

Chapter 13 / Capítulo 13

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.Ch13
Pages: 178-208



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

Cohort Studies

Estudios de Cohorte

Jorge Marquez Molina  

Isis Scarleth Funes Galindo  

Stanley Aguirre

Giovanni Callizaya-Macedo  

1 Hospital del Norte. Cochabamba, Bolivia.

2 Clínica municipal de emergencias. Talanga. Honduras.

3 Investigador independiente.

4 Hospital del Niño “Dr. Ovidio Aliaga Uría”. Departamento de Cirugía Pediátrica. La Paz, Bolivia.

ABSTRACT

Introduction: Cohort studies are among the most robust analytical observational designs for assessing the association between an exposure and the occurrence of an outcome. Their main strength lies in establishing the temporal sequence between these two events, allowing for the estimation of incidences and the quantification of risks using measures such as relative risk (RR), incidence rate, and hazard ratio (HR). **Development:** This chapter describes the methodological foundations of prospective, retrospective, and ambispective cohorts, emphasizing the importance of appropriate population selection, precise definition of the exposure and outcomes, longitudinal follow-up, and control for confounding factors. It addresses the main measures of frequency and association, including cumulative incidence, incidence rate, attributable risk, RR, and HR, complemented by survival analysis tools such as Kaplan-Meier curves and Cox regression. Criteria for evaluating the internal and external validity and clinical relevance of the findings are also highlighted. **Conclusions:** Cohort studies provide fundamental evidence for understanding the natural history of diseases and estimating the impact of risk and protective factors. Their proper interpretation requires assessing the time frame, the quality of follow-up, the control of confounding factors, and the true magnitude of the observed effects. Beyond statistical results, these studies allow for the translation of evidence into clinical decisions and public health strategies aimed at prevention and improving healthcare.

Keywords: Cohort Studies; Relative Risk; Hazard Ratio; Survival Analysis; Confounding Factors

RESUMEN

Introducción: Los estudios de cohorte constituyen uno de los diseños observacionales analíticos más sólidos para evaluar la asociación entre una exposición y la ocurrencia de un desenlace. Su principal fortaleza radica en establecer la secuencia temporal entre ambos eventos, permitiendo estimar incidencias y cuantificar riesgos mediante medidas como el riesgo relativo (RR), la tasa de incidencia y el hazard ratio (HR). **Desarrollo:** El capítulo describe los fundamentos metodológicos de las cohortes prospectivas, retrospectivas y ambispectivas, enfatizando la importancia de una adecuada selección de la población, la definición precisa de la exposición y los desenlaces, el seguimiento longitudinal y el control de factores de confusión. De esta manera, aborda las principales medidas de frecuencia y asociación, incluyendo incidencia acumulada, tasa de incidencia, riesgo atribuible, RR y HR, complementadas con herramientas de análisis de supervivencia como las curvas de Kaplan-Meier y la regresión de Cox. También se destacan los criterios para evaluar la validez interna, externa y la relevancia clínica de los hallazgos. **Conclusiones:** Los estudios de cohorte proporcionan evidencia fundamental para comprender la historia natural de las enfermedades y estimar el impacto de factores de riesgo y protección. Su adecuada interpretación exige valorar la temporalidad, la calidad del seguimiento, el control de la confusión y la magnitud real de los efectos observados. Más allá de los resultados estadísticos, estos estudios permiten traducir la evidencia en decisiones clínicas y estrategias de salud pública orientadas a la prevención y al mejoramiento de la atención sanitaria.

Palabras clave: Estudios de Cohorte; Riesgo Relativo; Hazard Ratio; Análisis de Supervivencia; Factores de Confusión

1. CASO CLÍNICO

En una comunidad rural del altiplano boliviano, donde la mayoría de los hogares utilizan leña como principal fuente de energía para cocinar, un médico internista nota un número creciente de pacientes con tos crónica, disnea progresiva y diagnóstico de enfermedad pulmonar obstructiva crónica (EPOC). Surge la pregunta clínica: ¿la exposición prolongada al humo de leña aumenta el riesgo de desarrollar EPOC en comparación con la no exposición?

