociación exposición-enfermedad se estima ajustando por la puntuación de riesgo de enfermedad, en vez de las covariables individuales [80, 81].

1.2.2.4. Grado de predisposición a recibir un tratamiento: Análisis con una variable de propensión a la exposición.

El grado o índice de predisposición/propensión (GP) también combina un gran número de posibles factores de confusión en una sola variable. Se trata de una técnica empleada para tratar cuestiones relacionadas con el sesgo de selección, la confusión por indicación o la endogeneidad, habituales en estudios farmacoepidemiológicos. Por definición, en estos estudios los grupos de exposición están desequilibrados por potenciales variables de confusión; una opción es ajustar por ellas en modelos multivariantes, pero cuando el número de covariables es elevado y el número de eventos bajo, los efectos pueden estar sesgados y los

errores estándar no ser correctos [82]. En esta situación, los métodos basados en el GP permiten equilibrar posibles variables de confusión entre los grupos.

El GP se define como la probabilidad de exposición a un determinado tratamiento condicionada a las covariables observadas de un individuo [83, 84]. Los efectos del tratamiento no están sesgados si los grupos de tratamiento y la variable de resultado son condicionalmente independientes dadas las covariables. Esta hipótesis de independencia no se cumplirá cuando haya variables no incluidas como covariables de GP que se correlacionan con los eventos de resultado y la selección del tratamiento. El método más común para calcular los valores (o puntuaciones) de propensión es el modelo de regresión logística, de modo que la variable dependiente es el tratamiento y las covariables son los factores determinantes de la selección del tratamiento; el evento de estudio no forma parte del modelo para generar los valores de propensión. El ajuste de un modelo predictivo permite determinar la probabilidad de que cada sujeto reciba el tratamiento. Cada valor generado contiene la información de todas las covariables que dieron lugar al GP; así, si dos individuos presentan puntuaciones similares indica que globalmente tienen características basales similares [85] (Fig. 5).

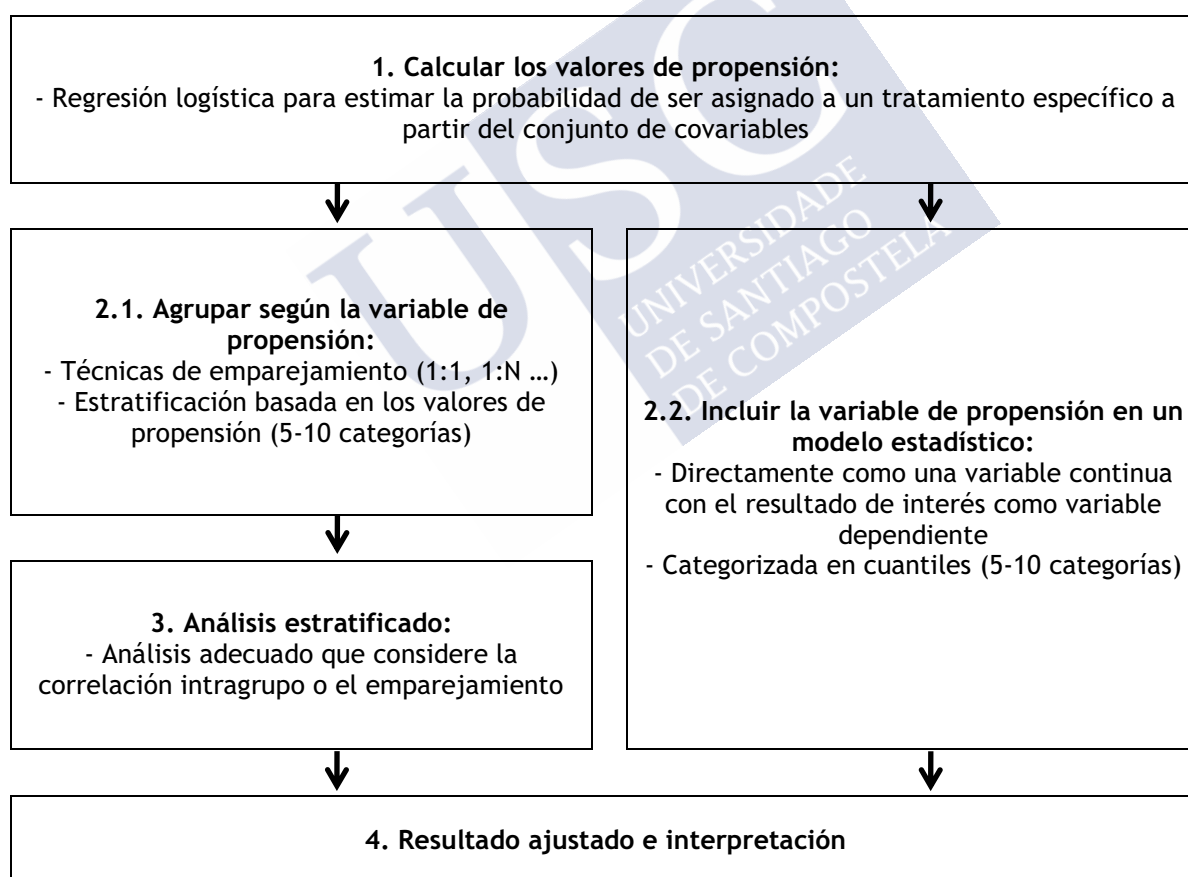


Figura 5.

Aplicaciones de los valores de propensión para el control de la confusión medida. *Etapa de diseño del estudio:* El uso de los valores de propensión en el emparejamiento se centra en la búsqueda de la correspondencia más próxima, basada en el valor de propensión, entre un individuo del grupo tratado y un sujeto(s) del grupo de comparación, lo que tiende a equilibrar todas las covariables observadas. Para garantizar un mayor solapamiento de las covariables solo se mantienen los pares de sujetos que son comparables, excluyéndose los sujetos en los que el tratamiento está contraindicado o raramente indicado en la muestra de

control (baja propensión), y en los que el tratamiento siempre está indicado en la muestra de tratamiento (alta propensión). **Etapa de análisis de los datos:** Los valores de propensión pueden emplearse para agrupar los sujetos de los grupos de comparación en niveles de estratificación (por ejemplo, todos los pacientes con un valor de propensión para recibir un determinado tratamiento entre 0 y 0,10 se juntan en un mismo estrato, el siguiente estrato contiene pacientes con un valor de propensión entre 0,10 y 0,20), de modo que los efectos del tratamiento se pueden comparar directamente dentro de cada estrato. Los métodos de regresión incluyen el GP como una covariable (continua o categorizada en cuantiles), además de una variable para determinar el efecto del tratamiento recibido (exógena), con el resultado de interés como variable dependiente (por ejemplo, la reacción adversa a un medicamento).

Para la construcción del modelo debe tenerse en cuenta que (i) la inclusión de variables claramente relacionadas con la exposición, pero no con el resultado, disminuye la precisión del efecto (aumenta la varianza); no obstante, debe considerarse incluir variables que están solo débilmente relacionadas con la selección del tratamiento, ya que potencialmente pueden reducir el sesgo en mayor medida de lo que aumentan la varianza [86, 87]. Mientras que (ii) las variables relacionadas con los resultados siempre deben ser incluidas en el GP, al margen de su fuerza de asociación en la selección de tratamiento; este es el caso de las variables de confusión (relacionadas con la exposición y el resultado) [88].

Debido a que en la ecuación del GP no se considera el principio de parsimonia, pueden incluirse más covariables e interacciones que en las técnicas habituales de regresión. Este problema se acentúa cuando hay pocos eventos de resultado, dada la restricción de al menos 8-10 eventos por covariable [89]. No obstante, si la exposición no es poco frecuente, con las técnicas de GP es posible ajustar por un gran número de covariables incluso si los resultados son poco frecuentes.

El *modelo de propensión de alta dimensión* [90] trata de identificar empíricamente un gran número de posibles factores de confusión, o variables sucedáneas (*proxies*) de los mismos, en bases de datos sanitarias [91]. A excepción del GP de alta dimensión, estos modelos no superan a los métodos multivariantes convencionales en cuanto al ajuste por los factores de confusión identificados [92]. Por otro lado, cuando se trabaja con bases secundarias los datos no se recogen con fines de investigación, y es habitual un desequilibrio entre los factores no medidos que puede violar el supuesto de independencia de tratamiento [93]. En esta circunstancia, el GP no aporta ventaja alguna sobre las técnicas de ajuste clásicas. La *calibración del grado de propensión* [94] es una técnica que combina métodos de emparejamiento con valores de propensión y modelos de regresión del error de medida para hacer frente a la confusión por variables no medidas en el estudio principal, mediante el uso de medidas de covarianza adicionales observadas en un estudio de validación [95].

1.2.2.5. Análisis de sensibilidad y ajuste externo.

El concepto básico del análisis de sensibilidad es hacer suposiciones fundadas acerca de posibles factores de confusión residual y cuantificar su efecto en la estimación de la medida de asociación. Si se dispone de fuentes con datos fiables, se pueden sustituir estas aproximaciones por estimaciones empíricas, y luego utilizarlas para el ajuste externo de la asociación. Estos análisis pueden utilizarse para evaluar cuantitativamente el sesgo de confusión en estudios con bases de datos.

1.2.2.5.1. Análisis simples de sensibilidad de factores de confusión a partir de datos agregados [96].

Pueden llevarse a cabo análisis a partir de suposiciones estructurales y seleccionando parámetros para un único factor de confusión que explicaría el resultado observado en el estudio bajo una hipótesis alternativa, en general nula. La ventaja de estos análisis para limitar la posible confusión de factores individuales es que se pueden realizar fácilmente a partir de información procedente de la literatura, o la opinión de expertos, sobre la distribución en la población de un factor de confusión no medido y su asociación con la exposición y el resultado, no siendo necesario recoger datos adicionales [97].

1.2.2.5.2. Ajustes externos que abordan la distribución conjunta de varios factores de confusión a partir de datos individuales.

Para hacer frente a la confusión conjunta de múltiples covariables se necesita información adicional sobre la distribución conjunta de estas covariables. Este procedimiento requiere datos a nivel individual de los estudios de validación. Estos estudios se clasifican en internos o externos, según los datos provengan de individuos incluidos o no en la población principal del estudio.

- *Estudios de validación interna:* Se basan en datos obtenidos para un subgrupo de la población principal del estudio. El resto de los participantes presentan valores desconocidos para todos los factores de confusión, evaluados solo en los sujetos del estudio de validación. Utilizando un método como la imputación múltiple para sustituir los valores ausentes del factor de confusión en el conjunto de datos principal se puede controlar la confusión no medida. Las principales ventajas de estos estudios son una mayor representatividad del estudio principal y la posible mejora de eficiencia basada en el sobremuestreo de la exposición/resultados de interés [98, 99]. No obstante, dependen de la viabilidad de recoger datos adicionales de los participantes, pudiendo no ser posible contactar con las personas cuyos datos se recogen para fines administrativos debido a cuestiones de privacidad o regulatorias. También debe considerarse la temporalidad y el sesgo de recuerdo.

- *Estudios de validación externa:* Los datos ya están recogidos y se pueden volver a utilizar para varios estudios principales que abordan múltiples hipótesis. Sin embargo, a menudo no contienen exactamente las mismas medidas que las utilizadas en el estudio principal y carecen de información sobre el resultado de interés. La calibración del grado de propensión para incorporar información sobre variables no disponibles en datos sanitarios puede ser útil para ajustar por factores de confusión no medidos. Estos estudios son bastante accesibles debido a que no son específicos de una determinada hipótesis, pero su uso en farmacoepidemiología es limitado ya que la información de fuentes externas sobre la utilización de medicamentos suele ser transversal (cuestionarios autocumplimentables o entrevistas), lo que conlleva un potencial de mala clasificación y un número limitado de usuarios incidentes.

1.2.2.6. Análisis con variables instrumentales.

En los estudios observacionales el sesgo se produce como resultado de una correlación entre una variable particular y el término de error en la ecuación de regresión; esto se conoce como endogeneidad, y se relaciona estrechamente con el concepto de confusión residual. La endogeneidad a menudo se presenta en forma de sesgo de selección de la muestra; es el caso de la selección no aleatoria del tratamiento, debiéndose a variables no medidas que están correlacionadas con el término de error de la ecuación [100].

El método de la variable instrumental (VI) [101] para abordar la endogeneidad se basa en la búsqueda de al menos una variable (instrumento) que se correlacione con la variable explicativa (endógena), pero no con el término de error (Fig. 6). Una limitación práctica importante es la dificultad para identificar buenos instrumentos [102]; con frecuencia los instrumentos utilizados tienden a correlacionarse débilmente con la variable a la que están destinadas a servir como instrumento, con el término de error, o con ambos, lo que reduce la eficiencia estadística y conduce a estimaciones sesgadas [103].

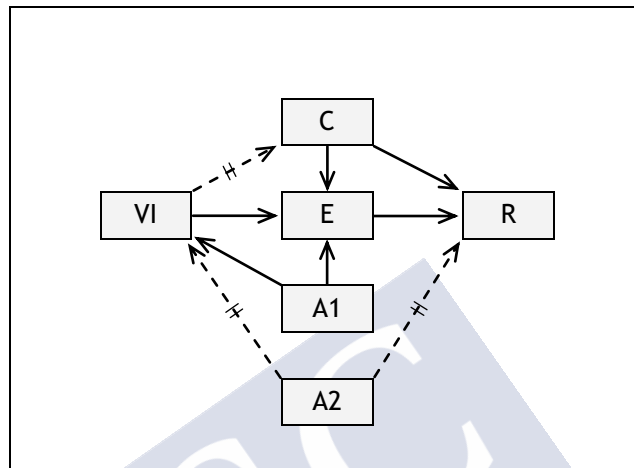


Figura 6.

Características de una variable instrumental (VI). (1) La VI se asocia con la exposición E , directamente o a través de una causa común $A1$. (2) La VI no se asocia con el resultado R (ni directamente ni a través de una causa común $A2$); solo afecta al resultado a través de la exposición. (3) La VI no está asociada con las variables de confusión C .

1.2.2.7. Control de la confusión tiempo-dependiente en el análisis.

En los estudios farmacoepidemiológicos los efectos de los medicamentos son frecuentemente dependientes del tiempo y están afectados por factores de confusión tiempo-dependientes que se ven afectados por el tratamiento. Los métodos estándar (estratificación, regresión) suelen ser suficientes para ajustar por la confusión, excepto en el caso de variables de confusión tiempo-dependientes. En particular, la confusión por indicación suele ser variable en el tiempo. En esta situación, los métodos estadísticos estándar pueden estar sesgados y se deben contemplar métodos alternativos, como los modelos estructurales marginales o la estimación-g.

1.2.2.7.1. Estimación-g.

Es un método utilizado para estimar los efectos conjuntos de tratamientos variables en el tiempo, basándose en ideas de los métodos de VI. El procedimiento permite el ajuste apropiado del efecto de una exposición variable en el tiempo en presencia de factores de confusión tiempo-dependientes que están influenciados por la exposición [104]. Aun siendo útil para los resultados de tiempo de supervivencia, los resultados de medición continuos y los resultados de recuento de Poisson, la estimación-g logística no puede ser convenientemente empleada para estimar el efecto del tratamiento sobre los resultados dicotómicos, salvo que el resultado sea raro [22].

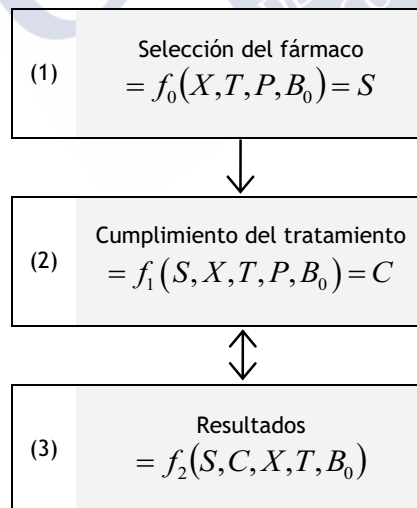
1.2.2.7.2. Modelos estructurales marginales.

El uso de modelos estructurales marginales es más parecido a los modelos estándar, y puede ser una alternativa a la estimación-g, permitiendo una mejor adaptación de la confusión en estas situaciones [105, 106].

Se han desarrollados modelos estructurales marginales que utilizan la ponderación por probabilidad inversa de tratamiento y que parecen estimar consistentemente los efectos de una exposición tiempo-dependiente en presencia de factores de confusión tiempo-dependientes. En ocasiones, el análisis estándar de Cox no puede detectar un beneficio neto clínicamente significativo del tratamiento debido a que no ajusta adecuadamente para covariables tiempo-dependientes que son simultáneamente factores de confusión y variables intermedias [107]; en este caso, mediante un modelo de supervivencia estructural marginal sí sería posible [108].

1.2.2.8. Otras técnicas.

En los métodos estadísticos comentados hasta ahora se utilizan generalmente variables ficticias para evaluar los efectos del tratamiento. Aunque los modelos multivariantes intentan controlar por otras variables observables (y en el caso de las VI, no observables), en última instancia miden una diferencia media esperada en la variable dependiente entre los grupos de tratamiento. Los *modelos de ecuaciones estructurales* pueden ofrecer más detalles [22]. Estos modelos consideran la posibilidad de que el tratamiento para una enfermedad puede caracterizarse por los siguientes procesos: (i) la elección del fármaco; (ii) el cumplimiento del tratamiento o los patrones de utilización; y (iii) los resultados (Fig. 7). El principal inconveniente en el uso de estos modelos es que no contienen una simple variable ficticia que proporcione la magnitud, el signo y la significación estadística del efecto del tratamiento estimado.



f_0 - f_2 : relaciones entre los procesos (lineales o no lineales)
 X : vector de variables explicativas que incluye las características del paciente
 T : vector del tratamiento recibido por los pacientes en el periodo previo y las condiciones basales de salud
 P : vector de variables que mide características del proveedor
 B_0 : características basales de salud del individuo

Figura 7.

Modelo de ecuaciones estructurales. La selección del fármaco se produce en primer lugar en la secuencia de acontecimientos (ecuación 1). A continuación, los pacientes generan patrones de cumplimiento a la terapia

(ecuación 2) que a su vez influyen en los resultados observados (ecuación 3); también es posible que los resultados puedan retroalimentar los patrones de consumo de los medicamentos. Tanto los patrones de consumo como los resultados pueden estar influidos por factores no observados. Los patrones de consumo y los resultados pueden estar correlacionados con las variables no observadas asociadas con la selección del fármaco. Los patrones de variables tiempo-dependientes y variables omitidas pueden conducir a estimaciones sesgadas de los parámetros. El método de estimación adecuado para el sistema de ecuaciones depende críticamente de la estructura de la covarianza de los términos de error de las tres ecuaciones (si éstos no están correlacionados, se puede estimar cada ecuación de forma independiente).

1.2.3. Conclusiones.

Al evaluar los efectos de los medicamentos en estudios que utilizan bases de datos secundarias surgen diversas limitaciones, resultando especialmente preocupante el fenómeno de confusión. No obstante, existen mecanismos para el control de la mayoría de ellas. Es necesario tener en cuenta estas estrategias de control para el diseño y/o análisis e interpretación de los estudios con estas fuentes de datos.

Entre las principales estrategias de control disponibles para la confusión medida en la fase de diseño, se incluyen la restricción de nuevos usuarios y la consideración de solo casos. Y para el control de la confusión no medida, la restricción a un grupo de comparación activo y el diseño de casos cruzados. En la fase de análisis de datos, los mecanismos para el control de la confusión medida incluyen técnicas tradicionales como la estratificación y la regresión. Se han propuesto varios métodos para hacer frente a los factores de confusión no medidos, que pueden clasificarse según se utilicen o no datos adicionales no incluidos en el conjunto de datos original. Algunos métodos que no se basan en la recogida de datos adicionales incluyen el análisis con variables instrumentales o una variable de propensión. Otros métodos utilizan información adicional sobre factores de confusión que se mide en un conjunto externo de datos y que puede servir como un estudio de validación.

Un encuadre cuidadoso de la pregunta de investigación, con el diseño metodológico adecuado y la aplicación de técnicas de análisis estadístico, puede producir resultados válidos y mejorar las inferencias sobre los efectos de los medicamentos en estudios farmacoepidemiológicos que utilizan bases de datos secundarias. En cualquier caso, aunque se haya intentado controlar toda la variedad de sesgos descrita, nunca se puede descartar la posibilidad de que persistan factores de confusión que sesguen los resultados. Por ello, la evaluación de las asociaciones entre fármacos y sus efectos es limitada, habiéndose estimado que solo uno de cada cinco resultados positivos obtenidos de estudios observacionales es realmente cierto [109].



2. OBJETIVOS.

2.1. OBJETIVO GENERAL.

Identificar y describir las limitaciones de las bases de datos administrativas y registros médicos (datos secundarios) cuando se utilizan con fines de investigación en farmacoepidemiología, y valorar la necesidad de su complementación o validación con datos *ad hoc* (primarios).

2.2. OBJETIVO ESPECÍFICOS.

2.2.1. Revisar publicaciones científicas, indexadas en las principales bases de datos bibliográficas médicas, que pongan de manifiesto potenciales sesgos en estudios farmacoepidemiológicos que utilizan bases de datos secundarias, y proponer estrategias para la mitigación de dichos sesgos.

2.2.2. Revisar los diferentes tipos de fuentes de datos disponibles para resumir los resultados sobre seguridad de medicamentos, así como los diversos métodos de síntesis, centrándose en la técnica del metanálisis.

2.2.3. Averiguar si el tipo de fuente de datos utilizada, según sea su origen primario o secundario, introduce heterogeneidad en los metanálisis de la asociación entre diversos fármacos y sus respectivos efectos adversos.

2.2.4. Evaluar el grado en que el origen de los datos, primario o secundario, se explora como una fuente de heterogeneidad en los metanálisis de estudios observacionales publicados en una selección de revistas de alto impacto.



3. MÉTODOS.

3.1. EL SESGO EN ESTUDIOS FARMACOEPIDEMIOLÓGICOS QUE UTILIZAN BASES DE DATOS SECUNDARIAS: UNA REVISIÓN EXPLORATORIA (OBJETIVO 1 - ANEXO 1).

Se llevó a cabo una “*scoping review*”, que es una estrategia metodológica que permite resumir los resultados de una investigación exploratoria. En este tipo de revisión, a diferencia de otras revisiones sistemáticas, la aplicación de filtros de calidad no es una prioridad inicial [110]. El estudio fue realizado en base a la guía metodológica para la realización de una *scoping review* del Instituto Joanna Briggs [111] y la guía de extensión para *scoping reviews* de PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*) [112].

3.1.1. Fuentes de datos y estrategia de búsqueda.

Se realizó una búsqueda automatizada en bases de datos bibliográficas, con una búsqueda inicial en MEDLINE, complementada posteriormente por EMBASE y *Web of Science*. Para evitar la duplicación de resultados, en EMBASE y *Web of Science* se utilizó la opción que permite excluir revistas indexadas en MEDLINE. La misma estrategia de búsqueda de texto libre se aplicó en las tres bases de datos: (*clinical–data* OR health–data* OR medical–data* OR prescription–data* OR administrative–data* OR epidemiologic–data* OR health–claim* OR administrative–claim* OR insurance–claim* OR claims–data* OR health–record* OR medical–record**) AND (*confounding OR bias* OR missing–data OR misclassification*) AND (*observational OR epidemiolog* OR pharmacovigilance OR challenge**) AND *drug*, del 1 de enero de 2000 al 1 de enero de 2018. Todos los tipos de diseño de investigación fueron tenidos en cuenta. No se consideró adecuado añadir términos restrictivos de MeSH (*Medical Subject Headings*) según el tipo de publicación, pues se comprobó que esto conducía a una reducción excesiva de la sensibilidad de búsqueda.

Una vez identificadas las referencias, los títulos y los resúmenes, cuando estuvieron disponibles, se utilizaron como filtro de selección preliminar y, cuando se consideró potencialmente relevante, se recuperó el texto completo de los artículos. Otras referencias relevantes fueron identificadas mediante la comprobación cruzada manual de las listas de referencias de los artículos seleccionados y el uso de la opción “artículos relacionados”. Este cribado exhaustivo fue realizado por dos revisores (GP-R, AF). Se discutieron las discrepancias entre ambos revisores para lograr un consenso. En caso de un posible desacuerdo, se designó a un tercer revisor (BT).

3.1.2. Selección de artículos y extracción de datos.

En la revisión se incluyeron artículos de opinión, revisiones metodológicas, análisis/reanálisis y estudios de simulación, así como cartas al editor o retracciones, cuyo objetivo principal, descrito en sus resúmenes, fue destacar la existencia de algún tipo de sesgo en estudios farmacoepidemiológicos que utilizaron bases de datos secundarias.

Para reducir el número de referencias identificadas y simplificar así la presentación de los resultados, se consideraron los siguientes criterios de exclusión que clasificaron las referencias descartadas en subgrupos: (i) el objetivo principal del artículo fue describir, comparar, evaluar, validar o desarrollar una estrategia de control para un sesgo o limitación conocidos (por ejemplo, método analítico, diseño del estudio, algoritmo, marco de trabajo); (ii) se estimó una medida (por ejemplo, una asociación tratamiento-efecto) o se identificaron factores de riesgo para una enfermedad, mencionando la existencia de sesgo como una limitación del estudio, independientemente de que se utilizaran o no estrategias para su control; (iii) el estudio tenía características diferentes de las indicadas anteriormente (por ejemplo, estudios con objetivos diferentes, no basados en bases de datos secundarias, sin algún fármaco involucrado, sin sesgo mencionado) o era una comunicación/ponencia sin un resumen/texto completo disponible.

Dos revisores (GP-R y AF) desarrollaron conjuntamente un formulario de datos para determinar qué variables debían extraerse. Una persona (GP-R) extrajo la información de los artículos (es decir, el primer autor, la fecha de publicación, la categoría bajo la cual se indexó la revista –si la revista se indexó bajo más de una categoría, se consideró la categoría bajo la cual estaba mejor clasificada–, el tipo de artículo, el tipo de sesgo(s) mencionado(s)) y cuando fue necesario, los artículos fueron revisados y validados por revisores adicionales como una forma de control de calidad (AF y BT). Los tres revisores discutieron los resultados y actualizaron continuamente el formulario de datos.

La síntesis incluyó tanto el análisis cuantitativo (es decir, la tendencia de publicación de los artículos identificados/incluidos y el análisis de frecuencia de los sesgos mencionados) como el análisis cualitativo (es decir, el análisis de contenido) de los componentes del propósito de la investigación.

3.2. CONCLUSIONES DIVERGENTES A PARTIR DEL MISMO METANÁLISIS EN ESTUDIOS DE SEGURIDAD DE MEDICAMENTOS. RELEVANCIA DEL TIPO DE FUENTE DE DATOS CONSIDERADA (OBJETIVO 3 - ANEXO 3).

3.2.1. Criterios de inclusión.

Para utilizar ejemplos de la vida real de metanálisis sobre seguridad de medicamentos que pudiesen ilustrar la importancia de tener en cuenta el origen de los datos en los resultados, se seleccionaron metanálisis de estudios observacionales que cumplieron los siguientes criterios (ver diagrama de flujo de la estrategia de búsqueda en suplemento 1, anexo 3): (i) el consumo de medicamentos de venta libre (sin receta) de uso frecuente fue la exposición principal o un tratamiento concomitante, ya que una de las limitaciones de la utilización de bases de datos de origen secundario en los estudios farmacoepidemiológicos es la dificultad para encontrar información sobre la exposición a los fármacos de venta libre; (ii) el metanálisis se publicó recientemente (entre 2013 y 2015); (iii) los eventos clínicos evaluados en los estudios individuales del metanálisis tuvieron períodos cortos de inducción y latencia (de unos pocos días a unos pocos meses), para minimizar el riesgo de sesgo de recuerdo cuando se utilizaron fuentes de datos primarias; y (iv) el metanálisis incluyó al menos dos estudios que utilizaron datos primarios o secundarios, de manera que fuese posible evaluar la heterogeneidad en cada estrato. Los fármacos de venta libre considerados en la estrategia de búsqueda fueron el paracetamol, el omeprazol y el ibuprofeno. Se tuvo en cuenta la posibilidad de que estos fármacos pudiesen desempeñar el papel de una potencial variable de confusión cuando se consumieron de forma concomitante (por ejemplo, con los inhibidores selectivos de la recaptación de serotonina (ISRSs)), y que los estudios basados en datos secundarios pudiesen ser propensos a sesgos de clasificación errónea en la exposición debido a la falta de registro de la parte de su consumo que corresponde a los medicamentos de venta libre.

3.2.2. Evaluación de la exposición y el resultado.

Se consideró “datos primarios” la información sobre la exposición al fármaco recogida por los investigadores directamente del paciente, mediante entrevistas (personales o telefónicas) o cuestionarios autoadministrados. Los “datos secundarios” son datos recogidos originalmente para fines distintos a los del estudio en cuestión e incluidos en bases de datos de prescripción y dispensación de medicamentos (ver las características de los estudios incluidos en suplemento 1, anexo 3).

Cuando el evento clínico (por ejemplo, hemorragia gastrointestinal alta) tuvo dos resultados distintos con diferente etiología (por ejemplo, variceal y no variceal), se excluyeron los estudios que no diferenciaron ambos resultados. Nuestro propósito fue mitigar la heterogeneidad que pudiese surgir de la inclusión de enfermedades de diferente etiología.

3.2.3. Extracción de datos.

Los datos se extrajeron directamente de las tablas o figuras de los metanálisis, no de los estudios individuales. Cuando los metanálisis incluyeron tanto estudios observacionales como ensayos clínicos aleatorizados, solo los estudios observacionales fueron considerados para el análisis. Excepcionalmente, se incluyeron también aquellos estudios en los que el fármaco de venta libre se evaluó de forma observacional como un fármaco concomitante con el tratamiento aleatorizado (por ejemplo, tratamiento antiplaquetario dual).

3.2.4. Análisis de datos.

Se utilizó el inverso de la varianza para calcular las estimaciones combinadas de razones de riesgos (RR), razones de *odds* (OR) o razones de riesgos instantáneos (HR), y sus intervalos de confianza (IC) del 95% en cada grupo (fuentes de datos primarias o fuentes de datos secundarias). Se utilizó un modelo de efectos aleatorios y se midió la heterogeneidad de los resultados mediante el estadístico I^2 . Los gráficos de bosque se generaron mediante el programa informático *Review Manager 5.3* (the Cochrane Collaboration, Oxford, UK).



3.3. UTILIZACIÓN DE FUENTES DE DATOS Y HETEROGENEIDAD EN METANÁLISIS DE ESTUDIOS OBSERVACIONALES SOBRE EFECTOS DE MEDICAMENTOS: UN ESTUDIO DE LAS PRINCIPALES REVISTAS MÉDICAS (OBJETIVO 4 - ANEXO 4).

3.3.1. Selección de metanálisis y proceso de recogida de datos.

En la revisión se incluyeron revistas de medicina general e interna con un factor de impacto >15 según *Web of Science* [113]. Este método se ha utilizado ampliamente para evaluar la calidad y las tendencias de publicación en revistas médicas [114-116]. La razón es que los metanálisis publicados en revistas de alto impacto: (i) probablemente se realicen e informen rigurosamente debido al exhaustivo proceso editorial [115, 117]; y, (ii) en general, ejercen una mayor influencia en la práctica médica debido al importante papel desempeñado por estas revistas en la difusión de las nuevas evidencias médicas [117, 118]. Se realizaron búsquedas en MEDLINE en mayo de 2018 con los términos de búsqueda “*meta-analysis*” como tipo de publicación y “*drug*” en cualquier campo, entre el 1 de enero de 2012 y el 7 de mayo de 2018, en *the New England Journal of Medicine* (NEJM), *Lancet*, *Journal of the American Medical Association* (JAMA), *British Medical Journal* (BMJ), *JAMA Internal Medicine* (JAMA Intern Med), *Annals of Internal Medicine* (Ann Intern Med) y *Nature Reviews Disease Primers* (Nat Rev Dis Primers).

Dos investigadores (GP-R, AF) evaluaron de forma independiente la elegibilidad de las publicaciones. Se revisaron los resúmenes y, cuando se consideró oportuno, se recuperaron los artículos completos. Los artículos se excluyeron si cumplían alguna de las siguientes condiciones: (i) no ser un metanálisis de estudios publicados, (ii) no evaluar los efectos de un medicamento, (iii) el metanálisis incluir únicamente ensayos clínicos aleatorizados (con el fin de considerar los estudios observacionales), (iv) el metanálisis incluir menos de dos estudios observacionales (ya que con un solo estudio no habría sido posible calcular una medida combinada). Cuando un metanálisis incluyó tanto estudios observacionales como ensayos clínicos, solo se consideraron los estudios observacionales.

Previamente se desarrolló un formulario de extracción de datos para obtener la información de los artículos. Dos investigadores (GP-R, AF) extrajeron y registraron la información de forma independiente y resolvieron las discrepancias cotejando el artículo original. En caso necesario, se pidió a un tercer autor (BT) que resolviese los desacuerdos entre los investigadores.

Cuando estuvieron disponibles, se extrajeron los siguientes datos de cada metanálisis elegible: primer autor, año de publicación, revista, exposición a fármacos y resultados; número de estudios individuales incluidos en el metanálisis según el tipo de fuente de datos utilizada (primaria frente a secundaria), tanto para la evaluación de la exposición como para la evaluación de los resultados; y variables relacionadas con la exposición y los resultados incluidas en los análisis de sensibilidad, subgrupos o metarregresión. Los datos se extrajeron directamente de las tablas, las figuras, el texto y el material suplementario de los metanálisis, no de los estudios individuales.

