

Chapter 12 / Capítulo 12

Critical reading of scientific articles on
health: A practical guide (Spanish Edition)
ISBN: 978-9915-9928-3-9
DOI: 10.62486/978-9915-9928-3-9.Ch12
Pages: 156-177



© 2026 The authors. This is an open access
book chapter distributed under the terms of
the Creative Commons Attribution (CC BY)
4.0 License.

Case-Control Studies

Estudios de casos y controles

Mauricio Alejandro Paz del Rio  

Isaura Oberson-Santander  

María Valeria Canedo-Sánchez  

1 Clínica Los Andes. La Paz, Bolivia.

2 Hôpitaux Universitaires de Genève (HUG). Genève, Suisse.

3 Ministerio de Salud y Deportes. Dirección General del Sistema Nacional de Salud. Tarija, Bolivia.

ABSTRACT

Introduction: Case-control studies are an analytical observational design used to investigate the association between an exposure and a previously occurring outcome. Their usefulness is particularly relevant in rare or long-latency diseases, allowing for the efficient exploration of risk factors by comparing individuals with the disease (cases) and without it (controls). **Development:** This design uses the odds ratio (OR) as the primary measure of association, retrospectively evaluating the frequency of exposure in both groups. This chapter highlights fundamental methodological aspects such as the appropriate selection of cases and controls, the objective measurement of exposure, the control of confounding factors, and the identification of biases, including selection, information, and survival biases. It emphasizes the critical interpretation of the OR, confidence intervals, and analyses adjusted using logistic regression. **Conclusions:** Case-control studies represent a valuable tool for generating epidemiological evidence and formulating etiological hypotheses. However, the validity of their findings depends on the methodological rigor applied and a critical interpretation that considers the potential biases and limitations inherent in the design. Their appropriate use contributes to the development of preventive strategies and evidence-based decision-making.

Keywords: Case-Control Studies; Risk Factors; Odds Ratio; Epidemiological Biases; Evidence-Based Medicine

RESUMEN

Introducción: Los estudios de casos y controles constituyen un diseño observacional analítico orientado a investigar la asociación entre una exposición y un desenlace previamente ocurrido. Su utilidad es especialmente relevante en enfermedades raras o de larga latencia, permitiendo explorar factores de riesgo de manera eficiente mediante la comparación de individuos con la enfermedad (casos) y sin ella (controles). **Desarrollo:** Este diseño utiliza el odds ratio (OR) como medida principal de asociación, evaluando retrospectivamente la frecuencia de exposición en ambos grupos. El capítulo destaca aspectos metodológicos fundamentales como la adecuada selección de casos y controles, la medición objetiva de la exposición, el control de factores de confusión y la identificación de sesgos, entre ellos los de selección, información y supervivencia. De esta manera, enfatiza la interpretación crítica del OR, los intervalos de confianza y los análisis ajustados mediante regresión logística. **Conclusiones:** los estudios de casos y controles representan una herramienta valiosa para generar evidencia epidemiológica y formular hipótesis etiológicas. Sin embargo, la validez de sus hallazgos depende del rigor metodológico aplicado y de una interpretación crítica que considere los posibles sesgos y limitaciones inherentes al diseño. Su adecuada utilización contribuye al desarrollo de estrategias preventivas y a la toma de decisiones fundamentadas en evidencia científica.

Palabras clave: Estudios de Casos y Controles; Factores de Riesgo; Odds Ratio; Sesgos Epidemiológicos; Medicina Basada en la Evidencia

1. CASO CLÍNICO

En el hospital de tercer nivel de la ciudad de La Paz, el servicio de oncología identificó un incremento inusual en la frecuencia de diagnósticos de cáncer gástrico en pacientes menores de 55 años. Durante la discusión clínica de estos casos, se observó que varios de los pacientes procedían de zonas periurbanas donde el abastecimiento de agua para consumo humano dependía principalmente de pozos sin sistemas adecuados de tratamiento o filtración. Se presentó un paciente de 48 años, agricultor, con un cuadro clínico de seis semanas de evolución caracterizado por pérdida ponderal superior a 10 kg, dolor epigástrico urente persistente y episodios recurrentes de vómitos. La endoscopia digestiva alta evidenció una lesión ulcerada localizada en el antro gástrico, mientras que el estudio histopatológico de la biopsia confirmó el diagnóstico de adenocarcinoma gástrico.

Durante la anamnesis, el paciente refirió que el agua utilizada para consumo en su comunidad presentaba frecuentemente un “sabor metálico” característico. Asimismo, indicó que, aunque algunos habitantes realizaban procesos de ebullición previos a su consumo, esta práctica no se efectuaba de manera sistemática ni uniforme en toda la población.

A partir de estas observaciones, el equipo de salud formuló las siguientes preguntas de investigación:

- ¿Existe una asociación entre la exposición crónica a concentraciones elevadas de nitratos presentes en el agua de consumo y el desarrollo de cáncer gástrico?
- ¿Factores como el tabaquismo, el consumo de alcohol, los antecedentes familiares de cáncer gástrico u otras exposiciones ambientales podrían actuar como variables de confusión o modificadores del efecto de dicha asociación?

Con el propósito de explorar estas hipótesis, los investigadores identificaron un estudio de casos y controles diseñado para evaluar la posible relación entre la exposición a nitratos y el riesgo de cáncer gástrico.

Los casos estuvieron constituidos por pacientes con diagnóstico reciente de cáncer gástrico confirmado mediante estudio histopatológico, atendidos en el mismo centro hospitalario. Los controles correspondieron a individuos sin antecedente de cáncer gástrico ni patología digestiva mayor, procedentes de la misma población de referencia y seleccionados aleatoriamente en la comunidad. El objetivo principal fue determinar si la frecuencia de exposición a agua potencialmente contaminada con nitratos era mayor entre los casos que entre los controles y, en consecuencia, evaluar si dicha exposición podría contribuir a explicar el incremento observado de cáncer gástrico en adultos jóvenes de esta región.

2. OBJETIVOS

Los estudios de casos y controles tienen como finalidad principal explorar asociaciones entre una exposición y un desenlace. A diferencia de otros diseños, no buscan estimar directamente la incidencia o el riesgo absoluto, sino comparar la frecuencia de exposición en personas con la enfermedad (casos) frente a aquellas que no la presentan (controles).

De manera general, los objetivos de un estudio de casos y controles se pueden organizar en tres niveles:

- Determinar si existe una asociación estadísticamente significativa entre una exposición sospechosa y un desenlace.
- Medir dicha asociación a través del odds ratio (OR), que estima cuántas veces es más (o menos) probable que los casos hayan estado expuestos en comparación con los controles.

OBJETIVOS SECUNDARIOS

- Identificar factores de confusión que puedan alterar la relación entre exposición y desenlace, ajustando el OR mediante técnicas estadísticas (p. ej., regresión logística).
- Explorar modificadores de efecto, es decir, variables que cambian la magnitud de la asociación (por ejemplo, edad, sexo, tabaquismo, dieta).
- Evaluar si los hallazgos se mantienen consistentes en diferentes subgrupos de la población.

OBJETIVOS ADICIONALES SEGÚN CONTEXTO

- Generar hipótesis etiológicas para futuras investigaciones de mayor nivel de evidencia (p. ej., cohortes o ensayos clínicos).
- Valorar la magnitud potencial del impacto en salud pública si la asociación fuese causal (ejemplo: agua contaminada y cáncer gástrico).
- Diseñar estrategias de prevención primaria o secundaria basadas en factores de riesgo modificables.

3. INTRODUCCIÓN

El estudio de casos y controles es un diseño observacional, analítico y generalmente retrospectivo, cuyo propósito es evaluar la relación entre una exposición y un desenlace. A diferencia de los estudios de cohortes, que siguen a los sujetos en el tiempo, en este diseño se parte del desenlace ya ocurrido y se investiga hacia atrás para conocer si los pacientes estuvieron expuestos o no al factor de interés.

Este tipo de estudio resulta especialmente útil en situaciones donde:

- El desenlace es infrecuente o tiene una larga latencia (por ejemplo, cáncer, enfermedades raras).
- La cohorte prospectiva sería inviable por requerir muchos años de seguimiento o grandes recursos.
- Se busca generar hipótesis etiológicas que puedan guiar investigaciones posteriores de mayor nivel de evidencia.

En términos metodológicos, los casos y controles comparan la frecuencia de exposición entre dos grupos:

- Casos: personas que presentan la enfermedad o desenlace de interés.
- Controles: personas libres de la enfermedad, seleccionadas de la misma población donde aparecieron los casos.