Para responder, se diseña un estudio de cohorte prospectivo:

- Se reclutan 1,200 adultos sin EPOC al inicio del estudio (600 expuestos al humo de leña diariamente por ≥ 10 años y 600 no expuestos que utilizan gas natural o electricidad).
- Se los sigue durante 12 años, evaluando anualmente la función pulmonar mediante espirometría y registrando aparición de EPOC.
- Durante el seguimiento, algunos participantes cambian de fuente de energía, lo que obliga a registrar la exposición de forma dinámica.
- El desenlace principal es la incidencia acumulada de EPOC, y se planea calcular el riesgo relativo (RR), la tasa de incidencia (TI) y el hazard ratio (HR) mediante modelos de Cox.

2. OBJETIVOS

Los estudios de cohorte son diseños investigativos longitudinales con el propósito de responder preguntas clínicas u epidemiológicas estableciendo la secuencia temporal en tiempo entre la exposición y el desenlace sobre los individuos expuestos y no expuestos.

2.1 OBJETIVOS ESPECIFICOS

- Evaluar la asociación entre exposición y desenlace.
- Estimar la incidencia de la enfermedad en cada grupo.
- Calcular medidas de asociación: Riesgo Relativo, Tasa de Incidencia, Hazard Ratio.

2.2 FINALIDAD DE LOS ESTUDIOS DE COHORTE

- Determinar la evaluación de la causalidad.
- Cuantificar el impacto de la exposición: Permite calcular el riesgo atribuible (RA) y la fracción atribuible poblacional (FAP).
- Generar evidencia informativa para la prevención.
- Analizar desenlaces secundarios.

3. INTRODUCCIÓN

Los estudios de cohorte son estudios observacionales, longitudinales, analíticos, generalmente prospectivos que constituyen uno de los pilares de la epidemiología clínica porque permiten observar cómo evoluciona una población en relación con una exposición específica a lo largo del tiempo. A diferencia de los estudios transversales, que solo muestran una “fotografía” de la realidad, las cohortes proporcionan una película en movimiento, siguiendo a los individuos desde que están libres del desenlace hasta que éste ocurre. En términos metodológicos, una cohorte se define como un grupo de personas que comparte una característica común en un momento inicial (por ejemplo, estar expuestos al humo de leña, como en nuestro caso clínico) y que se sigue en el tiempo para evaluar la aparición de un desenlace (EPOC). El rasgo distintivo es que la exposición se mide antes de que ocurra el desenlace, lo que garantiza la secuencia temporal, requisito fundamental para aproximarse a inferencias causales.

Importancia clínica y epidemiológica:

- Los estudios de cohorte permiten estimar la incidencia de enfermedades, una medida clave para entender cuán frecuente es que aparezca un desenlace en personas inicialmente sanas.
- Al comparar expuestos y no expuestos, se obtiene el RR y otras métricas (tasa de incidencia, hazard ratio), que son las medidas de asociación más usadas en investigación clínica.
- Una sola cohorte puede analizar múltiples desenlaces para una misma exposición (por

ejemplo, humo de leña → EPOC, hipertensión pulmonar, cáncer pulmonar, mortalidad).

- De igual forma, permite estudiar tanto factores de riesgo como de protección (ejemplo: uso de cocina a gas como medida protectora frente a EPOC).

En países de ingresos medios y bajos, especialmente en zonas rurales, los estudios de cohorte son herramientas valiosas para documentar la carga de enfermedad relacionada con exposiciones ambientales y ocupacionales. El caso clínico que plantea la asociación entre humo de leña y EPOC refleja un problema real de salud pública en Bolivia y América Latina, donde más del 30% de los hogares rurales aún utilizan biomasa como principal fuente de energía.

Para el personal de salud, la lectura crítica de cohortes es esencial porque:

- Aclara el riesgo real de que un paciente desarrolle una enfermedad tras estar expuesto a un factor.
- Informa decisiones preventivas: identificar factores de riesgo modificables permite orientar recomendaciones prácticas (dejar de fumar, cambiar hábitos dietéticos, reducir exposición ambiental).
- Fortalece la medicina basada en evidencia (MBE): una cohorte bien diseñada ofrece evidencia de alta calidad entre los estudios observacionales analíticos, especialmente cuando se complementa con análisis estadísticos robustos que controlan confusores.