3.3.2. Evaluación de la exposición y el resultado.

Se consideró “datos primarios” la información sobre la exposición al fármaco recogida directamente por los investigadores mediante entrevistas –personales o telefónicas– o mediante cuestionarios autoadministrados. El origen de los datos también se consideró

primario cuando se utilizaron métodos de diagnóstico objetivos para la determinación de la exposición al fármaco (por ejemplo, análisis de sangre). Los “datos secundarios” son los datos que inicialmente se recogieron para fines distintos de los del estudio en cuestión y que se incluyeron en bases de datos de prescripción de medicamentos (por ejemplo, registros de prescripciones, historiales médicos) y dispensación (por ejemplo, registros de farmacia informatizados, bases de datos de seguros médicos). Con respecto a la evaluación de resultados, los datos se consideraron primarios cuando se dispuso de una confirmación objetiva que los avalase (por ejemplo, confirmada por un diagnóstico médico *ad hoc* individual, pruebas de laboratorio o resultados por imagen). Estos criterios se basan en los utilizados habitualmente en la evaluación del riesgo de sesgo para los estudios observacionales [119-122].



4. RESULTADOS.

4.1. EL SESGO EN ESTUDIOS FARMACOEPIDEMIOLÓGICOS QUE UTILIZAN BASES DE DATOS SECUNDARIAS: UNA REVISIÓN EXPLORATORIA (OBJETIVO 1 - ANEXO 1).

La figura 8 muestra el proceso de selección de artículos. Se incluyeron un total de 117 artículos. La búsqueda automatizada dio lugar a la identificación de 863 referencias no duplicadas, que se redujeron a 56 tras la aplicación de los criterios de exclusión. El proceso de selección manual incorporó otras 61 referencias.

4.1.1. Tendencia de publicación.

La figura 9 muestra una suavización polinómica de la frecuencia con la que se publicaron los artículos incluidos en la revisión desde el año 2000. Se observa una tendencia creciente, de manera que casi la mitad (45,3%, 53/117) de los artículos fueron publicados durante los últimos 5 años completos de esta revisión. Existe una tendencia temporal similar de las referencias identificadas mediante la búsqueda automatizada cuando se ajusta por el número de citas indexadas añadidas a MEDLINE durante cada año [123], lo que sugiere que los criterios de restricción considerados no introdujeron sesgo de selección. Una ligera disminución en 2017 puede deberse a las características inherentes al proceso de indexación en las bases de datos bibliográficas, o al hecho de que las referencias más recientes han tenido menos tiempo para ser citadas y, en consecuencia, es menos probable que sean identificadas por la búsqueda manual de referencias cruzadas.

Parece haber una gran variedad de disciplinas interesadas en artículos sobre las limitaciones potenciales de las bases de datos secundarias (Fig. 10a). Las categorías de revistas médicas más utilizadas fueron “Salud pública, ambiental y ocupacional” (24,8%, 29/117 artículos incluidos) y “Farmacología y farmacia” (14,5%, 17/117). En general, se observa la misma tendencia de publicación a lo largo del tiempo cuando se estratifica por disciplina (Fig. 10b).

4.1.2. Principales sesgos mencionados en los artículos incluidos en la revisión.

En la tabla 5 se enumeran los artículos que mencionan las categorías o subcategorías de los sesgos más habitualmente descritos en estudios observacionales que utilizan bases de datos farmacoepidemiológicas. El sesgo de confusión como tal, o en cualquiera de sus diversas formas de presentación, fue la categoría de sesgo más frecuentemente mencionada (63,2%, 74/117 artículos incluidos), mientras que el sesgo de confusión por indicación fue la subcategoría más frecuente (32,5%, 38/117), seguida por el sesgo de confusión no medida/residual (28,2%, 33/117). También se mencionó la existencia de confusión tiempo-dependiente y el sobreajuste debido a una elección inadecuada de las variables en el modelo estadístico (sesgo debido a una especificación errónea de las variables de control).

Asimismo, se mencionó algún tipo de sesgo de selección y medición en el 47,0% (55/117) y 46,2% (54/117) de los artículos incluidos, respectivamente. El sesgo debido a la

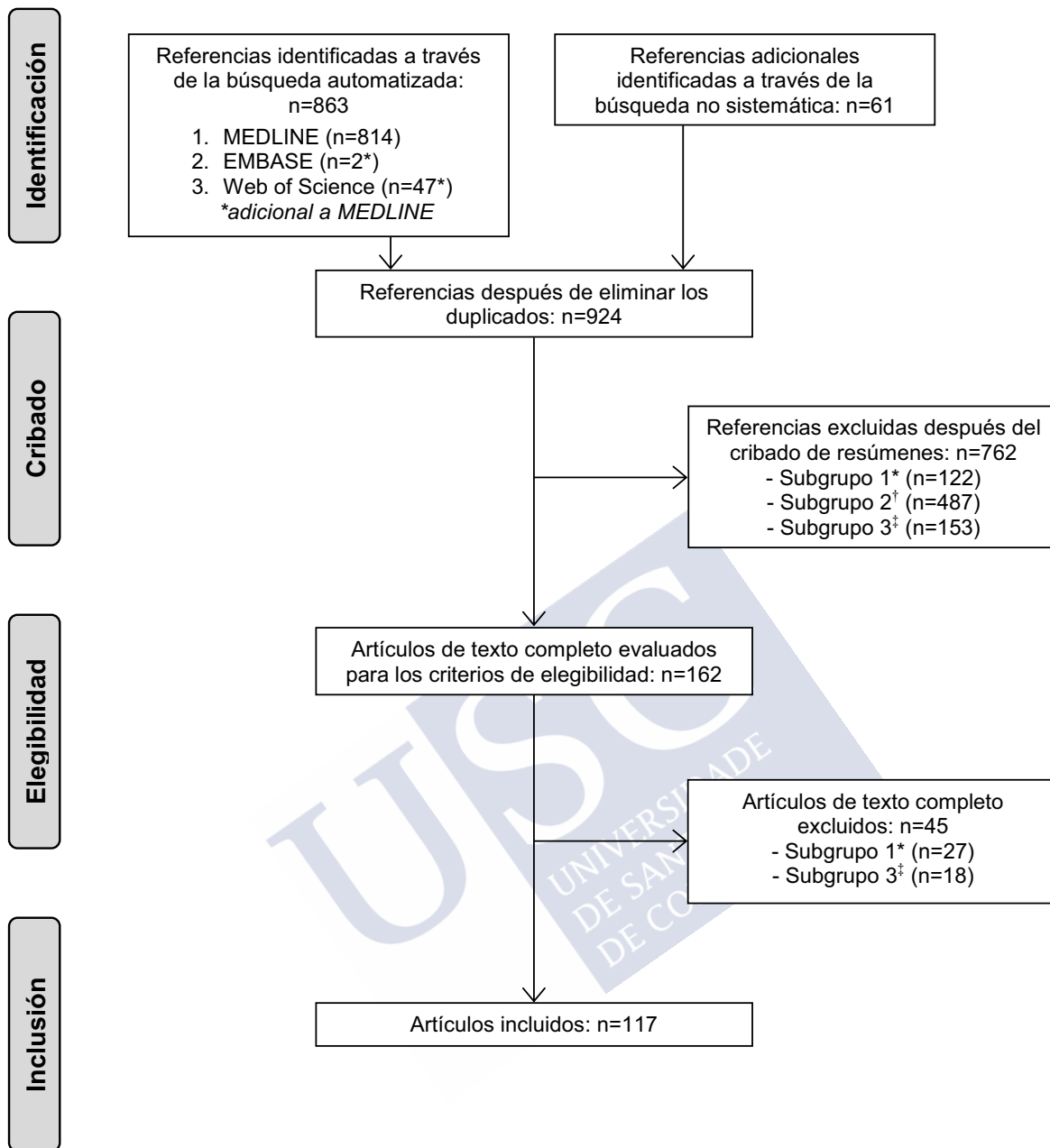


Figura 8.

Diagrama de flujo del proceso de selección de artículos. * Subgrupo 1: Su objetivo principal fue describir, comparar, evaluar, validar o desarrollar una estrategia de control para un sesgo o limitación conocido. † Subgrupo 2: Se estimó una medida o se identificaron factores de riesgo para una enfermedad, mencionando la existencia de sesgo como una limitación del estudio, independientemente de que se utilizaran o no estrategias para su control. ‡ Subgrupo 3: Tenía características diferentes a las indicadas anteriormente o era una comunicación/ponencia sin resumen/texto completo disponible.

falta de datos y el sesgo del usuario prevalente fueron los sesgos de selección reportados con mayor frecuencia (38,2%, 21/55 y 21,8%, 12/55, respectivamente); además, también se describieron otras formas de sesgo, como el sesgo protopático, la censura informativa, los riesgos competitivos y el sesgo del acceso diferencial a la atención sanitaria. La clasificación errónea de la exposición o del resultado fueron las causas más habituales de sesgo de

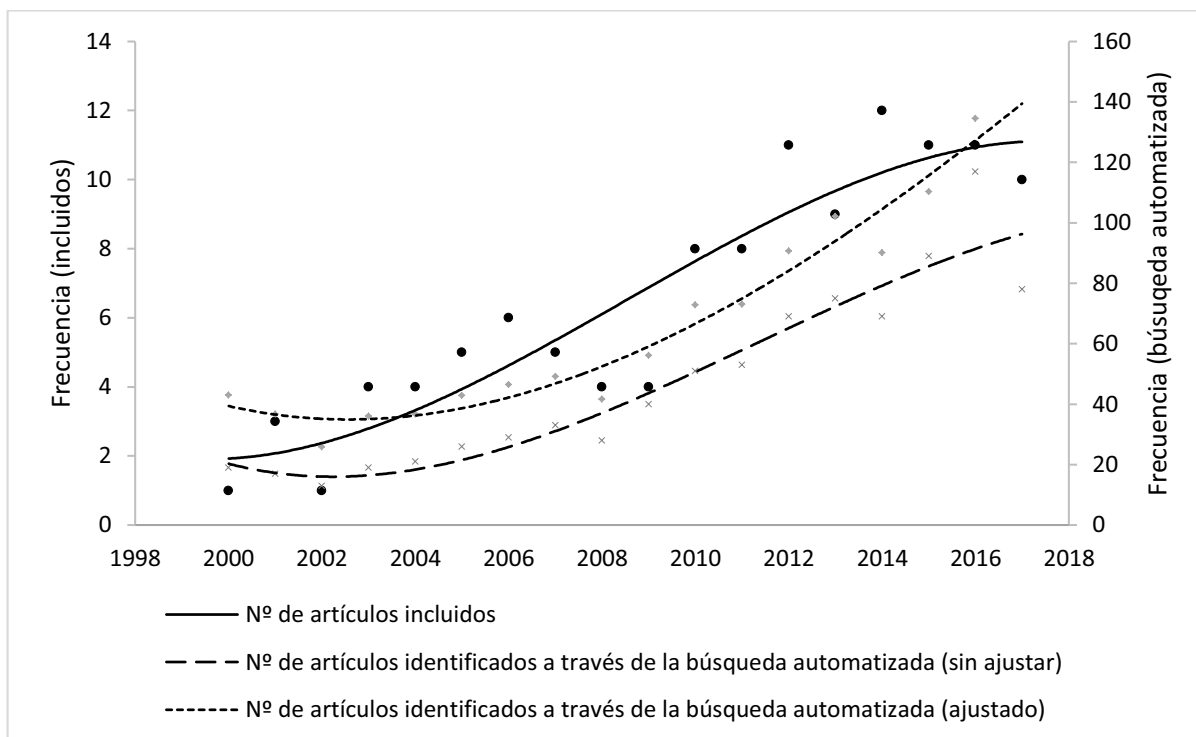


Figura 9.

Cronología de la publicación de los 117 artículos incluidos en la revisión (eje Y izquierdo) y las 863 referencias identificadas mediante la búsqueda automatizada (eje Y derecho) sin ajustar y ajustado por el número de citas indexadas añadidas a MEDLINE.

medición (51,9%, 28/54 y 61,1%, 33/54 respectivamente). También se citó la ambigüedad temporal y la clasificación errónea de los factores de confusión.

Aunque puede considerarse estrictamente como un subconjunto de las tres categorías más grandes (es decir, confusión, sesgo de selección o sesgo de medición), la última categoría en estudiarse fue la categoría del sesgo relacionado con el tiempo, como el sesgo de “tiempo inmortal”, que resultó ser el sesgo más mencionado (25,6%, 30/117) después de la confusión por indicación, la confusión no medida/residual y la clasificación errónea del resultado (28,2%, 33/117). También se describió el sesgo de tiempo inconmensurable, el sesgo de ventana temporal y el sesgo de desfase temporal. La figura 11 muestra la frecuencia de cada sesgo mencionado en los artículos incluidos, así como las categorías generales, estratificada por períodos de 6 años.

El suplemento (ver suplemento 1, anexo 1) contiene los datos extraídos de los artículos revisados en orden descendente de fecha de publicación por categoría de campo de investigación bajo la cual se indexó la revista. Los artículos también se clasificaron según el tipo de contenido, incluyendo, en cada caso, las categorías o subcategorías de sesgo mencionadas.

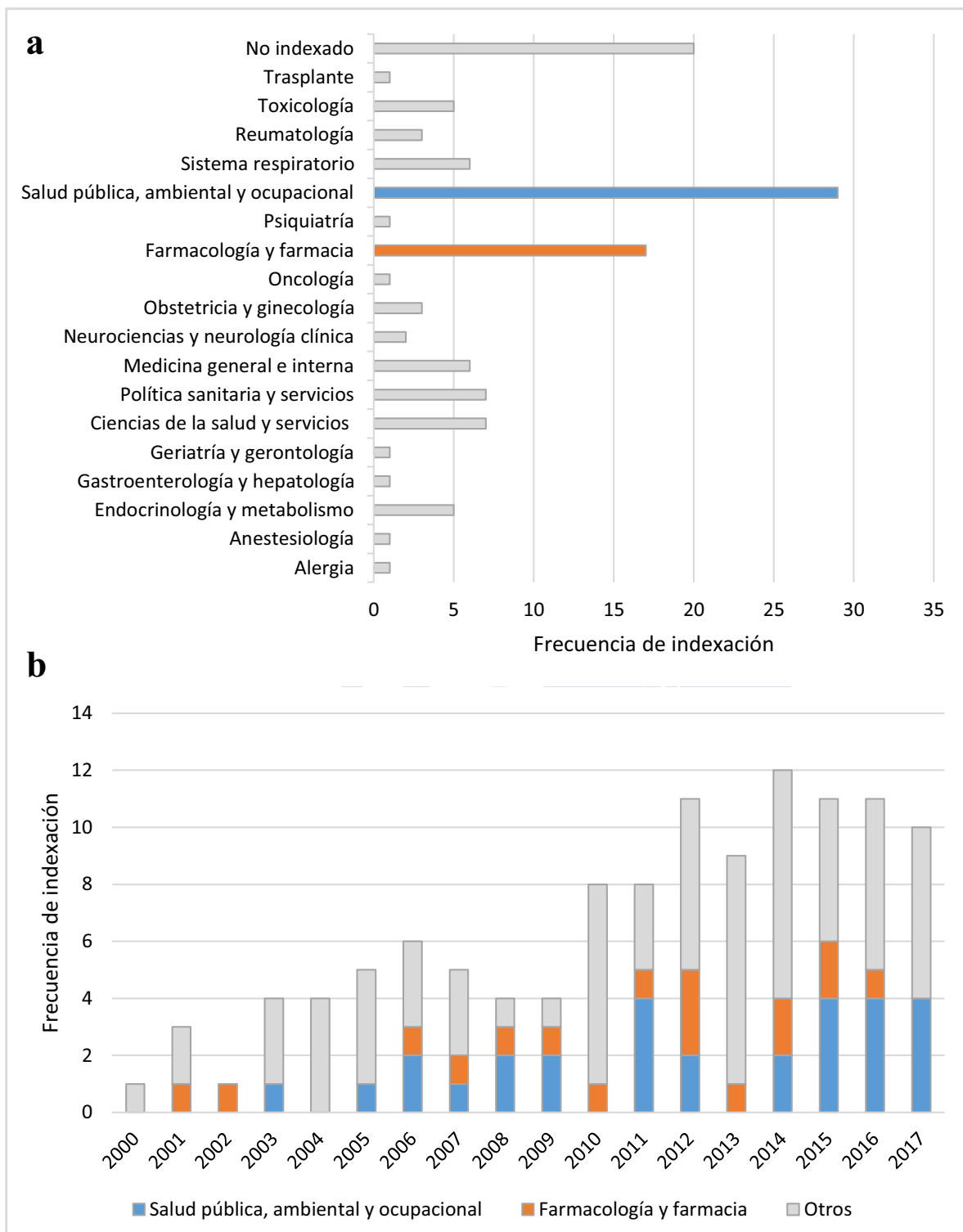


Figura 10.

(a) Distribución de los artículos incluidos entre las distintas disciplinas médicas. **(b)** Cronología de los artículos incluidos según las disciplinas indexadas más frecuentes.

Tabla 5. Artículos que mencionan los sesgos más comunes descritos en estudios observacionales con bases de datos farmacoepidemiológicas.

Categoría/ Subcategoría	Descripción del sesgo	Referencias (n=117)	Porcentaje (%)
Confusión	La medida de la asociación entre el tratamiento y el resultado está distorsionada por el efecto de una o más variables, que también son factores de riesgo para el resultado de interés	1-3, 8, 10, 21, 27, 32, 37, 39, 42, 65, 124-185	63.2
Confusión por indicación*	La condición clínica que determinó la prescripción del tratamiento está asociada con el efecto, actuando como un factor de confusión (por ejemplo, un peor estado de la enfermedad al inicio del estudio: confusión por la gravedad de la enfermedad)	1, 10, 27, 32, 39, 42, 65, 124, 126, 128, 130, 131, 133, 134, 136, 140-142, 144, 146, 147, 151, 153, 156, 158, 159, 161, 164, 166, 170, 171, 173-176, 179, 180, 184	32.5
Confusión tiempo-dependiente	Una variable que puede variar con el tiempo actúa como factor de confusión entre la exposición actual y el resultado, y como intermediario entre la exposición anterior y la actual	2, 27, 125, 136, 140, 142, 151	6.0
Confusión no medida/residual	No hay suficiente información sobre todos los factores de confusión relevantes conocidos, desconocidos o difíciles de medir (por ejemplo, fragilidad). Si no se puede controlar completamente el efecto de confusión, el efecto de confusión residual de algunos factores permanece en el efecto final que se observa	1-3, 8, 10, 21, 37, 39, 42, 124-127, 130, 133, 135-137, 141, 143, 148, 150, 154, 158, 161, 163, 172, 175, 177, 178, 180, 182, 185	28.2
Efecto del usuario sano/ adherente	El acceso a los recursos de atención sanitaria está asociado con un mayor nivel educativo y un comportamiento de búsqueda de salud. Además, los pacientes que cumplen con el tratamiento durante períodos prolongados de tiempo tienden a estar más sanos	3, 8, 39, 135, 141, 172	5.1
Sesgo de selección	La población de la muestra del estudio no es representativa de la población objetivo a la que se extrapolarán los resultados	2, 3, 8, 27, 32, 39, 125, 127, 128, 131, 132, 134, 135, 137-140, 142, 146, 148-150, 152-158, 160-163, 165, 166, 168, 169, 171, 181-183, 186-199	47.0
Sesgo protopático	El tratamiento se asocia a estadios subclínicos de la enfermedad (una manifestación precoz de la condición clínica bajo estudio, todavía no diagnosticada, da lugar a la prescripción del tratamiento)	125, 140, 142, 155	3.4
Pérdidas durante el seguimiento (censura informativa)	El mecanismo que desencadena la discontinuidad del tratamiento se asocia con el riesgo de observar el resultado de interés	140, 142, 161	2.6
Depleción de susceptibles (sesgo del usuario prevalente)	La inclusión de usuarios prevalentes en lugar de usuarios incidentes implica una verificación insuficiente de los efectos adversos que se producen al comienzo del tratamiento (los susceptibles al efecto adverso han interrumpido el tratamiento)	3, 27, 32, 127, 134, 140, 142, 146, 153, 156, 161, 196	10.3
Falta de datos	En los análisis multivariados, como los modelos de regresión, las observaciones que carecen de uno o más de los valores de una variable incluida en el modelo tienden a ser eliminadas	2, 8, 131, 137, 138, 154, 157, 161, 163, 181-183, 186, 187, 189-192, 194, 195, 199	17.9

Categoría/ Subcategoría	Descripción del sesgo	Referencias (n=117)	Porcentaje (%)
Sesgo de medición	Los datos sobre las exposiciones reales, los resultados y otras variables se registran en forma de indicadores (medidas observadas) que no reflejan con precisión la realidad	1-3, 8, 10, 11, 42, 55, 131, 132, 135, 137, 138, 140-142, 148, 152, 154, 157, 159, 160, 162, 163, 165, 168, 169, 175, 181-184, 186, 187, 190, 191, 193-195, 197, 199-212	46.2
Sesgo de mala clasificación	La asociación entre el tratamiento y el resultado está distorsionada por errores sistemáticos, debido a la forma en que se miden las variables de interés en los grupos de comparación	1-3, 8, 10, 11, 42, 55, 131, 132, 135, 137, 138, 140-142, 148, 152, 154, 157, 159, 160, 163, 165, 168, 175, 181-184, 186, 187, 190, 191, 193-195, 197, 200-212	43.6
Mala clasificación de la exposición	La medida de la exposición de un tratamiento dado no es un reflejo exacto de su uso real (por ejemplo, medición defectuosa, incumplimiento del tratamiento, uso inapropiado de las ventanas de tiempo)	1-3, 42, 55, 131, 135, 137, 138, 140-142, 148, 163, 165, 168, 175, 184, 186, 193-195, 200, 202, 204, 206, 207, 212	23.9
Mala clasificación del resultado	Error en el diagnóstico (por ejemplo, ambigüedad clínica, codificación no uniforme)	1-3, 8, 10, 11, 42, 131, 135, 137, 138, 140-142, 148, 157, 159, 165, 168, 181-183, 187, 190, 193, 197, 201, 203, 205, 208-211	28.2
Sesgo relacionado con el tiempo	El tiempo de seguimiento y el estado de exposición no se tienen en cuenta adecuadamente en las fases de diseño del estudio o de análisis	3, 11, 27, 32, 60, 62-65, 127, 130, 131, 134, 140, 142, 146, 148, 152, 153, 156, 159, 173, 174, 179, 188, 213-223	30.8
Sesgo de tiempo inmortal	Se incluye en el seguimiento un período de tiempo (inmortal) durante el cual el evento de estudio no puede ocurrir, o se excluye del análisis debido a una definición incorrecta del inicio del seguimiento	3, 11, 27, 32, 60, 62, 63, 65, 127, 130, 131, 134, 140, 142, 146, 148, 153, 156, 159, 173, 174, 179, 214, 215, 218-223	25.6
Sesgo de tiempo inconmensurable	Se ignora un período de tiempo (inconmensurable) durante el seguimiento y, por lo tanto, se clasifica erróneamente como período no expuesto, ya que las prescripciones ambulatorias que definen la exposición no pueden ocurrir (por ejemplo, enfermedades crónicas graves que requieren el uso extensivo de medicamentos y hospitalizaciones múltiples)	64, 188, 213, 216	3.4
Sesgo de ventana temporal	El uso de ventanas de tiempo de diferente duración entre los casos y los controles para definir las exposiciones dependientes del tiempo, impide que los sujetos tengan la misma oportunidad de recibir prescripciones	65, 134, 217	2.6
Sesgo de desfase temporal	Se realizan comparaciones de los tratamientos administrados en diferentes estadios de la enfermedad, lo que intrínsecamente introduce un sesgo relacionado con la duración y la progresión de la enfermedad	65	0.9

* A veces también denominado sesgo de canalización.

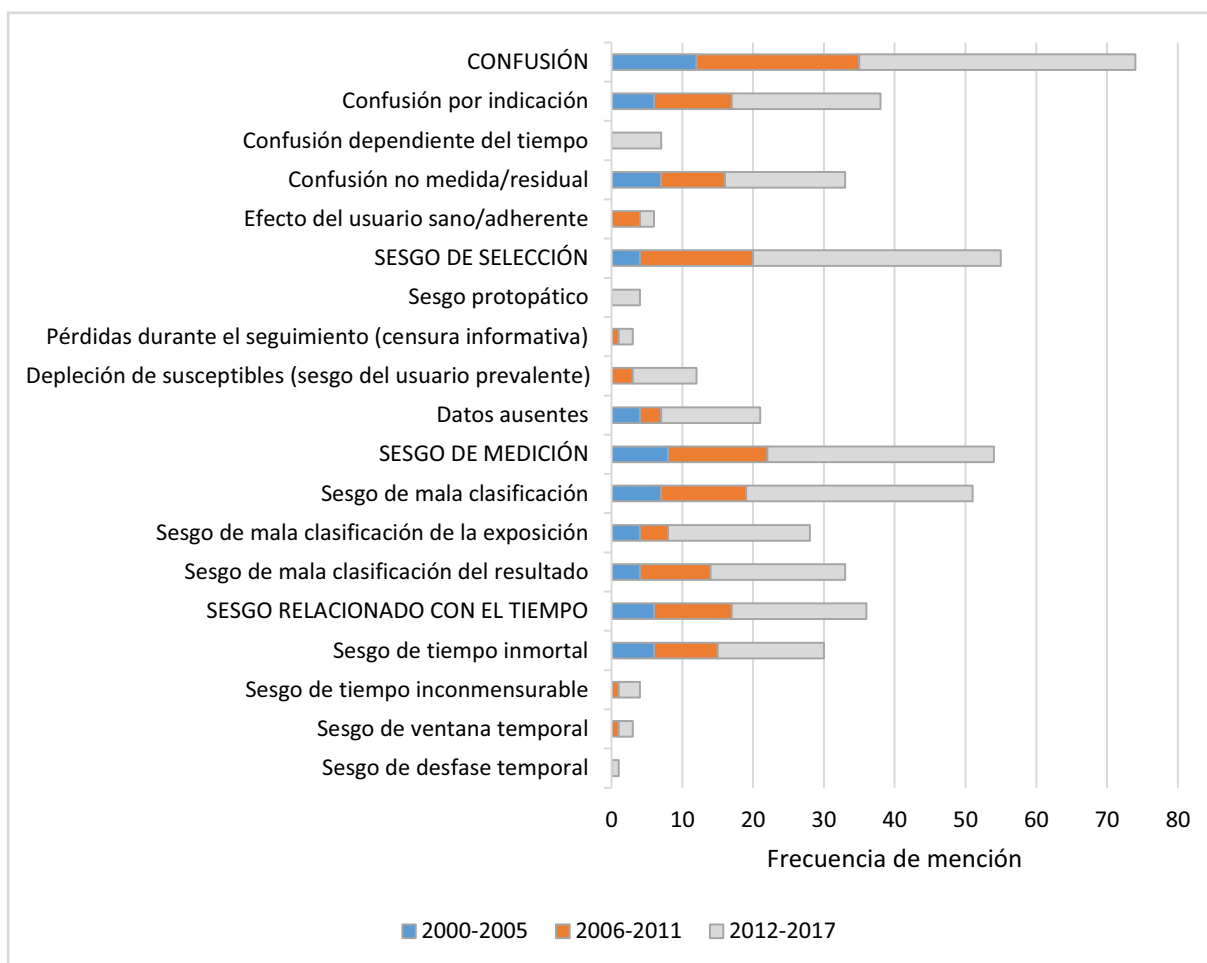


Figura 11.

Frecuencia de los sesgos mencionados en los artículos incluidos, estratificada por períodos de tiempo.

4.2. CONCLUSIONES DIVERGENTES A PARTIR DEL MISMO METANÁLISIS EN ESTUDIOS DE SEGURIDAD DE MEDICAMENTOS. RELEVANCIA DEL TIPO DE FUENTE DE DATOS CONSIDERADA (OBJETIVO 3 - ANEXO 3).

Se incluyeron cuatro metanálisis que cumplieron los criterios de inclusión.

4.2.1. Inhibidores selectivos de la recaptación de serotonina y sangrado gastrointestinal alto [224].

Se excluyeron los estudios que no diferenciaron entre la hemorragia variceal y la hemorragia no variceal. Se observó que al estratificar por la fuente de datos no hubo aumento del riesgo en el grupo de datos primarios ni heterogeneidad (OR = 1,19; IC del 95%: 0,90 a 1,58), mientras que el aumento del riesgo fue alto en el grupo de datos secundarios (OR = 1,81; IC del 95%: 1,50 a 2,17), con heterogeneidad alta. El análisis original, que no distinguió entre las diferentes fuentes de datos, concluyó que hubo un aumento importante del riesgo (OR = 1,71; IC del 95%: 1,44 a 2,04) (Fig. 12).

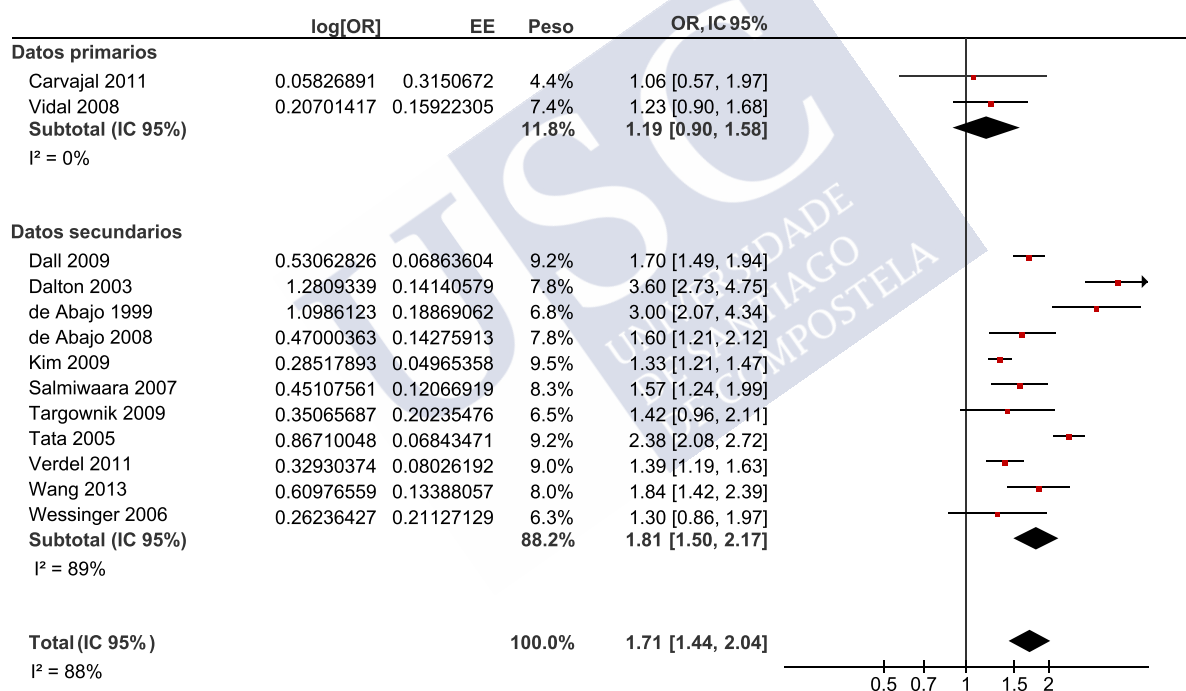


Figura 12.

Inhibidores selectivos de la recaptación de serotonina y efecto sobre el sangrado gastrointestinal alto. (Adaptado de [224], figura 2). OR, razón de odds; EE, error estándar; IC, intervalo de confianza. (Reproducida con autorización de la revista *Drug Safety*).

4.2.2. Inhibidores de la bomba de protones junto con tratamiento antiplaquetario dual y eventos adversos cardiovasculares mayores.

Los resultados del metanálisis realizado por Melloni et al. [225], que evaluó el efecto del uso concomitante de un inhibidor de la bomba de protones (IBP) y tratamiento antiplaquetario dual en la aparición de eventos cardiovasculares, sugieren un aumento moderado del riesgo (HR = 1,35; IC del 95%: 1,18 a 1,54). La estratificación por el origen de

los datos sugiere que este aumento solo está presente si la fuente son datos secundarios. Por el contrario, los resultados para el grupo de datos primarios sugieren ausencia de significación estadística y disminución de la heterogeneidad (Fig. 13).