La medida de asociación característica de este diseño es el odds ratio (OR). Cuando la enfermedad es poco frecuente, el OR puede aproximar razonablemente al riesgo relativo, lo que se conoce como la rare disease assumption. Sin embargo, si la enfermedad no es rara, el OR tiende a sobreestimar la magnitud de la asociación, lo que debe considerarse al interpretar resultados. Ejemplo aplicado al caso clínico: En el hospital de La Paz, se observa un incremento inusual de cáncer gástrico en adultos jóvenes. Para investigar, se seleccionan como casos a pacientes diagnosticados recientemente con cáncer gástrico confirmado por biopsia, y como controles a sujetos sanos de la misma comunidad sin historia de neoplasias gástricas. El interés principal es determinar si la exposición a agua con alto contenido de nitratos está asociada al desarrollo de cáncer gástrico. Este ejemplo ilustra por qué el diseño de casos y controles es particularmente adecuado: el desenlace (cáncer gástrico en <55 años) es relativamente raro, el tiempo de latencia es prolongado y no sería factible ni eficiente esperar la aparición de nuevos casos mediante una cohorte prospectiva.

4. METODOLOGÍA: COMO LEER CRÍTICAMENTE UN ESTUDIO DE CASOS Y CONTROLES

Al enfrentarse a un artículo científico de casos y controles, la sección de Métodos es la columna vertebral sobre la cual descansa la validez del estudio. Un lector crítico no debe limitarse a comprender lo que hicieron los investigadores, sino que debe preguntarse si lo hicieron de la manera adecuada, si justificaron sus elecciones metodológicas y si esas decisiones permiten confiar en los resultados. La lectura crítica de esta sección implica recorrer una serie de elementos clave, cada uno con posibles fortalezas o riesgos de sesgo.

4.1 POBLACIÓN FUENTE Y PERIODO DE ESTUDIO

El primer aspecto que debe revisarse es la definición de la población fuente. Un estudio de casos y controles es válido solo si los casos y los controles provienen de la misma población en riesgo. Por ello, es necesario que el artículo especifique con claridad cuál fue el marco poblacional y en qué periodo temporal se realizó la investigación. Cuando esta información no está bien delimitada, los resultados pierden aplicabilidad y pueden generar conclusiones engañosas.

En nuestro caso clínico, por ejemplo, si los investigadores de La Paz no precisaran que tanto casos como controles provienen de las zonas periurbanas abastecidas por pozos de agua sin filtración, el lector debería cuestionar de inmediato la comparabilidad de ambos grupos.

4.2 DEFINICIÓN DE CASO

Una vez identificada la población, corresponde examinar cómo fueron definidos los casos. La pregunta crítica es si los autores utilizaron criterios diagnósticos estandarizados, válidos y reproducibles. Además, conviene verificar si los casos fueron incidentes o prevalentes.

- Los casos incidentes, son aquellos pacientes diagnosticados recientemente, dentro de un periodo definido por el estudio. Su inclusión permite acercarse de manera más fiel a la secuencia causal exposición–enfermedad, y además reduce el riesgo del llamado sesgo de supervivencia o sesgo de Neyman que aparece cuando solo se incluyen pacientes que han sobrevivido lo suficiente para ser detectados, lo que excluye a quienes murieron tempranamente o tuvieron formas más agresivas de la enfermedad.
- Casos prevalentes son los pacientes que ya presentaban la enfermedad antes del inicio del estudio, sin importar cuándo ocurrió el diagnóstico. Aunque suelen ser más fáciles de reclutar —pues basta con identificar a quienes ya están en seguimiento o en registros hospitalarios—, su uso tiene limitaciones importantes. Al incluir pacientes con distinta duración de enfermedad, se corre el riesgo de seleccionar preferentemente a aquellos con formas más benignas o de curso lento, lo que introduce un sesgo que tiende a infraestimar la fuerza de la asociación. Además, en los casos prevalentes puede resultar más difícil establecer con precisión la temporalidad entre la exposición y la aparición de la enfermedad, ya que la exposición pudo haber cambiado después del inicio de la patología.

En el ejemplo de cáncer gástrico en La Paz, un estudio de buena calidad debería trabajar con casos incidentes, es decir, pacientes recién diagnosticados y confirmados histológicamente en los últimos meses o años del periodo de estudio. Si, en cambio, se incluyeran como casos a pacientes que llevan más de cinco años conviviendo con el cáncer, la muestra podría estar compuesta por sobrevivientes con características distintas (mayor acceso a tratamiento, tumores menos agresivos), lo que sesgaría la interpretación de la relación con la exposición a nitratos.

Tabla 1: Comparación entre casos incidentes y casos prevalentes en estudios de casos y controles: definiciones, ventajas y limitaciones.

Tipo de caso	Definición	Ventajas	Desventajas / Sesgos
Casos incidentes	Pacientes diagnosticados recientemente, durante el periodo definido por el estudio.	- Reflejan mejor la secuencia causal exposición–enfermedad.	- Requieren vigilancia activa y mayor tiempo de reclutamiento.
		- Reducen el riesgo de sesgo de supervivencia (Neyman).	- Pueden ser más costosos de identificar.
		- Permiten analizar factores asociados al inicio de la enfermedad.	
Casos prevalentes	Pacientes que ya presentaban la enfermedad antes del inicio del estudio, independientemente de cuándo ocurrió el diagnóstico.	- Más fáciles de reclutar (hospitales, registros).	- Mayor riesgo de sesgo de supervivencia: sobrerrepresentan casos de evolución larga y subrepresentan los de curso letal o rápido.
		- Útiles para enfermedades crónicas de larga duración.	- Dificultan establecer la temporalidad entre exposición y enfermedad.
			- Pueden infraestimar o distorsionar la fuerza de la asociación.

Fuente: Elaboración propia

4.3 DEFINICIÓN DE CONTROL

En paralelo, es fundamental evaluar la selección de los controles, probablemente el punto más crítico de este diseño. Los controles deben representar de forma fiel la población de

donde surgieron los casos, es decir, aquellos que podrían haber desarrollado la enfermedad si hubiesen estado expuestos. Cuando los investigadores seleccionan controles hospitalarios con otras enfermedades, el riesgo de sesgo de selección aumenta, especialmente por el conocido sesgo de Berkson: los controles hospitalizados suelen tener exposiciones distintas a la población general, lo cual distorsiona las asociaciones. Volviendo a nuestro caso clínico, sería mucho más válido que los controles se eligieran al azar en la comunidad de origen de los pacientes, y no entre enfermos hospitalizados, ya que, por ejemplo, si se elige enfermos renales, también pueden estar relacionados con exposición a nitratos.

Tabla 2. Comparación entre controles hospitalarios y controles poblacionales en estudios de casos y controles: definiciones, ventajas y principales sesgos.

Tipo de control	Definición	Ventajas	Desventajas / Sesgos principales
Controles hospitalarios	Pacientes seleccionados de hospitales o centros de salud, que no presentan la enfermedad de interés.	- Más fáciles y económicos de reclutar.	- Sesgo de Berkson: los pacientes hospitalarios no representan a la población general y pueden tener exposiciones asociadas a su motivo de hospitalización.
		- Frecuentemente más dispuestos a participar (ya están en contacto con el sistema sanitario).	- Pueden compartir factores de riesgo con los casos (ej. consumo de agua contaminada si provienen de la misma comunidad).
		- Mejor calidad de información clínica (historias médicas disponibles).	- Limitan la validez externa.
Controles poblacionales	Individuos seleccionados de la misma comunidad o área geográfica que generó los casos, sin la enfermedad de interés.	- Representan de manera más fiel la población en riesgo.	- Más costosos y complejos de reclutar (encuestas, registros comunitarios).
		- Mayor validez externa.	- Menor disponibilidad de información clínica.
		- Menor riesgo de sesgos relacionados con enfermedades distintas a la de estudio.	- Pueden tener menor tasa de participación, generando sesgo si quienes aceptan difieren de quienes rechazan.

Fuente: Elaboración propia

4.4 TIPO DE MUESTREO EMPLEADO

Otro elemento central es el tipo de muestreo empleado. No todos los estudios de casos y controles se diseñan de la misma manera, y esto determina qué parámetro puede estimarse de forma válida. Existen diferentes tipos de muestreo

- En los diseños de tipo acumulativo, los controles se seleccionan al final del periodo de estudio y el odds ratio obtenido puede aproximar el riesgo relativo únicamente cuando la enfermedad es rara.
- En los diseños de muestreo por densidad de incidencia, los controles se eligen en paralelo a la aparición de cada caso y, en ese contexto, el odds ratio estima de manera directa una razón de tasas.
- En el diseño case-cohort permite aproximarse al riesgo relativo cuando se conoce bien la base de población.

Un lector crítico debe identificar cuál de estos métodos fue usado y si los autores lo explicaron con claridad. En el ejemplo del cáncer gástrico, un artículo sólido debería indicar que se utilizó muestreo por densidad, seleccionando controles en el mismo periodo en que aparecían los casos, lo que asegura comparabilidad en el tiempo.

4.5 MEDICIÓN DE EXPOSICIÓN

La medición de la exposición es otro punto neurálgico. El lector debe preguntarse: ¿cómo se midió la exposición?, ¿se emplearon cuestionarios, entrevistas, registros médicos o pruebas objetivas? El sesgo de recuerdo es un problema frecuente en este tipo de estudios, ya que los pacientes con la enfermedad tienden a recordar o reportar la exposición con mayor detalle que los controles. Además, cuando los investigadores saben qué participante es caso o control al

momento de medir la exposición, aparece el sesgo del observador. Por esta razón, se valoran más los estudios que utilizan métodos objetivos y ciegos.

En el caso clínico propuesto, la diferencia entre un artículo débil y uno sólido es clara: no es lo mismo preguntar al paciente si cree que el agua estaba contaminada, que medir de manera estandarizada la concentración de nitratos en muestras de agua del domicilio de cada participante.

4.6 FACTORES DE CONFUSIÓN

Otro componente clave en la lectura crítica es la gestión de los factores de confusión que distorsiona la asociación entre la exposición y el desenlace (como edad o sexo). Un estudio válido debe identificar de manera explícita los factores de confusión. El lector debe valorar si los autores seleccionaron confusores de forma justificada, por ejemplo, utilizando diagramas causales (DAGs), y no únicamente en función de pruebas estadísticas superficiales. Además, conviene analizar si se aplicó pareamiento y, de ser así, si este se manejó de manera correcta en el análisis. El pareamiento por edad, sexo o lugar de residencia puede ser útil para equilibrar grupos, pero cuando se realiza en exceso —el llamado *overmatching*— puede borrar la asociación real.

En el caso del cáncer gástrico, un estudio bien diseñado podría parear casos y controles por edad y sexo, ajustando luego en el análisis por tabaquismo, alcoholismo y dieta salada, factores reconocidos como potenciales confusores (porque igual podrían producir cáncer gástrico).

4.7 TAMAÑO MUESTRAL

La sección de metodología también debe explicar cómo se determinó el tamaño muestral. Un lector crítico debe comprobar si se calculó previamente el número de casos y controles necesarios para obtener resultados con suficiente poder estadístico. La ausencia de un cálculo muestral puede indicar un riesgo de resultados poco confiables. En el ejemplo de La Paz, sería tranquilizador leer que el estudio planificó incluir 150 casos y 300 controles (relación 1:2) para detectar un odds ratio de al menos 2 con un poder del 80% y un nivel de confianza del 95%.

4.8 PLAN DE ANÁLISIS

Finalmente, el plan de análisis debe estar claramente detallado. El lector debe confirmar si los autores explicaron de antemano cómo calcularon los odds ratios, cómo ajustarían por confusión, si explorarían interacciones o modificación de efecto y si consideraron análisis de sensibilidad, como el cálculo de un E-value para estimar la posible influencia de confusión no medida. La transparencia en el plan analítico ayuda a prevenir la manipulación de resultados y otorga credibilidad al artículo.

En el caso clínico, un estudio sólido informaría que se calculará primero un odds ratio crudo, seguido de un modelo de regresión logística ajustado por tabaquismo, alcoholismo y dieta, además de un análisis estratificado para explorar si el tabaco modifica el efecto de la exposición a nitratos.

La lectura crítica de la metodología en un estudio de casos y controles exige un recorrido atento y sistemático. Desde la definición de la población y de los casos, hasta la elección de los controles, la medición de la exposición y el manejo de confusores, cada detalle puede marcar la diferencia entre un estudio confiable y uno sesgado. El lector que aprende a formular estas preguntas logra no solo interpretar con precisión los resultados, sino también valorar si estos pueden o no aplicarse en la práctica clínica.

5. EVALUACIÓN DE LA VALIDEZ Y RELEVANCIA

Una vez entendida la metodología, el siguiente paso en la lectura crítica de un estudio de casos y controles es evaluar si los resultados son válidos y si tienen relevancia clínica y epidemiológica. La validez se divide en dos grandes dimensiones: validez interna, que responde a la pregunta “¿los resultados del estudio son confiables dentro de la población analizada?”, y validez externa, que pregunta “¿estos resultados pueden aplicarse a otros contextos o poblaciones?”.

5.1 VALIDEZ INTERNA

Para juzgar la validez interna, el lector debe analizar si el estudio evitó los principales sesgos y si controló de manera adecuada la confusión. En un diseño de casos y controles, los sesgos más frecuentes se verán en el acápite 7. La validez interna se refuerza también con decisiones metodológicas como el uso de casos incidentes en lugar de prevalentes, la aplicación de análisis estratificados para verificar la consistencia de los resultados y la realización de análisis de sensibilidad (por ejemplo, calculando un E-value para explorar el posible impacto de confusores no medidos).

5.2 VALIDEZ EXTERNA

Una vez establecida la confianza en los resultados dentro del estudio, corresponde preguntarse si esos hallazgos pueden extrapolarse a otras poblaciones. La validez externa depende de cuán representativa sea la población fuente y de cuán similares sean los determinantes de salud en otros contextos. En el ejemplo, un estudio realizado en comunidades periurbanas de La Paz puede ser muy válido para entender la relación entre nitratos y cáncer gástrico en esa región, pero el lector debe cuestionar si los hallazgos pueden aplicarse a zonas rurales donde el agua proviene de manantiales de montaña, o a países donde la dieta y las fuentes de exposición a nitratos son diferentes. No es lo mismo interpretar un estudio en una comunidad de altura, con costumbres alimentarias específicas, que en una ciudad costera de Asia donde el consumo de pescados salados y nitrosaminas es mucho mayor.

5.3 RELEVANCIA CLÍNICA Y EPIDEMIOLÓGICA

Más allá de la validez, un lector crítico debe evaluar si la asociación encontrada tiene significado clínico y de salud pública. Un odds ratio elevado pero basado en un estudio pequeño, con intervalos de confianza muy amplios, puede carecer de utilidad práctica. Por el contrario, una asociación moderada pero consistente en varios estudios, ajustada por confusores y replicada en diferentes poblaciones, puede ser altamente relevante para la prevención. En el caso del cáncer gástrico, demostrar una asociación robusta entre exposición a nitratos y riesgo de enfermedad podría tener repercusiones inmediatas en políticas de salud, tales como la regulación de calidad del agua potable, campañas de educación sanitaria y vigilancia ambiental. Por ello, al leer un artículo de casos y controles, el lector debe pensar siempre más allá del valor estadístico: debe preguntarse si los hallazgos pueden contribuir a mejorar la práctica clínica o las decisiones de salud pública.

6. INTERPRETACIÓN ESTADÍSTICA

La interpretación estadística en casos y controles comienza aceptando un límite fundamental del diseño: no estima directamente riesgos absolutos, tasas de incidencia ni diferencias de riesgo. Su lenguaje nativo es la razón de momios u odds ratio (OR). Por eso, leer bien un artículo significa comprobar que lo que se informa y lo que se concluye sí es coherente con lo que el diseño puede ofrecer.

6.1 ANTES DE MIRAR NÚMEROS ¿QUÉ SÍ Y QUE NO PUEDE INTERPRETARSE?

Un estudio de casos y controles compara cuán frecuente fue la exposición entre quienes desarrollaron la enfermedad (casos) y quienes no (controles). Si la enfermedad es rara en la población, el OR puede aproximar el riesgo relativo (RR). Si no es rara, el OR tiende a sobreestimar la magnitud del efecto y no debe traducirse como “veces más riesgo” sin aclaraciones. La lectura crítica exige verificar que el artículo no presente RR ni riesgo absoluto como si se hubieran medido directamente.

6.2 DE “ODDS” A “ODDS RATIO”: EL MÍNIMO QUE SE DEBE DOMINAR

El odds (momio) de exposición en un grupo es la razón entre “expuestos” y “no expuestos”

dentro de ese grupo. El odds ratio (OR) compara esos momios entre casos y controles. Se interpreta así:

- “En los casos, la probabilidad relativa de haber estado expuesto, en comparación con los controles, fue X veces mayor (o menor).”

Ejemplo: Siguiendo con el caso, se obtuvieron los siguientes resultados:

Tabla 3. Distribución de casos y controles según exposición a nitratos en el caso clínico (tabla 2×2).

	Expuestos a nitratos	No expuestos	Total
Casos (cáncer gástrico)	40	10	50
Controles	20	30	50

Fuente: Elaboración propia con apoyo de Inteligencia Artificial

ODDS RATIO ELEVADO

Cuando el OR es mayor a 1, indica que la exposición fue más frecuente entre los casos que entre los controles.

- Un OR moderado, por ejemplo 1.5 o 2.0, sugiere que los casos tuvieron entre 1.5 y 2 veces más probabilidad de estar expuestos que los controles.
- Un OR alto, como 5.0 o más, indica una asociación fuerte. Sin embargo, el lector debe recordar que un OR grande no siempre implica un efecto clínicamente importante si los intervalos de confianza son muy anchos o si hay sesgos en la selección o medición.

Tabla 4. Odds Ratio de exposición a nitratos.

Exposición	OR (IC95%)	Valor p
Agua con nitratos	6.0 (2.5–14.7)	<0.001

Fuente: Elaboración propia con apoyo de Inteligencia Artificial

Ejemplo del caso clínico: un OR de 6.0 para exposición a nitratos significa que los pacientes con cáncer gástrico tenían seis veces más probabilidad de haber estado expuestos a agua con alto nitrato respecto de los controles. Esto es una asociación fuerte y, si el intervalo de confianza es estrecho, sugiere una señal robusta. Lo que no debe decirse: “Seis veces más riesgo de cáncer” (salvo que se justifique enfermedad rara y diseño compatible).