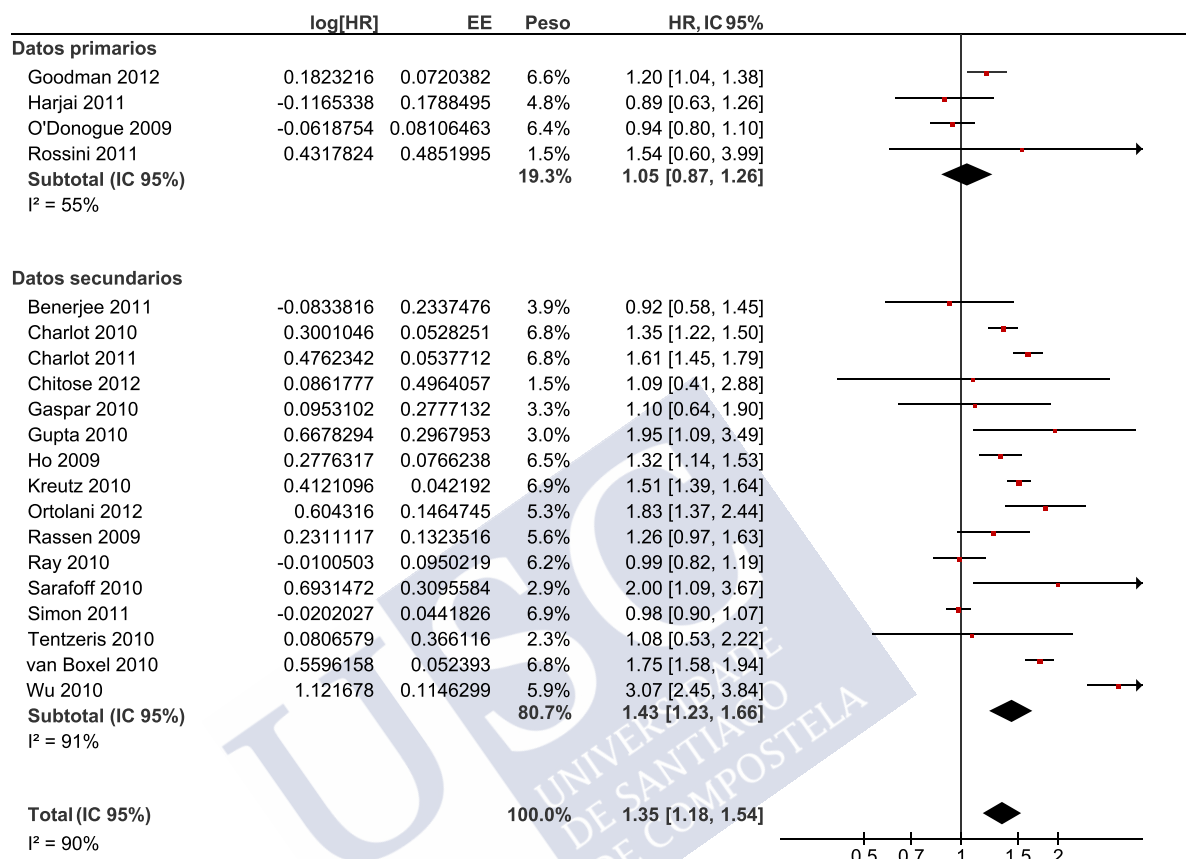


Figura 13.

Cualquier inhibidor de la bomba de protones (+ tratamiento antiplaquetario dual) y efecto sobre el resultado compuesto de eventos isquémicos (mortalidad cardiovascular, infarto de miocardio no fatal o accidente cerebrovascular) a 1 año. (Adaptado de [225], figura 2A). HR, razón de riesgos instantáneos; EE, error estándar; IC, intervalo de confianza. (Reproducida con autorización de la revista *Drug Safety*).

4.2.3. Fármacos antiinflamatorios no esteroideos no aspirina e infarto de miocardio.

El metanálisis más reciente, publicado por Varas-Lorenzo et al. [226], que evalúa la relación entre los fármacos antiinflamatorios no esteroideos (AINEs) individuales y el infarto de miocardio, se basa en estudios que solo utilizaron datos secundarios. Por lo tanto, se utilizó un metanálisis más completo publicado anteriormente sobre el mismo tema [227]. Este metanálisis no mostró aumento significativo del riesgo (RR = 1,08; IC del 95%: 0,95 a 1,22). Sin embargo, nuestro análisis de subgrupos muestra que el efecto combinado de los AINEs entre los estudios que utilizaron datos primarios es fuertemente protector (RR = 0,57; IC del 95%: 0,34 a 0,96), mientras que el efecto entre los estudios que utilizaron datos secundarios

es el de un factor de riesgo (RR = 1,15; IC del 95%: 1,03 a 1,28) (Fig. 14). De nuevo, en el primer grupo no se detectó heterogeneidad ($I^2 = 12\%$).

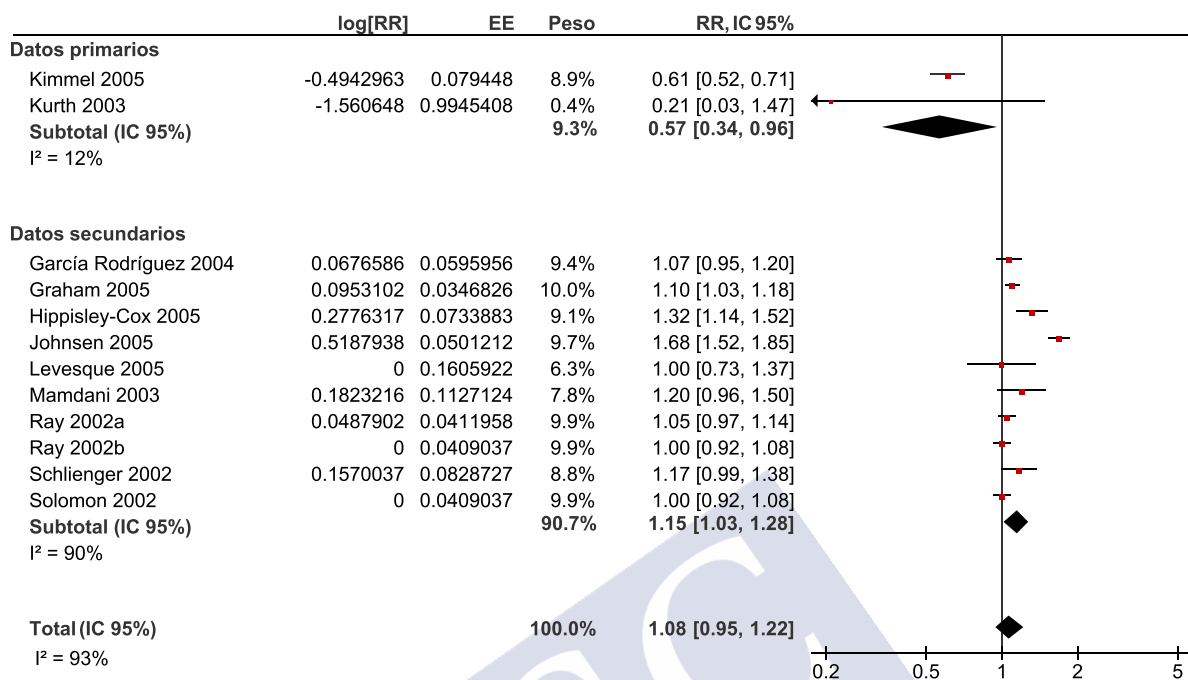


Figura 14.

Cualquier fármaco antiinflamatorio no esteroideo no aspirina y su efecto sobre el infarto de miocardio. (Adaptado de [227], tabla 2). RR, razón de riesgos; EE, error estándar; IC, intervalo de confianza. (Reproducida con autorización de la revista *Drug Safety*).

4.2.4. Paracetamol en el embarazo y asma infantil.

Cheelo et al. [228] realizaron un metanálisis sobre el uso del paracetamol durante el embarazo y el riesgo posterior de asma en los primeros años de vida. El análisis de subgrupos por la fuente de datos muestra una divergencia con respecto a los resultados del análisis combinado original; el efecto combinado entre los estudios con datos primarios muestra un menor aumento del riesgo (OR = 1,23; IC del 95%: 1,06 a 1,42) en comparación con el análisis global, y entre los estudios con datos secundarios, un mayor aumento del riesgo (OR = 1,53; IC del 95%: 1,33 a 1,75) (Fig. 15).

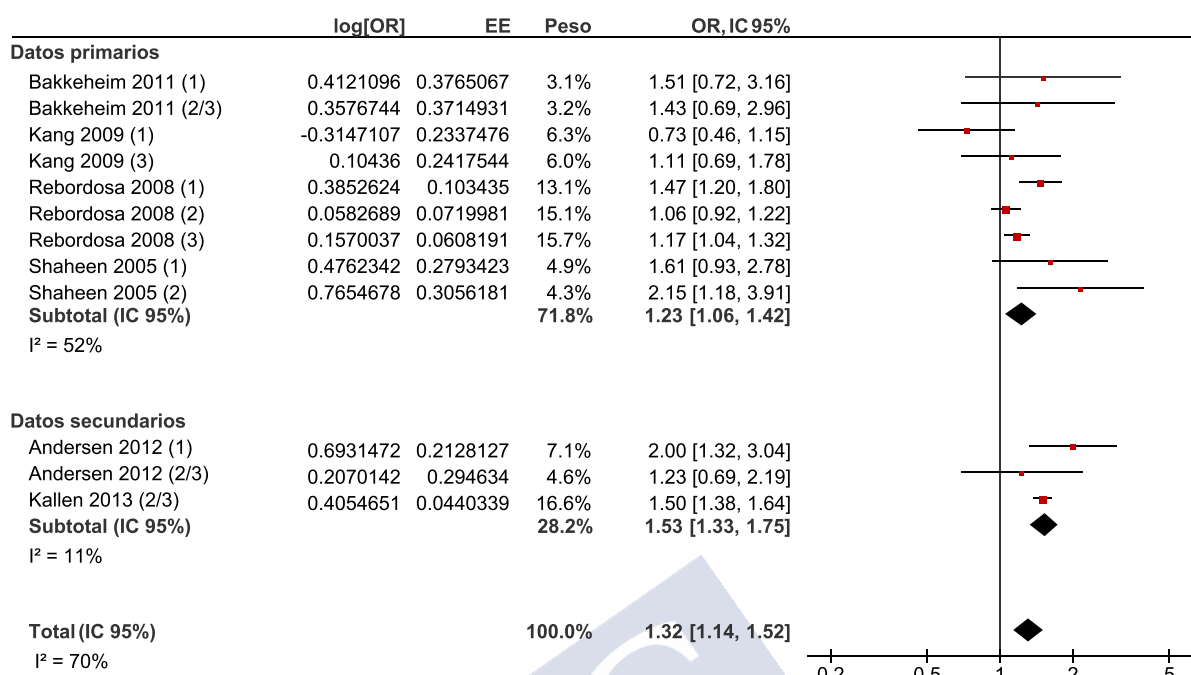


Figura 15.

Paracetamol en el embarazo y efecto sobre el asma infantil. (Adaptado de [228], figuras 2 y 3). OR, razón de odds; EE, error estándar; IC, intervalo de confianza. (Reproducida con autorización de la revista *Drug Safety*).

4.3. UTILIZACIÓN DE FUENTES DE DATOS Y HETEROGENEIDAD EN METANÁLISIS DE ESTUDIOS OBSERVACIONALES SOBRE EFECTOS DE MEDICAMENTOS: UN ESTUDIO DE LAS PRINCIPALES REVISTAS MÉDICAS (OBJETIVO 4 - ANEXO 4).

La búsqueda en MEDLINE obtuvo 217 artículos de las principales revistas médicas generales (3 de NEJM, 46 de Lancet, 26 de JAMA, 85 de BMJ, 19 de JAMA Intern Med, 38 de Ann Intern Med y 0 de Nat Rev Dis Primers) (Fig. 16). Se excluyeron un total de 194 artículos (ver lista de artículos excluidos con motivos de exclusión en suplemento 1, anexo 4), dejando 23 artículos para examinar [229-251]. Las características generales de los 23 metanálisis incluidos se describen en la tabla 6.

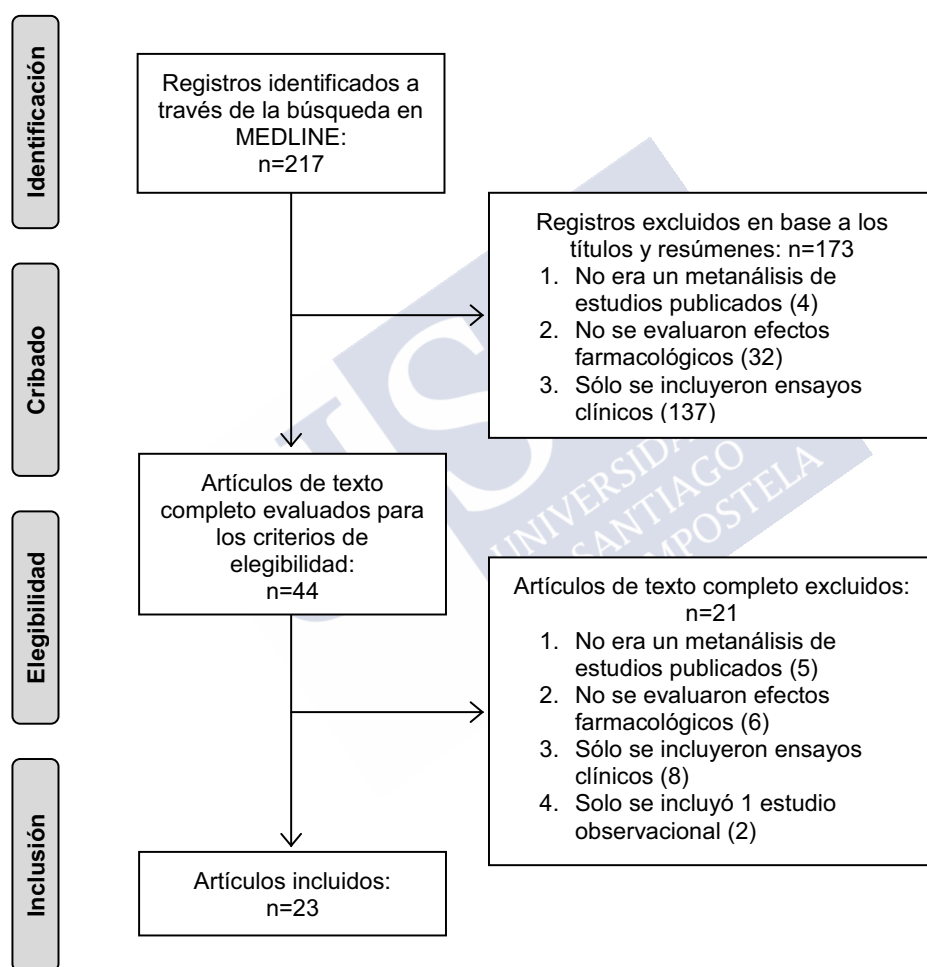


Figura 16.
Diagrama de flujo de los resultados de la búsqueda de literatura.

4.3.1. Fuente de datos de la exposición y el resultado.

La tabla 7 resume la evidencia sobre el tipo de fuente de datos incluida en cada metanálisis, de acuerdo con la información presentada en las tablas de extracción de datos del artículo. La información se evaluó teniendo en cuenta el diseño del estudio. Solo ocho

Tabla 6. Características de los 23 metanálisis incluidos.

Metanálisis			Variables	
Primer autor	Año	Revista	Exposición	Resultado
Weiss J [229]	2017	Ann Intern Med	Fármacos antihipertensivos	Resultados sobre daños: Deterioro cognitivo, calidad de vida, caídas, fracturas, síncope, estado funcional, hipotensión, lesión renal aguda, carga de medicación, abstinencia debido a eventos adversos
Bally M [230]	2017	BMJ	AINEs	Infarto de miocardio
Sordo L [231]	2017	BMJ	Tratamiento de sustitución de opiáceos (metadona, buprenorfina)	Mortalidad por cualquier causa y por sobredosis
Tariq R [232]	2017	JAMA Intern Med	Supresores de ácido gástrico	Infección recurrente por <i>Clostridium difficile</i>
Maruthur N [233]	2016	Ann Intern Med	Monoterapia para la diabetes (tiazolidinedionas, metformina, sulfonilureas, inhibidores DPP-4, inhibidores SGLT-2, agonistas de los receptores GLP-1) o combinaciones basadas en metformina	Mortalidad por cualquier causa, resultados macrovasculares y microvasculares, resultados intermedios (hemoglobina A1c, peso corporal, presión arterial sistólica, frecuencia cardíaca), hipoglucemia, efectos secundarios gastrointestinales, infecciones micóticas genitales, insuficiencia cardíaca congestiva
Paul S [234]	2016	Ann Intern Med	Profilaxis antiviral	Resultado primario: Reactivación del virus de la hepatitis B (VHB)
				Resultados secundarios: Hepatitis relacionada con el VHB, quimioterapia interrumpida, insuficiencia hepática aguda, mortalidad
Li L [235]	2016	BMJ	Inhibidores de la dipeptidil peptidasa-4	Insuficiencia cardíaca
				Admisiones hospitalarias por insuficiencia cardíaca
Molnar AO [236]	2015	BMJ	Fármacos inmunosupresores genéricos	Supervivencia del paciente, supervivencia del aloinjerto, rechazo agudo, eventos adversos, bioequivalencia.
Ziff OJ [237]	2015	BMJ	Digoxina	Resultado primario: Mortalidad por cualquier causa
				Resultados secundarios: Mortalidad cardiovascular; ingreso hospitalario por cualquier causa, causas cardiovasculares e insuficiencia cardíaca; accidente cerebrovascular, infarto de miocardio incidente
CGESOC [238]	2015	Lancet	Terapia hormonal (estrógeno, progestágeno)	Cáncer de ovario
Bellemain-Appaix A [239]	2014	BMJ	Tienopiridinas (clopidogrel)	Resultado primario: Mortalidad por cualquier causa, hemorragia grave Resultados secundarios: Eventos cardiovasculares mayores e infarto de miocardio, accidente cerebrovascular, revascularización urgente, trombosis de <i>stent</i>

Metanálisis			Variables	
Primer autor	Año	Revista	Exposición	Resultado
Grigoriadis S [240]	2014	BMJ	Antidepresivos (ISRSs)	Hipertensión pulmonar persistente del recién nacido
Li L [241]	2014	BMJ	Tratamientos basados en incretina	Pancreatitis
Kalil AC [242]	2014	JAMA	Vancomicina (CMI)	Mortalidad por cualquier causa
Stegeman BH [243]	2013	BMJ	Anticonceptivos orales combinados	Trombosis venosa
Maneiro JR [244]	2013	JAMA Intern Med	Agentes biológicos (abatacept, adalimumab, etanercept, golimumab, infliximab, rituximab)	Influencia de los AABs: sobre la eficacia en enfermedades inflamatorias inmunomediadas (artritis reumatoide, artritis idiopática juvenil, enfermedad inflamatoria intestinal, espondilitis anquilosante, psoriasis, artritis psoriásica u otras espondiloartropatías), en reacciones de hipersensibilidad, y sobre la concentración de fármacos biológicos; efecto del tratamiento concomitante en el desarrollo de AABs
Hartling L [245]	2012	Ann Intern Med	Antipsicóticos	Resultados primarios: Mejora de los síntomas básicos de la enfermedad (síntomas positivos y negativos y psicopatología general), eventos adversos: diabetes mellitus, muerte, discinesia tardía, síndrome metabólico mayor
				Resultados secundarios: Resultados funcionales, uso del sistema de atención sanitaria; tasas de respuesta, remisión y recaída; cumplimiento de la medicación, calidad de vida relacionada con la salud, otros resultados orientados al paciente (por ejemplo, satisfacción del paciente), otros eventos adversos: síntomas extrapiramidales, aumento de peso
Hsu J [246]	2012	Ann Intern Med	Antivirales (oseltamivir, zanamivir, amantadina, rimantadina)	Mortalidad, hospitalización, ingreso en la unidad de cuidados intensivos, ventilación mecánica e insuficiencia respiratoria, duración de la hospitalización, duración de los signos y síntomas, tiempo para retornar a la actividad normal, complicaciones, eventos adversos críticos: trastornos psicóticos graves, encefalitis, accidente cerebrovascular, convulsiones; eventos adversos importantes: dolor en las extremidades, fasciculaciones clónicas, debilidad corporal, cambios dermatológicos (urticaria o erupción cutánea); eliminación del virus de la gripe, aparición de resistencia antiviral
Caldeira D [247]	2012	BMJ	IECAs y ARAs	Incidencia de neumonía
				Mortalidad relacionada con la neumonía
MacArthur GJ [248]	2012	BMJ	Sustitución de opiáceos, desintoxicación de metadona	Infección por el VIH entre las personas que se inyectan drogas

Metanálisis			Variables	
Primer autor	Año	Revista	Exposición	Resultado
Mantha S [249]	2012	BMJ	Anticoncepción con progestina sola	Eventos tromboembólicos venosos
Silvain J [250]	2012	BMJ	Enoxaparina, heparina no fraccionada	Resultado primario: Mortalidad, hemorragia grave Resultados secundarios: resultado compuesto de eventos isquémicos (muerte o infarto de miocardio), complicaciones del infarto de miocardio, hemorragia leve
McKnight RF [251]	2012	Lancet	Litio	Función renal, función tiroidea, función paratiroidea, trastornos del cabello, trastornos de la piel, peso corporal, teratogenicidad

Abreviaturas: AABs, anticuerpos contra agentes biológicos; IECAs, inhibidores de la enzima convertidora de angiotensina; Ann Intern Med, *Annals of Internal Medicine*; ARAs, antagonistas de los receptores de angiotensina; BMJ, *British Medical Journal*; DPP-4, dipeptidil peptidasa-4; GLP-1, péptido-1 similar al glucagón; JAMA, *Journal of the American Medical Association*; CMI, concentración mínima inhibitoria; AINEs, antiinflamatorios no esteroideos; SGLT-2, cotransportador sodio-glucosa tipo-2; ISRSs, inhibidores selectivos de la recaptación de serotonina.

metanálisis [230, 233, 235, 240, 241, 243, 247, 250] informaron sobre la fuente de datos, tres de ellos [240, 243, 247] reportaron fuentes mixtas tanto para la evaluación de la exposición como los resultados. Cinco metanálisis [230, 233, 235, 241, 250] reportaron solo fuentes secundarias para la evaluación de la exposición, tres de ellos [230, 233, 250] reportando también solo fuentes secundarias para la evaluación de los resultados, mientras que en los otros dos [235, 241] se informaron solo fuentes primarias y mixtas para la evaluación de los resultados respectivamente.

4.3.2. Fuente de datos en el análisis de la heterogeneidad.

Todos los metanálisis menos dos [229, 251] realizaron análisis de subgrupos y/o de sensibilidad. Aunque tres de ellos [232, 243, 245] consideraron los métodos de evaluación de los resultados –tipo de prueba diagnóstica utilizada para la infección por *Clostridium difficile*, método de confirmación del diagnóstico de trombosis venosa y tipo de escala para la evaluación de los síntomas de psicosis, respectivamente– como variables de estratificación, solo el segundo se refirió al origen de los datos. Solo cinco metanálisis [231, 237, 242, 244, 248] incluyeron análisis de metarregresión para describir la heterogeneidad, ninguno de los cuales consideró la fuente de los datos como una variable explicativa. En la tabla 8 se muestran otros resultados acerca de la inclusión de la fuente de datos como variable en el análisis de la heterogeneidad.

Finalmente, se evaluó si la influencia del origen de los datos en las conclusiones de los metanálisis había sido discutida por sus respectivos autores. Se encontró que solo en cuatro metanálisis [230, 240, 241, 243] se señalaron limitaciones derivadas del tipo de fuente de datos utilizada.

Tabla 7. Reporte de la fuente de datos en las tablas de extracción de datos de los metanálisis incluidos.

Metanálisis (MA)	Evaluación de la exposición							Evaluación del resultado						
	Fuente de datos presentada en el MA	Estudios de cohorte (n)			Estudios de casos y controles (n)			Fuente de datos presentada en el MA	Estudios de cohorte (n)			Estudios de casos y controles (n)		
		1 ^{aria}	2 ^{aria}	NR	1 ^{aria}	2 ^{aria}	NR		1 ^{aria}	2 ^{aria}	NR	1 ^{aria}	2 ^{aria}	NR
Weiss J [229] Harms outcomes	No	-	-	-	-	-	-	No	-	-	-	-	-	-
Bally M [230]	Sí	0	3 [†]	0	0	1	0	Sí	0	3 [†]	0	0	1	0
Sordo L [231]	No*	-	-	-	-	-	-	No*	-	-	-	-	-	-
Tariq R [232]	No* [‡]	-	-	-	-	-	-	No*	-	-	-	-	-	-
Maruthur N [233]	Sí [§]	0	3	0	-	-	-	Sí [§]	0	3	0	-	-	-
Paul S [234]	No*	-	-	-	-	-	-	No*	-	-	-	-	-	-
Li L [235] Heart failure	Sí	0	1	2	0	0	1	Sí	1	0	2	0	0	1
Li L [235] Hospital admissions for heart failure	Sí	0	0	6	0	0	2	Sí	3	0	3	0	0	2
Molnar AO [236]	No*	-	-	-	-	-	-	No*	-	-	-	-	-	-
Ziff OJ [237]	No*	-	-	-	-	-	-	No*	-	-	-	-	-	-
CGESOC [238]	No	-	-	-	-	-	-	No	-	-	-	-	-	-
Bellemain-Appaix A [239]	No*	-	-	-	-	-	-	No*	-	-	-	-	-	-
Grigoriadis S [240]	Sí	2	3	0	1	1	0	Sí	4	1	0	2	0	0
Li L [241]	Sí	0	1	2	0	1	1	Sí	1	2	0	0	0	2
Kalil AC [242]	No	-	-	-	-	-	-	No	-	-	-	-	-	-
Stegeman BH [243]	Sí	0	9	0	8	8	1	Sí	4	5	0	5	12	0
Maneiro JR [244]	No*	-	-	-	-	-	-	No*	-	-	-	-	-	-
Hartling L [245]	No	-	-	-	-	-	-	No	-	-	-	-	-	-
Hsu J [246]	No*	-	-	-	-	-	-	No*	-	-	-	-	-	-
Caldeira D [247]	Sí	2	2	7	0	7	1	Sí	0	1	10	3	1	4
MacArthur GJ [248]	No*	-	-	-	-	-	-	No*	-	-	-	-	-	-
Mantha S [249]	No	-	-	-	-	-	-	No	-	-	-	-	-	-
Silvain J [250]	Sí	0	7	0	-	-	-	Sí	0	7	0	-	-	-
McKnight RF [251]	No	-	-	-	-	-	-	No	-	-	-	-	-	-

Abreviaturas: 1^{aria}, Número de estudios individuales en cada MA basados en fuentes de datos primarias; 2^{aria}, Número de estudios individuales en cada MA basados en fuentes de datos secundarias; NR, Número de estudios individuales en cada MA con fuentes de datos no reportadas.

* Aunque el metanálisis muestra los resultados de la evaluación de la calidad metodológica basada en una escala estandarizada, no indica el tipo de fuente de datos utilizada para cada estudio observacional individual incluido en el metanálisis.

† Cohorte con análisis caso-control anidado.

‡ El metanálisis reporta que la mayoría de los estudios observacionales incluidos evaluaron la exposición a los fármacos a través de una revisión de los registros médicos.

§ El metanálisis solo reporta datos de estudios observacionales de alta calidad.

Tabla 8. Inclusión de la fuente de datos como variable en el análisis de la heterogeneidad de los metanálisis incluidos.

Metanálisis	Análisis de subgrupos/ sensibilidad				Análisis de metarregresión			
	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida
Weiss J [229] Resultados sobre daños	-	-	-	No	-	-	-	No
Bally M [230]	Momento de la exposición a los AINEs, dosis y duración del tratamiento, tratamiento farmacológico concomitante	Comorbilidades	Modelo estadístico alternativo, motivo de exclusión	No	-	-	-	No
Sordo L [231]	Intervalo de tiempo dentro y fuera del tratamiento de sustitución de opiáceos	-	Modelo estadístico alternativo	No	Proveedor del tratamiento, prevalencia de la inyección de opiáceos, dosis promedio de metadona	-	Edad media, porcentaje de hombres, ubicación, porcentaje de inducción de pacientes hospitalizados, porcentaje de pérdida durante el seguimiento, punto intermedio del período de seguimiento	No

Metanálisis	Análisis de subgrupos/ sensibilidad				Análisis de metarregresión			
	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida
Tariq R [232]	Tipo de supresor de ácido gástrico (IBP y H2B reportados juntos, PPI solo o H2B solo)	Definición del caso (intervalo de tiempo de recurrencia: dentro de 60 días frente a 90 días), tipo de ensayo diagnóstico utilizado para la infección por <i>Clostridium difficile</i>	Diseño del estudio, entorno del estudio (pacientes hospitalizados frente a ambulatorios), ajuste de datos	No	-	-	-	No
Maruthur N [233]	Modo de terapia	-	-	No	-	-	-	No
Paul S [234] Resultado primario	-	Infección crónica o resuelta por el virus de la hepatitis B	Subtipo de tumor y quimioterapia, modelo estadístico alternativo, calidad de diseño	No	-	-	-	No
Paul S [234] Resultados secundarios	-	-	Modelo estadístico alternativo, calidad del diseño	No	-	-	-	No
Li L [235]	Tipo de control, modo de terapia, fármacos individuales	-	Duración del seguimiento, tipo de diseño	No	-	-	-	No

Metanálisis	Análisis de subgrupos/ sensibilidad				Análisis de metarregresión			
	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida
Molnar AO [236]	-	-	Tipo de diseño	No	-	-	-	No
Ziff OJ [237] Resultado primario	-	-	Ajuste de datos, tipo de población	No	Diferencia entre los brazos de digoxina y de control al inicio del estudio: Diabetes, hipertensión, diuréticos, fármacos antiarrítmicos	-	Puntuación de sesgo resumida, variable de nivel de estudio de línea de base: Año de publicación, edad, sexo, infarto de miocardio previo	No
Ziff OJ [237] Resultados secundarios	-	-	-	No	-	-	-	No
CGESOC [238]	Duración del uso en usuarias actuales y pasadas de terapia hormonal, tipos de terapia hormonal	Histología tumoral y potencial maligno del tumor	Diseño del estudio, región geográfica, edad en la que se utilizó por primera vez la terapia hormonal, edad de la menarquia, número de partos, uso de anticonceptivos orales, altura, índice de masa corporal, consumo de alcohol, tabaquismo, madre o hermana con cáncer de ovario o de mama, histerectomía	No	-	-	-	No

Metanálisis	Análisis de subgrupos/ sensibilidad				Análisis de metarregresión			
	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida
Bellemain-Appaix A [239]	Dosis de clopidogrel	Tipos de intervención coronaria percutánea	Tipo de diseño	No	-	-	-	No
Grigoriadis S [240]	Momento de la exposición a los ISRSs	-	Diseño del estudio, malformaciones congénitas, control, aspiración de meconio	No	-	-	-	No
Li L [241]	Tipo de agentes de la incretina, tipo de control, modo de terapia, agentes individuales de incretina	-	Duración del seguimiento, medida del efecto alternativo, modelo estadístico alternativo	No	-	-	-	No
Kalil AC [242]	Diferentes cortes de CMI, tipo de ensayo	Mortalidad hospitalaria o dentro de 30 días	Año de publicación, calidad del diseño	No	Corte de CMI de vancomicina, exposición a vancomicina en los 6 meses anteriores, niveles pico de vancomicina, proporción de pacientes que recibieron tratamiento con vancomicina	Mortalidad de control, puntuación APACHE II, puntuación de Charlson, duración de la bacteriemia, proporción de pacientes con endocarditis, proporción de pacientes ubicados en la UCI	Edad	No

Metanálisis	Análisis de subgrupos/ sensibilidad				Análisis de metarregresión			
	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida
Stegeman BH [243]	Generación de progestágeno utilizado en anticonceptivos orales combinados, píldoras anticonceptivas orales combinadas	Método de confirmación del diagnóstico	Fuente de financiación, diseño del estudio	Sí (resultado)	-	-	-	No
Maneiro JR [244]	Tipo de agente biológico, tratamiento concomitante (monoterapia frente a terapia combinada), uso previo de inhibidores del FNT	Tipo de enfermedad	Duración del seguimiento, calidad de los datos, diseño del estudio, nivel de evidencia de los estudios	No	Tipo de agente biológico, uso previo de inhibidores del FNT, método de medición de anticuerpos, tipo de anticuerpo monoclonal anti-FNT	Tipo de enfermedad, tiempo de duración de la enfermedad, tiempo para evaluar la respuesta	Edad y sexo de los pacientes, número de participantes, duración del seguimiento, calidad de los datos, diseño del estudio, nivel de evidencia de los estudios	No
Hartling L [245] Resultados primarios	Tipo de comparación de fármacos	Tipo de escala para la evaluación de los síntomas y la calidad de vida	-	No	-	-	-	No
Hartling L [245] Resultados secundarios	-	-	-	No	-	-	-	No

Metanálisis	Análisis de subgrupos/ sensibilidad				Análisis de metarregresión			
	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida
Hsu J [246]	Fármacos individuales, dosis de antivirales, momento del tratamiento	-	Ajuste de datos, gripe confirmada, tipo de gripe A frente a B, gripe pandémica frente a estacional, severidad de la gripe, edad, embarazo, riesgo inicial (por ejemplo, inmunocomprometido), contexto, conflicto de financiación	No	-	-	-	No
Caldeira D [247] Incidencia de neumonía	-	-	Diseño del estudio, accidente cerebrovascular previo, insuficiencia cardíaca, enfermedad renal crónica, pacientes no asiáticos	No	-	-	-	No
Caldeira D [247] Mortalidad relacionada con neumonía	-	-	Diseño del estudio	No	-	-	-	No

Metanálisis	Análisis de subgrupos/ sensibilidad				Análisis de metarregresión			
	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida	Variables relacionadas con la exposición	Variables relacionadas con el resultado	Otras variables	Tipo de fuente de datos incluida
MacArthur GJ [248]	Duración de la exposición al tratamiento de sustitución de opiáceos	-	Ajuste de datos, región geográfica, lugar de reclutamiento, incentivos económicos, porcentaje de mujeres participantes, porcentaje de personas pertenecientes a minorías étnicas	No	Exposición al tratamiento de mantenimiento con metadona solo al inicio del estudio	-	Inclusión solo de los estudios con menor riesgo de sesgo, inclusión solo de los estudios que midieron una tasa de incidencia, exclusión de los estudios que no se ajustaron a los factores de confusión	No
Mantha S [249]	Vía de administración	-	Ajuste de datos	No	-	-	-	No
Silvain J [250]	Vía de administración	-	Tipos de intervención coronaria percutánea, publicación del estudio, tamaño del estudio, calidad del diseño	No	-	-	-	No
McKnight RF [251]	-	-	-	No	-	-	-	No

Abreviaturas: APACHE, *Acute Physiology and Chronic Health Evaluation*; CMI, concentración mínima inhibitoria; ISRSs, inhibidores selectivos de la recaptación de serotonina; FNT, factor de necrosis tumoral; UCI, unidad de cuidados intensivos.