ODDS RATIO BAJO

Cuando el OR es menor a 1, se interpreta como que la exposición es menos frecuente en los casos que en los controles. En términos clínicos, puede sugerir un efecto protector.

- Un OR de 0.7 significa que la exposición se observó 30% menos entre los casos que entre los controles.
- Un OR de 0.3 indicaría una asociación protectora más marcada.

Tabla 5. Odds Ratio de consumo de frutas.

Exposición	OR (IC95%)	Valor p
Consumo de frutas	0.4 (0.2–0.9)	0.03

Fuente: Elaboración propia con apoyo de Inteligencia Artificial

Ejemplo aplicado: si el estudio encontrara que el consumo habitual de frutas frescas tiene un OR de 0.4 frente al cáncer gástrico, esto sugeriría un posible efecto protector de la dieta, siempre que el intervalo de confianza no cruce 1.0.

¿QUÉ PASA CUANDO EL OR ESTÁ EN 1?

El valor clave de referencia en la interpretación es 1.0, que representa “no asociación” (los odds de exposición son iguales en casos y controles).

- Si el IC95% incluye el 1.0, significa que no puede descartarse el azar como explicación de los

resultados. En ese caso, la asociación no es estadísticamente significativa, aunque el OR sea mayor o menor que 1.

- Si el IC es muy amplio, además, la estimación es poco precisa: no se puede afirmar con confianza si existe o no una relación.

Tabla 6. Odds Ratio de exposición a tabaco.

Exposición	OR (IC95%)	Valor p
Tabaquismo	1.3 (0.8–2.2)	0.03

Fuente: Elaboración propia con apoyo de Inteligencia Artificial

Ejemplo aplicado: supongamos que, al analizar el hábito tabáquico, el estudio reporta un OR de 1.3 con un IC95% de 0.8 a 2.2. El valor puntual (1.3) sugiere una ligera mayor probabilidad de exposición en casos, pero como el intervalo incluye 1.0, la asociación no es estadísticamente significativa. En lectura crítica, esto se traduce en “no hay evidencia suficiente para afirmar una relación entre el tabaco y cáncer gástrico”.

Reglas de oro:

- OR > 1 (IC95% excluye 1.0): la exposición podría estar asociada a mayor probabilidad de enfermedad (posible factor de riesgo).
- OR < 1 (IC95% excluye 1.0): la exposición podría estar asociada a menor probabilidad de enfermedad (posible factor protector).
- OR ≈ 1 o IC95% incluye 1.0: no se demuestra asociación; puede deberse a azar, sesgo, confusión o falta de poder estadístico.

6.3 INTERVALOS DE CONFIANZA Y VALOR P: SEÑAL, PRECISIÓN Y AZAR

El intervalo de confianza (IC) es, probablemente, el recurso más valioso para interpretar los resultados de un estudio de casos y controles. Mientras que el OR nos da una estimación puntual de la asociación, el IC nos dice qué tan precisa es esa estimación y en qué rango de valores se podría mover la “verdadera asociación” en la población.

El valor p, por su parte, indica la probabilidad de obtener los resultados observados (o más extremos) si en realidad no hubiera asociación. Sin embargo, el valor p no mide magnitud ni relevancia clínica; simplemente ayuda a decidir si el resultado podría explicarse por azar. Por eso, la lectura crítica debe priorizar el IC por encima del p, aunque ambos se complementen.

Ejemplo con caso clínico: En el estudio de cáncer gástrico y nitratos, se presenta la siguiente tabla

Tabla 7: Odds Ratio en diferentes escenarios de exposición a nitratos.

Exposición	OR (IC95%)	Valor p
Agua con nitratos	6.0 (2.5–14.7)	<0.001
Tabaquismo	1.3 (0.8–2.2)	0.01
Consumo de frutas	0.4 (0.2–0.9)	0.03

Fuente: Elaboración propia

El OR crudo de agua con nitratos fue de 6.0.

- Si el IC95% es 2.5 a 14.7, esto significa que, con un 95% de confianza, la verdadera asociación se encuentra en algún punto entre 2.5 y 14.7.
- Como el IC no incluye 1.0, el resultado es estadísticamente significativo.
- El intervalo es amplio, lo que refleja incertidumbre en la magnitud exacta: ¿el efecto real es 2.5 veces o 14 veces mayor? Ambos valores son posibles.

El valor p (<0.001) refuerza que el resultado es muy poco compatible con el azar, pero lo más importante sigue siendo el rango del IC.

En todos los escenarios se interpretaría de esta forma:

- OR = 6.0, IC95% 2.5–14.7, asociación fuerte y significativa. El efecto existe, pero su magnitud exacta es incierta.
- OR = 1.3, IC95% 0.8–2.2, el OR sugiere ligera asociación positiva, pero como el IC incluye 1.0, no es significativa. Interpretación crítica: “No se encontró evidencia suficiente de asociación”.
- OR = 0.4, IC95% 0.2–0.9, asociación negativa significativa, posible efecto protector. Interpretación crítica: el consumo frecuente de frutas es un posible factor protector contra el cáncer

gástrico.

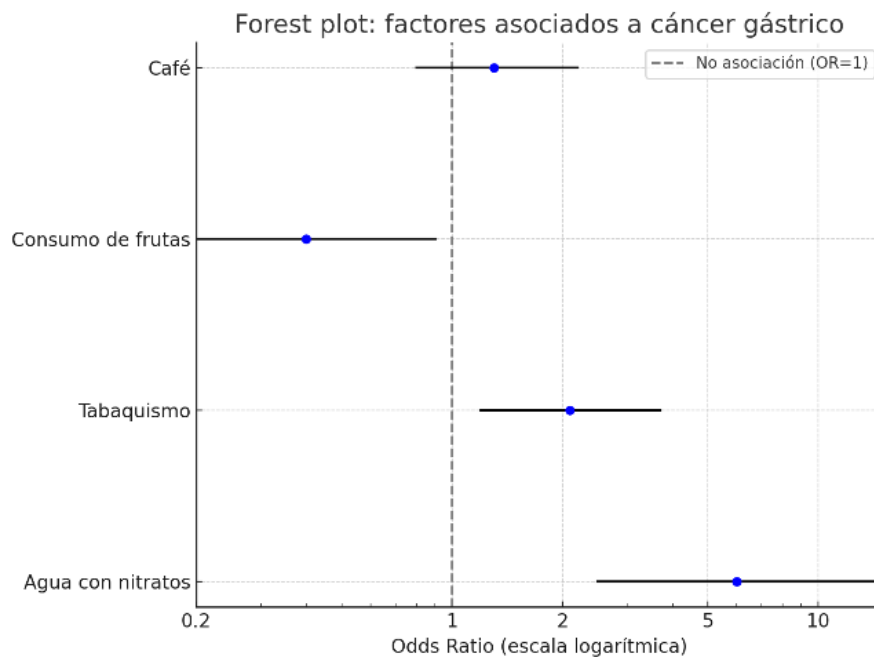
- Todos cuentan con el valor p (<0.05) que refuerza que los resultados son poco compatibles con el azar.

6.4 INTERPRETACIÓN DE FOREST PLOT (GRÁFICO DE BOSQUE):

Representa visualmente cada OR como un punto (●) y su IC95% como una línea horizontal.

- La línea vertical en 1.0 representa “no asociación”.
- Si el intervalo cruza esa línea, el resultado no es significativo.
- Cuanto más corta la línea, más precisa la estimación.

Figura 1 Forest plot de los factores asociados al riesgo de cáncer gástrico en el caso clínico.



Fuente: Elaboración propia

Para su interpretación de un Forest plot basado en nuestro caso clínico.

- Eje X: representan el OR en escala logarítmica
 - Eje Y: las variables de intervención a estudiar como posible factor de riesgo (agua con nitratos, tabaquismo, consumo de frutas, café)
 - La línea vertical punteada en OR = 1.0 (sin asociación).
 - Cada punto azul representa el OR puntual.
 - Las barras negras muestran los rangos del IC95%.
 - A la derecha de línea punteada 1.0 = asociaciones positivas (factores de riesgo).
 - A la izquierda de línea punteada 1.0 = asociaciones negativas (posibles factores protectores).
- Ejemplo del Forest plot: Su interpretación sería de esta forma:
- Agua con nitratos: ● muy a la derecha, con línea 2.5–14.7 → asociación fuerte y significativa.
 - Tabaquismo: ● también a la derecha, línea corta 1.2–3.7 → asociación moderada y precisa.
 - Consumo de frutas: ● a la izquierda, línea 0.2–0.9 → efecto protector significativo.
 - Café: ● cerca del 1, línea 0.8–2.2 → no significativo.

Reglas prácticas para la interpretación:

- El OR puntual da una idea del efecto, pero nunca debe leerse aislado.
- El IC95% indica precisión: Intervalo estrecho = estimación confiable, intervalo amplio = mucha incertidumbre.
- El 1.0 dentro del IC = resultado no significativo.
- El valor p solo complementa: nunca sustituye al IC.

6.5 CAMBIOS DEL OR CUANDO SE AJUSTA

Cuando un artículo reporta un OR ajustado, significa que la asociación entre una exposición y una enfermedad fue estimada controlando estadísticamente otras variables que podrían distorsionar el resultado, en otras palabras, el OR ajustado intenta responder a la pregunta:

“¿Cuál es la asociación entre la exposición y la enfermedad si todos los grupos comparados tuvieran características similares en edad, sexo, tabaquismo u otros factores relevantes?”

“Ajustar” no significa modificar artificialmente los datos. Significa que el análisis estadístico —habitualmente mediante regresión logística en estudios de casos y controles— estima el efecto de cada exposición manteniendo constantes las demás variables incluidas en el modelo.

Por ejemplo, si el consumo de agua con nitratos está asociado con cáncer gástrico, pero las personas expuestas también fuman más, parte del efecto observado en el OR crudo podría deberse al tabaquismo, por ello el OR ajustado permite estimar la asociación independiente de los nitratos, descontando el efecto del tabaco y otras variables.

En estudios bien hechos, los autores describen los confusores preespecificados (edad, sexo, tabaquismo, alcoholismo, dieta salada, *H. pylori*, nivel socioeconómico) y muestran el OR ajustado.

Tabla 8. Odds Ratios crudos y ajustados de factores de riesgo para cáncer gástrico en el caso clínico.

Exposición	OR crudo (IC95%)	OR ajustado (IC95%)
Agua con nitratos	6.0 (2.5–14.7)	3.2 (1.8–5.6)
Tabaquismo	2.5 (1.3–4.8)	1.9 (1.1–3.4)
Consumo de alcohol	1.8 (1.0–3.2)	1.4 (0.8–2.6)

Fuente: Elaboración propia

Interpretación:

- Agua con nitratos: el OR crudo de 6.0 sugería una asociación muy fuerte. Tras ajustar por tabaquismo y alcohol, baja a 3.2. Esto indica que parte del efecto se explicaba por confusión, pero la asociación persiste, y es significativa.
- Tabaquismo: también disminuye levemente tras el ajuste, pero sigue siendo significativo.
- Consumo de alcohol: el OR crudo parecía mostrar riesgo, pero tras el ajuste ya no es significativo (IC95% incluye 1.0). Esto muestra cómo el ajuste puede cambiar la conclusión.

6.6 ESTRATIFICACIÓN Y MANTEL-HAENSZEL (IDEA INTERPRETATIVA)

En algunos estudios de casos y controles, los autores no se limitan a presentar un Odds Ratio global. En lugar de ello, dividen el análisis en subgrupos o estratos, por ejemplo fumadores y no fumadores, hombres y mujeres, o distintos grupos etarios. Este procedimiento se denomina estratificación y tiene un objetivo claro: explorar si una tercera variable está confundiendo la asociación principal o si, por el contrario, está modificando el efecto de la exposición.

Para comprenderlo mejor, retomemos el caso clínico del cáncer gástrico y la exposición a agua con nitratos. Supongamos que el OR crudo fue de 6.0. El lector podría concluir inicialmente que la exposición aumenta seis veces la probabilidad de enfermedad. Sin embargo, si el tabaquismo está asociado tanto con el consumo de agua contaminada como con el cáncer gástrico, podría estar inflando esa estimación (tabla 8).

Si el análisis se realiza por separado en fumadores y no fumadores, pueden aparecer dos escenarios distintos.

- En el primer escenario, los OR en fumadores y en no fumadores resultan muy similares entre sí, por ejemplo 3.1 y 3.3. En este caso, el efecto es consistente en ambos grupos, pero claramente menor que el OR crudo de 6.0. Esto indica que el tabaquismo estaba distorsionando la asociación inicial. Al separar los datos por estratos, el efecto real se aproxima a 3.2. Aquí se habla de confusión. El método de Mantel-Haenszel permite entonces combinar esos OR estratificados similares en una única estimación ajustada, obteniendo un OR “combinado” que refleja la asociación controlando esa variable.
- En el segundo escenario, los resultados por estrato son muy diferentes entre sí, por ejemplo un OR de 8.0 en fumadores y de 1.5 en no fumadores. Aquí el efecto no es uniforme: la exposición parece mucho más dañina en fumadores que en no fumadores. Esto ya no es confusión, sino modificación de efecto o interacción. En esta situación no tiene sentido combinar los OR, porque no existe un único efecto verdadero; el impacto de la exposición

depende del subgrupo. Desde el punto de vista clínico, esto es incluso más relevante, pues sugiere que ciertos pacientes son particularmente vulnerables.

El método de Mantel–Haenszel se utiliza precisamente cuando los OR por estrato son relativamente homogéneos. Antes de combinar los resultados, el investigador debe verificar que las estimaciones no difieran de manera importante entre sí. Si son similares, puede calcularse un OR ajustado global. Si son muy distintos, se debe reportar el efecto por separado para cada estrato.

Para el lector crítico, la interpretación práctica es la siguiente: si los OR estratificados son parecidos entre sí pero distintos del OR crudo, probablemente existe confusión. Si los OR estratificados difieren marcadamente entre sí, probablemente existe modificación de efecto. Reconocer esta diferencia es esencial para no interpretar de forma errónea los resultados y para comprender si una variable está distorsionando la asociación o realmente está cambiando su magnitud en determinados subgrupos.

6.7 REGRESIÓN LOGÍSTICA

En casos y controles, la regresión logística sirve para ajustar confusores y explorar interacciones. El lector debe fijarse en:

- Qué variables se ajustaron y por qué (mejor si lo justifican con diagramas causales; peor si “entró lo que dio $p < 0,05$ ”).
- OR ajustados con IC95% (no basta el p).
- Aclaraciones sobre no colapsabilidad: el OR puede cambiar al ajustar incluso sin confusión; es propiedad matemática del OR y no implica necesariamente “sesgo”.
- Medidas de desempeño: evitar interpretar pseudoR^2 como “% de varianza explicada” (no lo es). Es preferible que el artículo muestre calibración (qué tan bien encajan probabilidades predichas con observadas) y, si construye un modelo predictivo, que informe discriminación (AUC/ROC). En estudios puramente etiológicos, AUC es secundaria; lo crucial es el OR ajustado bien interpretado.

Para entenderlo se utilizará la siguiente tabla:

Tabla 9. Resultados de la regresión logística multivariada: factores de riesgo y protección asociados a cáncer gástrico.

Variable	OR crudo (IC95%)	Valor p	OR ajustado (IC95%)	Valor p
Agua con nitratos	6.0 (2.5–14.7)	<0.001	3.2 (1.8–5.6)	<0.001
Tabaquismo	2.5 (1.3–4.8)	0.01	1.9 (1.1–3.4)	0.02
Consumo de alcohol	1.8 (1.0–3.2)	0.05	1.4 (0.8–2.6)	0.18
Consumo de frutas	0.5 (0.3–0.9)	0.03	0.6 (0.3–1.1)	0.08

Fuente: Elaboración propia

INTERPRETACIÓN PASO A PASO:

- Agua con nitratos: el OR crudo de 6.0 sugería una asociación fuerte; tras ajustar por tabaquismo, alcohol y frutas, disminuye a 3.2, pero sigue siendo significativa. Esto significa que parte de la asociación se debía a confusión, pero la exposición sigue mostrando un efecto robusto.
- Tabaquismo: mantiene significancia tras el ajuste, aunque con menor magnitud.
- Consumo de alcohol: pierde significancia al ajustar, lo que indica que su aparente efecto estaba explicado por otros factores.
- Consumo de frutas: parecía protector en el análisis crudo, pero tras el ajuste la asociación ya no alcanza significancia estadística.

Consejo para el lector crítico: Al leer una tabla de regresión logística, lo más importante es comparar OR crudo vs OR ajustado. Si el OR cambia de manera relevante después del ajuste, es señal de confusión. Si se mantiene estable, la asociación es más consistente. Y si pierde significancia, probablemente la relación observada era espuria.

Tabla 10. Estratificación (Mantel–Haenszel) vs regresión logística en estudios de casos y controles

Aspecto	Estratificación / Mantel–Haenszel	Regresión logística
¿Qué hace?	Divide el análisis en subgrupos (estratos) según una variable potencialmente confusora (ej. fumadores vs no fumadores).	Analiza simultáneamente varias variables en un modelo matemático para estimar el efecto independiente de cada una.
¿Qué problema intenta resolver?	Identificar confusión o modificación de efecto.	Controlar múltiples confusores y estimar OR ajustados.
¿Cuándo se usa?	Cuando hay una o pocas variables importantes que pueden confundir la asociación.	Cuando existen varias variables que deben ajustarse al mismo tiempo.
¿Cómo detecta confusión?	Si los OR por estrato son similares entre sí pero diferentes del OR crudo.	Si el OR ajustado difiere sustancialmente del OR crudo.
¿Cómo detecta modificación de efecto?	Si los OR por estrato son muy distintos entre sí.	Si el término de interacción es significativo en el modelo.
Resultado final	OR combinado de Mantel–Haenszel (si hay homogeneidad entre estratos).	OR ajustado multivariado.
Ventaja principal	Método simple y transparente; fácil de entender visualmente.	Permite ajustar múltiples variables simultáneamente y evaluar interacciones complejas.
Limitación	Se vuelve impráctico con muchas variables.	Requiere mayor conocimiento estadístico y supuestos del modelo.