5. DISCUSIÓN.

5.1. EL SESGO EN ESTUDIOS FARMACOEPIDEMIOLÓGICOS QUE UTILIZAN BASES DE DATOS SECUNDARIAS: UNA REVISIÓN EXPLORATORIA (OBJETIVO 1 - ANEXO 1).

Esta es la primera revisión estructurada conocida que explora los sesgos potenciales en estudios observacionales que utilizan bases de datos farmacoepidemiológicas. Los resultados de esta revisión sugieren que existe una preocupación creciente en la literatura científica acerca de la identificación, descripción y control de dichos sesgos. Esto no debe pasarse por alto, ya que los estudios observacionales que utilizan bases de datos epidemiológicas ofrecen actualmente una excelente oportunidad para la investigación médica. Los resultados de estos estudios deben ser válidos y aplicables a la toma de decisiones sobre seguridad y eficacia. Por lo tanto, es de suma importancia que se tengan debidamente en cuenta estos sesgos para garantizar que se controlan correctamente.

El sesgo de confusión como tal, o en cualquiera de sus diversas formas de presentación, se menciona en casi dos tercios de los artículos incluidos en la revisión (ver Tabla 5 para consultar referencias). El control adecuado de los factores de confusión plantea un desafío en los estudios que utilizan bases de datos de atención sanitaria, ya que éstas no fueron diseñadas para realizar estudios epidemiológicos. La ausencia o mala calidad de los datos sobre posibles factores de confusión en las bases de datos secundarias (por ejemplo, medicamentos de venta libre, fragilidad del sujeto, hábito de fumar) es un fenómeno frecuente [37, 143, 168, 252], que dificulta o incluso imposibilita el ajuste por tales factores, a fin de controlar la confusión [39].

Cuando se hayan recogido datos sobre variables de confusión, los artículos revisados proponen diferentes métodos de control: (i) en la etapa de diseño, mediante la aplicación de criterios de restricción, métodos de emparejamiento o la implementación de un diseño de nuevo usuario (ver más adelante, depleción de susceptibles); y (ii) en la etapa de análisis, mediante la estratificación de los pacientes entre los grupos de tratamiento de acuerdo con los factores relevantes, o técnicas de regresión multivariada, incluyendo estos factores de confusión como variables independientes en los modelos de regresión. En los casos en que el número de variables es muy alto, puede ser interesante ajustar por la puntuación de riesgo de enfermedad [80] o la puntuación de propensión a recibir tratamiento [82, 253].

Entre los estudios que tratan el tema de la confusión en farmacoepidemiología, el tipo de confusión más comúnmente descrito es la **confusión por indicación para el tratamiento** (la decisión de tratamiento se asocia con una indicación, que a su vez es un factor de riesgo para la enfermedad), mencionada en un tercio de los artículos revisados (Tabla 5). La confusión por indicación, también conocida como sesgo de canalización, está estrechamente relacionada con el sesgo de selección [171]. Algunos métodos analíticos de control útiles propuestos incluyen la separación de los efectos de un fármaco tomado en diferentes momentos [30], el análisis de sensibilidad para factores de confusión no medidos (ver más adelante) y el uso de variables instrumentales [101]. De acuerdo con la literatura

revisada, parece existir un acuerdo general de que los métodos convencionales para el control de los factores de confusión son inadecuados para controlar la **confusión dependiente del tiempo** (mencionada en el 6.0% de los artículos revisados, ver tabla 5). La estimación-g [104] y los modelos estructurales marginales [105] son métodos alternativos para lograr dicho control.

Más de una cuarta parte de los artículos incluidos en la revisión consideran la ausencia de datos de calidad para el control de posibles variables de confusión como una limitación importante de los estudios farmacoepidemiológicos observacionales que utilizan bases de datos secundarias (Tabla 5). Las estrategias propuestas para el control de las **variables no medidas** incluyen la realización de análisis de sensibilidad y el uso de información externa a la base de datos [96, 254, 255]. También se han utilizado técnicas de variables instrumentales, medidas *proxy* y puntuaciones de propensión, excluyendo del análisis a sujetos tratados y no tratados con valores extremos [256]. En la etapa de diseño, los diseños de casos cruzados, en los que cada participante recibe todos los tratamientos que se están investigando, pero en momentos diferentes [71], y la restricción a un grupo de comparación activa, pueden ser útiles. El diseño del comparador activo emula el diseño de un ensayo controlado aleatorizado de comparación directa. En lugar de utilizar un grupo de no usuarios, el fármaco de interés se compara con otro fármaco comúnmente utilizado para la misma indicación. Al asegurar que los grupos de tratamiento tengan características similares, este diseño ayuda potencialmente a mitigar los factores de confusión medidos y no medidos [257]. En cualquier caso, con la excepción de los diseños cruzados, en los que el orden en el que un participante del estudio recibe los tratamientos es aleatorio, el control de las variables no medidas nunca será óptimo o, en el mejor de los casos, nunca se podrá garantizar que lo sea. Pero incluso en este caso, el diseño cruzado puede verse afectado por confusión dependiente del tiempo.

En este contexto, Hernán ha propuesto un nuevo enfoque basado en el uso de datos observacionales de una gran base de datos de salud para emular un hipotético ensayo aleatorio (*target trial*) [258]. Aunque el ensayo diana emulado ayuda a evitar las dificultades metodológicas habituales, el ajuste apropiado por los factores de confusión dependientes del tiempo sigue siendo crítico [259].

A diferencia de los ensayos clínicos, una ventaja de los estudios farmacoepidemiológicos observacionales en los que las poblaciones de estudio se construyen sobre la base de grandes bases de datos de asistencia sanitaria, es la inclusión de **pacientes frágiles**. Sin embargo, algunos autores han argumentado que debido al hecho de que la fragilidad es difícil de medir y un fuerte factor de riesgo para resultados desfavorables, conducirá a una confusión residual no medida y, posiblemente, a resultados paradójicos [34, 38]. La fragilidad es un ejemplo de una variable de confusión no medida [37, 143].

Alrededor del 5% de los artículos revisados mencionan el **efecto del usuario sano** (Tabla 5), que consiste en un tipo de confusión generada porque los pacientes con conductas más saludables generalmente demandan atención médica con mayor frecuencia para tratamientos preventivos o enfermedades crónicas asintomáticas. Estos pacientes también es más probable que se adhieran mejor al tratamiento. En consecuencia, parte de la aparente eficacia/seguridad del tratamiento se debe, no al tratamiento en sí, sino a los comportamientos más saludables que se asocian con quienes lo toman [39, 40]. En estudios observacionales que

utilizan bases de datos farmacoepidemiológicas, estos tipos de comportamiento rara vez se miden, lo que hace muy difícil controlar su efecto [260].

Casi la mitad de los artículos incluidos en la revisión mencionan algún tipo de sesgo de selección. Dentro de esta categoría cabe destacar el **sesgo protopático**. Aunque este sesgo en concreto no se menciona demasiado en nuestra revisión (3,4%, ver tabla 5), posiblemente porque es inusual que el tratamiento se asocie con estados subclínicos y/o síntomas tempranos de la enfermedad, el impacto de este sesgo puede ser importante. Sin embargo, controlar el sesgo protopático no es fácil, ya que no es un sesgo de confusión y, por lo tanto, las técnicas de ajuste no resultan útiles. En este caso, debemos recurrir a la restricción del grupo de exposición a pacientes con indicaciones que no están relacionadas con los estados iniciales de la enfermedad en estudio. Otra opción para controlar el sesgo protopático es utilizar el concepto de tiempo demora (*lag-time*) para definir la ventana etiológica en la que se evalúa la exposición al fármaco [261].

El consumo de medicamentos en condiciones reales está sujeto a variaciones importantes (por ejemplo, variaciones en la dosis, interrupciones del tratamiento, abandonos), especialmente en el tratamiento de enfermedades crónicas. Esta variabilidad puede deberse a cambios en la enfermedad (aumento o disminución de la gravedad) o en el efecto del fármaco (eventos adversos o interacciones). El enfoque tradicional a través de un análisis “por intención de tratar”, en el que se censura a los sujetos que interrumpen su tratamiento durante el seguimiento, puede introducir sesgos, ya que los sujetos censurados (perdidos durante el seguimiento) tienen sistemáticamente un mayor o menor riesgo de desarrollar el resultado [140, 142]. En la práctica, esta **censura informativa** (mencionada en solo el 2,6% de los artículos revisados, ver tabla 5) conduce a un sesgo de selección. Por ejemplo, si los efectos clínicos esperados no se cumplen, o el tratamiento se suspende o se modifica. El sesgo consiste en seleccionar para el análisis los datos de los pacientes para los cuales el tratamiento produce el resultado esperado [46]. Este sesgo puede identificarse mediante un análisis de sensibilidad. En este sentido, el uso de bases de datos representa una ventaja importante, ya que la información sobre el resultado puede estar disponible incluso cuando se haya suspendido el tratamiento. Para controlar el sesgo introducido por una exposición al fármaco que varía con el tiempo, podría ser útil considerar esa exposición como una variable dependiente del tiempo en un modelo de regresión multivariante apropiado. También se han propuesto procedimientos basados en la ponderación por probabilidad inversa de censura [262].

A juzgar por el número de artículos que lo mencionan (10.3%), se le ha dado mayor importancia a otro tipo de sesgo de selección conocido como **depleción de susceptibles**, el cual es causado por la inclusión en el estudio tanto de usuarios tratados prevalentes como incidentes (Tabla 5). Los usuarios prevalentes (“supervivientes” del primer período de tratamiento) pueden no tener el mismo riesgo de un evento adverso que los usuarios incidentes (nuevos), es decir, aquellos que toleran el fármaco continúan usándolo y aquéllos que no lo toleran (susceptibles al evento adverso) han dejado de usarlo. Este sesgo puede evitarse en la fase de diseño del estudio limitando el seguimiento a los nuevos usuarios [48]. El diseño de nuevo usuario permite medir los factores de confusión potenciales justo antes del inicio del seguimiento. De esta manera, estos factores de confusión no se verán afectados por el tratamiento. El ajuste por las diferencias entre grupos de tratamiento utilizará los valores basales de los factores de confusión [49].

Aparte de asegurar un ajuste apropiado de la confusión, el diseño de nuevo usuario reduce potencialmente el sesgo de tiempo inmortal (ver abajo) cuando se combina con el diseño de comparador activo, mediante la implementación de definiciones similares de la fecha índice en todos los grupos de comparación [257]. El diseño de nuevo usuario combinado con el diseño de comparador activo también puede reducir la confusión por indicación y otras características no medidas del paciente (por ejemplo, fragilidad, usuario sano) en la fase de diseño [263].

Como sugieren los resultados, uno de los mayores desafíos en el análisis de los datos observacionales es el problema de la **falta de datos** [23], que se menciona en casi uno de cada cinco artículos incluidos en la revisión (Tabla 5). Si la probabilidad de que falte una observación es independiente tanto de los datos observados como de los ausentes, se supone que los casos completos son una muestra aleatoria del conjunto de datos completo (es decir, “ausencia completamente aleatoria” [264]). En este caso, la eliminación de casos con datos ausentes puede dar estimaciones no segadas. Sin embargo, en el análisis multivariado, las observaciones (o sujetos) se eliminan cuando faltan datos de una variable incluida en el modelo. Como consecuencia, las observaciones con valores perdidos pueden llevar a una reducción sustancial del tamaño de la muestra. Si esta falta de información se asocia a una característica importante, como la gravedad o la fragilidad, se produce un efecto equivalente al sesgo de selección.

A veces se supone que la probabilidad de que falte una observación puede ser predicha por variables que se miden previamente, pero que no dependen de variables no medidas (es decir, “ausencia aleatoria” [264]). Es decir, la probabilidad de abandono dependerá de los valores observados. Aunque el análisis estándar de los casos disponibles está potencialmente sesgado en este caso, existen métodos que pueden proporcionar un análisis válido, pero éstos requieren un modelo estadístico adicional apropiado.

En ambas circunstancias descritas anteriormente, los métodos basados en la verosimilitud (por ejemplo, modelos mixtos), en los que los datos ausentes pueden estimarse utilizando la distribución condicional de las otras variables, pueden ser útiles para controlar el sesgo [265]. Existen técnicas alternativas, como la imputación múltiple, que preservan la variabilidad natural de los datos [79] e incorporan la incertidumbre por falta de datos [266], con lo que se obtienen resultados similares. La ponderación por probabilidad inversa (cuando los casos completos se ponderan por la inversa de su probabilidad de ser un caso completo) es también un método comúnmente utilizado para reducir este sesgo. Mientras que la imputación múltiple requiere un modelo para la distribución de los datos que faltan dados los datos observados, la ponderación inversa de la probabilidad requiere un modelo para la probabilidad de que un individuo sea un caso completo [267]. En cualquier caso, es importante que se incluyan en el modelo todas las covariables de las que depende la falta de datos.

Por el contrario, si el hecho de que una observación esté ausente se predice mediante variables no medidas, como el resultado de interés (es decir, “ausencia no aleatoria”, a veces llamada “falta de respuesta no ignorable” o “ausencia informativa”), entonces ninguna técnica estadística puede dar estimaciones no segadas. Cuando la falta de datos no puede ser modelada empíricamente, el enfoque recomendado es realizar análisis de sensibilidad para determinar el grado de ausencia [268].

Después de la confusión por indicación y la confusión no medida/residual, los resultados muestran que el sesgo más frecuentemente descrito en los estudios que utilizan bases de datos secundarias de asistencia sanitaria es el debido a errores sistemáticos de clasificación que distorsionan la asociación entre el tratamiento y el resultado. La **clasificación errónea de la exposición o del resultado**, que se menciona en casi la mitad de los artículos incluidos en la revisión (Tabla 5), puede dar lugar a sesgos de medición y heterogeneidad [55, 193, 252]. Para evitarlo, se debe realizar primero un estudio de validación de estas variables, seguido de la realización de un análisis de sensibilidad o la aplicación de técnicas de regresión [269]. Los registros médicos se consideran normalmente el estándar de oro o la referencia para las variables de resultados intermedios y finales, pero muestran limitaciones en el registro de todos los medicamentos tomados por los pacientes [27]. Si bien los registros de dispensación son más detallados en la medición de la exposición (aunque no registran el consumo de venta libre o el copago a nivel individual), carecen de variables de resultado [1, 2, 21, 270]. Por lo tanto, es importante vincular ambos tipos de fuentes de datos [271, 272] y considerar, cuando sea necesario, el uso de datos adicionales recogidos expresamente con fines de investigación [37, 177, 189] para evitar errores que puedan generar conclusiones equivocadas [273, 274].

La última categoría de sesgo identificada fue la relacionada con el tiempo. No obstante, debe tenerse en cuenta que el mecanismo que subyace en la generación de un sesgo relacionado con el tiempo puede estar estrechamente relacionado con las otras categorías principales descritas (es decir, sesgo de confusión, selección o medición). Con diferencia, el sesgo temporal más frecuentemente descrito es el **sesgo de tiempo inmortal**, que se menciona en uno de cada cuatro artículos revisados (Tabla 5). El sesgo de tiempo inmortal (donde el seguimiento incluye un período de tiempo durante el cual el evento de estudio no puede ocurrir, o se excluye del análisis debido a una definición incorrecta del inicio del seguimiento) resurgió con varios estudios observacionales que reportaron efectos sorprendentemente beneficiosos de algunos fármacos [275, 276] y está siendo descrito cada vez más en estudios de cohortes a partir de bases de datos farmacoepidemiológicas [60, 62, 63]. Suissa advierte sobre el riesgo de obtener conclusiones absurdas si se utilizan métodos inadecuados en el análisis de datos [60, 63, 219-224]. Para evitarlo, se debe considerar todo el tiempo de seguimiento, incluido el que precede al inicio de la exposición; y la exposición durante el tiempo inmortal se debe clasificar correctamente [277]. Aplicando un modelo de Cox con exposiciones dependientes del tiempo se pueden obtener estimaciones más fiables [63, 218, 278].

La aparición de grandes bases de datos de asistencia sanitaria ha provocado cambios drásticos en la farmacoepidemiología. Debido a la captura rutinaria y automatizada de los datos de prescripción y dispensación de medicamentos que se utilizan con fines administrativos, junto con la implementación de historias clínicas electrónicas, las bases de datos secundarias han generado enormes posibilidades y expectativas acerca de su potencial. Esto ocurre, además, en un momento en que se reconoce que los ensayos clínicos no pueden responder a preguntas sobre la eficacia y la seguridad de los tratamientos en la práctica clínica.

En apariencia, las bases de datos secundarias ofrecen la posibilidad de realizar estudios rápidamente, a bajo coste, con enormes tamaños de muestra, datos objetivos y seguimiento a largo plazo. No obstante, sus limitaciones no deben ser ignoradas. Los sesgos

inherentes a este tipo de fuentes de datos más descritos en la literatura son la confusión por indicación, la confusión no medida/residual, la clasificación errónea del resultado y el sesgo de tiempo inmortal.

Para controlar estas situaciones se debe considerar la aplicación de diseños metodológicos apropiados y técnicas de análisis estadístico. Las principales estrategias se resumen en la tabla 9. En general, antes de iniciar una investigación utilizando bases de datos secundarias, los investigadores deben evaluar en detalle las fuentes de datos disponibles, centrándose en el propósito para el que fueron creadas, y así tomar conciencia de su potencial de sesgo. La vinculación de los registros médicos con las bases de datos administrativas es útil para minimizar el riesgo de sesgo, así como para complementar o validar los datos secundarios con datos primarios (es decir, recopilados a partir de métodos *ad hoc*) cuando se cuestiona la integridad o la calidad de los datos originales.

5.1.1. Limitaciones del método.

Esta *scoping review* presenta las limitaciones inherentes a este tipo de diseño de estudio. A diferencia de las revisiones sistemáticas clásicas, cuyo objetivo es proporcionar respuestas a una pregunta de investigación claramente definida, es menos probable que este tipo de estudios exploren preguntas de investigación muy específicas ni, en consecuencia, evalúen la calidad de los estudios incluidos [279]. En este sentido, no se puede descartar un potencial sesgo del revisor en la evaluación de los criterios de restricción, ya que éstos no se basan en una calidad mensurable de las referencias identificadas. Sin embargo, no creemos que esto obstaculice el propósito y las conclusiones de la revisión.

Debido a la naturaleza exploratoria de esta revisión, su propósito no fue obtener toda la evidencia disponible sobre un tema específico, sino más bien de un subconjunto de la literatura sobre un tema amplio (sesgo en los estudios farmacoepidemiológicos observacionales basados en fuentes de datos secundarias), en el que pueden ser aplicables muchos diseños de estudio diferentes (ensayos de opinión, revisiones metodológicas, análisis, cartas al editor o retracciones). Aunque se utilizó una estrategia de búsqueda amplia, es posible que se hayan omitido algunos estudios relevantes. Por lo tanto, no se puede descartar la existencia de algún sesgo de selección. Además, la estrategia de búsqueda en sí misma, diseñada intencionadamente para identificar artículos que destacan las limitaciones de las bases de datos secundarias, no permite una comparación no sesgada con los artículos que pueden mostrar la ventaja de las bases de datos secundarias.

Dadas las limitaciones anteriores, y el hecho de que la información sobre el sesgo se obtuvo a partir de la descripción proporcionada por los autores originales, otra limitación estaría relacionada con la cuantificación de cada tipo de sesgo, que debe ser interpretada como una medida aproximada del impacto del sesgo en la literatura publicada (es decir, qué es de lo que más se habla), pero no como la probabilidad de ocurrencia (o detección) del sesgo en la población de estudios farmacoepidemiológicos que utilizan bases de datos secundarias, ya que puede estar influenciada por la facilidad de descripción de ese sesgo específico, o por el interés que el sesgo haya suscitado en los estudios de los autores más prolíficos en el campo (por ejemplo, el sesgo de tiempo inmortal). Por lo tanto, es posible que exista cierto grado de clasificación errónea de algunos sesgos.

Tabla 9. Principales estrategias de control de sesgos en estudios observacionales de bases de datos farmacoepidemiológicas.

Categoría	Estrategias de control
Confusión	
Confusión medida	<ul style="list-style-type: none"> - Análisis multivariante - Restricción* - Estratificación - Emparejamiento - Diseño de usuarios nuevos - Puntuación de propensión - Ensayos aleatorios simples a gran escala - Metaanálisis de ensayos clínicos <p><i>* Confusión por indicación: Restringir el grupo no tratado a una población con la misma indicación, o limitar la participación a pacientes sin factor de riesgo para el efecto que pudo haber condicionado el tratamiento</i></p>
Confusión tiempo-dependiente	<ul style="list-style-type: none"> - Estimación-g - Modelos estructurales marginales
Confusión no medida	<ul style="list-style-type: none"> - Diseño de casos cruzados - Exclusión asimétrica de pacientes con valores extremos de puntuaciones de propensión - Variables instrumentales - Medidas sucedáneas (<i>proxy</i>) - Restricción (grupo de comparación activo) - Análisis de sensibilidad - Estudio de validación + ajuste externo
Sesgo de selección	
Sesgo protopático	<ul style="list-style-type: none"> - Restricción (por ejemplo, restringir el grupo no tratado a una población con la misma indicación, o restringir el grupo tratado a una población con una indicación que no sea un estadio subclínico de la enfermedad) - Excluir de la ventana etiológica un período de tiempo específico antes de la fecha de diagnóstico de la enfermedad (<i>lag-time</i>)
Pérdidas durante el seguimiento (censura informativa)	<ul style="list-style-type: none"> - Inclusión de las variables que afectan a la censura y los tiempos de evento en el modelo de regresión multivariante - Ponderación por probabilidad inversa de censura - Análisis de sensibilidad
Depleción de susceptibles (sesgo del usuario prevalente)	<ul style="list-style-type: none"> - Diseño de usuarios nuevos - Metaanálisis de ensayos clínicos
Falta de datos	<ul style="list-style-type: none"> - Sustituir cada observación ausente por un valor medio a partir de los valores observados de la variable, o el valor predicho a partir de un modelo de regresión - Métodos de imputación (por ejemplo, imputación múltiple) - Métodos basados en la verosimilitud - Ponderación por probabilidad inversa

Categoría	Estrategias de control
Sesgo de medición	
Sesgo de mala clasificación	<ul style="list-style-type: none"> - Estudio de validación (exposición/resultado/factores de confusión) + (análisis de sensibilidad/técnicas de control de la mala clasificación mediante regresión multivariante)
Sesgo relacionado con el tiempo	
Sesgo de tiempo inmortal	<ul style="list-style-type: none"> - Análisis de datos con procedimientos que tengan en cuenta la exposición tiempo-dependiente dentro de una cohorte - Trasladar el inicio del tratamiento al final del período de tiempo inmortal en ambos grupos
Sesgo de tiempo inconmensurable	<ul style="list-style-type: none"> - Análisis de los datos teniendo en cuenta el período de exposición variable en el tiempo
Sesgo de ventana temporal	<ul style="list-style-type: none"> - Tener en cuenta la duración del tratamiento en la selección de los controles - Análisis dependiente del tiempo
Sesgo de desfase temporal	<ul style="list-style-type: none"> - Comparar pacientes en el mismo estadio de la enfermedad



5.2. DISCUSIÓN METODOLÓGICA ACERCA DE LAS DIFERENTES TÉCNICAS DE METANÁLISIS PARA RESUMIR LA EVIDENCIA SOBRE SEGURIDAD DE MEDICAMENTOS (OBJETIVO 2 - ANEXO 2).

Los ensayos clínicos aleatorizados (ECAs) se consideran el estándar de oro para evaluar la efectividad de los fármacos, aunque su capacidad para detectar problemas de seguridad es limitada. Esta limitación se debe al hecho de que los ECAs evalúan los medicamentos en muestras de pequeño tamaño durante un período de seguimiento relativamente corto. En consecuencia, se conoce poco sobre el riesgo de eventos adversos que pueden ser atribuibles a medicamentos que, a pesar de tener una baja incidencia, pueden, sin embargo, ser de gran relevancia para la salud pública cuando son graves y el medicamento en cuestión es consumido por millones de personas [280].

Además, las poblaciones de los estudios controlados aleatorizados tienden a estar muy estrechamente definidas y excluyen a los pacientes que reciben otros tratamientos, lo que hace imposible identificar los problemas de seguridad en condiciones de la vida real, como las interacciones con fármacos concomitantes o las reacciones farmacológicas a edades tempranas o avanzadas [281]. Para resolver los problemas de validez externa de los ECAs se ha sugerido la realización de ensayos simples de gran tamaño (equivalentes a ensayos pragmáticos) y estudios epidemiológicos observacionales que utilizan grandes bases de datos electrónicas. Sin embargo, pocos ensayos simples de gran tamaño han abordado los daños del tratamiento [282], probablemente debido a barreras operativas, financieras, éticas o científicas, como la utilidad limitada en situaciones en las que el evento adverso es muy raro o tiene un período de latencia largo [282]. Aunque son mucho más frecuentes los estudios con bases de datos secundarias cuando se trata de evaluar los daños debido a que el acceso a dicha información resulta simple en apariencia, la experiencia indica que estas bases de datos deben ser abordadas con cautela debido a importantes limitaciones potenciales [3].

El metanálisis (MA) es una herramienta potente para abordar las cuestiones anteriores. Mientras que los MAs de eficacia farmacológica se llevan a cabo sistemáticamente cuando se alcanza un número suficiente de ECAs para la misma relación fármaco-resultado, los MAs de seguridad farmacológica (es decir, aquellos en los que el objetivo es evaluar el riesgo de reacciones adversas a medicamentos (RAMs)) no son tan sencillos de llevar a cabo. Estos MAs son particularmente relevantes cuando los estudios existentes (observacionales o experimentales) tienen un poder estadístico insuficiente y/o son inconsistentes con el fin de evaluar el riesgo de las RAM que tienen una incidencia baja o muy baja. Recientemente, en un informe de *Council for International Organisations of Medical Sciences (CIOMS Working Group X)* [283] se pusieron a disposición de la comunidad médica directrices destinadas a mejorar el desarrollo y la interpretación de los MAs sobre seguridad de medicamentos y, por extensión, el proceso de toma de decisiones.

A continuación se complementan estas directrices, discutiendo los diversos tipos de técnicas para sintetizar la información sobre seguridad de medicamentos procedente de diferentes fuentes de datos, y resumiendo sus respectivas ventajas y limitaciones.

5.2.1. Fuentes de datos utilizadas para resumir los resultados de seguridad de medicamentos.

5.2.1.1. Ensayos clínicos aleatorizados y estudios no aleatorizados.

La calidad de los datos para resumir determinará necesariamente la validez de los resultados. Idealmente, los datos de más calidad serán los extraídos de los ECAs, ya que la asignación aleatoria del tratamiento garantiza un cierto grado de control de los factores de confusión (conocidos, desconocidos o no medibles), asegurando un reparto equilibrado entre los grupos de tratamiento y de comparación.

Aun así, los datos de los ECAs pueden ser inadecuados para el propósito de estudiar los daños del tratamiento porque: (i) no todos los ECAs informan sobre eventos adversos [284], (ii) la duración del estudio es limitada y, en muchos casos, más corta que la de la práctica clínica habitual o que el tiempo de inducción y latencia necesario para que se desarrolle una RAM, (iii) las poblaciones de los estudios son altamente seleccionadas [285] (se excluye a los pacientes con mayor riesgo de desarrollar RAMs debido a la presencia de comorbilidades), y (iv) el número de pérdidas durante el seguimiento puede ser considerable en algunos ensayos (cuanto más largo sea el ensayo, mayor será la pérdida) y puede ser diferente entre los grupos, en algunos casos estando vinculadas dichas pérdidas a la RAM (censura informativa), algo que puede sesgar los resultados a menos que el análisis se realice por intención de tratar [286]. Además, una consecuencia de la reducción del tamaño de la muestra, incluso después de combinar varios ECAs, es que la potencia estadística puede ser insuficiente para detectar algunos daños del tratamiento.

La falta de cumplimiento del tratamiento en el ECA también puede estar relacionada con: (i) un peor estado de salud al inicio del estudio, (ii) una peor evolución de la enfermedad en todo el ensayo, o (iii) una peor tolerancia al tratamiento. Cualquiera de estas situaciones puede aumentar el riesgo de una RAM [287]. Un efecto similar, aunque en dirección opuesta, es el efecto de la depleción de susceptibles [47]. Es más probable que los pacientes propensos a una RAM determinada presenten el evento en la etapa inicial del tratamiento y, por lo tanto, se retiren del ensayo. Por lo tanto, solo los pacientes menos susceptibles continuarían tomando el medicamento, lo que llevaría a una aparente reducción del riesgo de RAM a lo largo del tiempo.

Idealmente, para combinar los datos sobre daños de varios ECAs, cada estudio debe haber recopilado un registro completo de datos sobre todos los eventos indeseables para todos los participantes del ensayo (y no solo los eventos adversos atribuidos específicamente al tratamiento), incluyendo cualquier posible pérdida y retirada relacionada con el evento adverso. La realidad muestra, sin embargo, que hay inconsistencias en la manera de registrar los eventos adversos. Esta información a veces se registra sistemáticamente, pero también es frecuente que los datos sobre daños se registren como resultados secundarios y no de forma predefinida [288], lo que puede, a su vez, generar clasificaciones erróneas y sesgos en la evaluación de la seguridad del medicamento y contribuir a la variabilidad de los resultados. Para poder controlar o tener en cuenta la heterogeneidad de los resultados de los daños, los ensayos tendrían que informar sobre cómo se recopilaban los eventos adversos en cada estudio (por ejemplo, mediante el contacto con los pacientes por teléfono, el examen de sus registros médicos, o el acceso a sus registros de urgencias o mortalidad hospitalarios).

Por otro lado, los estudios no aleatorizados (ENAs) se basan en datos extraídos de la práctica clínica habitual, una característica que confiere una mayor validez externa. Además, debido a que la incidencia de la RAM en los sujetos no expuestos es similar a la de la población general (lo que puede no ocurrir en todos los ECAs con poblaciones seleccionadas), este tipo de estudio debería permitir al investigador calcular un “número necesario para dañar (NND)” más cercano al del mundo real.

Sin embargo, uno de los mayores inconvenientes de estos diseños, al menos desde un punto de vista teórico, es su incapacidad para adaptarse a los factores de confusión importantes (por ejemplo, fragilidad) que pueden ser desconocidos o no medibles o, alternativamente, ser conocidos, pero no haber sido registrados o haber sido medidos de manera inapropiada. Una de las fuentes más importantes de confusión en los estudios sobre los daños del tratamiento es la confusión por indicación para el tratamiento (la decisión de tratamiento se asocia con una indicación, que a su vez es un factor de riesgo para la enfermedad). Aunque el ajuste estadístico por las variables de confusión es posible, puede estar sujeto a cierto grado de arbitrariedad (por ejemplo, categorización de las variables de confusión, criterios para su inclusión en el modelo). Todo esto disminuye la homogeneidad de los resultados. Otro problema es que los estudios que utilizan datos observacionales tienden a asociarse con un mayor grado de sesgo de publicación; es decir, los ECAs son más fáciles de publicar porque se considera que tienen un mayor grado de evidencia.

A pesar de las limitaciones de los ENAs, se ha comprobado que cuando estos estudios están bien diseñados, el análisis de los datos obtenidos en condiciones clínicas reales permite obtener estimaciones similares a las de los ECAs [289]. Sin embargo, debido a la ausencia de aleatorización, existe una tendencia a descartar los resultados obtenidos en los ENAs [290].

Algunos estudios indican que cuando hay datos suficientes para comparar los efectos de ambos tipos de estudios, los resultados de los ENAs (y por lo tanto los resultados de su MA) son a menudo más conservadores que los de los ECAs al estimar los riesgos absolutos y relativos [291]. La imprecisión derivada de la medición de la exposición al fármaco del estudio, de la notificación de posibles efectos adversos, o de la presencia de factores potenciales de confusión, puede diluir los hallazgos y así subestimar los resultados de estos estudios [292]. También debe tenerse en cuenta que, en presencia de cualquier sesgo, y debido a que los ENAs pueden tener un tamaño de muestra muy grande (especialmente aquellos que utilizan bases de datos secundarias), el aumento de la potencia estadística puede llevar a conclusiones erróneas. Así, los inconvenientes de las pruebas de significación, ya conocidos para otras cuestiones epidemiológicas, se ponen de relieve en el contexto del MA: un tamaño de muestra grande siempre conduce a valores p significativos, incluso si la significación clínica es escasa o inexistente, ya que el valor p es directamente proporcional al tamaño de la muestra.

Los resultados obtenidos con los diferentes diseños de ENAs (estudios de cohortes, estudios de casos y controles) se verán influidos por las diferentes fuentes de sesgo, lo que puede dar lugar a un aumento de la heterogeneidad de los efectos en un MA posterior. El hecho de que los ENAs utilicen criterios diferentes para ajustar por los factores de confusión también puede ser una fuente de heterogeneidad. Por otra parte, es común que los ENAs reporten varios conjuntos de medidas de efecto, cada uno de los cuales se ajusta en función de

diferentes covariables [293], lo que dificulta la decisión sobre las medidas que deben combinarse.

Por lo tanto, ambos enfoques, experimental y no experimental, proporcionan información que puede ser útil en el estudio de los daños del tratamiento. Aunque existe cierta controversia con respecto a cuál es la mejor evidencia para evaluar los daños del tratamiento médico [292], ambos enfoques pueden proporcionar información complementaria: los ECAs pueden ser útiles en el estudio de las RAMs frecuentes que se producen poco después del inicio del tratamiento, mientras que los ENAs serán más apropiados en el caso de eventos raros o eventos que solo pueden medirse a largo plazo en la práctica clínica habitual. A este respecto, se ha sugerido que en las revisiones sistemáticas de eventos adversos el análisis no debe limitarse a determinados tipos de diseño, sino que, por el contrario, debe utilizar la información de todos los diseños disponibles. De esta manera, se mejorará la generalización sin disminuir la validez [294].