Fuente: Elaboración propia

6.8 MODIFICACIÓN DEL EFECTO (INTERACCIÓN): CUANDO LA ASOCIACIÓN “DEPENDIE DE...”

No todo cambio del OR entre subgrupos es confusión; a veces el efecto realmente varía. Si el OR de nitratos es 2.0 en no fumadores y 6.0 en fumadores, el tabaquismo podría potenciar el efecto de los nitratos. El artículo debe declararlo como interacción y, idealmente, mostrar un término de interacción en el modelo (interpretación clínica: el riesgo atribuible a la exposición no es igual para todos).

Lectura práctica:

- Si los OR por estratos son parecidos → combinar (confusión).
- Si son claramente distintos → considerar modificación de efecto y reportar por subgrupos.

6.9 ESTUDIOS PAREADOS: COMO LEER EL OR “CONDICIONAL”

Cuando los autoresorean casos y controles (p. ej., por edad y sexo), el análisis correcto no usa el 2×2 simple, sino el OR de pares discordantes o una regresión logística condicional.

Ejemplo: imaginemos que, en el estudio de cáncer gástrico y nitratos, cada caso fue apareado con un control de la misma edad y sexo. Esto genera pares de individuos, y la comparación se hace “dentro de cada par”.

En estos diseños, el OR no se calcula con la tabla 2×2 clásica, sino con los pares discordantes:

- Pares donde el caso estaba expuesto y el control no (b).
- Pares donde el caso no estaba expuesto y el control sí (c).

Tabla 11. Distribución de pares concordantes y discordantes en un estudio de casos y controles apareados.

	Control expuesto	Control no expuesto
Caso expuesto	12 (ambos expuestos, no cuentan)	18 (discordante)
Caso no expuesto	9 (discordante)	21 (ambos no expuestos, no cuentan)

Fuente: Elaboración propia

Interpretación: En 60 pares, hubo 18 pares “Caso expuesto / Control no expuesto” y 9 pares “Caso no expuesto / Control expuesto”. El OR condicional es $18/9 = 2.0$.

- El resultado indica que, dentro de los pares apareados por edad y sexo, los casos tenían el doble de probabilidad de estar expuestos a agua con nitratos que sus controles.
- Como el análisis fue condicional, la asociación se estimó independientemente de la edad y el

sexo, porque esas variables ya estaban controladas por diseño.

Cuando el artículo dice que aplicó “regresión logística condicional”, significa que analizó los datos respetando el apareamiento. El lector debe fijarse en que se reporten OR condicionales (y no simplemente OR crudos) cuando hubo apareamiento, de lo contrario los resultados podrían estar mal estimados.

6.10 OTRAS PRUEBAS QUE SUELEN APARECER

CHI CUADRADO O FISHER (TABLAS 2×2 NO PAREADAS)

Cuando se presenta una tabla simple de exposición/no exposición entre casos y controles, los autores suelen usar la prueba de Chi cuadrado para evaluar si las diferencias observadas podrían deberse al azar. Si el número de observaciones es pequeño (por ejemplo, cuando alguna celda tiene menos de 5 casos), se usa la prueba exacta de Fisher, que es más confiable en muestras pequeñas.

Ejemplo: En el estudio, si de 50 casos, 40 estaban expuestos y en los 50 controles solo 20 lo estaban, un Chi cuadrado o Fisher mostrará si esa diferencia es estadísticamente significativa. El OR nos dice la magnitud de la asociación (ej. OR=6.0), mientras que Chi cuadrado/Fisher nos dicen si esa diferencia podría explicarse solo por azar (<0.001).

MCNEMAR (PARA ESTUDIOS APAREADOS)

Cuando los casos y controles fueron apareados (pareados), por ejemplo por edad y sexo, el análisis se centra en los pares discordantes: aquellos en los que el caso estuvo expuesto y el control no, o viceversa.

La prueba de McNemar evalúa si la diferencia en estos pares discordantes es significativa. Va de la mano con el cálculo del OR condicional.

Ejemplo: Si en 60 pares, 18 pares muestran exposición solo en el caso y 9 pares exposición solo en el control, el OR condicional es 2.0. La prueba de McNemar dirá si esa diferencia (18 vs 9) es suficientemente grande para no atribuirla al azar.

PRUEBAS DE INTERACCIÓN (MODELOS MULTIVARIABLES)

En los modelos de regresión logística, los investigadores pueden añadir un término de interacción entre dos variables. Esto sirve para saber si el efecto de una exposición cambia dependiendo de otra variable.

Ejemplo: En nuestro estudio, el efecto de los nitratos podría ser diferente en fumadores y no fumadores. Si el modelo muestra un término de interacción significativo, eso significa que el tabaquismo modifica el efecto de los nitratos.

- En no fumadores: OR=2.5.
- En fumadores: OR=6.5.

Esto no es confusión, sino modificación de efecto real. El lector debe interpretar que el riesgo asociado a nitratos es mucho más alto en fumadores, y por lo tanto, los resultados deben reportarse por separado en cada subgrupo.

Tabla 12. Principales pruebas estadísticas utilizadas en estudios de casos y controles y su interpretación.

Prueba	¿Qué evalúa?	¿Cuándo se usa?	¿Cómo interpretarlo?
Chi cuadrado	Si la distribución de exposición difiere entre casos y controles.	Tablas 2×2 con muestras grandes.	Si el valor $p < 0.05$, la diferencia en exposición no se debe solo al azar. Complementa al OR/IC95%.
Exacta de Fisher	Igual que Chi cuadrado, pero más precisa con números pequeños.	Tablas 2×2 con frecuencias bajas (<5 en alguna celda).	Si $p < 0.05$, la asociación es significativa aunque la muestra sea pequeña.
Mantel-Haenszel	Calcula un OR “combinado” ajustado por una variable de estratificación.	Cuando se sospecha confusión y se dividen los datos en estratos (ej. fumadores vs no fumadores).	Si los OR por estrato son similares, se reporta el OR combinado. Si son muy distintos, hay modificación de efecto.

Prueba	¿Qué evalúa?	¿Cuándo se usa?	¿Cómo interpretarlo?
McNemar	Si existe diferencia entre pares discordantes en estudios apareados.	Estudios de casos y controles apareados (1:1 o más).	Un $p < 0.05$ indica que la exposición fue más frecuente en casos que en controles dentro de los pares.
Pruebas de interacción (en regresión)	Si el efecto de la exposición cambia según otra variable.	Modelos multivariantes (ej. regresión logística).	Un término de interacción significativo indica modificación de efecto real, no confusión. Ejemplo: nitratos con efecto más fuerte en fumadores que en no fumadores.

Fuente: Elaboración propia

Consejo para el lector crítico: Cuando un artículo presenta estas pruebas, lo importante no es memorizar fórmulas, sino reconocer qué pregunta responde cada una. Así, la lectura crítica se vuelve más ordenada: Chi cuadrado/Fisher dicen si hay diferencia, Mantel-Haenszel muestra un OR ajustado por estratos, McNemar evalúa discordancia en pares apareados, y las pruebas de interacción responden si el efecto depende de otro factor.

6.11 ERRORES DE INTERPRETACIÓN FRECUENTES (Y CÓMO EVITARLOS)

Tabla 13. Errores frecuentes en la interpretación de estudios de casos y controles y cómo evitarlos

Error frecuente	Consecuencia	Cómo evitarlo
Traducir el OR como Riesgo Relativo (RR) sin que la enfermedad sea rara o sin justificación	Exagera la fuerza de la asociación	Recordar: $OR \approx RR$ solo cuando la enfermedad es poco frecuente
Confiar únicamente en el valor p	Puede llevar a conclusiones engañosas; ignora la magnitud y precisión del efecto	Revisar siempre el IC95% y su amplitud junto al p
Ignorar el diseño del estudio (ej. apareamiento)	Análisis inadecuado; resultados sesgados	Verificar si hubo apareamiento y buscar OR condicional
Ajustar por todas las variables o por las "significativas"	Riesgo de ajustar por colisionadores y distorsionar la asociación	Seleccionar variables de ajuste a priori con criterio causal
Olvidar la precisión de la muestra	OR extremos con IC muy amplios → resultados inciertos	Valorar el tamaño de muestra y la amplitud del IC95%

Fuente: Elaboración propia

7. SESGOS Y LIMITACIONES

Uno de los mayores desafíos al leer un estudio de casos y controles es identificar los sesgos que pueden distorsionar los resultados. Un sesgo es un error sistemático que afecta la validez interna del estudio y puede llevar a sobreestimar, infraestimar o incluso invertir la asociación real entre exposición y enfermedad. El lector crítico debe entrenarse para reconocerlos en la sección de métodos y discusión del artículo.