Cuando existen datos aleatorios y no aleatorios, puede ser más conveniente realizar un análisis estratificado según el tipo de diseño. Los estudios se pueden mostrar en un gráfico de bosque (*forest plot*) para evaluar visualmente las discrepancias en los efectos de los estudios e identificar fuentes potenciales de heterogeneidad.

5.2.1.2. Estudios publicados y no publicados.

La validez de un MA puede estar limitada por la falta de información sobre eventos adversos. Esto puede ocurrir en el caso de estudios publicados en los que no se han medido o reportado los eventos adversos [295], y también en los casos en los que estos estudios permanecen sin publicar.

El hecho de que los eventos adversos no hayan sido medidos o reportados podría interpretarse como que el evento no hubiese ocurrido. El sesgo de notificación puede existir si se omiten ciertos resultados, o solo se informan parcialmente, porque: (i) los resultados obtenidos fueron diferentes de los esperados, (ii) los resultados no alcanzaron significación estadística, o (iii) los autores no consideraron los eventos adversos como clínicamente relevantes. Este sesgo puede afectar a las conclusiones de las revisiones sistemáticas, dando así una falsa sensación de ausencia de daño [296]. En este contexto, también se ha descrito que los estudios que no reportan eventos adversos tienen menor probabilidad de ser incluidos en el MA debido a la dificultad analítica que plantea la manipulación de celdas con ceros [297].

Otro de los grandes desafíos a la hora de combinar los resultados de los daños es el acceso a todos los estudios realizados, ya sean observacionales o experimentales. El sesgo de publicación describe el fenómeno por el cual la naturaleza y la dirección de los resultados de un estudio individual influyen en su publicación o no publicación [298]. La autocensura de los autores (en los casos en que creen que sus resultados no son suficientemente notables), las decisiones editoriales por parte de editores o revisores de revistas y los intereses corporativos pueden, entre otras razones, influir negativamente en la decisión de publicar un estudio determinado [298]. Existen diferentes situaciones que pueden contribuir a que un estudio se incluya o se excluya de un MA y, por lo tanto, dar lugar a sesgos relacionados con la publicación. El sesgo puede ocurrir incluso si el estudio se publica en una revista pequeña y no indexada, o si se imponen restricciones de idioma (sesgo de idioma) [299]. La inclusión de

un estudio en un MA también puede estar influenciada por la velocidad relativa con la que es publicado. El grado de importancia o interés médico despertado por los resultados de un estudio puede acelerar o retrasar (sesgo de demora) su publicación. La publicación también puede retrasarse si el artículo en cuestión no es aceptado por la revista de primera elección del investigador [300].

La situación inversa al sesgo de publicación también puede ocurrir cuando un solo evento adverso se contabiliza más de una vez en el mismo MA como resultado del recuento múltiple de estudios o subgrupos de pacientes [301]. Todo esto puede alterar sustancialmente las conclusiones de un MA y, por extensión, cualquier decisión clínica y/o regulatoria resultante.

Los resultados de los ECAs se publican más fácilmente si son estadísticamente significativos [302]. Esto puede explicar el hecho de que los hallazgos de seguridad del medicamento puedan estar sesgados hacia una toxicidad significativa. Por lo tanto, se recomienda que las revisiones sistemáticas y los MAs que abordan los daños del tratamiento incluyan datos de estudios no publicados [303-305]. Algunos autores han enfatizado el hecho de que se dispone de más información sobre los eventos adversos a partir de fuentes de datos no publicados que a partir de fuentes de datos publicados, y que en la versión no publicada del mismo estudio se informa de un mayor número de eventos adversos que en la publicada [306].

Para acceder a dicha información se pueden realizar búsquedas en los registros y depósitos de ensayos clínicos y/o en los sitios web de compañías farmacéuticas y agencias reguladoras. Por ejemplo, desde 2007, ciertos ensayos con productos regulados por *US Food and Drug Administration* (FDA) están obligados a reportar un resumen de los resultados, incluyendo los eventos adversos de los medicamentos, al sitio web *ClinicalTrials.gov* [307] en el plazo de un año a partir de la finalización del ensayo (aunque todavía hay evidencia de informes selectivos en los ensayos registrados [308]). En este sentido, recientemente se ha implementado una nueva guía regulatoria para una mayor transparencia en la notificación de los resultados de los estudios [309], de manera que los eventos adversos graves y fatales deben ser reportados en su totalidad, así como los métodos utilizados en la monitorización de dichos eventos.

También puede resultar útil utilizar motores de búsqueda genéricos para recuperar la literatura gris [310]. Sin embargo, no siempre es posible acceder a todos los datos debido a la falta de voluntad de los autores de los estudios no publicados para proporcionar datos.

La inclusión de estudios no publicados tiene la ventaja de reducir un posible sesgo de publicación. Sin embargo, no está exenta de dificultades, a saber: (i) una menor reproducibilidad de la búsqueda bibliográfica, que puede minimizarse si los autores del MA incluyen información detallada sobre cómo rastrearon la información (sitios web, estrategias de búsqueda, fechas), y (ii) la inclusión de datos de estudios no publicados puede introducir sesgos, ya que los estudios encontrados pueden ser una muestra no representativa de todos los estudios no publicados (por ejemplo, en términos de su calidad metodológica) [311]. En este sentido, un análisis de subgrupos que considere estudios publicados frente a no publicados podría ser de interés.

Uno de los métodos tradicionales para la evaluación del sesgo de publicación consiste en la interpretación visual del gráfico de embudo (*funnel plot*), en el que la medida del efecto se representa con su correspondiente medida de precisión (varianza, varianza inversa, tamaño de la muestra, etc.) y en el que una forma asimétrica indicaría la presencia de sesgo de publicación. Sin embargo, debido a su naturaleza subjetiva, se ha comprobado que los investigadores tienen una capacidad limitada para identificar correctamente los gráficos de embudo de los MAs que son propensos al sesgo de publicación [312]. Además, la presencia de sesgo no siempre implica asimetría en este gráfico. Estas limitaciones pueden explicar un uso menos frecuente de este método en los MAs recientes. Algunos autores prefieren complementar esta evaluación visual de la asimetría con pruebas de hipótesis formales como la prueba de asimetría de Egger [313].

Finalmente, es importante mencionar que las agencias reguladoras de medicamentos están facultadas para solicitar todos los resultados de los estudios primarios a las compañías farmacéuticas. Por lo tanto, en un contexto regulatorio, la consideración de los datos publicados puede ser menos relevante, y no es probable que las conclusiones del MA obtenidas por dichas agencias (muy consideradas para el desarrollo de directrices y recomendaciones) puedan verse afectadas por la ausencia de los datos originales.

También es importante recordar que la falta de pruebas sobre los daños del tratamiento en una revisión sistemática puede explicarse por múltiples factores. En consecuencia, para evitar una falsa sensación de seguridad (error de tipo II), algunos autores recomiendan que la evaluación de los daños se realice con el mismo rigor que la de los estudios de eficacia [314, 315], y proponen el uso de herramientas específicas para ello.

5.2.1.3. Fuentes de datos primarias y secundarias.

Un punto importante a tener en cuenta al combinar estudios es la fuente de los datos. La investigación primaria, diseñada de acuerdo con sus necesidades específicas, proporciona datos más precisos y actualizados. Por “datos primarios” se entiende la información recogida por los investigadores a través de sus propios instrumentos, como entrevistas o cuestionarios. Por el contrario, los “datos secundarios” son datos recogidos para otros fines (por ejemplo, administrativos o clínicos); los investigadores reutilizan la información y le dan un nuevo propósito, ya que es más fácil y menos costoso que recogerla desde el comienzo. El modo en que se tiene en cuenta la fuente de los datos como criterio de heterogeneidad puede cambiar las conclusiones de un MA [273].

La tendencia creciente hacia la recogida y el registro informatizado de datos está incrementando la utilización de grandes conjuntos de datos. Esta tendencia es alimentada por el creciente reconocimiento de que los ECAs, aunque esenciales, no pueden responder a preguntas importantes relacionadas con los daños del tratamiento en la práctica clínica diaria [1].

El potencial de investigación que ofrece el uso de bases de datos secundarias preexistentes en farmacoepidemiología es indiscutible. Esta fuente se caracteriza por [316]: (i) su alta representatividad, (ii) un bajo riesgo de sesgo de selección debido a la no participación, (iii) una muestra de gran tamaño que puede ser seguida ininterrumpidamente durante largos períodos de tiempo con pocas pérdidas, algo que puede prevenir sesgos como el de causalidad inversa, y (iv) la disponibilidad de información precisa sobre la exposición a los fármacos

prescritos, que previene los sesgos de memoria y del entrevistador. El uso de bases de datos secundarias también ofrece ventajas de carácter práctico, en la medida en que permiten el estudio rápido y a bajo coste de los efectos de los fármacos en las condiciones de la práctica clínica habitual en poblaciones más heterogéneas que las de los ECAs. Aunque el potencial de las bases de datos automatizadas para estudios epidemiológicos es ampliamente reconocido, su uso, sin embargo, conlleva numerosas dificultades [1, 12, 316]. Los registros médicos se consideran normalmente el estándar de oro para las variables de resultados intermedios y finales, pero muestran limitaciones en el registro de todos los medicamentos consumidos por los pacientes [2]. Aunque los registros de dispensación son más detallados en la medición de la exposición (aunque no registran el uso de medicamentos de venta libre a nivel individual), carecen de variables de resultado [2, 21].

5.2.2. Diferentes tipos de síntesis.

La enorme cantidad de publicaciones existentes hoy en día hace necesario resumir la evidencia científica sobre los daños del tratamiento. Se han definido diferentes tipos de revisiones bibliográficas [14, 317]. Esta sección describe brevemente las características básicas de cada una de estas técnicas, y luego se centra en la técnica del MA, con el fin de discutir los aspectos más relevantes y/o controvertidos de su uso en el ámbito de la seguridad de los medicamentos.

5.2.2.1. Síntesis cualitativa.

5.2.2.1.1. *Revisiones narrativas tradicionales.*

Permiten obtener información actualizada sobre un tema de investigación específico con relativa rapidez y sirven para ofrecer una visión general. Sin embargo, la selección de los trabajos incluidos es subjetiva y no sistemática, y tiende a citar literatura que refuerza las ideas preconcebidas [317].

5.2.2.1.2. *Revisiones sistemáticas.*

A diferencia de las revisiones narrativas, las revisiones sistemáticas incluyen una búsqueda exhaustiva y reproducible de los estudios individuales. La selección de los estudios se basa en criterios de elegibilidad bien definidos. Con frecuencia se incluye una evaluación de la calidad de los estudios, que a menudo es exigida por las revistas. A veces, debido a diferencias considerables entre los estudios, las revisiones sistemáticas resumen los resultados cualitativa pero no cuantitativamente. Sin embargo, siempre que sea posible, se debe realizar un MA como parte de la revisión sistemática. Como cualquier otro estudio, un MA debe realizarse de acuerdo con un plan de análisis preestablecido [318].

Una “*scoping review*” es un tipo particular de revisión sistemática que se ha vuelto cada vez más popular como una forma de síntesis de conocimiento. Este tipo de estudio “*aborda una pregunta de investigación exploratoria dirigida a trazar conceptos clave, tipos de evidencia y lagunas en la investigación relacionados con un área definida mediante la búsqueda, selección y síntesis sistemática del conocimiento existente*” [110]. A diferencia de otras revisiones sistemáticas, la aplicación de filtros de calidad no es una prioridad inicial en este tipo de revisión.

5.2.2.2. Síntesis cuantitativa.

5.2.2.2.1. *Metanálisis de datos agregados.*

En general, se basan en la agrupación de las medidas de efecto reportadas en la literatura, con lo que se prescindiría de la cooperación de los autores de los estudios originales para recopilar los datos. Los resultados se sintetizan cuantitativamente, utilizando técnicas estadísticas para calcular las estimaciones del efecto combinado. Esta es la técnica convencional de MA.

Este tipo de análisis también se denomina “MA de datos de resumen” o “MA de resultados” [319]. El término “datos agregados” se refiere a la información promediada de todos los individuos de un estudio (como la estimación del efecto medio del tratamiento y su error estándar).

5.2.2.2.2. *Metanálisis de datos de pacientes individuales (MA-DPI).*

Este tipo de estudios requiere la colaboración de los autores de los estudios originales –publicados y, en ocasiones, no publicados– para obtener las bases de datos que contienen los datos registrados para cada participante. Sin embargo, no siempre es posible acceder a todas las bases de datos, ya sea porque los autores o los promotores no están dispuestos a proporcionarlas, o porque un estudio no cumple con los criterios de definición de las variables relevantes.

Una vez que las bases de datos estén disponibles, deberán adaptarse a un formato común, de modo que puedan combinarse para formar una única base de datos, para la que se definirán criterios de inclusión específicos, o bien se modificarán las definiciones existentes de las variables de exposición y de confusión [320].

5.2.2.2.3. *Metanálisis de datos de pacientes individuales planificados prospectivamente [317].*

Los estudios individuales previstos para ser incluidos en un futuro MA-DPI están diseñados conjuntamente para que compartan una gran parte del protocolo. Por lo tanto, los objetivos del estudio, los criterios de inclusión, las variables y los métodos de recogida de datos coinciden en gran medida, lo que a su vez facilita el análisis y la interpretación de los resultados.

Este tipo de estudio no es sistemático en su naturaleza, ya que no abarca toda la evidencia disponible sobre una cuestión de investigación determinada. De hecho, este tipo de análisis, en el que solo se consideran los estudios que comparten el protocolo, se asemeja más a un estudio multicéntrico que a un MA convencional. Dado que los estudios individuales pueden analizarse y publicarse por separado, puede haber todavía un grado de heterogeneidad de los efectos debido a las pequeñas diferencias en el diseño de los estudios o en las poblaciones de estudio.

5.2.2.2.4. *Metanálisis en red (metanálisis de tratamiento múltiple o de comparación de tratamiento mixto) [318].*

Los MAs convencionales permiten a los investigadores resumir la evidencia de estudios independientes que comparan los mismos tratamientos, como en los ensayos de fármacos nuevos que utilizan tratamientos actuales o antiguos como comparadores. Sin embargo, no permiten evaluar la efectividad relativa o los daños de varios tratamientos que no

han sido comparados directamente. En este sentido, tener comparaciones directas entre los nuevos fármacos resulta extremadamente útil para la práctica clínica.

El MA en red es un método para comparar más de dos intervenciones, algunas de las cuales pueden no haber sido comparadas directamente en el mismo estudio. Con este fin, las comparaciones indirectas del tratamiento entre los ensayos se realizan mediante un comparador común [321, 322], lo que permite identificar el tratamiento más efectivo y seguro. La validez de estas comparaciones depende de que los estudios midan los mismos efectos relativos del tratamiento y de que las diferencias observadas no se deban al azar [323].

La mayoría de los MAs en red publicados hasta la fecha han incorporado datos de ECAs solamente. Sin embargo, la inclusión de estudios no aleatorizados podría resolver algunas de las limitaciones del MA en red, como períodos de seguimiento cortos, tamaños de muestra pequeños y poblaciones altamente seleccionadas. Sin embargo, esta práctica también puede aumentar el riesgo de sesgos como el de confusión, que a veces no son evaluables ni detectables [323].

5.2.3. ¿Cuándo se pueden realizar metanálisis y revisiones sistemáticas?

Las relaciones entre las exposiciones y los resultados se evalúan a menudo de manera no convencional. Muchos estudios no proporcionan medidas de efecto o datos suficientes para calcularlas. Esto ocurre a menudo en los estudios clínicos. En tal caso, es imposible proporcionar una estimación combinada, y entonces puede resultar útil una revisión sistemática narrativa. Cuando hay pocos estudios, realizar un MA puede ser inútil o incluso contraproducente. La presentación de una medida de efecto combinado de solo dos o tres estudios puede dar una falsa sensación de rigor científico.

Cuando hay heterogeneidad excesiva, es aconsejable realizar análisis de subgrupos y mostrar los resultados estratificados por variables clave. Sin embargo, esto no siempre es posible si persiste un alto grado de heterogeneidad dentro de los subgrupos. La presentación de los resultados en forma de una revisión sistemática sin MA a menudo se considera la mejor solución. La presencia de sesgo de publicación bien podría ser un caso en el que se requiera una revisión sistemática sin MA. Por otro lado, existen técnicas como el método de recorte y relleno (*trim-and-fill method*) [324] que permiten predecir el número de estudios que faltan y su efecto. Aun así, es aconsejable abstenerse de presentar resultados metanalíticos si el número de estudios que faltan es demasiado grande.

5.2.4. Metanálisis sobre seguridad de medicamentos: justificación, ventajas y limitaciones.

5.2.4.1. Metanálisis de datos agregados.

El aumento drástico en el uso de la técnica de MA para el estudio de los daños del tratamiento se debe a que es una herramienta rápida y de bajo coste para estimar el riesgo global de un evento adverso que los estudios individuales son incapaces de estimar [325], ya sea porque los estudios individuales tienen un tamaño de muestra insuficiente o porque la incidencia del evento adverso es demasiado baja. El MA también puede ser útil cuando los resultados de los estudios individuales son inconsistentes [318].

Las técnicas estadísticas más ampliamente utilizadas para combinar los datos de estudio utilizan un enfoque de efectos fijos o de efectos aleatorios. Se ha criticado la realización de MAs sin tener plenamente en cuenta las hipótesis estadísticas subyacentes e inferencias necesarias para este tipo de análisis [326, 327]. En el caso de los modelos de efectos fijos, se supone que el efecto del tratamiento es constante en todos los estudios y que la variación entre estudios del efecto se debe exclusivamente al azar. El modelo de efectos aleatorios considera una fuente adicional de variación entre los estudios, que puede deberse, entre otras razones, a diferencias en el diseño de los estudios, el período de seguimiento o la selección de pacientes. El resultado final es que el modelo de efectos aleatorios tiende a producir estimaciones más conservadoras (intervalos de confianza más amplios) del efecto combinado. No obstante, el uso de modelos de efectos aleatorios no exime en absoluto de la necesidad de un estudio adecuado de la heterogeneidad y sus fuentes, ya que estos modelos no pueden explicar la presencia de una heterogeneidad importante. Cuando existe una heterogeneidad mínima de las estimaciones puntuales entre los estudios, ambas técnicas producen resultados similares.

A pesar de su uso generalizado, el MA sigue siendo una técnica controvertida en la investigación epidemiológica [317, 328-330]. Los argumentos más consistentes contra el MA son la presencia de confusión, sesgo de selección y marcada heterogeneidad en el diseño de los estudios y los procedimientos de recogida de datos [14]. Las principales limitaciones se deben al mal uso de la técnica; por ejemplo, cuando se ignora la diversidad de los estudios incluidos en el análisis o se aborda de forma inadecuada, o cuando no se tiene en cuenta la calidad de los datos, la variabilidad en las poblaciones de pacientes y/o los posibles sesgos subyacentes [331]. Algunos autores sostienen que el uso adecuado del MA consiste en mejorar la precisión de las estimaciones cuantitativas de los estados de salud en las poblaciones, y que no debe utilizarse para la realización de pruebas de hipótesis [332]. En el campo de la seguridad de los medicamentos, sin embargo, el MA de ENAs puede ser una herramienta fundamental para comprobar señales sobre cuestiones de seguridad; por ejemplo, cuando el número de eventos adversos en los estudios individuales es tan pequeño que no se puede calcular una medida de efecto [333]. En los ECAs, sin embargo, el tiempo de seguimiento suele ser demasiado corto para que ocurra el evento; en tal caso, un MA no tendría sentido.

Es una práctica estándar que los MAs sean actualizados acumulativamente con cualquier nueva evidencia que esté disponible. Sin embargo, al igual que se establecen métodos para controlar el error de tipo I en los análisis intermedios y las comparaciones múltiples de ECAs, estas correcciones también deben aplicarse en el MA acumulativo (utilizando un nivel de significación más estricto que el habitual de 0,05) [334].

Otro punto a tener en cuenta es que la validez externa de un MA no es mucho mayor que la validez de los estudios individuales, aunque la distribución geográfica global puede ser más dispersa. La validez externa está determinada por criterios de inclusión/exclusión, y si éstos son muy restrictivos, la validez externa del MA también será baja. Desde el punto de vista clínico, el interés real radica en determinar el mejor tratamiento para un paciente específico [330]. Por lo tanto, los resultados del MA solo son aplicables a las poblaciones que presentan características similares a las de las poblaciones incluidas en ese MA. En un MA, los estudios que tienen criterios de inclusión estrictos coexisten frecuentemente con otros cuyos criterios son más flexibles (además de tener diferentes resultados o tiempos de

seguimiento, entre otros factores). Es por ello por lo que algunos autores consideran el MA como una herramienta que sirve para evaluar la asociación entre un factor y un resultado, pero no la causalidad [13].

Además, en los MAs de datos agregados la información sobre eventos adversos dentro de subgrupos no suele estar disponible, lo que hace imposible realizar análisis por subgrupos. Incluso cuando se dispone de dicha información, el poder estadístico del MA de ECAs es generalmente inadecuado para detectar la modificación del efecto.

En los MAs de datos agregados, la heterogeneidad no debe ser vista como una molestia sino más bien como una característica intrínseca. Sería casi imposible que los estudios sobre poblaciones diferentes, con metodologías e investigadores diferentes, obtuviesen resultados idénticos [15]. Por lo tanto, la heterogeneidad debe considerarse la regla y no la excepción [335], y debe describirse cuidadosamente analizando los posibles factores responsables de su generación. De hecho, la existencia de heterogeneidad “aparente” es probablemente la mejor razón para realizar un MA. Si hubiese un veredicto unánime y todos los estudios sobre una relación particular exposición-resultado indicaran un mayor riesgo, entonces un MA sería de menor interés (si no completamente inútil).

La evaluación de la heterogeneidad mediante la comprobación de hipótesis es ineficaz, ya que las pruebas de heterogeneidad tienen un poder estadístico muy bajo. En cambio, el uso de gráficos como el gráfico de bosque (que evalúa visualmente las discrepancias en los efectos de los estudios e identifica las fuentes potenciales de heterogeneidad) y medidas más descriptivas como la proporción de varianza debida a la varianza entre estudios (R_I [17] e I^2 [336]), son más útiles porque, en lugar de implicar la aceptación o el rechazo de cualquier hipótesis de homogeneidad, la describen y la miden.

Una de las opciones disponibles para explicar la heterogeneidad es realizar una metarregresión, donde se busca la explicación de la respuesta media de todos los pacientes del estudio (variable dependiente) en una o más covariables que diferencian unos estudios de otros. Sin embargo, el efecto de estas covariables sobre la variable de respuesta en la metarregresión puede no ser el mismo que en los estudios, dando lugar a un efecto equivalente a la “falacia ecológica” o sesgo ecológico (es decir, las relaciones encontradas a nivel agregado no representan relaciones a nivel individual).

En el contexto de la seguridad de los medicamentos, los resultados de una metarregresión deben interpretarse con cautela, especialmente cuando se analiza la susceptibilidad de la población a los eventos adversos. Siempre es posible que la heterogeneidad no se deba a la susceptibilidad de la población, sino a que haya una o más variables de confusión que no fueron evaluadas y que muestran frecuencias diferentes entre las distintas poblaciones de estudio. La única forma de evitarlo es utilizar datos de pacientes individuales en un MA-DPI, en el que esta covariable se incluya como variable de ajuste [337].

5.2.4.2. Metanálisis de datos individuales.

En el MA-DPI, se combinan los datos primarios (o datos brutos) de todos los pacientes incluidos en cada uno de los estudios publicados y no publicados. Esta técnica permite responder a preguntas que no pudieron ser resueltas satisfactoriamente por los

estudios individuales, y proporciona una estimación más precisa de los efectos de la confusión y la interacción.

Algunos autores reconocen que el enfoque de MA-DPI es el estándar de oro [319, 338-344], gracias a sus numerosas ventajas sobre los análisis que se basan en datos agregados, a saber: (i) el uso de datos actualizados, (ii) un examen detallado de los datos y un manejo más apropiado de los datos ausentes en cada estudio [345], (iii) el potencial para transformar los datos en medidas comunes, (iv) la generación de análisis más flexibles y consistentes y, específicamente, (v) la realización de análisis más potentes por subgrupos. Además, esta técnica también ofrece (vi) la posibilidad de prevenir los sesgos habituales en el uso de datos agregados en la metarregresión [346], y (vii) la investigación de nuevas hipótesis relacionadas con las características individuales de los pacientes (para lo cual faltarían datos en los resultados publicados) [342].

Una de las principales ventajas sobre los MAs convencionales (con o sin metarregresión) es que permiten que los daños del tratamiento en subpoblaciones sean analizados de manera efectiva y no sesgada. Trabajar con datos individuales sirve para prevenir el riesgo de la falacia ecológica, riesgo que existe en el MA con metarregresión [318].

Para tener en cuenta la heterogeneidad, el MA-DPI puede utilizar potenciales variables de confusión y/o modificadores de efecto para cada uno de los sujetos del estudio, además de otras variables “ecológicas” (aquellas con el mismo valor para todos los participantes del estudio, tales como el tipo de estudio, los criterios de inclusión, las características del sistema de salud, las instalaciones de diagnóstico y los recursos). Esto permite lograr un efecto comparable al de la metarregresión en un MA convencional, con la ventaja adicional de disponer de los datos necesarios para ajustar las variables individuales.

Cuando se combinan estudios con diferentes tiempos de seguimiento en un MA-DPI, se pueden aplicar técnicas de análisis de datos que permiten evaluar la relación entre el fármaco y el evento adverso, teniendo en cuenta el tiempo.

El uso de datos de pacientes individuales (DPI) permite reducir el sesgo de publicación en los casos en que los datos provengan también de estudios no publicados. Existe la posibilidad de sesgos relacionados con la publicación que podrían afectar el análisis si los investigadores se limitasen a los estudios publicados [347]. Otro factor a tener en cuenta es el sesgo de disponibilidad de datos (donde la falta de disponibilidad de DPI está relacionada con los resultados del estudio) [347]. Podría surgir un sesgo a favor de la seguridad del tratamiento si los DPI fueron proporcionados únicamente por los autores de los estudios que detectaron menos problemas de seguridad, o alternativamente, si los autores de los estudios con resultados relacionados con los eventos adversos significativos no proporcionaron los datos de los pacientes para realizar nuevos análisis o ampliar el seguimiento. A este respecto, es importante incluir un diagrama de flujo que describa los criterios de búsqueda y los documentos seleccionados, a fin de mejorar la reproducibilidad. También podría estar presente un sesgo de selección del revisor (cuando los DPI se buscan entre un subgrupo de estudios que no representan la evidencia general) [347]. Este sesgo es particularmente importante en los casos en que no se utiliza un método sistemático de selección de estudios, como ocurre en las revisiones narrativas.

Aparte del sesgo de publicación descrito anteriormente en el caso de artículos ausentes, es esencial que los datos ausentes sean tenidos en cuenta en el MA-DPI. No todos los estudios registrarán necesariamente las mismas variables de exposición y/o efecto, por lo que existe un alto riesgo de que falten datos. Si dichas variables se incluyen en los modelos estadísticos, esto puede dar lugar a un sesgo equivalente al sesgo de selección, debido a que los participantes de los estudios que no pudieron evaluar tales variables no serán considerados en el análisis. Por consiguiente, es aconsejable realizar un análisis descriptivo de los datos que faltan para cada estudio y controlarlos utilizando una técnica como la imputación múltiple, que preserva la variabilidad de los datos. El informe final de los resultados debe especificar cuántos sujetos fueron incluidos en el análisis y el porcentaje del total inicial que esto representa.

Para el análisis estadístico del MA-DPI es necesario considerar la estructura jerárquica (o multinivel) de los datos, donde el primer nivel contiene los individuos como unidades de análisis, mientras que el segundo nivel es aquél en el que la unidad de análisis es el estudio. En las estimaciones de las respuestas, además de la variación entre individuos, se debe tener en cuenta una segunda fuente de variación: la variación entre estudios [348].

Para la elección de un modelo se puede adoptar un marco frecuentista (en el que se maximiza la función de verosimilitud) o un marco bayesiano (en el que la función de verosimilitud se combina con información previa). Debido a la complejidad de estas técnicas, se recomienda la colaboración de expertos en la materia. Un trabajo reciente de Debray et al. revisa en profundidad los métodos de análisis del MA-DPI [349].

Aunque el MA-DPI resulta atractivo desde un punto de vista científico, puede presentar limitaciones éticas/legales en algunos países en los que el intercambio de datos a nivel individual no sea legal. Para sortear este obstáculo, se han descrito métodos sencillos para el análisis de datos combinados [350].

Debido a que tanto el MA de datos agregados como el MA-DPI comparten muchas de sus limitaciones, debería ser un requisito que el MA-DPI incluyese un análisis con datos agregados ("MA-DPI de dos etapas"*) como forma de análisis de sensibilidad adicional, y determinar así la posible presencia no solo de (i) sesgo de publicación o (ii) heterogeneidad, sino también de (iii) confusión por estudio, un sesgo que puede surgir cuando los datos de

* En un MA-DPI de una etapa los DPI de todos los estudios se agrupan para realizar un único análisis teniendo en cuenta la agrupación de los participantes dentro de los estudios. Se debe especificar un modelo que se adapte al tipo de datos que se resumen (por ejemplo, un modelo de regresión lineal, o de Cox, según la naturaleza de las variables del estudio) y también a los supuestos estadísticos del metanálisis (efectos fijos o aleatorios). Aunque este método requiere una separación cuidadosa de la variabilidad dentro del estudio y entre los estudios [338], no es infrecuente que se ignoren las diferencias entre los estudios y que los datos se analicen como si perteneciesen a un solo estudio (es decir, un análisis de "megaensayo").

Por el contrario, en un MA-DPI de dos etapas los DPI se analizan primero por separado para cada estudio utilizando el método estadístico apropiado, obteniendo así datos agregados. Estos datos se agrupan en una segunda etapa utilizando el modelo más adecuado para el metanálisis de datos agregados (asumiendo efectos fijos o aleatorios del tratamiento en todos los estudios). Este método, por lo tanto, puede ser considerado como un MA clásico, con la ventaja de utilizar los mismos criterios para analizar los estudios individuales. Debido a que este método ignora la información a nivel de paciente, es más adecuado para aquellos casos en los que el interés radica exclusivamente en determinar el efecto combinado global del tratamiento en relación con las covariables a nivel de estudio (es decir, las covariables que son idénticas para cada paciente en un estudio dado) [319].

pacientes de diferentes estudios se combinan como si procediesen de la misma muestra[†]. Además, en este tipo de análisis también es aconsejable determinar la influencia de cada estudio en el resultado final.

Los resultados de un solo estudio generalmente no son suficientes para tomar decisiones sobre el perfil de seguridad de un medicamento determinado. Los estudios epidemiológicos, y los ensayos clínicos aleatorizados en particular, carecen del tamaño de muestra necesario para identificar los riesgos de reacciones adversas graves y poco frecuentes entre los consumidores de medicamentos. Además, el poder estadístico de estos estudios es insuficiente para determinar el perfil de seguridad de un medicamento en ciertas subpoblaciones, evaluar su efecto dosis-riesgo o evaluar la influencia del tiempo de uso en sus riesgos asociados.

La síntesis sistemática de los datos disponibles puede proporcionar información decisiva para los fines antes mencionados y es considerada por muchos autores como el método de referencia para obtener conocimientos acerca de la seguridad de los medicamentos. Aunque el metanálisis basado en datos agregados es el método tradicional y más extendido para la evaluación de los daños del tratamiento, no está exento de controversia. Por el contrario, los MA-DPI no han generado mucho debate y a menudo son considerados el estándar de oro. Aun así, comparten muchas limitaciones con los metanálisis de datos agregados, y también la desventaja de requerir la verificación de varios supuestos estadísticos específicos. Por ello, los MA-DPI deben ser diseñados, analizados, reportados e interpretados con cautela. La tabla 10 incluye las cuestiones más importantes que se deben tener en cuenta. En general, se debe realizar un análisis de sensibilidad para determinar si los resultados son sensibles a la técnica utilizada.

[†] La confusión por estudio [287] (un tipo de falacia ecológica; ver “paradoja de Simpson” [351, 352]) aparece cuando se agrupan varios estudios y se obtienen resultados contraintuitivos [353]. Por ejemplo, si la mayoría de los estudios sobre un fármaco determinado coinciden en que éste se comporta como un factor de riesgo para un evento adverso, pero cuando estos estudios se agrupan en un MA-DPI, el mismo fármaco se muestra como un factor protector [354].

Este sesgo surge en MA-DPI porque se considera que los datos provienen de una sola muestra y, por lo tanto, no se ponderan. Sin embargo, los sujetos de los respectivos estudios (o subgrupos) incluidos en el análisis pueden diferir significativamente en características importantes (por ejemplo, el riesgo inicial). Tales diferencias se pueden deber al hecho de que: (i) existe un desequilibrio considerable entre los estudios debido al tamaño desigual de los grupos, (ii) las poblaciones incluidas en los estudios con placebo y control activo son muy diferentes, (iii) existen diferencias importantes en el diseño del reclutamiento de pacientes y el estadio de la enfermedad; o bien, (iv) no se tienen en cuenta las variaciones en la duración del seguimiento del estudio [355].

La técnica de metanálisis de datos agregados (o metanálisis de dos etapas) combina los resultados tras la ponderación de las unidades de estudio, evitando así el riesgo de confusión por estudio [355]. Es por ello por lo que algunos autores proponen que en el proceso de registro de un medicamento se realice un metanálisis de datos agregados para evaluar los riesgos del medicamento, en lugar de realizar un MA-DPI [355]. Otra consecuencia de los diferentes riesgos iniciales entre los estudios es que el NND calculado considerando los datos como si fuesen todos extraídos de un solo estudio (“*treat-as-one-trial method*”) puede estar sesgado en los casos en que los grupos no estén equilibrados [356].

Tabla 10. Resumen de las principales cuestiones a considerar en metanálisis de datos de pacientes individuales sobre daños relacionados con medicamentos.

Fase del estudio	Cuestión
Concepción	<p>Registrar el diseño del estudio y asegurar la visibilidad de toda la información detallada sobre:</p> <ul style="list-style-type: none"> • definición de los criterios de estudio y de inclusión de los pacientes • elección de la variable principal, tiempo de seguimiento • análisis por subgrupos a realizar • plan de análisis (por protocolo/intención de tratamiento), pruebas estadísticas que se utilizarán y ajuste para comparaciones múltiples • las variables de confusión que deben tenerse en cuenta y los criterios para su inclusión en los modelos • tratamiento de los datos que faltan • plan de difusión de los resultados
	<p>Incluir datos sobre estudios no publicados mediante búsquedas realizadas en:</p> <ul style="list-style-type: none"> • registros y repositorios de ensayos clínicos • sitios web de compañías farmacéuticas y agencias reguladoras • motores de búsqueda genéricos (literatura gris)
	<p>Incluir información detallada sobre cómo se obtuvieron los datos (sitios web, estrategias de búsqueda, fechas)</p>
	<p>Informar cómo se recogieron las posibles reacciones adversas a los medicamentos en cada estudio (por ejemplo, poniéndose en contacto con los pacientes por teléfono, examinando sus historias clínicas o accediendo a sus registros de urgencias o mortalidad en el hospital)</p>
	<p>Evaluar la calidad de los estudios que comprenden el metanálisis</p> <p>Para ensayos clínicos aleatorios, verificar si:</p> <ul style="list-style-type: none"> • todos los ensayos clínicos aleatorios reportan eventos adversos • la duración del estudio es similar a la de la práctica clínica habitual o al tiempo de inducción y de latencia necesario para que se desarrollen los efectos adversos • se excluyen los pacientes con mayor riesgo de desarrollar eventos adversos debido a la presencia de comorbilidades • las pérdidas durante el seguimiento son diferentes entre los grupos de comparación y están relacionadas con los eventos adversos • el análisis se realiza por intención de tratar <p>Para los estudios no aleatorizados, se debe verificar si se tiene en cuenta la fuente de datos (primaria frente a secundaria)</p>
Análisis	<p>Para evaluar la seguridad de los fármacos es preferible realizar un metanálisis de datos agregados o un metanálisis de datos individuales en dos etapas</p>
	<p>Realizar un análisis de heterogeneidad y/o un análisis estratificado que considere: (i) estudios publicados frente a no publicados, y (ii) fuente de datos</p>
	<p>Realizar un análisis de sensibilidad para evaluar: (i) sesgo de publicación, (ii) heterogeneidad, y (iii) confusión por estudio (mediante un metanálisis convencional de datos agregados o un metanálisis de datos individuales en dos etapas)</p>
	<p>Tener en cuenta los datos ausentes en cada estudio, preservando la variabilidad de los datos (por ejemplo, imputación múltiple)</p>
	<p>Tener en cuenta la naturaleza multinivel de los datos (agrupación de pacientes dentro de los estudios) en el análisis estadístico</p>

Fase del estudio	Cuestión
Reporte	Incluir un diagrama de flujo que explique las razones para excluir estudios no publicados
	Describir los datos ausentes para cada estudio
	Mostrar el número (y porcentaje) de pacientes en cada estudio incluido en el modelo estadístico final
	Metanálisis con análisis multinivel: incluir una medida de heterogeneidad entre los estudios
	Mostrar los resultados del análisis de sensibilidad



5.3. CONCLUSIONES DIVERGENTES A PARTIR DEL MISMO METANÁLISIS EN ESTUDIOS DE SEGURIDAD DE MEDICAMENTOS. RELEVANCIA DEL TIPO DE FUENTE DE DATOS CONSIDERADA (OBJETIVO 3 - ANEXO 3).

Los resultados muestran que es de suma importancia tener en cuenta la fuente de datos como fuente potencial de heterogeneidad y variable modificadora del efecto en los metanálisis farmacoepidemiológicos. De lo contrario, se llegaría a conclusiones que podrían ser erróneas.

La contradicción entre los resultados de la versión revisada de los metanálisis y los resultados originales obtenidos sin tener en cuenta la fuente de los datos se debe a la presencia de sesgos no controlados en los estudios farmacoepidemiológicos. Los ensayos clínicos aleatorizados sobre la misma relación exposición-enfermedad que la evaluada en los metanálisis de los estudios observacionales son extremadamente escasos o inexistentes. Por lo tanto, no se puede utilizar un estándar de oro para determinar qué fuente de datos es más fiable.

Si el resultado se espera a largo plazo después de la exposición, como el cáncer y las enfermedades neurodegenerativas, los datos secundarios son probablemente más fiables, ya que los registros proporcionan información sobre la mayoría de los factores de confusión. Por el contrario, el uso de datos secundarios en el estudio de los efectos a corto plazo de los medicamentos, especialmente en el caso de los medicamentos de venta libre, puede ser problemático.

Los sesgos que con más frecuencia explican la divergencia de los resultados son los de confusión (especialmente la confusión por indicación) y la clasificación errónea de la exposición.

La confusión por indicación se produce cuando los primeros síntomas inespecíficos, que podrían ser manifestaciones tempranas de la enfermedad, llevan al consumo del fármaco evaluado en el estudio, exagerando así su efecto. Por ejemplo, una madre podría usar paracetamol para mitigar el dolor y la fiebre debidos a infecciones del tracto respiratorio, mientras que las infecciones son un factor de riesgo de asma en su descendencia [228]. También es posible que los sujetos enfermos hayan consumido otros medicamentos que incrementaron su riesgo de enfermedad. Esto puede explicar los resultados en el metanálisis de los ISRSs y el sangrado gastrointestinal alto. Es probable que los estudios que utilizaron datos secundarios sean más susceptibles de presentar confusión debido al hecho de que los consumidores de ISRSs son probablemente más propensos a consumir otros medicamentos, como los AINEs, fármacos que se asocian frecuentemente con la hemorragia [357]. En general, las bases de datos secundarias no evalúan la ingesta de medicamentos de venta libre. También es posible que algún tercer factor, como la depresión, una afección que aumenta el riesgo de síntomas digestivos menores [358], haya aumentado la vulnerabilidad de los sujetos deprimidos a los fármacos gastrointestinales lesivos. Otros posibles factores de confusión difíciles de evaluar en las bases de datos secundarias son la presencia de *Helicobacter pylori* [359] y el consumo de alcohol, tabaco y café [360].

Existen varios métodos para mitigar el efecto de la confusión en farmacoepidemiología, como los que utilizan variables instrumentales [101] o puntuaciones

de propensión [85]. Otros métodos miden los factores de confusión en una base de datos externa y realizan un estudio de validación [254].

Por otro lado, en los estudios farmacoepidemiológicos es probable que se produzca un sesgo de clasificación errónea que pueda subestimar o sobrestimar el efecto. En las bases de datos administrativas (por ejemplo, las bases de datos de dispensación farmacéutica o de seguros médicos), la fecha y la cantidad del medicamento dispensado se miden con precisión. Sin embargo, la información sobre prescripción y dispensación no refleja necesariamente el consumo real, lo que puede causar una clasificación errónea de la exposición [55]. No tener en cuenta la parte del consumo de medicamentos debida a los medicamentos de venta sin receta puede introducir sesgos, ya que ese consumo puede estar relacionado con el nivel socioeconómico y/o con un peor acceso al sistema de salud, factores de riesgo independientes de un resultado negativo de algunas enfermedades, como el infarto de miocardio. Otra consecuencia de la falta de registro de los datos relacionados con los medicamentos de venta libre es el menor grado de control de las interacciones con otros fármacos. Esto puede causar confusión residual si se tiene en cuenta que los IBPs inhiben la enzima que activa el clopidogrel [361, 362].

También pueden producirse errores de clasificación en aquellos estudios que evalúan la exposición de forma retrospectiva a partir de fuentes de datos de origen primario. Este es el caso de los dos estudios incluidos en el metanálisis de los AINEs y el infarto de miocardio, en los que es probable que se produjese una clasificación errónea, ya que las exposiciones mensuales se extrapolaron a todo el año y se utilizaron diferentes intervalos de tiempo para los casos y controles [227]. Este tipo de estudio puede ser propenso a un mayor riesgo de sesgo de clasificación errónea debido al sesgo de recuerdo cuando la ventana etiológica de la exposición al fármaco es remota y la población del estudio es de edad avanzada o susceptible de presentar problemas de memoria, como los observados en la demencia.

Otra fuente de discrepancia en los metanálisis seleccionados para este trabajo es la alta frecuencia de datos ausentes en las fuentes secundarias, especialmente en los registros clínicos.

El origen de los datos, según sea la fuente de información primaria o secundaria, debe ser tenido en cuenta en cualquier metanálisis que considere fármacos como factor de exposición. Se debe realizar siempre un primer análisis estratificado por los estudios que utilizan fuentes de datos primarias y los estudios que utilizan fuentes de datos secundarias. Cualquier discrepancia en el signo o la magnitud de la relación exposición-enfermedad debe ser suficiente para presentar resultados separados para cada subgrupo y abstenerse de calcular una estimación general del efecto combinado. Este análisis de subgrupos debe planificarse a priori en cualquier metanálisis relacionado con fármacos y mencionarse en la fase de registro del metanálisis.

El metanálisis es el diseño de estudio que conduce a la mejor evidencia posible en epidemiología, superando la alcanzada por un ensayo clínico aleatorizado [13], y probablemente represente la herramienta más valiosa para evaluar la seguridad de los medicamentos. Por lo tanto, es importante que las conclusiones estén exentas de debilidades debidas al origen de los datos. Además, cualquier metanálisis en farmacoepidemiología puede tener consecuencias decisivas después de su publicación a nivel del paciente, ya que la elección entre varios medicamentos es, en primer lugar, una cuestión individual; a nivel regulatorio, ya que las decisiones se toman generalmente una vez se dispone de una cantidad suficiente de información, por ejemplo, cuando se publica un metanálisis metodológicamente

sólido; y a nivel de salud pública, ya que el metanálisis se utiliza para decidir sobre la intervención que resulta más efectiva.

5.3.1. Limitaciones del método.

El objetivo de este trabajo no fue reunir una muestra representativa de metanálisis. Simplemente se intentó seleccionar aquellos metanálisis que cumplieran una serie de criterios de inclusión y en los que se observó un efecto relevante en el análisis de subgrupos por fuente de datos (primaria frente a secundaria). Aunque pueden existir otros metanálisis en los que los efectos no son tan evidentes como los observados en este trabajo, creemos que los ejemplos similares a los nuestros no son escasos. Aunque la ausencia de interacción significativa en el total de casos puede ser vista como una limitación, no creemos que sea una condición necesaria considerar los cambios en los resultados como clínicamente relevantes.

Además, hay que tener en cuenta la posible influencia que el diseño del estudio puede ejercer sobre los resultados divergentes que presentamos (ver metanálisis de subgrupos por diseño del estudio en suplemento 2, anexo 3). A diferencia de la estratificación por diseño, la estratificación por el origen de los datos sugiere una modificación en la conclusión del metanálisis en el ejemplo de los AINEs [227], así como una disminución sustancial de la heterogeneidad en el ejemplo de los ISRSs [224]. Sin embargo, no podemos descartar la posibilidad de que el diseño explique parte de la variabilidad atribuida a la fuente de datos en los otros dos ejemplos, ya que existe una superposición considerable entre ambas clasificaciones. Debido a que cada criterio de estratificación tiene su propio fundamento, consideramos que pueden ser complementarios.

5.4. UTILIZACIÓN DE FUENTES DE DATOS Y HETEROGENEIDAD EN METANÁLISIS DE ESTUDIOS OBSERVACIONALES SOBRE EFECTOS DE MEDICAMENTOS: UN ESTUDIO DE LAS PRINCIPALES REVISTAS MÉDICAS (OBJETIVO 4 - ANEXO 4).

Los resultados de esta investigación sugieren que el origen de los datos, primario o secundario, está poco explorado como fuente de heterogeneidad y modificador del efecto en los metanálisis sobre efectos de medicamentos publicados en revistas de medicina general de alto impacto. Unos pocos metanálisis reportaron la fuente de los datos y solo uno [243] de los artículos incluidos en nuestra revisión comparó y discutió los resultados del metanálisis considerando las diferentes fuentes de datos.

Aunque es habitual considerar el diseño de los estudios individuales (es decir, estudios de casos y controles, de cohortes o experimentales) en el análisis de la heterogeneidad de un metanálisis [18, 363], el tipo de fuente de datos (primaria frente a secundaria) raramente se utiliza para este propósito [20, 364]. De hecho, las directrices actuales para la presentación de informes para los metanálisis, como MOOSE (*Meta-analysis Of Observational Studies in Epidemiology*) [121] o PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*) [365, 366], no recomiendan que los autores reporten específicamente el origen de los datos. Esto se debe probablemente a la estrecha relación que existe entre el diseño del estudio y el tipo de fuente de datos utilizada, a pesar de que cada criterio tiene su propio fundamento. La realización de este análisis adicional es una tarea sencilla que no implica ningún coste adicional. De lo contrario, pueden surgir conclusiones divergentes [273].

Las conclusiones sobre los efectos de un medicamento que se derivan de estudios basados exclusivamente en datos de fuentes secundarias pueden ser arriesgadas, entre otras razones, porque no se recoge información sobre el consumo de medicamentos de venta libre (es decir, medicamentos que las personas pueden comprar sin receta médica) [252] y/o el desembolso económico que debe realizar el propio paciente en relación con el coste de medicamentos de venta con receta médica [55]. En el contexto de la atención médica y de los seguros de salud, este concepto suele referirse a los deducibles, copagos o coseguros. La figura 17 muestra el modelo que proponemos para describir la relación entre los diferentes registros de datos según su origen, incluyendo la posible pérdida de información (susceptible de ser registrada solo a través de la investigación primaria).

Si no se tienen en cuenta estas situaciones, puede producirse un sesgo en la medición de la exposición [55, 252]. El consumo de un medicamento puede subestimarse cuando solo se utilizan datos de prescripción como fuente secundaria sin considerar adicionalmente el consumo no registrado, como el consumo de venta libre (por ejemplo, anticonceptivos orales [243, 367]), que tal vez solo esté disponible en una base de datos primaria. Alternativamente, esto puede ocurrir cuando se utilizan datos con fines de facturación (por ejemplo, con un fin de reembolso por parte de una compañía de seguros) para la investigación clínica, si no se considera el copago (Fig. 17). Es poco probable que se registre la parte de la factura médica que la compañía de seguros no cubre y que el individuo debe pagar por su cuenta. En este escenario tampoco se dispondrá de datos sobre la venta de medicamentos de venta libre.

La situación inversa también puede ocurrir, y el consumo ser sobrestimado, cuando solo se utilizan datos de prescripción, si el farmacéutico no dispensa el medicamento prescrito, o cuando se utilizan datos de dispensación, si el medicamento no es realmente

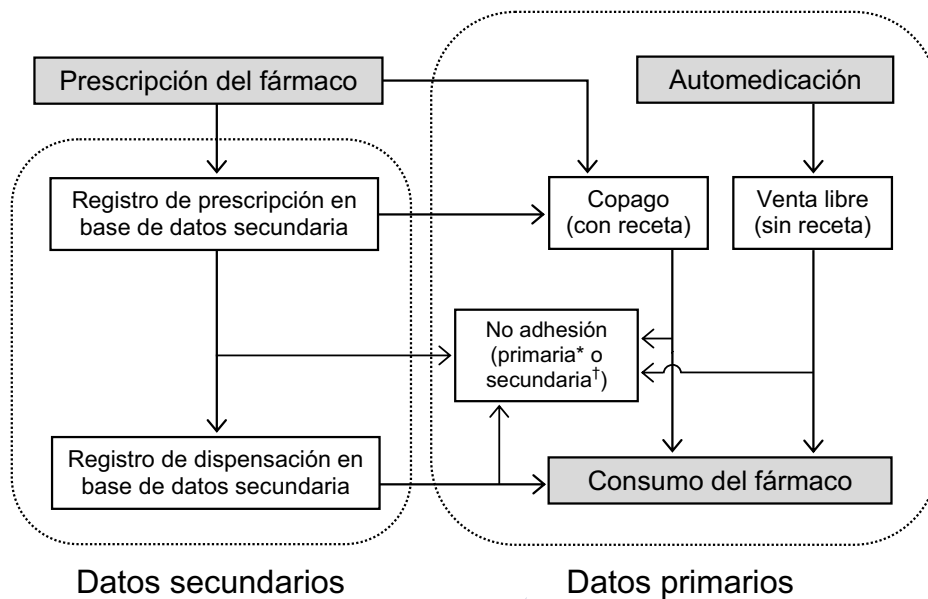


Figura 17.

Modelo conceptual de registro de datos individuales. * Nunca dispensado. † Ausencia de dispensación de prescripciones sucesivas entre pacientes con adherencia primaria, o adhesión secundaria inadecuada.

consumido por el paciente. Mientras que la falta de adhesión primaria ocurre cuando no se dispensa el medicamento después de la primera prescripción, la falta de adhesión secundaria se refiere a la ausencia de dispensación de prescripciones sucesivas entre pacientes con adherencia primaria, o a la adhesión secundaria inadecuada (es decir, $\geq 20\%$ del tiempo sin la medicación adecuada) [368] (Fig. 17). En algunas enfermedades, la adhesión a la medicación es muy baja [369-372], con porcentajes de falta de adhesión primaria (medicación nunca dispensada) que superan el 30% [373]. Debe tenerse en cuenta que el impacto de la falta de adhesión varía de un medicamento a otro. Por lo tanto, debe ser definido y medido en el contexto de una terapia particular [374].

Además, no tener en cuenta la parte del consumo debida a los gastos de venta libre y/o copago puede dar lugar a confusión, ya que esa variable puede estar relacionada con el nivel socioeconómico y/o el potencial de acceso al sistema de salud [375], que son factores de riesgo independientes de los resultados adversos de algunos medicamentos (por ejemplo, el infarto de miocardio [230, 237, 239, 250]). Dada la presencia de planes de salud altamente deducibles y la alta tasa de coseguro para algunos medicamentos, la coparticipación en los costes puede disuadir a los pacientes clínicamente vulnerables de iniciar el tratamiento con los medicamentos esenciales, afectando así negativamente al cumplimiento del paciente [376, 377].

La clasificación errónea de los resultados también puede dar lugar a sesgos de medición y heterogeneidad [378]. Esto ocurre, por ejemplo, en el metanálisis que evalúa la relación entre los anticonceptivos orales combinados y el riesgo de trombosis venosa [243]. En los estudios sin confirmación objetiva del resultado, las mujeres fueron clasificadas erróneamente con independencia del uso de anticonceptivos. Esto condujo a una clasificación errónea no diferencial que puede haber subestimado la relación fármaco-resultado, especialmente cuando se analiza la tercera generación de progestágeno: razón de riesgos (RR)

de datos primarios = 6,2 (intervalo de confianza [IC] del 95%: 5,2 a 7,4), RR de datos secundarios = 3,0 (IC del 95%: 1,7 a 5,4) [243].

Por un lado, los registros médicos son considerados a menudo como la mejor fuente de información para las variables de resultado. Sin embargo, presentan limitaciones importantes en el registro de los medicamentos que toman los pacientes [2]. Por otra parte, los registros de dispensación muestran datos más detallados sobre la medición de la exposición a los fármacos. Sin embargo, no registran el consumo de medicamentos de venta libre o el desembolso económico que debe realizar el propio paciente a nivel individual [55, 252], aparte de ofrecer datos poco fiables sobre las variables de resultado [2, 270].

La existencia de heterogeneidad en un metanálisis no es en absoluto un síntoma de mala calidad, sino una característica que debe ser descrita. Idealmente, ninguna fuente de heterogeneidad en un metanálisis debe ser considerada por separado, sino como parte de un conjunto interconectado de variables potenciales que deben ser estudiadas. En particular, el origen de los datos, primario o secundario, no se explora suficientemente como fuente de heterogeneidad en los metanálisis sobre efectos de medicamentos, ni siquiera en los publicados en revistas de medicina general de alto impacto. Por lo tanto, los autores deben incluir sistemáticamente la fuente de datos como una variable adicional en los análisis de subgrupos y de sensibilidad, o en los análisis de metarregresión, y discutir su influencia en los resultados del metanálisis. Asimismo, revisores, editores y futuras guías también deben considerar el origen de los datos como una causa potencial de heterogeneidad en los metanálisis de estudios observacionales que incluyan estudios con datos primarios y estudios con datos secundarios. De lo contrario, se pueden obtener conclusiones erróneas, con efectos negativos en las decisiones clínicas y regulatorias.

5.4.1. Limitaciones del método.

La primera limitación de esta investigación es que sus hallazgos pueden no ser aplicables a revistas no incluidas en nuestra revisión, como las revistas con bajo factor de impacto. A pesar del uso generalizado de la métrica del factor de impacto [379], este método tiene debilidades inherentes [380, 381]. Sin embargo, es probable que los metanálisis publicados en revistas de medicina general de alto impacto se realicen y reporten con mayor rigor debido a su mayor disponibilidad de recursos y procedimientos [115, 117]. Por ello, es esperable que la calidad de los datos reportados en los artículos publicados en otras revistas menos conocidas sea similar. Otra limitación estaría relacionada con el período de búsqueda limitado. En este sentido, y dado que la tendencia general es la mejora de la metodología de los metanálisis publicados [382, 383], no encontramos ninguna razón para sospechar que las conclusiones adversas puedan ser diferentes antes del período 2012-2018. Aunque excede el objetivo de esta investigación, una última limitación puede ser la incapacidad de reanalizar los metanálisis incluidos estratificando por el tipo de fuente de datos, ya que el diseño de nuestro estudio restringe las conclusiones a los datos publicados de los metanálisis, los cuales fueron reportados de manera insuficiente, o el número de estudios individuales en cada estrato fue insuficiente para calcular una medida combinada (Tabla 7).

6. CONCLUSIONES GENERALES.

6.1.1. Antes de iniciar una investigación utilizando bases de datos secundarias, los investigadores deben evaluar en detalle las fuentes de datos disponibles, centrándose en el propósito para el que fueron creadas, y así tomar conciencia de su potencial de sesgo.

6.1.2. La vinculación de los registros médicos con bases de datos administrativas es útil para minimizar el riesgo de sesgo, así como la complementación o validación de los datos secundarios con datos primarios (es decir, recopilados a partir de métodos *ad hoc*) cuando se cuestiona la integridad o la calidad de los datos originales.

6.2.1. Los metanálisis de datos de pacientes individuales a menudo son considerados el estándar de oro. Aun así, comparten muchas limitaciones con los metanálisis de datos agregados, y también la desventaja de requerir la verificación de varios supuestos estadísticos específicos. Por ello, los metanálisis de datos de pacientes individuales (MA-DPI) deben ser diseñados, analizados, reportados e interpretados con cautela. Se debe realizar un análisis de sensibilidad para determinar si los resultados son sensibles a la técnica utilizada.

6.3.1. El origen de los datos, según sea la fuente de información primaria o secundaria, debe ser tenido en cuenta en cualquier metanálisis que considere fármacos como factor de exposición. Se debe realizar siempre un primer análisis estratificado por los estudios que utilizan fuentes de datos primarias y los estudios que utilizan fuentes de datos secundarias.

6.3.2. Cualquier discrepancia en el signo o la magnitud de la relación exposición-enfermedad debe ser suficiente para presentar resultados separados para cada subgrupo y abstenerse de calcular una estimación general del efecto combinado. Este análisis de subgrupos debe planificarse *a priori* en cualquier metanálisis relacionado con fármacos y mencionarse en la fase de registro del metanálisis.

6.4.1. La existencia de heterogeneidad en un metanálisis no es en absoluto un síntoma de mala calidad, sino una característica que debe ser descrita. Idealmente, ninguna fuente de heterogeneidad en un metanálisis debería ser considerada por separado, sino como parte de un conjunto interconectado de variables potenciales que deben ser estudiadas.

6.4.2. En particular, el origen de los datos, primario o secundario, no se explora suficientemente como fuente de heterogeneidad en los metanálisis sobre efectos de medicamentos, ni siquiera en los publicados en revistas de medicina general de alto impacto. Por lo tanto, los autores deben incluir sistemáticamente la fuente de datos como una variable adicional en los análisis de subgrupos y de sensibilidad, o en los análisis de metarregresión, y discutir su influencia en los resultados del metanálisis.



7. BIBLIOGRAFÍA.

1. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol*. 2005;58:323–37.
2. Takahashi Y, Nishida Y, Asai S. Utilization of health care databases for pharmacoepidemiology. *Eur J Clin Pharmacol*. 2012;68:123–9.
3. Ray WA. Improving automated database studies. *Epidemiology*. 2011;22:302–4.
4. Weiss NS. The new world of data linkages in clinical epidemiology: are we being brave or foolhardy? *Epidemiology*. 2011;22:292–4.
5. European Commission. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). 2016. Disponible en: <http://eur-lex.europa.eu/eli/reg/2016/679/oj>. Acceso el 20 febrero de 2019.
6. U.S. Department of Health and Human Services. Code of Federal Regulations. Title 45 Public Welfare. Part 46 Protection of Human Subjects. 2016. Disponible en: <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html>. Acceso el 20 de febrero de 2019.
7. Knottnerus JA, Tugwell P. Requirements for utilizing health care-based data sources for research. *J Clin Epidemiol*. 2011;64:1051–3.
8. Terris DD, Litaker DG, Koroukian SM. Health state information derived from secondary databases is affected by multiple sources of bias. *J Clin Epidemiol*. 2007;60:734–41.
9. Schneeweiss S. Understanding secondary databases: a commentary on “sources of bias for health state characteristics in secondary databases”. *J Clin Epidemiol*. 2007;60:648–50.
10. Moore TJ, Furberg CD. Electronic health data for postmarket surveillance: a vision not realized. *Drug Saf*. 2015;38:601–10.
11. Gagne JJ. Restrictive reimbursement policies: bias implications for claims-based drug safety studies. *Drug Saf*. 2014;37:771–6.
12. van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. *J Clin Epidemiol*. 2012;65:126–31.
13. Berlin JA, Golub RM. Meta-analysis as evidence: building a better pyramid. *JAMA*. 2014;312:603–5.
14. Blettner M, Schlattmann P. *Meta-analysis in epidemiology*. In: Ahrens W, Pigeot I, editors. *Handbook of epidemiology*. Berlin: Springer; 2005. pp. 829–59.
15. Higgins JPT. Heterogeneity in meta-analysis should be expected and appropriately quantified. *Int J Epidemiol*. 2008;37:1158–60.
16. Stoto MA. Drug safety meta-analysis: promises and pitfalls. *Drug Saf*. 2015;38:233–43.
17. Takkouche B, Khudyakov P, Costa-Bouzas J, et al. Confidence intervals for heterogeneity measures in meta-analysis. *Am J Epidemiol*. 2013;178:993–1004.

18. Egger M, Davey Smith G, Schneider M. *Systematic reviews of observational studies*. In: Egger M, Davey Smith G, Altman DG, editors. *Systematic reviews in health care: meta-analysis in context*. 2nd ed. London: BMJ Publishing Group; 2001. pp. 211–27.
19. Petitti DB. *Exploring heterogeneity*. In: Petitti DB, editor. *Meta-analysis, decision analysis, and cost-effectiveness analysis. Methods for quantitative synthesis in medicine*. 2nd ed. New York: Oxford University Press; 2000. pp. 213–227.
20. Madigan D, Ryan PB, Schuemie M, et al. Evaluating the impact of database heterogeneity on observational study results. *Am J Epidemiol*. 2013;178:645–51.
21. Hennessy S. Use of health care databases in pharmacoepidemiology. *Basic Clin Pharmacol Toxicol*. 2006;98:311–3.
22. Cox E, Martin BC, Van Staa T, et al. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: The ISPOR good research practices for retrospective database analysis task force report—Part II. *Value Health*. 2009;12:1053–61.
23. Bayley KB, Belnap T, Savitz L, et al. Challenges in using electronic health record data for CER: experience of 4 learning organizations and solutions applied. *Med Care*. 2013;51(8 Suppl 3):S80–6.
24. Kush RD, Helton E, Rockhold FW, et al. Electronic health records, medical research, and the Tower of Babel. *N Engl J Med*. 2008;358:1738–40.
25. Stürmer T, Jonsson Funk M, Poole C, et al. Nonexperimental comparative effectiveness research using linked healthcare databases. *Epidemiology*. 2011;22:298–301.
26. Berger ML, Mamdani M, Atkins D, et al. Good research practices for comparative effectiveness research: defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: The ISPOR good research practices for retrospective database analysis task force report—Part I. *Value Health*. 2009;12:1044–52.
27. Ali A. Methodological challenges in observational research: A pharmacoepidemiological perspective. *Br J Pharm Res*. 2013;3:161–75.
28. Delgado-Rodríguez M, Llorca J. Bias. *J Epidemiol Community Health*. 2004;58:635–41.
29. Miettinen O. Confounding and effect modification. *Am J Epidemiol*. 1974;100:350–3.
30. Joffe MM. Confounding by indication: the case of calcium channel blockers. *Pharmacoepidemiol Drug Saf*. 2000;9:37–41.
31. Sjoding MW, Luo K, Miller MA, et al. When do confounding by indication and inadequate risk adjustment bias critical care studies? A simulation study. *Crit Care*. 2015;19:195.
32. Hudson M, Suissa S. Avoiding common pitfalls in the analysis of observational studies of new treatments for rheumatoid arthritis. *Arthritis Care Res (Hoboken)*. 2010;62:805–10.
33. Poses RM, Smith WR, McClish DK, et al. Controlling for confounding by indication for treatment. Are administrative data equivalent to clinical data? *Med Care*. 1995;33:AS36–46.
34. Fewell Z, Davey Smith G, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *Am J Epidemiol* 2007;166:646–55.
35. Stram DO, Huberman M, Wu AH. Is residual confounding a reasonable explanation for the apparent protective effects of beta-carotene found in epidemiologic studies of lung cancer in smokers? *Am J Epidemiol*. 2002;155:622–8.
36. Brenner H, Blettner M. Controlling for continuous confounders in epidemiologic research. *Epidemiology*. 1997;8:429–34.

37. Schneeweiss S, Setoguchi S, Brookhart MA, et al. Assessing residual confounding of the association between antipsychotic medications and risk of death using survey data. *CNS Drugs*. 2009;23:171–80.
38. Glynn RJ, Knight EL, Levin R, et al. Paradoxical relations of drug treatment with mortality in older persons. *Epidemiology*. 2001;12:682–9.
39. Brookhart MA, Stürmer T, Glynn RJ, et al. Confounding control in healthcare database research: challenges and potential approaches. *Med Care*. 2010;48(6 Suppl):S114–20.
40. Shrank WH, Patrick AR, Brookhart MA. Healthy user and related biases in observational studies of preventive interventions: a primer for physicians. *Gen Intern Med*. 2012;26:546–50.
41. Andersen M, Brookhart MA, Glynn RJ, et al. Practical issues in measuring cessation and re-initiation of drug use in databases [Abstract]. *Pharmacoepidemiol Drug Saf*. 2008;17(Suppl 1):S27.
42. Kiri VA. A pathway to improved prospective observational post-authorization safety studies. *Drug Saf*. 2012;35:711–24.
43. Salas M, Hofman A, Stricker BH. Confounding by indication: an example of variation in the use of epidemiologic terminology. *Am J Epidemiol*. 1999;149:981–3.
44. Horwitz RI, Feinstein AR. The problem of “protopathic bias” in case-control studies. *Am J Med*. 1980;68:255–8.
45. Gerhard T. Bias: Considerations for research practice. *Am J Health Syst Pharm*. 2008;65:2159–68.
46. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–25.
47. Moride Y, Abenhaim L. Evidence of the depletion of susceptibles effect in non-experimental pharmacoepidemiologic research. *J Clin Epidemiol*. 1994;47:731–7.
48. Danaei G, Tavakkoli M, Hernán MA. Bias in observational studies of prevalent users: Lessons for comparative effectiveness research from a meta-analysis of statins. *Am J Epidemiol*. 2012;175:250–62.
49. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol*. 2003;158:915–20.
50. Johnson ES, Bartman BA, Briesacher BA, et al. The incident user design in comparative effectiveness research. *Pharmacoepidemiol Drug Saf*. 2013;22:1–6.
51. Johnson ML, Crown W, Martin BC, et al. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: The ISPOR good research practices for retrospective database analysis task force report—Part III. *Value Health*. 2009;12:1062–73.
52. Hernán MA, Cole SR. Causal diagrams and measurement bias. *Am J Epidemiol*. 2009;170:959–62; discussion 963–4.
53. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet*. 2002;359:248–52.
54. Hartzema AG, Perfetto EM. *Sources and effects of drug exposure and unintended effect misclassification in pharmacoepidemiologic studies*. In: Hartzema AG, Porta MS, Tilson HH, eds. *Pharmacoepidemiology*. 2nd ed. Cincinnati: OH: Harvey Whitney Books Co.; 1991.