7.1 SESGO DE SELECCIÓN

Aparece cuando los casos y los controles no provienen de la misma población fuente. Esto hace que las diferencias observadas no reflejen una verdadera asociación, sino la forma en que fueron elegidos los participantes.

- Ejemplo aplicado: si los casos de cáncer gástrico se reclutan en La Paz pero los controles provienen de otra ciudad con agua tratada, la comparación es inválida. Igualmente, si los controles se seleccionan solo entre pacientes hospitalizados por enfermedades renales (también asociadas a nitratos), se produce el sesgo de Berkson, que tiende a diluir la asociación real.

7.2 SESGO DE INFORMACIÓN

Ocurre cuando la medición de la exposición o del desenlace no es precisa o difiere entre casos y controles. En estudios retrospectivos, el sesgo de recuerdo es frecuente: los casos tienden

a recordar más su historia de exposición (ej. agua con mal sabor, consumo de alimentos salados) que los controles. También puede presentarse sesgo del observador, cuando el investigador sabe quién es caso y quién es control y mide la exposición con cierta expectativa.

- Ejemplo aplicado: si la exposición se evaluara solo preguntando al paciente si creía que el agua estaba contaminada, los casos tenderían a recordarlo con más detalle, exagerando la asociación. Si en cambio se midieran niveles de nitratos en laboratorio, el riesgo de este sesgo se reduciría.

7.3 SESCO DE SUPERVIVENCIA (NEYMAN)

Aparece cuando el estudio incluye casos prevalentes en lugar de incidentes. Como los casos prevalentes son aquellos que han sobrevivido cierto tiempo con la enfermedad, suelen representar formas más benignas.

- Ejemplo aplicado: si en el estudio se incluyen pacientes diagnosticados con cáncer gástrico hace 5 años y que aún viven, se excluiría a los pacientes con tumores agresivos que fallecieron rápidamente. El resultado sería una subestimación de la asociación entre nitratos y cáncer.

7.4 SESCO POR SOBREEMPAJEAMIENTO (OVERMATCHING)

El emparejamiento puede ser útil para equilibrar grupos, pero si se aplica en exceso puede eliminar diferencias reales y hacer que la asociación desaparezca.

- Ejemplo aplicado: si los investigadores aparearan casos y controles no solo por edad y sexo, sino también por características relacionadas con la exposición (ej. zona exacta de residencia y tipo de agua consumida), ambos grupos tendrían prácticamente la misma probabilidad de exposición, borrando la señal real del efecto de los nitratos.

7.5 CONFUSIÓN RESIDUAL

Incluso con buen diseño, algunos confusores pueden no medirse o no ajustarse de manera adecuada. Esto se llama confusión residual.

- Ejemplo aplicado: si el estudio no midiera adecuadamente la infección por *Helicobacter pylori*, un factor de riesgo fuerte para cáncer gástrico, la asociación entre nitratos y cáncer podría estar parcialmente distorsionada.

7.6 LIMITACIONES INHERENTES AL DISEÑO

Más allá de los sesgos específicos, el lector debe recordar las limitaciones propias de los estudios de casos y controles:

- No permiten calcular incidencia ni riesgo absoluto. Solo se puede estimar un OR.
- La temporalidad puede ser difícil de establecer. Aun cuando el OR sea alto, no siempre queda claro si la exposición precedió al inicio de la enfermedad.
- Susceptibilidad a sesgos de recuerdo y selección. Esto obliga a ser especialmente crítico al evaluar cómo se midió la exposición y cómo se eligieron los controles.
- Generalización limitada. Sus resultados suelen aplicarse con cautela a otras poblaciones, salvo que los contextos sean muy similares.

8. CHECKLIST

Los checklistson andamios que ayudan al lector a verificar, de forma ordenada, si un artículo reporta lo esencial y si lo hace con suficiente calidad. A nivel internacional, las guías más empleadas son:

- **STROBE (EQUATOR)**: pauta de reporte para estudios observacionales (incluye una lista específica para casos y controles, además de la combinada) y un documento de Explicación y Elaboración que aclara el porqué de cada ítem. Es el estándar de referencia para autores, revisores y lectores.
- **JBICriticalAppraisalChecklist (2020)**: lista de evaluación crítica enfocada en riesgo de sesgo

y validez del estudio, muy útil para lectores y para revisiones sistemáticas.

- **CASP Case-Control Checklist (edición vigente, 2024):** 11 preguntas prácticas para valorar validez, resultados y aplicabilidad. Gran herramienta docente y clínica.
 - **RECORD / RECORD-PE (extensión de STROBE):** si el estudio usa datos rutinarios (registros, bases administrativas, historias clínicas electrónicas) o es farmacoepidemiológico, añade ítems de transparencia, enlace de bases, calidad de datos y trazabilidad del código.
- A partir de estas guías, se propone el siguiente checklist práctico para lectura crítica:

Tabla 13. Checklist para la lectura crítica de estudios de casos y controles

Sección	Pregunta clave	Fuente principal
Identificación y objetivo	¿El título/resumen dejan claro que es un estudio de casos y controles y cuál es la pregunta principal?	STROBE
	¿El artículo justifica el diseño casos-controles (desenlace raro, latencia, factibilidad)?	STROBE-E&E
Métodos (validez interna)	¿Población fuente y periodo claramente definidos? ¿Casos y controles provienen de la misma población en riesgo?	STROBE, JBI
	¿Definición de casos estandarizada (idealmente incidentes)? ¿Se evitó sesgo de supervivencia (Neyman)?	STROBE-E&E, JBI
	¿Selección de controles adecuada? ¿Se evitó sesgo de Berkson?	STROBE-E&E, JBI, CASP
	¿Tipo de muestreo declarado y congruente con la medida reportada (OR vs IRR/RR)?	STROBE-E&E
	¿Medición de la exposición válida y objetiva/cegada? Si fue autorreporte, ¿se discutió sesgo de recuerdo?	STROBE, JBI, CASP
	Si hubo pareamiento: ¿se explican variables, criterios y análisis condicional? ¿Se evitó el overmatching?	STROBE-E&E, JBI
	¿Tamaño muestral y supuestos justificados (prevalencia, OR detectable, poder)?	STROBE
	¿Plan analítico preespecificado (OR crudo, ajuste por confusión, modificación de efecto, análisis de sensibilidad)?	STROBE-E&E, CASP
Resultados	¿Tablas 2x2 y OR con IC95% y p-valores presentados claramente (crudos y ajustados)?	STROBE
	Si hubo pareamiento: ¿se reportó OR condicional/logística condicional? Si estratificación: ¿OR de Mantel-Haenszel y homogeneidad?	STROBE-E&E
	Si se usaron datos rutinarios (HCE, registros), ¿se informó procedencia, códigos, enlaces y calidad?	RECORD/RECORD-PE
Discusión y aplicabilidad	¿Se identificaron sesgos plausibles (selección, recuerdo, supervivencia, overmatching) y su dirección probable?	CASP, JBI
	¿Se discutió validez externa y aplicabilidad clínica/salud pública?	STROBE
	¿Conclusiones prudentes? Evitar traducir OR a RR sin enfermedad rara y evitar inferencias causales sin soporte.	STROBE-E&E, CASP
Uso rápido (60 segundos)	1. Revisar título/resumen: diseño, población, exposición/desenlace.	Adaptado de STROBE, CASP, JBI
	2. Métodos: casos incidentes, controles comparables, exposición medida objetivamente, tipo de muestreo.	
	3. Resultados: OR con IC95% (¿cruza 1.0?), ajustes (¿cambian mucho?), análisis condicional si hubo pareamiento.	
	4. Discusión: sesgos reconocidos y aplicabilidad; si datos rutinarios, buscar ítems RECORD.	

Fuente: Elaboración propia

9. UTILIDAD

El diseño de casos y controles, pese a sus limitaciones, conserva un lugar central en la investigación clínica y epidemiológica. Su mayor fortaleza radica en la eficiencia: permite explorar asociaciones entre exposiciones y enfermedades sin necesidad de seguir a grandes cohortes durante años. Esto lo convierte en una herramienta particularmente útil en tres escenarios:

Enfermedades raras o de baja incidencia: Cuando el desenlace es infrecuente —como ciertos cánceres, enfermedades autoinmunes poco comunes o patologías raras, los estudios de cohorte requerirían decenas de miles de participantes para acumular suficientes casos. El diseño

de casos y controles, en cambio, comienza reclutando directamente los casos ya ocurridos, lo que lo vuelve mucho más factible.