55. Gamble JM, McAlister FA, Johnson JA, et al. Quantifying the impact of drug exposure misclassification due to restrictive drug coverage in administrative databases: a simulation cohort study. *Value Health*. 2012;15:191–7.
56. Andrade SE, Kahler KH, Frech FF, et al. Methods for evaluation of medication adherence and persistence using automated databases. *Pharmacoepidemiol Drug Saf*. 2006;15:565–74; discussion 575–7.
57. Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *J Clin Epidemiol*. 2004;57:131–41.
58. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on Methodological Standards in Pharmacoepidemiology (Revision 3). EMA/95098/2010. Disponible en: http://www.encepp.eu/standards_and_guidances. Acceso el 20 febrero de 2019.
59. van Staa TP, Abenham L, Leufkens H. A study of the effects of exposure misclassification due to the time-window design in pharmacoepidemiologic studies. *Clin Epidemiol*. 1994;47:183–9.
60. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf*. 2007;16:241–9.
61. Rothman KJ, Suissa S. Exclusion of immortal person-time. *Pharmacoepidemiol Drug Saf*. 2008;17:1036.
62. Lévesque LE, Hanley JA, Kezouh A, et al. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ*. 2010;340:b5087.
63. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol*. 2008;167:492–9.
64. Suissa S. Immeasurable time bias in observational studies of drug effects on mortality. *Am J Epidemiol*. 2008;168:329–35.
65. Suissa S, Azoulay L. Metformin and the risk of cancer: time-related biases in observational studies. *Diabetes Care*. 2012;35:2665–73.
66. Hall GC, Sauer B, Bourke A, et al. Guidelines for Good Database Selection and use in 2011. *Pharmacoepidemiol Drug Saf*. 2012;21:1–10.
67. Perrio M, Waller PC, Shakir SAW. An analysis of the exclusion criteria used in observational pharmacoepidemiological studies. *Pharmacoepidemiol Drug Saf*. 2006;16:329–36.
68. Schneeweiss S, Patrick AR, Stürmer T, et al. Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results. *Med Care*. 2007;45(10 Suppl 2):S131–42.
69. MacMahon S, Collins R. Reliable assessment of the effects of treatment on mortality and major morbidity, II: observation studies *Lancet*. 2001;357:455–62.
70. Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol*. 1991;133:144–53.
71. Delaney JA, Suissa S. The case-crossover study design in pharmacoepidemiology. *Stat Methods Med Res*. 2009;18:53–65.
72. Carracedo-Martínez E, Tobías A, Saez M, et al. [Case-crossover design: Basic essentials and applications]. *Gac Sanit*. 2009;23:161–5.
73. Suissa S. The case-time-control design. *Epidemiology*. 1995;6:248–53.
74. Greenland S. Confounding and exposure trends in case-crossover and case-time-control designs. *Epidemiology*. 1996;7:231–9.

75. Schneeweiss S, Stürmer T, Maclure M. Case-crossover and case-time-control designs as alternatives in pharmacoepidemiologic research. *Pharmacoepidemiol Drug Saf.* 1997;6 Suppl 3:S51–9.
76. Weldeselassie YG, Whitaker HJ, Farrington CP. Use of the self-controlled case-series method in vaccine safety studies: review and recommendations for best practice. *Epidemiol Infect.* 2011;139:1805–17.
77. Maclure M, Fireman B, Nelson JC, et al. When should case-only designs be used for safety monitoring of medicinal products? *Pharmacoepidemiol Drug Saf.* 2012;21 Suppl 1:50–61.
78. Xu S, Zhang L, Nelson JC, et al. Identifying optimal risk windows for selfcontrolled case series studies of vaccine safety. *Stat Med.* 2011; 30:742–52.
79. Moodie EE, Delaney JA, Lefebvre G, et al. Missing confounding data in marginal structural models: a comparison of inverse probability weighting and multiple imputation. *Int J Biostat.* 2008;4:Article 13.
80. Arbogast PG, Ray WA. Use of disease risk scores in pharmacoepidemiologic studies. *Stat Methods Med Res.* 2009;18:67–80.
81. Tadrus M, Gagne JJ, Stürmer T, et al. Disease risk score as a confounder summary method: systematic review and recommendations. *Pharmacoepidemiol Drug Saf.* 2013;22;122–9.
82. Cepeda MS, Boston R, Farrar JT, et al. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am J Epidemiol.* 2003;158:280–7.
83. D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med.* 1998;17:2265–81.
84. Joffe MM, Rosebaum PR. Invited Commentary: Propensity Scores. *Am J Epidemiol.* 1999;150:327–33.
85. Glynn RJ, Schneeweiss S, Stürmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. *Basic Clin Pharmacol Toxicol.* 2006;98:253–9.
86. Baser O. Too much ado about propensity score models? Comparing methods of propensity score matching. *Value Health.* 2006;9:377–85.
87. Rubin DB, Thomas N. Combining propensity score matching with additional adjustment for prognostic covariates. *J Am Stat Assoc.* 2000;95:573–85.
88. Brookhart MA, Schneeweiss S, Rothman KJ, et al. Variable selection for propensity score models. *Am J Epidemiol.* 2006;163:1149–56.
89. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol.* 1996;49:1373–9.
90. Schneeweiss S, Rassen JA, Glynn RJ, et al. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology.* 2009;20:512–22.
91. Rassen JA, Schneeweiss S. Using high-dimensional propensity scores to automate confounding control in a distributed medical product safety surveillance system. *Pharmacoepidemiol Drug Saf.* 2012;21 Suppl 1:41–9.
92. Shah BR, Laupacis A, Hux JE, et al. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. *J Clin Epidemiol.* 2005;58:550–9.
93. Seeger JD, Kurth T, Walker AM. Use of propensity score technique to account for exposure-related covariates: an example and lesson. *Med Care.* 2007;45(10 Suppl 2):S143–8.

94. Stürmer T, Schneeweiss S, Rothman KJ, et al. Performance of propensity score calibration--a simulation study. *Am J Epidemiol.* 2007;165:1110–8.
95. Stürmer T, Schneeweiss S, Avorn J, et al. Adjusting effect estimates for unmeasured confounding with validation data using propensity score calibration. *Am J Epidemiol.* 2005;162:279–89.
96. Stürmer T, Glynn RJ, Rothman KJ, et al. Adjustments for unmeasured confounders in pharmacoepidemiologic database studies using external information. *Med Care.* 2007; 45:S158–65.
97. Schneeweiss S, Glynn RJ, Tsai EH, et al. Adjusting for unmeasured confounders in pharmacoepidemiologic claims data using external information: the example of COX2 inhibitors and myocardial infarction. *Epidemiology.* 2005;16:17–24.
98. Stürmer T, Brenner H. Flexible matching strategies to increase power and efficiency to detect and estimate gene-environment interactions in case-control studies. *Am J Epidemiol.* 2002;155:593–602.
99. Hanley JA, Csizmadi I, Collet JP. Two-stage case-control studies: precision of parameter estimates and considerations in selecting sample size. *Am J Epidemiol.* 2005;162:1225–34.
100. Heckman J. Sample selection as a specification error. *Econometrica.* 1979;47:153–61.
101. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiol Drug Saf.* 2010;19:537–54.
102. Martens EP, Pestman WR, de Boer A, et al. Instrumental variables: application and limitations. *Epidemiology.* 2006;17:260–7.
103. Crown WH, Henk HJ, Vanness DJ. Some cautions on the use of instrumental variables estimators in outcomes research: how bias in instrumental variables estimators is affected by instrument strength, instrument contamination, and sample size. *Value Health.* 2011;14:1078–84.
104. Witteman JC, D'Agostino RB, Stijnen T, et al. G-estimation of causal effects: isolated systolic hypertension and cardiovascular death in the Framingham Heart Study. *Am J Epidemiol.* 1998;148:390–401.
105. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11:550–60.
106. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology.* 2000;11:561–70.
107. Cole SR, Hernán MA, Robins JM, et al. Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. *Am J Epidemiol.* 2003;158:687–94.
108. Sanni Ali M, Groenwold RH, Pestman WR, et al. Time-dependent propensity score and collider-stratification bias: an example of beta(2)-agonist use and the risk of coronary heart disease. *Eur J Epidemiol.* 2013;28:291–9.
109. Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005;2:e124.
110. Colquhoun HL, Levac D, O'Brien KK, et al. Scoping reviews: time for clarity in definition, methods, and reporting. *J Clin Epidemiol.* 2014;67:1291–4.
111. Peters MD, Godfrey CM, Khalil H, et al. Guidance for conducting systematic scoping reviews. *Int J Evid Based Healthc.* 2015;13:141–6.
112. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med.* 2018;169:467–73.

113. InCites Journal Citation Reports. Science citation index expanded - medicine. General & Internal Thomson Reuters. Disponible en: <https://jcr.incites.thomsonreuters.com>. Acceso el 20 de febrero de 2019.
114. Faggion CM Jr, Bakas NP, Wasiak J. A survey of prevalence of narrative and systematic reviews in five major medical journals. *BMC Med Res Methodol*. 2017;17:176.
115. Hopewell S, Ravaud P, Baron G, et al. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. *BMJ*. 2012;344:e4178.
116. Blanc X, Collet TH, Auer R, et al. Publication trends of shared decision making in 15 high impact medical journals: a full-text review with bibliometric analysis. *BMC Med Inform Decis Mak*. 2014;14:71.
117. Rehal S, Morris TP, Fielding K, et al. Non-inferiority trials: are they inferior? A systematic review of reporting in major medical journals. *BMJ Open*. 2016;6:e012594.
118. Callahan M, Wears RL, Weber E. Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *JAMA*. 2002;287:2847–50.
119. Kim SY, Park JE, Lee YJ, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *J Clin Epidemiol*. 2013;66:408–14.
120. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa: Univ of Ottawa; 2009. Disponible en: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Acceso el 20 de febrero de 2019.
121. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. *JAMA*. 2000;283:2008–12.
122. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Epidemiology*. 2007;18:800–4.
123. U.S. National Library of Medicine. Citations Added to MEDLINE® by Fiscal Year. 2016. Disponible en: https://www.nlm.nih.gov/bsd/stats/cit_added.html. Acceso el 20 de febrero de 2019.
124. Weinstein RB, Ryan P, Berlin JA, et al. Channeling in the use of nonprescription paracetamol and ibuprofen in an electronic medical records database: evidence and implications. *Drug Saf*. 2017;40:1279–92.
125. Pottegård A, Friis S, Stürmer T, et al. Considerations for pharmacoepidemiological studies of drug-cancer associations. *Basic Clin Pharmacol Toxicol*. 2018;122:451–9.
126. Melamed A, Rauh-Hain JA, Schorge JO. Clinical outcomes research in gynecologic oncology. *Gynecol Oncol*. 2017;146:653–60.
127. Dong YH, Alcusky M, Maio V, et al. Evidence of potential bias in a comparison of β blockers and calcium channel blockers in patients with chronic obstructive pulmonary disease and acute coronary syndrome: results of a multinational study. *BMJ Open*. 2017;7:e012997.
128. Bourbeau J, Aaron SD, Barnes NC, et al. Evaluating the risk of pneumonia with inhaled corticosteroids in COPD: retrospective database studies have their limitations SA. *Respir Med*. 2017;123:94–7.

129. Macías Saint-Gerons D, de la Fuente Honrubia C, de Andrés Trelles F, et al. [Future perspective of pharmacoepidemiology in the “Big Data Era” and the growth of information sources]. *Rev Esp Salud Publica*. 2016;90:e1–e7.
130. Hudson M, Tascilar K, Suissa S. Comparative effectiveness research with administrative health data in rheumatoid arthritis. *Nat Rev Rheumatol*. 2016;12:358–66.
131. Bérard A, Wisner KL, Hultsch S, et al. Field studies versus database studies on the risks and benefits of medication use during pregnancy: distinct pieces of the same puzzle. *Reprod Toxicol*. 2016;60:123–8.
132. Haneuse S. Distinguishing selection bias and confounding bias in comparative effectiveness research. *Med Care*. 2016;54:e23–9.
133. Fillion KB, Eberg M, Ernst P. Confounding by drug formulary restriction in pharmacoepidemiologic research. *Pharmacoepidemiol Drug Saf*. 2016;25:278–86.
134. Golozar A, Liu S, Lin JA, et al. Does metformin reduce cancer risks? Methodologic considerations. *Curr Diab Rep*. 2016;16:4.
135. Willis AW. Using Administrative data to examine health disparities and outcomes in neurological diseases of the elderly. *Curr Neurol Neurosci Rep*. 2015;15:75.
136. Swanson SA, Hernandez-Diaz S, Palmsten K et al. Methodological considerations in assessing the effectiveness of antidepressant medication continuation during pregnancy using administrative data. *Pharmacoepidemiol Drug Saf*. 2015;24:934–42.
137. Heinze G, Wallisch C, Kainz A et al. Chances and challenges of using routine data collections for renal health care research. *Nephrol Dial Transplant*. 2015;30 Suppl 4:iv68–75.
138. Moulis G, Lapeyre-Mestre M, Palmaro A, et al. French health insurance databases: what interest for medical research? *Rev Med Interne*. 2015;36:411–7.
139. Jensen ET, Cook SF, Allen JK, et al. Enrollment factors and bias of disease prevalence estimates in administrative claims data. *Ann Epidemiol*. 2015;25:519–525.e2.
140. Patorno E, Garry EM, Patrick AR, Schneeweiss S, Gillet VG, Zorina O, et al. Addressing limitations in observational studies of the association between glucose-lowering medications and all-cause mortality: a review. *Drug Saf*. 2015;38:295–310.
141. Funk MJ, Landi SN. Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. *Curr Epidemiol Rep*. 2014;1:175–85.
142. Patorno E, Patrick AR, Garry EM, et al. Observational studies of the association between glucose-lowering medications and cardiovascular outcomes: addressing methodological limitations. *Diabetologia*. 2014;57:2237–50.
143. Kim DH, Schneeweiss S. Measuring frailty using claims data for pharmacoepidemiologic studies of mortality in older adults: evidence and recommendations. *Pharmacoepidemiol Drug Saf*. 2014;23:891–901.
144. Datta R, Kleinman K, Rifas-Shiman S, et al. Confounding by indication affects antimicrobial risk factors for methicillin-resistant *Staphylococcus aureus* but not vancomycin-resistant enterococci acquisition. *Antimicrob Resist Infect Control*. 2014;3:19.
145. Schneeweiss S. Learning from big health care data. *N Engl J Med*. 2014;370:2161–3.
146. Yang X, Chan JC. Metformin and the risk of cancer in type 2 diabetes: methodological challenges and perspectives. *Ann Transl Med*. 2014;2:52.
147. Zhang J, Curtis JR. Considerations in using registry and health plan data for studying pregnancy in rheumatic diseases. *Curr Opin Rheumatol*. 2014;26:315–20.

148. Gavrielov-Yusim N, Friger M. Use of administrative medical databases in population-based research. *J Epidemiol Community Health*. 2014;68:283–7.
149. Paxton C, Niculescu-Mizil A, Saria S. Developing predictive models using electronic medical records: challenges and pitfalls. *AMIA Annu Symp Proc*. 2013;2013:1109–15.
150. Gallego B, Dunn AG, Coiera E. Role of electronic health records in comparative effectiveness research. *J Comp Eff Res*. 2013;2:529–32.
151. Ryan PB, Madigan D, Stang PE, et al. Medication-wide association studies. *CPT Pharmacometrics Syst Pharmacol*. 2013;2:e76.
152. Hersh WR, Weiner MG, Embi PJ, et al. Caveats for the use of operational electronic health record data in comparative effectiveness research. *Med Care*. 2013;51(8 Suppl 3):S30–7.
153. Yang X, Weng J. Increased cancer risk with drug use among patients with diabetes: are the biased methods the culprit? *J Diabetes Investig*. 2012;3:479–80.
154. Hershman DL, Wright JD. Comparative effectiveness research in oncology methodology: observational data. *J Clin Oncol*. 2012;30:4215–22.
155. Suling M, Pigeot I. Signal detection and monitoring based on longitudinal healthcare data. *Pharmaceutics*. 2012;4:607–40.
156. Yang XL, Ma RC, So WY, et al. Addressing different biases in analysing drug use on cancer risk in diabetes in non-clinical trial settings—what, why and how? *Diabetes Obes Metab*. 2012;14:579–85.
157. Nelson JC, Cook AJ, Yu O, et al. Challenges in the design and analysis of sequentially monitored postmarket safety surveillance evaluations using electronic observational health care data. *Pharmacoepidemiol Drug Saf*. 2012;21 Suppl 1:62–71.
158. Fung V, Brand RJ, Newhouse JP, et al. Using medicare data for comparative effectiveness research: opportunities and challenges. *Am J Manag Care*. 2011;17:488–96.
159. Zhang J, Yun H, Wright NC, et al. Potential and pitfalls of using large administrative claims data to study the safety of osteoporosis therapies. *Curr Rheumatol Rep*. 2011;13:273–82.
160. Hernán MA. With great data comes great responsibility: publishing comparative effectiveness research in epidemiology. *Epidemiology*. 2011;22:290–1.
161. McNeil JJ, Piccenna L, Ronaldson K, et al. The value of patient-centred registries in phase IV drug surveillance. *Pharmaceut Med*. 2010;24:281–8.
162. Ehrenstein V, Sørensen HT, Bakkevig LS, et al. Medical databases in studies of drug teratogenicity: methodological issues. *Clin Epidemiol*. 2010;2:37–43.
163. Martin-Latry K, Bégaud B. Pharmacoepidemiological research using French reimbursement databases: yes we can! *Pharmacoepidemiol Drug Saf*. 2010;19:256–65.
164. Giezen TJ, Mantel-Teeuwisse AK, Leufkens HG. Pharmacovigilance of biopharmaceuticals: challenges remain. *Drug Saf*. 2009;32:811–7.
165. Harpe SE. Using secondary data sources for pharmacoepidemiology and outcomes research. *Pharmacotherapy*. 2009;29:138–53.
166. Giordano SH, Kuo YF, Duan Z, et al. Limits of observational data in determining outcomes from cancer therapy. *Cancer*. 2008;112:2456–66.
167. Pigeot I, Ahrens W. Establishment of a pharmacoepidemiological database in Germany: methodological potential, scientific value and practical limitations. *Pharmacoepidemiol Drug Saf*. 2008;17:215–23.
168. Strom BL. Methodologic challenges to studying patient safety and comparative effectiveness. *Med Care*. 2007;45(10 Suppl 2):S13–5.

169. Cramer JA, Silverman SL, Gold DT. Methodological considerations in using claims databases to evaluate persistence with bisphosphonates for osteoporosis. *Curr Med Res Opin.* 2007;23:2369–77.
170. Hughes MD, Williams PL. Challenges in using observational studies to evaluate adverse effects of treatment. *N Engl J Med.* 2007;356:1705–7.
171. Smeeth L, Douglas I, Hubbard R. Commentary: we still need observational studies of drugs—they just need to be better. *Int J Epidemiol.* 2006;35:1310–1.
172. de Vries F, de Vries C, Cooper C, et al. Reanalysis of two studies with contrasting results on the association between statin use and fracture risk: the General Practice Research Database. *Int J Epidemiol.* 2006;35:1301–8.
173. Suissa S. Observational studies of inhaled corticosteroids in chronic obstructive pulmonary disease: misconstrued immortal time bias. *Am J Respir Crit Care Med.* 2006;173:464; author reply 464–5.
174. Etminan M, Gill S, Fitzgerald M, et al. Challenges and opportunities for pharmacoepidemiology in drug-therapy decision making. *J Clin Pharmacol.* 2006;46:6–9.
175. Ray WA. Observational studies of drugs and mortality. *N Engl J Med.* 2005;353:2319–21.
176. Holbrook A, Grootendorst P, Willison D, et al. Can current electronic systems meet drug safety and effectiveness requirements? *AMIA Annu Symp Proc.* 2005:335–9.
177. Schneeweiss S, Wang PS. Claims data studies of sedative-hypnotics and hip fractures in older people: exploring residual confounding using survey information. *J Am Geriatr Soc.* 2005;53:948–54.
178. Schneeweiss S, Wang PS. Association between SSRI use and hip fractures and the effect of residual confounding bias in claims database studies. *J Clin Psychopharmacol.* 2004;24:632–8.
179. Kiri VA, Vestbo J, Pride NB, et al. Inhaled steroids and mortality in COPD: bias from unaccounted immortal time. *Eur Respir J.* 2004;24:190–1; author reply 191–2.
180. Bannwarth B. Gastrointestinal safety of paracetamol: is there any cause for concern? *Expert Opin Drug Saf.* 2004;3:269–72.
181. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Qual Saf Health Care.* 2003;12 Suppl 2:ii58–63.
182. Strom BL. Data validity issues in using claims data. *Pharmacoepidemiol Drug Saf.* 2001;10:389–92.
183. Sørensen HT, Johnsen SP, Nørgård B. Methodological issues in using prescription and other databases in pharmacoepidemiology. *Nor Epidemiol.* 2001;11:13–8.
184. Hallas J. Pharmacoepidemiology – current opportunities and challenges. *Nor Epidemiol.* 2001;11:7–12.
185. Skegg DC. Pitfalls of pharmacoepidemiology. *BMJ.* 2000;321:1171–2.
186. Cepeda MS, Fife D, Denarié M, et al. Quantification of missing prescriptions in commercial claims databases: results of a cohort study. *Pharmacoepidemiol Drug Saf.* 2017;26:386–92.
187. Camplain R, Kucharska-Newton A, Cuthbertson CC, et al. Misclassification of incident hospitalized and outpatient heart failure in administrative claims data: the Atherosclerosis Risk in Communities (ARIC) study. *Pharmacoepidemiol Drug Saf.* 2017;26:421–8.
188. Palmaro A, Moulis G, Despas F, et al. Overview of drug data within French health insurance databases and implications for pharmacoepidemiological studies. *Fundam Clin Pharmacol.* 2016;30:616–24.

189. Haneuse S, Bogart A, Jazic I, et al. Learning about missing data mechanisms in electronic health records-based research: a survey-based approach. *Epidemiology*. 2016;27:82–90.
190. Lanes S, Brown JS, Haynes K, et al. Identifying health outcomes in healthcare databases. *Pharmacoepidemiol Drug Saf*. 2015;24:1009–16.
191. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. 2015;44:827–36.
192. Weil G, Motamed C, Eghiaian A, et al. The use of a clinical database in an anesthesia unit: focus on its limits. *J Clin Monit Comput*. 2015;29:163–7.
193. de Groot MC, Klungel OH, Leufkens HG, et al. Sources of heterogeneity in case–control studies on associations between statins, ACE–inhibitors, and proton pump inhibitors and risk of pneumonia. *Eur J Epidemiol*. 2014;29:767–75.
194. Li X, Stürmer T, Brookhart MA. Evidence of sample use among new users of statins: implications for pharmacoepidemiology. *Med Care*. 2014;52:773–80.
195. Lauffenburger JC, Balasubramanian A, Farley JF, et al. Completeness of prescription information in US commercial claims databases. *Pharmacoepidemiol Drug Saf*. 2013;22:899–906.
196. Maciejewski ML. Potential bias in medication adherence studies of prevalent users. *Health Serv Res*. 2013;48:1468–86.
197. Grimes DA. Epidemiologic research using administrative databases: garbage in, garbage out. *Obstet Gynecol*. 2010;116:1018–9.
198. Velthove KJ, Leufkens HG, Souverein PC, et al. Testing bias in clinical databases: methodological considerations. *Emerg Themes Epidemiol*. 2010;7:2.
199. Oostenbrink R, Moons KG, Bleeker SE, et al. Diagnostic research on routine care data: prospects and problems. *J Clin Epidemiol*. 2003;56:501–6.
200. Wade RL, Patel JG, Hill JW, et al. Estimation of missed statin prescription use in an administrative claims dataset. *J Manag Care Spec Pharm*. 2017;23:936–42.
201. Czwikla J, Jobski K, Schink T. The impact of the lookback period and definition of confirmatory events on the identification of incident cancer cases in administrative data. *BMC Med Res Methodol*. 2017;17:122.
202. Pauly NJ, Talbert JC, Brown J. Low-cost generic program use by Medicare beneficiaries: implications for medication exposure misclassification in administrative claims data. *J Manag Care Spec Pharm*. 2016;22:741–51.
203. Mazzali C, Paganoni AM, Ieva F, et al. Methodological issues on the use of administrative data in healthcare research: the case of heart failure hospitalizations in Lombardy region, 2000 to 2012. *BMC Health Serv Res*. 2016;16:234.
204. Hampp C, Greene P, Pinheiro SP. Use of prescription drug samples in the USA: a descriptive study with considerations for pharmacoepidemiology. *Drug Saf*. 2016;39:261–70.
205. Winterstein AG, Kubilis P, Bird S, et al. Misclassification in assessment of diabetogenic risk using electronic health records. *Pharmacoepidemiol Drug Saf*. 2014;23:875–81.
206. Skurtveit S, Selmer R, Tverdal A, et al. Drug exposure: inclusion of dispensed drugs before pregnancy may lead to underestimation of risk associations. *J Clin Epidemiol*. 2013;66:964–72.
207. Gamble JM, McAlister FA, Johnson JA, et al. Restrictive drug coverage policies can induce substantial drug exposure misclassification in pharmacoepidemiologic studies. *Clin Ther*. 2012;34:1379–86.

208. van Walraven C, Bennett C, Forster AJ. Administrative database research infrequently used validated diagnostic or procedural codes. *J Clin Epidemiol*. 2011;64:1054–9.
209. Hoover KW, Tao G, Kent CK, et al. Epidemiologic research using administrative databases: garbage in, garbage out. *Obstet Gynecol*. 2011;117:729; author reply 729–30.
210. Dore DD, Chaudhry S, Hoffman C, et al. Stratum-specific positive predictive values of claims for acute pancreatitis among commercial health insurance plan enrollees with diabetes mellitus. *Pharmacoepidemiol Drug Saf*. 2011;20:209–13.
211. Lanes SF, de Luise C. Bias due to false-positive diagnoses in an automated health insurance claims database. *Drug Saf*. 2006;29:1069–75.
212. Ray WA, Thapa PB, Gideon P. Misclassification of current benzodiazepine exposure by use of a single baseline measurement and its effects upon studies of injuries. *Pharmacoepidemiol Drug Saf*. 2002;11:663–9.
213. Palmaro A, Boucherie Q, Dupouy J, et al. Immeasurable time bias due to hospitalization in medico-administrative databases: which impact for pharmacoepidemiological studies? *Pharmacoepidemiol Drug Saf*. 2017;26:544–53.
214. Targownik LE, Suissa S. Understanding and avoiding immortal-time bias in gastrointestinal observational research. *Am J Gastroenterol*. 2015;110:1647–50.
215. Matok I, Azoulay L, Yin H, et al. Immortal time bias in observational studies of drug effects in pregnancy. *Birth Defects Res A Clin Mol Teratol*. 2014;100:658–62.
216. Cook EA, Schneider KM, Chrischilles E, et al. Accounting for unobservable exposure time bias when using medicare prescription drug data. *Medicare Medicaid Res Rev*. 2013;3. pii: mmrr.003.04.a01.
217. Suissa S, Dell'aniello S, Vahey S, et al. Time-window bias in case-control studies: statins and lung cancer. *Epidemiology*. 2011;22:228–31.
218. Kiri VA, Mackenzie G. Re: “Immortal time bias in pharmacoepidemiology”. *Am J Epidemiol*. 2009;170:667–8; author reply 668–9.
219. Suissa S, Ernst P. Bias in observational study of the effectiveness of nasal corticosteroids in asthma. *J Allergy Clin Immunol*. 2005;115:714–9.
220. Suissa S. Inhaled steroids and mortality in COPD: bias from unaccounted immortal time. *Eur Respir J*. 2004;23:391–5.
221. Sin DD, Man SF, Tu JV. Inhaled glucocorticoids in COPD: immortal time bias. *Am J Respir Crit Care Med*. 2003;168:126–7.
222. Pride NB, Vestbo J, Soriano JB, et al. Inhaled glucocorticoids in COPD: immortal time bias. *Am J Respir Crit Care Med*. 2003;168:127.
223. Suissa S. Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: immortal time bias in observational studies. *Am J Respir Crit Care Med*. 2003;168:49–53.
224. Jiang HY, Chen HZ, Hu XJ, et al. Use of selective serotonin reuptake inhibitors and risk of upper gastrointestinal bleeding: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. 2015;13:42–50.
225. Melloni C, Washam JB, Jones WS, et al. Conflicting results between randomized trials and observational studies on the impact of proton pump inhibitors on cardiovascular events when coadministered with dual antiplatelet therapy: systematic review. *Circ Cardiovasc Qual Outcomes*. 2015;8:47–55.

226. Varas-Lorenzo C, Riera-Guardia N, Calingaert B, et al. Myocardial infarction and individual nonsteroidal anti-inflammatory drugs meta-analysis of observational studies. *Pharmacoepidemiol Drug Saf.* 2013;22:559–70.
227. Hernández-Díaz S, Varas-Lorenzo C, García Rodríguez LA. Non-steroidal antiinflammatory drugs and the risk of acute myocardial infarction. *Basic Clin Pharmacol Toxicol.* 2006;98:266–74.
228. Cheelo M, Lodge CJ, Dharmage SC, et al. Paracetamol exposure in pregnancy and early childhood and development of childhood asthma: a systematic review and meta-analysis. *Arch Dis Child.* 2015;100:81–9.
229. Weiss J, Freeman M, Low A, et al. Benefits and harms of intensive blood pressure treatment in adults aged 60 years or older: a systematic review and meta-analysis. *Ann Intern Med.* 2017;166: 419–29.
230. Bally M, Dendukuri N, Rich B, et al. Risk of acute myocardial infarction with NSAIDs in real world use: bayesian meta-analysis of individual patient data. *BMJ.* 2017;357:j1909.
231. Sordo L, Barrio G, Bravo MJ, et al. Mortality risk during and after opioid substitution treatment: systematic review and meta-analysis of cohort studies. *BMJ.* 2017;357:j1550.
232. Tariq R, Singh S, Gupta A, et al. Association of Gastric Acid Suppression with Recurrent *Clostridium difficile* infection: a systematic review and meta-analysis. *JAMA Intern Med.* 2017;177:784–91.
233. Maruthur NM, Tseng E, Hutfless S, et al. Diabetes medications as monotherapy or metformin-based combination therapy for type 2 diabetes: a systematic review and meta-analysis. *Ann Intern Med.* 2016;164:740–51.
234. Paul S, Saxena A, Terrin N, et al. Hepatitis B virus reactivation and prophylaxis during solid tumor chemotherapy: a systematic review and meta-analysis. *Ann Intern Med.* 2016;164:30–40.
235. Li L, Li S, Deng K, et al. Dipeptidyl peptidase-4 inhibitors and risk of heart failure in type 2 diabetes: systematic review and meta-analysis of randomised and observational studies. *BMJ.* 2016;352:i610.
236. Molnar AO, Fergusson D, Tsampalieros AK, et al. Generic immunosuppression in solid organ transplantation: systematic review and meta-analysis. *BMJ.* 2015;350:h3163.
237. Ziff OJ, Lane DA, Samra M, et al. Safety and efficacy of digoxin: systematic review and meta-analysis of observational and controlled trial data. *BMJ.* 2015;351 h4451. Erratum in: *BMJ* 2015;351:h4937.
238. Collaborative Group On Epidemiological Studies Of Ovarian Cancer, Beral V, Gaitskell K, et al. Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies. *Lancet.* 2015;385:1835–42.
239. Bellemain-Appaix A, Kerneis M, O'Connor SA, et al. Reappraisal of thienopyridine pretreatment in patients with non-ST elevation acute coronary syndrome: a systematic review and meta-analysis. *BMJ.* 2014;349:g6269.
240. Grigoriadis S, Vonderporten EH, Mamisashvili L, et al. Prenatal exposure to antidepressants and persistent pulmonary hypertension of the newborn: systematic review and meta-analysis. *BMJ.* 2014;348:f6932.
241. Li L, Shen J, Bala MM, et al. Incretin treatment and risk of pancreatitis in patients with type 2 diabetes mellitus: systematic review and meta-analysis of randomised and non-randomised studies. *BMJ.* 2014;348:g2366.

242. Kalil AC, Van Schooneveld TC, Fey PD, et al. Association between vancomycin minimum inhibitory concentration and mortality among patients with *Staphylococcus aureus* bloodstream infections: a systematic review and meta-analysis. *JAMA*. 2014;312:1552–64.
243. Stegeman BH, de Bastos M, Rosendaal FR, et al. Different combined oral contraceptives and the risk of venous thrombosis: systematic review and network meta-analysis. *BMJ*. 2013;347:f5298.
244. Maneiro JR, Salgado E, Gomez-Reino JJ. Immunogenicity of monoclonal antibodies against tumor necrosis factor used in chronic immune-mediated inflammatory conditions: systematic review and meta-analysis. *JAMA Intern Med*. 2013;173:1416–28.
245. Hartling L, Abou-Setta AM, Dursun S, et al. Antipsychotics in adults with schizophrenia: comparative effectiveness of first-generation versus second-generation medications: a systematic review and meta-analysis. *Ann Intern Med*. 2012;157:498–511.
246. Hsu J, Santesso N, Mustafa R, et al. Antivirals for treatment of influenza: a systematic review and meta-analysis of observational studies. *Ann Intern Med*. 2012;156:512–24.
247. Caldeira D, Alarcão J, Vaz-Carneiro A, et al. Risk of pneumonia associated with use of angiotensin converting enzyme inhibitors and angiotensin receptor blockers: systematic review and meta-analysis. *BMJ*. 2012;345:e4260.
248. MacArthur GJ, Minozzi S, Martin N, et al. Opiate substitution treatment and HIV transmission in people who inject drugs: systematic review and meta-analysis. *BMJ*. 2012;345:e5945.
249. Mantha S, Karp R, Raghavan V, et al. Assessing the risk of venous thromboembolic events in women taking progestin-only contraception: a meta-analysis. *BMJ*. 2012;345:e4944.
250. Silvain J, Beygui F, Barthélémy O, et al. Efficacy and safety of enoxaparin versus unfractionated heparin during percutaneous coronary intervention: systematic review and meta-analysis. *BMJ*. 2012;344:e553.
251. McKnight RF, Adida M, Budge K, et al. Lithium toxicity profile: a systematic review and meta-analysis. *Lancet*. 2012;379:721–8.
252. Cohen JM, Wood ME, Hernandez-Diaz S, et al. Agreement between paternal self-reported medication use and records from a national prescription database. *Pharmacoepidemiol Drug Saf*. 2018;27:413–21.
253. Brookhart MA, Schneeweiss S, Rothman KJ, et al. Variable selection for propensity score models. *Am J Epidemiol*. 2006;163:1149–56.
254. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf*. 2006;15:291–303.
255. Lunt M, Glynn RJ, Rothman KJ, et al. Propensity score calibration in the absence of surrogacy. *Am J Epidemiol*. 2012;175:1294–302.
256. Stürmer T, Rothman KJ, Avorn J, et al. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution—a simulation study. *Am J Epidemiol*. 2010;172:843–54.
257. Yoshida K, Solomon DH, Kim SC. Active-comparator design and new-user design in observational studies. *Nat Rev Rheumatol*. 2015;11:437–41.
258. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183:758–64.
259. Zhang Y, Thamer M, Kaufman J, et al. Comparative effectiveness of two anemia management strategies for complex elderly dialysis patients. *Med Care*. 2014;52 Suppl 3:S132–9.

260. Stürmer T, Jonsson Funk M, Poole C, et al. Nonexperimental comparative effectiveness research using linked healthcare databases. *Epidemiology*. 2011;22:298–301.
261. Tamim H, Monfared AA, LeLorier J. Application of lag–time into exposure definitions to control for protopathic bias. *Pharmacoepidemiol Drug Saf*. 2007;16:250–8.
262. Rebolj Kodre A, Pohar Perme M. Informative censoring in relative survival. *Stat Med*. 2013;32:4791–802.
263. Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep*. 2015;2:221–8.
264. Rubin DB. Inference and missing data. *Biometrika*. 1976;63:581–92.
265. Ibrahim JG, Molenberghs G. Missing data methods in longitudinal studies: a review. *Test (Madr)*. 2009;18:1–43.
266. Siddique J, Harel O, Crespi CM. Addressing missing data mechanism uncertainty using multiple-model multiple imputation: application to a longitudinal clinical trial. *Ann Appl Stat*. 2012;6:1814–37.
267. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res*. 2013;22:278–9.
268. Resseguier N, Giorgi R, Paoletti X. Sensitivity analysis when data are missing not-at-random. *Epidemiology*. 2011;22:282.
269. Kosinski AS, Flanders WD. Evaluating the exposure and disease relationship with adjustment for different types of exposure misclassification: a regression approach. *Stat Med*. 1999;18:2795–808.
270. Prada-Ramallal G, Takkouche B, Figueiras A. Summarising the evidence for drug safety: a methodological discussion of different meta-analysis approaches. *Drug Saf*. 2017;40:547–58.
271. Lin KJ, Schneeweiss S. Considerations for the analysis of longitudinal electronic health records linked to claims data to study the effectiveness and safety of drugs. *Clin Pharmacol Ther*. 2016;100:147–59.
272. Dokholyan RS, Muhlbaier LH, Falletta JM, et al. Regulatory and ethical considerations for linking clinical and administrative databases. *Am Heart J*. 2009;157:971–82.
273. Prada-Ramallal G, Takkouche B, Figueiras A. Diverging conclusions from the same meta-analysis in drug safety: source of data (primary versus secondary) takes a toll. *Drug Saf*. 2017;40:351–8.
274. Prada-Ramallal G, Roque F, Herdeiro MT, et al. Primary versus secondary source of data in observational studies and heterogeneity in meta-analyses of drug effects: a survey of major medical journals. *BMC Med Res Methodol*. 2018;18:97.
275. Donahue JG, Weiss ST, Livingston JM, et al. Inhaled steroids and the risk of hospitalization for asthma. *JAMA*. 1997;277:887–91.
276. Rochon PA, Tu JV, Anderson GM, et al. Rate of heart failure and 1-year survival for older people receiving low-dose beta-blocker therapy after myocardial infarction. *Lancet*. 2000;356:639–44.
277. Mantel N, Byar DP. Evaluation of response–time data involving transient states: an illustration using heart–transplant data. *J Am Stat Assoc*. 1974;69:81–6.
278. Karim ME, Gustafson P, Petkau J, et al. Comparison of statistical approaches for dealing with immortal time bias in drug effectiveness studies. *Am J Epidemiol*. 2016;184:325–35.

279. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005;8:19–32.
280. Bouvy JC, De Bruin ML, Koopmanschap MA. Epidemiology of adverse drug reactions in Europe: a review of recent observational studies. *Drug Saf.* 2015;38:437–53.
281. Clark DW, Coulter DM, Besag FM. Randomized controlled trials and assessment of drug safety. *Drug Saf.* 2008;31:1057–61.
282. Reynolds RF, Lem JA, Gatto NM, et al. Is the large simple trial design used for comparative, post-approval safety research? A review of a clinical trials registry and the published literature. *Drug Saf.* 2011;34:799–820.
283. CIOMS Working Group X. Evidence synthesis and meta-analysis for drug safety: report of CIOMS Working Group X. Geneva: Council for International Organizations of Medical Sciences (CIOMS); 2016.
284. Ioannidis JP, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. *JAMA.* 2001;285:437–43.
285. Chou R, Helfand M. Challenges in systematic reviews that assess treatment harms. *Ann Intern Med.* 2005;142:1090–9.
286. Nuesch E, Trelle S, Reichenbach S, et al. The effects of excluding patients from the analysis in randomized controlled trials: metaepidemiological study. *BMJ.* 2009;339:b3244.
287. Hammad TA, Pinheiro SP, Neyarapally GA. Secondary use of randomized controlled trials to evaluate drug safety: a review of methodological considerations. *Clin Trials.* 2011;8:559–70.
288. Singh S, Loke YK. Drug safety assessment in clinical trials: methodological challenges and opportunities. *Trials.* 2012;13:138.
289. Vandembroucke JP. When are observational studies as credible as randomised trials? *Lancet.* 2004;363:1728–31.
290. Vandembroucke JP. The HRT controversy: observational studies and RCTs fall in line. *Lancet.* 2009;373:1233–5.
291. Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ.* 2006;174:635–41.
292. Vandembroucke JP. What is the best evidence for determining harms of medical treatment? *CMAJ.* 2006;174:645–6.
293. Reeves BC, Deeks JJ, Higgins JPT, et al. *Chapter 13: Including non-randomized studies.* In: Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions version 5.1.0 (updated March 2011).* The Cochrane Collaboration; 2011. Disponible en: <https://training.cochrane.org/handbook>. Acceso el 20 de febrero de 2019.
294. Golder S, Loke YK, Bland M. Meta-analyses of adverse effects data derived from randomised controlled trials as compared to observational studies: methodological overview. *PLoS Med.* 2011;8:e1001026.
295. Bailey C, Peddie D, Wickham ME, et al. Adverse drug event reporting systems-a systematic review. *Br J Clin Pharmacol.* 2016;82:17–29.
296. Saini P, Loke YK, Gamble C, et al. Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews. *BMJ.* 2014;349:g6501.
297. Rücker G, Schwarzer G, Carpenter J, et al. Why add anything to nothing? The arcsine difference as a measure of treatment effect in meta-analysis with zero cells. *Stat Med.* 2009;28:721–38.

298. Song F, Eastwood AJ, Gilbody S, et al. Publication and related biases. *Health Technol Assess.* 2000;4:1–115.
299. Morrison A, Moulton K, Clark M, Polisena, et al. English-language restriction when conducting systematic review-based meta-analyses: systematic review of published studies. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009.
300. Sedgwick P. What is publication bias in a meta-analysis? *BMJ.* 2015;351:h4419.
301. Senn SJ. Overstating the evidence: double counting in meta-analysis and related problems. *BMC Med Res Methodol.* 2009;9:10.
302. Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev.* 2007:MR000005.
303. Kotecha D, Manzano L, Krum H, et al. Beta-Blockers in Heart Failure Collaborative Group. Effect of age and sex on efficacy and tolerability of b blockers in patients with heart failure with reduced ejection fraction: individual patient data meta-analysis. *BMJ.* 2016;353:i1855.
304. Reichenpfader U, Gartlehner G, Morgan LC, et al. Sexual dysfunction associated with second-generation antidepressants in patients with major depressive disorder: results from a systematic review with network meta-analysis. *Drug Saf.* 2014;37:19–31.
305. Veroniki AA, Straus SE, Ashoor HM, et al. Comparative safety and effectiveness of cognitive enhancers for Alzheimer's dementia: protocol for a systematic review and individual patient data network meta-analysis. *BMJ Open.* 2016;6:e010251.
306. Golder S, Loke YK, Wright K, Norman G. Reporting of adverse events in published and unpublished studies of health care interventions: a systematic review. *PLoS Med.* 2016;13:e1002127.
307. FDA Amendments Act. Public Law 110-85. 2007. Ref Type: Statute.
308. Law MR, Kawasumi Y, Morgan SG. Despite law, fewer than one in eight completed studies of drugs and biologics are reported on time on ClinicalTrials.gov. *Health Aff (Millwood).* 2011;30:2338–45.
309. Zarin DA, Tse T, Williams RJ, Carr S. Trial reporting in ClinicalTrials.gov—the final rule. *N Engl J Med.* 2016;375:1998–2004.
310. Aalaei-Andabili SH, Alavian SM. Important steps for a reliable meta-analysis. *Lancet Infect Dis.* 2012;12:663.
311. Sterne JAC, Egger M, Moher D. *Chapter 10: Addressing reporting biases.* In: Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions version 5.1.0* (updated March 2011). The Cochrane Collaboration; 2011. Disponible en: <https://training.cochrane.org/handbook>. Acceso el 20 de febrero de 2019.
312. Terrin N, Schmid CH, Lau J. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *J Clin Epidemiol.* 2005;58:894–901.
313. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315:629–34.
314. Loke YK, Mattishent K. If nothing happens, is everything all right? Distinguishing genuine reassurance from a false sense of security. *CMAJ.* 2015;187:15–6.
315. Chou R, Aronson N, Atkins D, et al. AHRQ series paper 4: assessing harms when comparing medical interventions: AHRQ and the effective health-care program. *J Clin Epidemiol.* 2010;63:502–12.

316. Figueiras A, Ferreira MT, Gestal JJ. *Farmacovigilancia*. In: Fernández-Crehuet J, Gestal JJ, Delgado M, et al., editors. Piédrola Gil. Medicina Preventiva y Salud Pública. 12a Edición. Barcelona: Elsevier Masson; 2015. pp. 1093–104.
317. Blettner M, Sauerbrei W, Schlehofer B, et al. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol*. 1999;28:1–9.
318. Nordmann AJ, Kasenda B, Briel M. Meta-analyses: what they can and cannot do. *Swiss Med Wkly*. 2012;142:w13518.
319. Riley RD, Simmonds MC, Look MP. Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. *J Clin Epidemiol*. 2007;60:431–9.
320. Friedenreich CM. Methods for pooled analyses of epidemiologic studies. *Epidemiology*. 1993;4:295–302.
321. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ*. 2005;331:897–900.
322. Song F, Loke YK, Walsh T, et al. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ*. 2009;338:b1147.
323. Cameron C, Fireman B, Hutton B, et al. Network meta-analysis incorporating randomized controlled trials and non-randomized comparative cohort studies for assessing the safety and effectiveness of medical treatments: challenges and opportunities. *Syst Rev*. 2015;4:147.
324. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56:455–63.
325. Lyman GH, Kuderer NM. The strengths and limitations of meta-analyses based on aggregate data. *BMC Med Res Methodol*. 2005;5:14.
326. Tobías A, Saez M, Kogevinas M. Meta-analysis of results and individual patient data in epidemiological studies. *J Mod Appl Stat Methods*. 2004;1:176–85.
327. Oakes M. *On meta-analysis*. In: Statistical inference. Chestnut Hill: Epidemiology Resources Inc; 1990. pp. 157–63.
328. Thompson SG, Pocock SJ. Can meta-analyses be trusted? *Lancet*. 1991;338:1127–30.
329. Feinstein AR. Meta-analysis: statistical alchemy for the 21st century. *J Clin Epidemiol*. 1995;48:71–9.
330. Egger M, Smith GD. Meta-analysis. Potentials and promise. *BMJ*. 1997;315:1371–4.
331. Ioannidis JP, Lau J. Pooling research results: benefits and limitations of meta-analysis. *Jt Comm J Qual Improv*. 1999;25:462–9.
332. Charlton BG. The uses and abuses of meta-analysis. *Fam Pract*. 1996;13:397–401.
333. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med*. 2007;356:2457–71.
334. Hu M, Cappelleri JC, Lan KK. Applying the law of iterated logarithm to control type I error in cumulative meta-analysis of binary outcomes. *Clin Trials*. 2007;4:329–40.
335. Berlin JA. Benefits of heterogeneity in meta-analysis of data from epidemiologic studies. *Am J Epidemiol*. 1995;142:383–7.
336. Higgins JP, Thompson SG. Quantifying heterogeneity in meta-analysis. *Stat Med*. 2002;21:1539–58.

337. Berlin JA, Santanna J, Schmid CH, et al. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med*. 2002;21:371–87.
338. Riley RD, Lambert PC, Staessen JA, et al. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Stat Med*. 2008;27:1870–93.
339. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ*. 2010;340:c221.
340. Simmonds MC, Higgins JP, Stewart LA, et al. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clin Trials*. 2005;2:209–17.
341. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet*. 1993;341:418–22.
342. Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. *Stat Med*. 1995;14:2057–79.
343. Stewart LA, Clarke M, Rovers M, et al. PRISMA-IPD Development Group. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA*. 2015;313:1657–65.
344. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof*. 2002;25:76–97.
345. Burgess S, White IR, Resche-Rigon M, et al. Combining multiple imputation and meta-analysis with individual participant data. *Stat Med*. 2013;32:4499–514.
346. Lambert PC, Sutton AJ, Abrams KR, et al. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *J Clin Epidemiol*. 2002;55:86–94.
347. Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *BMJ*. 2012;344:d7762.
348. Abo-Zaid G, Guo B, Deeks JJ, et al. Individual participant data meta-analyses should not ignore clustering. *J Clin Epidemiol*. 2013;66(865–873):e4.
349. Debray TP, Moons KG, van Valkenhoef G, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. *Res Synth Methods*. 2015;6:293–309.
350. Wolfson M, Wallace SE, Masca N, et al. DataSHIELD: resolving a conflict in contemporary bioscience--performing a pooled analysis of individual-level data without sharing the data. *Int J Epidemiol*. 2010;39:1372–82.
351. Simpson EH. The interpretation of interaction in contingency tables. *J R Statist Soc B*. 1951;2:238–41.
352. Dong J. On avoiding association paradoxes in contingency tables. *J Syst Sci Complex*. 1998;11:272–79.
353. Bravata DM, Olkin I. Simple pooling versus combining in meta-analysis. *Eval Health Prof*. 2001;24:218–30.
354. Bracken MB. Rosiglitazone and cardiovascular risk. *N Engl J Med*. 2007;357:937–8; author reply 939–40.
355. Lièvre M, Cucherat M, Leizorovicz A. Pooling, meta-analysis, and the evaluation of drug safety. *Curr Control Trials Cardiovasc Med*. 2002;3:6.
356. Altman DG, Deeks JJ. Meta-analysis, Simpson's paradox, and the number needed to treat. *BMC Med Res Methodol*. 2002;2:3.

357. Dickinson T, Malhi S, Painter S, et al. Association between SSRIs and upper gastrointestinal bleeding. Self treatment with non-steroidal drugs may be confounding factor. *BMJ*. 2000;320:1406.
358. Tougas G, Chen Y, Hwang P, et al. Prevalence and impact of upper gastrointestinal symptoms in the Canadian population: findings from the DIGEST study. *Domestic/International Gastroenterology Surveillance Study*. *Am J Gastroenterol*. 1999;94:2845–54.
359. Nagata N, Niikura R, Sekine K, et al. Risk of peptic ulcer bleeding associated with *Helicobacter pylori* infection, nonsteroidal anti-inflammatory drugs, low-dose aspirin, and antihypertensive drugs: a case-control study. *J Gastroenterol Hepatol*. 2015;30:292–8.
360. Tanskanen A, Tuomilehto J, Viinamäki H, et al. Joint heavy use of alcohol, cigarettes and coffee and the risk of suicide. *Addiction*. 2000;95:1699–704.
361. U.S. Food and Drug Administration. Public Health Advisory: updated safety information about a drug interaction between clopidogrel bisulfate (marketed as Plavix) and omeprazole (marketed as Prilosec and Prilosec OTC). November 17, 2009. Disponible en: <https://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/DrugSafetyInformationforHeathcareProfessionals/ucm190787.htm>. Acceso el 1 de diciembre de 2016.
362. European Medicines Agency. Public statement on possible interaction between clopidogrel and proton pump inhibitors. May 29, 2009. Disponible en: <https://www.ema.europa.eu/en/news/public-statement-possible-interaction-between-clopidogrel-proton-pump-inhibitors>. Acceso el 20 de febrero de 2019.
363. Glasziou PP, Sanders SL. Investigating causes of heterogeneity in systematic reviews. *Stat Med*. 2002;21:1503–11.
364. Seeger J, Daniel GW. *Commercial Insurance Databases*. In: Strom BL, Kimmel SE, Hennessy S, editors. *Pharmacoepidemiology*. 5th ed. Chichester: John Wiley & Sons; 2012. pp. 189–208.
365. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700.
366. Zorzela L, Loke YK, Ioannidis JP, et al. PRISMA harms checklist: improving harms reporting in systematic reviews. *BMJ*. 2016;352:i157.
367. Upadhyia KK, Santelli JS, Raine-Bennett TR, et al. Over-the-counter access to oral contraceptives for adolescents. *J Adolesc Health*. 2017;60:634–40.
368. Raebel MA, Schmittiel J, Karter AJ, et al. Standardizing terminology and definitions of medication adherence and persistence in research employing electronic databases. *Med Care*. 2013;51(8 Suppl 3):S11–21.
369. Wu AC, Butler MG, Li L, et al. Primary adherence to controller medications for asthma is poor. *Ann Am Thorac Soc*. 2015;12:161–6.
370. Fallis BA, Dhalla IA, Klemensberg J, et al. Primary medication non-adherence after discharge from a general internal medicine service. *PLoS One*. 2013;8:e61735.
371. Anderson KL, Dothard EH, Huang KE, et al. Frequency of primary nonadherence to acne treatment. *JAMA Dermatol*. 2015;151:623–6.
372. Fischer MA, Stedman MR, Lii J, et al. Primary medication non-adherence: analysis of 195,930 electronic prescriptions. *J Gen Intern Med*. 2010;25:284–90.

373. Tambllyn R, Eguale T, Huang A, et al. The incidence and determinants of primary nonadherence with prescribed medication in primary care: a cohort study. *Ann Intern Med.* 2014;160:441–50.
374. Kolandaivelu K, Leiden BB, O'Gara PT, et al. Non-adherence to cardiovascular medications. *Eur Heart J.* 2014;35:3267–76.
375. Kirkeby MJ, Hansen CD, Andersen JH. Socio-economic differences in use of prescribed and over-the-counter medicine for pain and psychological problems among Danish adolescents--a longitudinal study. *Eur J Pediatr.* 2014;173:1147–55.
376. Mukherjee K, Kamal KM. Sociodemographic determinants of out-of-pocket expenditures for patients using prescription drugs for rheumatoid arthritis. *Am Health Drug Benefits.* 2017;10:7–15.
377. Karter AJ, Parker MM, Solomon MD, et al. Effect of out-of-pocket cost on medication initiation, adherence, and persistence among patients with type 2 diabetes: The Diabetes Study of Northern California (DISTANCE). *Health Serv Res.* 2018;53:1227–47.
378. Leong A, Dasgupta K, Bernatsky S, et al. Systematic review and meta-analysis of validation studies on a diabetes case definition from health administrative records. *PLoS One.* 2013;8:e75256.
379. Garfield E. The history and meaning of the journal impact factor. *JAMA.* 2006;295:90–3.
380. Seglen PO. Why the impact factor of journals should not be used for evaluating research. *BMJ.* 1997;314:498–502.
381. Brown H. How impact factors changed medical publishing--and science. *BMJ.* 2007;334:561–4.
382. Gerber S, Tallon D, Trelle S, et al. Bibliographic study showed improving methodology of meta-analyses published in leading journals 1993-2002. *J Clin Epidemiol.* 2007;60:773–80.
383. Petropoulou M, Nikolakopoulou A, Veroniki AA, et al. Bibliographic study showed improving statistical methodology of network meta-analyses published between 1999 and 2015. *J Clin Epidemiol.* 2017;82:20–8.



8. APÉNDICE.

8.1. ANEXO 1. PRADA-RAMALLAL G, TAKKOUCHE B, FIGUEIRAS A. <i>BIAS IN PHARMACOEPIDEMIOLOGIC STUDIES USING SECONDARY HEALTH CARE DATABASES: A SCOPING REVIEW. BMC MED RES METHODOL. 2019 MAR 11;19(1):53. DOI: 10.1186/s12874-019-0695-Y.</i>	137
8.1.1. Suplemento 1. Detalles de los artículos incluidos en la revisión.	151
8.2. ANEXO 2. PRADA-RAMALLAL G, TAKKOUCHE B, FIGUEIRAS A. <i>SUMMARISING THE EVIDENCE FOR DRUG SAFETY: A METHODOLOGICAL DISCUSSION OF DIFFERENT META-ANALYSIS APPROACHES. DRUG SAF. 2017 JUL;40(7):547-558. DOI: 10.1007/s40264-017-0518-1.</i>	157
8.3. ANEXO 3. PRADA-RAMALLAL G, TAKKOUCHE B, FIGUEIRAS A. <i>DIVERGING CONCLUSIONS FROM THE SAME META-ANALYSIS IN DRUG SAFETY: SOURCE OF DATA (PRIMARY VERSUS SECONDARY) TAKES A TOLL. DRUG SAF. 2017 APR;40(4):351-358. DOI: 10.1007/s40264-016-0492-Z.</i>	169
8.3.1. Suplemento 1. Diagrama de flujo de la estrategia de búsqueda. Características de los estudios incluidos.	177
8.3.2. Suplemento 2. Metanálisis de subgrupos por diseño del estudio.	181
8.4. ANEXO 4. PRADA-RAMALLAL G, ROQUE F, HERDEIRO MT, TAKKOUCHE B, FIGUEIRAS A. <i>PRIMARY VERSUS SECONDARY SOURCE OF DATA IN OBSERVATIONAL STUDIES AND HETEROGENEITY IN META-ANALYSES OF DRUG EFFECTS: A SURVEY OF MAJOR MEDICAL JOURNALS. BMC MED RES METHODOL. 2018 SEP 27;18(1):97. DOI: 10.1186/s12874-018-0561-3.</i>	183
8.4.1. Suplemento 1. Artículos excluidos con los motivos de exclusión.	197
8.5. ANEXO 5. PERMISOS DE PUBLICACIÓN.	211



Prada-Ramallal G, Takkouche B, Figueiras A. *Bias in pharmacoepidemiologic studies using secondary health care databases: a scoping review*. BMC Med Res Methodol. 2019 Mar 11;19(1):53. doi: 10.1186/s12874-019-0695-y.
(Ver [anexo 1](https://link.springer.com/article/10.1186/s12874-019-0695-y): <https://link.springer.com/article/10.1186/s12874-019-0695-y>)





Prada-Ramallal G, Takkouche B, Figueiras A. *Summarising the evidence for drug safety: a methodological discussion of different meta-analysis approaches*. Drug Saf. 2017 Jul;40(7):547-558. doi: 10.1007/s40264-017-0518-1.
(Ver [anexo 2](https://link.springer.com/article/10.1007/s40264-017-0518-1): <https://link.springer.com/article/10.1007/s40264-017-0518-1>)





Prada-Ramallal G, Takkouche B, Figueiras A. *Diverging conclusions from the same meta-analysis in drug safety: source of data (primary versus secondary) takes a toll.* Drug Saf. 2017 Apr;40(4):351-358. doi: 10.1007/s40264-016-0492-z.
(Ver [anexo 3](https://link.springer.com/article/10.1007/s40264-016-0492-z): <https://link.springer.com/article/10.1007/s40264-016-0492-z>)





Prada-Ramallal G, Roque F, Herdeiro MT, Takkouche B, Figueiras A. *Primary versus secondary source of data in observational studies and heterogeneity in meta-analyses of drug effects: a survey of major medical journals*. BMC Med Res Methodol. 2018 Sep 27;18(1):97. doi: 10.1186/s12874-018-0561-3.
(Ver [anexo 4](https://link.springer.com/article/10.1186/s12874-018-0561-3): <https://link.springer.com/article/10.1186/s12874-018-0561-3>)





RightsLink - Your Account

SPRINGER NATURE LICENSE TERMS AND CONDITIONS

This Agreement between Mr. Guillermo Prada Ramallal ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4671390069023
License date	Sep 17, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Drug Safety
Licensed Content Title	Diverging Conclusions from the Same Meta-Analysis in Drug Safety: Source of Data (Primary Versus Secondary) Takes a Toll
Licensed Content Author	Guillermo Prada-Ramallal, Bahi Takkouche, Adolfo Figueiras
Licensed Content Date	Jan 1, 2016
Licensed Content Volume	40
Licensed Content Issue	4
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter
Will you be translating?	yes, without original language
Number of languages	1
Circulation/distribution	1 - 29
Author of this Springer Nature content	yes
Title	Las fuentes de datos en farmacoepidemiología como generadoras de sesgos y heterogeneidad en metanálisis
Institution name	Universidad de Santiago de Compostela
Expected presentation date	Oct 2019
Specific Languages	Spanish
Requestor Location	Mr. Guillermo Prada Ramallal Santiago de Compostela, A Coruña Spain Attn: Mr. Guillermo Prada Ramallal
Total	0.00 EUR
Terms and Conditions	

Springer Nature Customer Service Centre GmbH Terms and Conditions

This agreement sets out the terms and conditions of the licence (the **License**) between you and **Springer Nature Customer Service Centre GmbH** (the **Licensor**). By clicking 'accept' and completing the transaction for the material (**Licensed Material**), you also confirm your acceptance of these terms and conditions.

1. Grant of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, world-wide licence to reproduce the Licensed Material for the purpose specified in your order only. Licences are granted for the specific use requested in the order and for no other use, subject to the conditions below.

RightsLink - Your Account

1. 2. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

1. 3. If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Scope of Licence

2. 1. You may only use the Licensed Content in the manner and to the extent permitted by these Ts&Cs and any applicable laws.

2. 2. A separate licence may be required for any additional use of the Licensed Material, e.g. where a licence has been purchased for print only use, separate permission must be obtained for electronic re-use. Similarly, a licence is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence. Any content owned by third parties are expressly excluded from the licence.

2. 3. Similarly, rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

2. 4. Where permission has been granted **free of charge** for material in print, permission may also be granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

2. 5. An alternative scope of licence may apply to signatories of the [STM Permissions Guidelines](#), as amended from time to time.

3. Duration of Licence

3. 1. A licence for is valid from the date of purchase ('Licence Date') at the end of the relevant period in the below table:

Scope of Licence	Duration of Licence
Post on a website	12 months
Presentations	12 months
Books and journals	Lifetime of the edition in the language purchased

4. Acknowledgement

4. 1. The Licensor's permission must be acknowledged next to the Licensed Material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

5. Restrictions on use

5. 1. Use of the Licensed Material may be permitted for incidental promotional use and minor editing privileges e.g. minor adaptations of single figures, changes of format, colour and/or style where the adaptation is credited as set out in Appendix 1 below. Any other changes including but not limited to, cropping, adapting, omitting material that affect the meaning, intention or moral rights of the author are strictly prohibited.

5. 2. You must not use any Licensed Material as part of any design or trademark.

5. 3. Licensed Material may be used in Open Access Publications (OAP) before publication by Springer Nature, but any Licensed Material must be removed from OAP sites prior to final publication.

6. Ownership of Rights

RightsLink - Your Account

6. 1. Licensed Material remains the property of either Licensor or the relevant third party and any rights not explicitly granted herein are expressly reserved.

7. Warranty

IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL OR INDIRECT DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

8. Limitations

8. 1. **BOOKS ONLY:** Where 'reuse in a dissertation/thesis' has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

9. Termination and Cancellation

9. 1. Licences will expire after the period shown in Clause 3 (above).

9. 2. Licensee reserves the right to terminate the Licence in the event that payment is not received in full or if there has been a breach of this agreement by you.

Appendix 1 — Acknowledgements:**For Journal Content:**

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

For Advance Online Publication papers:

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)]

For Adaptations/Translations:

Adapted/Translated by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

Note: For any republication from the British Journal of Cancer, the following credit line style applies:

Reprinted/adapted/translated by permission from [the Licensor]: on behalf of Cancer Research UK: : [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

For Advance Online Publication papers:

Reprinted by permission from The [the Licensor]: on behalf of Cancer Research UK: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)]

For Book content:

Reprinted/adapted by permission from [the Licensor]: [Book Publisher (e.g. Palgrave Macmillan, Springer etc) [Book Title] by [Book author(s)] [COPYRIGHT] (year of publication)]

Other Conditions:

RightsLink - Your Account

Version 1.2

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.



RightsLink - Your Account

SPRINGER NATURE LICENSE TERMS AND CONDITIONS

This Agreement between Mr. Guillermo Prada Ramallal ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4671390342555
License date	Sep 17, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Drug Safety
Licensed Content Title	Summarising the Evidence for Drug Safety: A Methodological Discussion of Different Meta-Analysis Approaches
Licensed Content Author	Guillermo Prada-Ramallal, Bahi Takkouche, Adolfo Figueiras
Licensed Content Date	Jan 1, 2017
Licensed Content Volume	40
Licensed Content Issue	7
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter
Will you be translating?	yes, without original language
Number of languages	1
Circulation/distribution	1 - 29
Author of this Springer Nature content	yes
Title	Las fuentes de datos en farmacoepidemiología como generadoras de sesgos y heterogeneidad en metanálisis
Institution name	Universidad de Santiago de Compostela
Expected presentation date	Oct 2019
Specific Languages	Spanish
Requestor Location	Mr. Guillermo Prada Ramallal Santiago de Compostela, A Coruña Spain Attn: Mr. Guillermo Prada Ramallal
Total	0.00 EUR
Terms and Conditions	

Springer Nature Customer Service Centre GmbH Terms and Conditions

This agreement sets out the terms and conditions of the licence (the **License**) between you and **Springer Nature Customer Service Centre GmbH** (the **Licensor**). By clicking 'accept' and completing the transaction for the material (**Licensed Material**), you also confirm your acceptance of these terms and conditions.

1. Grant of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, world-wide licence to reproduce the Licensed Material for the purpose specified in your order only. Licences are granted for the specific use requested in the order and for no other use, subject to the conditions below.

RightsLink - Your Account

1. 2. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

1. 3. If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Scope of Licence

2. 1. You may only use the Licensed Content in the manner and to the extent permitted by these Ts&Cs and any applicable laws.

2. 2. A separate licence may be required for any additional use of the Licensed Material, e.g. where a licence has been purchased for print only use, separate permission must be obtained for electronic re-use. Similarly, a licence is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence. Any content owned by third parties are expressly excluded from the licence.

2. 3. Similarly, rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

2. 4. Where permission has been granted **free of charge** for material in print, permission may also be granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

2. 5. An alternative scope of licence may apply to signatories of the [STM Permissions Guidelines](#), as amended from time to time.

3. Duration of Licence

3. 1. A licence for is valid from the date of purchase ('Licence Date') at the end of the relevant period in the below table:

Scope of Licence	Duration of Licence
Post on a website	12 months
Presentations	12 months
Books and journals	Lifetime of the edition in the language purchased

4. Acknowledgement

4. 1. The Licensor's permission must be acknowledged next to the Licensed Material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

5. Restrictions on use

5. 1. Use of the Licensed Material may be permitted for incidental promotional use and minor editing privileges e.g. minor adaptations of single figures, changes of format, colour and/or style where the adaptation is credited as set out in Appendix 1 below. Any other changes including but not limited to, cropping, adapting, omitting material that affect the meaning, intention or moral rights of the author are strictly prohibited.

5. 2. You must not use any Licensed Material as part of any design or trademark.

5. 3. Licensed Material may be used in Open Access Publications (OAP) before publication by Springer Nature, but any Licensed Material must be removed from OAP sites prior to final publication.

6. Ownership of Rights

RightsLink - Your Account

6. 1. Licensed Material remains the property of either Licensor or the relevant third party and any rights not explicitly granted herein are expressly reserved.

7. Warranty

IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL OR INDIRECT DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

8. Limitations

8. 1. **BOOKS ONLY:** Where 'reuse in a dissertation/thesis' has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

9. Termination and Cancellation

9. 1. Licences will expire after the period shown in Clause 3 (above).

9. 2. Licensee reserves the right to terminate the Licence in the event that payment is not received in full or if there has been a breach of this agreement by you.

Appendix 1 — Acknowledgements:**For Journal Content:**

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

For Advance Online Publication papers:

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)]

For Adaptations/Translations:

Adapted/Translated by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

Note: For any republication from the British Journal of Cancer, the following credit line style applies:

Reprinted/adapted/translated by permission from [the Licensor]: on behalf of Cancer Research UK: : [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

For Advance Online Publication papers:

Reprinted by permission from The [the Licensor]: on behalf of Cancer Research UK: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)]

For Book content:

Reprinted/adapted by permission from [the Licensor]: [Book Publisher (e.g. Palgrave Macmillan, Springer etc) [Book Title] by [Book author(s)] [COPYRIGHT] (year of publication)]

Other Conditions:

RightsLink - Your Account

Version 1.2

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