Enfermedades de larga latencia: En condiciones donde existe un período prolongado entre la exposición y el desarrollo de la enfermedad como la exposición a carcinógenos y la aparición de cáncer, un estudio prospectivo implicaría décadas de seguimiento. El enfoque de casos y controles acorta este tiempo, permitiendo evaluar retrospectivamente las exposiciones relevantes.

Exploración inicial de asociaciones: Frecuentemente, los estudios de casos y controles sirven como punto de partida para identificar hipótesis etiológicas que luego se confirmarán con diseños más robustos (cohortes o ensayos clínicos). Al detectar señales tempranas, pueden orientar la agenda de investigación y las políticas de salud.

En la práctica clínica: Para el médico lector, la utilidad está en la lectura crítica: al encontrarse con un estudio de casos y controles en una revista, debe reconocer que sus resultados no prueban causalidad, pero sí pueden aportar evidencia sugestiva y coherente con la biología de la enfermedad y otros estudios. En temas como el papel del *Helicobacter pylori* en cáncer gástrico, el consumo de tabaco en EPOC o la asociación entre fármacos y efectos adversos poco frecuentes, este diseño ha sido crucial para establecer vínculos iniciales que luego transformaron la práctica clínica.

En salud pública: Los hallazgos de los estudios de casos y controles pueden tener un impacto inmediato en políticas de prevención. Identificar factores de riesgo modificables (ej. consumo de agua con nitratos, exposición a humo de leña, hábitos dietéticos) permite diseñar intervenciones comunitarias y campañas educativas incluso antes de contar con grandes estudios prospectivos.



10. AUTOEVALUACIÓN

PREGUNTAS DE SELECCIÓN MÚLTIPLE

1. En un estudio de casos y controles, ¿cuál es la medida de asociación característica y más adecuada?

- Riesgo relativo (RR)
- Incidencia acumulada
- Odds ratio (OR)
- Diferencia de riesgos

2. Un artículo reporta OR = 2.5 (IC95% 0.9-6.0) para exposición a pesticidas y leucemia infantil. ¿Cuál es la interpretación correcta?

- Existe asociación significativa porque el OR es > 1.

- b) No hay evidencia suficiente de asociación porque el IC95% incluye 1.
 - c) La exposición duplica el riesgo de leucemia en todos los casos.
 - d) El valor p debe ser <0.05 porque el OR es elevado.
3. El sesgo de Berkson ocurre típicamente cuando:
- a) Los casos son seleccionados de otra población distinta a los controles.
 - b) Los controles hospitalarios tienen exposiciones relacionadas con su motivo de ingreso.
 - c) Se incluyen casos prevalentes en lugar de incidentes.
 - d) El investigador conoce la exposición antes de clasificar los grupos.
4. Si en un estudio apareado 1:1 se encuentran 18 pares con caso expuesto/control no expuesto y 9 pares con caso no expuesto/control expuesto, el OR condicional es:
- a) 0.5
 - b) 1.0
 - c) 2.0
 - d) 3.0
5. ¿Cuál de las siguientes afirmaciones refleja mejor el valor de un OR ajustado frente al crudo?
- a) El OR ajustado es siempre más alto que el crudo.
 - b) El OR ajustado refleja el efecto de la exposición controlando por posibles confusores.
 - c) El OR crudo es suficiente y no se necesita ajustar.
 - d) El ajuste elimina toda posibilidad de sesgo.

PREGUNTAS ABIERTAS

6. Explique con sus palabras la diferencia entre un OR crudo y un OR ajustado. ¿Por qué es importante mirar ambos al leer un artículo?
7. Describa un ejemplo de cómo el sesgo de recuerdo puede distorsionar los resultados de un estudio de casos y controles.
8. ¿Por qué los casos incidentes son preferibles a los casos prevalentes en este tipo de diseño?
9. Imagine que un artículo muestra un OR de 0.4 para consumo de frutas y cáncer gástrico, con IC95% 0.2–0.9. ¿Cómo lo interpretaría críticamente?
10. ¿En qué situaciones el OR puede aproximar al riesgo relativo (RR), y qué debe tener en cuenta el lector al hacer esa interpretación?

RESPUESTAS:

1: c) Odds ratio (OR). El Odds Ratio es la medida característica de los estudios de casos y controles porque no es posible calcular incidencia ni riesgo directamente, ya que los grupos se definen por el desenlace y no por la exposición.

2: b) No hay evidencia suficiente de asociación porque el IC95% incluye 1. Aunque el OR es mayor que 1, el intervalo de confianza cruza el valor nulo (1.0), lo que indica que la asociación no es estadísticamente significativa al 95%.

3: b) Los controles hospitalarios tienen exposiciones relacionadas con su motivo de ingreso. El sesgo de Berkson ocurre cuando los controles hospitalarios no representan adecuadamente la población fuente, generando una distorsión en la estimación del OR.

4: c) 2.0. En estudios apareados, el OR condicional se calcula como la razón entre pares discordantes: $18/9 = 2.0$. Solo los pares discordantes aportan información a la estimación.

5: b) El OR ajustado refleja el efecto de la exposición controlando por posibles confusores. El ajuste permite estimar la asociación independiente de otras variables que podrían distorsionar el resultado.

6: El OR crudo muestra la asociación entre exposición y enfermedad sin controlar otras variables. El OR ajustado estima esa misma asociación pero controlando posibles factores de confusión (como edad, sexo o tabaquismo). Es importante mirar ambos porque si el OR cambia mucho tras el ajuste, probablemente existía confusión. Si casi no cambia, la asociación es más estable. La comparación permite evaluar la validez interna del estudio.

7: El sesgo de recuerdo ocurre cuando los casos recuerdan mejor o con mayor detalle exposiciones pasadas que los controles. Por ejemplo, en un estudio sobre pesticidas y leucemia infantil, los padres de niños enfermos pueden esforzarse más por recordar exposiciones

ambientales, mientras que los padres de niños sanos pueden no hacerlo con la misma intensidad. Esto puede sobreestimar la exposición en los casos y exagerar el OR.

8: Los casos incidentes son preferibles porque representan diagnósticos recientes y reducen el sesgo de supervivencia (sesgo de Neyman). Los casos prevalentes incluyen solo pacientes que han sobrevivido, lo que puede excluir formas graves o letales de la enfermedad y distorsionar la asociación. Además, los casos incidentes permiten evaluar mejor la temporalidad entre exposición y enfermedad.

9: Un OR de 0.4 indica que la exposición (consumo de frutas) se asocia con una reducción aproximada del 60% en la probabilidad de cáncer gástrico. Como el IC95% (0.2–0.9) no incluye 1.0, la asociación es estadísticamente significativa. Sin embargo, debe analizarse si el OR es ajustado y qué variables se controlaron, ya que podría existir confusión por factores como dieta global, nivel socioeconómico o tabaquismo.

10: El OR puede aproximarse al riesgo relativo cuando la enfermedad es poco frecuente en la población (generalmente <10%). En ese contexto, la diferencia matemática entre odds y riesgo es pequeña. Sin embargo, si la enfermedad es común, el OR puede sobreestimar el efecto real. El lector debe verificar la frecuencia del desenlace antes de interpretar el OR como si fuera un RR.

REFERENCIAS

- Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Epidemiology*. 2007;18(6):800–4. Available at: <https://doi.org/10.1097/EDE.0b013e3181577654>.
- Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *PLoS Med*. 2007;4(10):e297. Available at: <https://doi.org/10.1371/journal.pmed.0040297>.
- Greenland S, Thomas DC. On the need for the rare disease assumption in case-control studies. *Am J Epidemiol*. 1982;116(3):547–53. Available at: <https://doi.org/10.1093/oxfordjournals.aje.a113439>.
- Labrecque JA, Knol MJ. Do case-control studies always estimate odds ratios? *Int J Epidemiol*. 2020;49(5):1763–71. Available at: <https://doi.org/10.1093/ije/dyaa052>.
- VanderWeele TJ, Ding P. Sensitivity analysis in observational research: Introducing the E-value. *Ann Intern Med*. 2017;167(4):268–74. Available at: <https://doi.org/10.7326/M16-2607>.
- Pearce N. Analysis of matched case-control studies. *BMJ*. 2016;352:i969. Available at: <https://doi.org/10.1136/bmj.i969>.
- Schneeweiss S, Rassen JA, Brown JS, Rothman KJ, Happe L, Arlett P, et al. Graphical depiction of longitudinal study designs in health care databases. *Ann Intern Med*. 2019;170(6):398–406. Available from: <https://doi.org/10.7326/M18-3079>.
- Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ. *Modern epidemiology*. 4th ed. Philadelphia: Wolters Kluwer; 2021.
- Delgado-Rodríguez M, Llorca J. Bias. *J Epidemiol Community Health*. 2004;58(8):635–41. Available at: <https://doi.org/10.1136/jech.2003.008466>.
- Alexander LK. *Case-control studies: Basic concepts and rare disease assumption*. Chapel Hill: North Carolina Institute for Public Health, University of North Carolina; 2015. Available at: https://sph.unc.edu/wp-content/uploads/sites/112/2015/07/nciph_ERIC5.pdf.