4. METODOLOGÍA

La sección de metodología es el corazón de un estudio de cohorte y constituye la parte más importante a la hora de evaluarlo críticamente. Allí se decide si los resultados serán creíbles, válidos y aplicables a la práctica clínica. Al leer un artículo, el personal de salud debe preguntarse si la población fue bien seleccionada, si la exposición y los desenlaces fueron medidos de manera confiable, si el seguimiento fue suficiente y si el análisis estadístico fue el adecuado para la pregunta planteada. A continuación, se presentan los principales componentes metodológicos que deben evaluarse al leer un artículo de cohorte.

4.1 TIPOS DE ESTUDIOS DE COHORTE

Los estudios de cohorte pueden clasificarse en prospectivos, retrospectivos o ambispectivos según el momento en que se recolectan los datos y la forma en que se sigue a los participantes. Identificar correctamente el tipo de cohorte en un artículo es esencial, porque de ello dependerán la validez interna, el tiempo necesario, los costos y el riesgo de sesgos.

COHORTE PROSPECTIVA

Una cohorte prospectiva, también llamada concurrente, comienza en el presente: los investigadores reclutan a los participantes, miden su exposición antes de que ocurra el desenlace, y los siguen hacia adelante en el tiempo. En este diseño, la secuencia temporal es muy clara, ya que la exposición se registra antes de la aparición del evento.

Su fortaleza principal radica en que permite estandarizar la medición de la exposición y los desenlaces, con criterios diagnósticos definidos desde el inicio. Miden características en puntos de tiempo predeterminados para entender la relación entre la cohorte y resultado del estudio. Además, facilita el control de calidad de los datos y la verificación de la temporalidad. Por estas razones, suele considerarse el “estándar de oro” dentro de los estudios observacionales.

Sin embargo, tiene limitaciones importantes: es costoso, requiere un gran tamaño de muestra, bastante tiempo de seguimiento y suele enfrentar pérdidas de participantes que, si son diferenciales entre expuestos y no expuestos, introducen sesgos (de desgaste).

En nuestro caso clínico: reclutar en la actualidad a 600 adultos expuestos al humo de leña y 600 no expuestos en comunidades rurales del altiplano, asegurándose de que ninguno tenga EPOC al inicio. Luego se realiza un seguimiento anual por 12 años, con espirometrías estandarizadas, hasta observar cuántos desarrollan EPOC en cada grupo.

COHORTE RETROSPECTIVA

En una cohorte retrospectiva, también llamada histórica, los análisis de datos son en

el presente, pero las mediciones en la línea de base de los participantes y sus seguimientos ocurrieron en el pasado. Se utilizan expedientes clínicos, bases de datos hospitalarias, censos o registros administrativos para clasificar a los participantes según su exposición y determinar si desarrollaron o no el desenlace.

La fortaleza de este diseño es que es mucho más rápido y económico que una cohorte prospectiva. Permite obtener resultados en menor tiempo al revisar variables predictoras (datos pasados) para examinar eventos o resultados, útil cuando se dispone de registros de buena calidad.

No obstante, su limitación principal es la dependencia absoluta de los registros existentes. Si la exposición o el desenlace no fueron medidos de forma precisa o estandarizada, el riesgo de sesgo de información es elevado. Además, la temporalidad puede estar menos clara, ya que los datos se reconstruyen de manera retrospectiva.

En nuestro caso clínico: revisar historias clínicas y registros de salud entre 2010 y 2020 para identificar a personas que estuvieron expuestas al humo de leña y compararlas con aquellas que no lo estuvieron. A partir de esos registros, se determina cuántos desarrollaron EPOC en el período. La calidad del estudio dependerá de si los registros incluyeron espirometrías y exposición detallada al tipo de combustible utilizado.

COHORTE AMBISPECTIVA

La cohorte ambispectiva, también llamada mixta, combina elementos de los dos diseños anteriores. Se inicia utilizando datos retrospectivos ya disponibles para clasificar la exposición y recoger parte de la información histórica, y a partir de ese punto se continúa con un seguimiento prospectivo hacia adelante.

La fortaleza de este diseño es que aprovecha la rapidez de los datos históricos, pero al mismo tiempo permite continuar la observación con métodos estandarizados en el presente. De esta forma, combina eficiencia con mayor control sobre la calidad de la información.

La limitación, sin embargo, es que hereda los problemas de ambas modalidades: puede tener registros incompletos del pasado y enfrenta pérdidas de seguimiento en la etapa prospectiva.

En nuestro caso clínico: los investigadores podrían revisar los expedientes de un hospital rural de los últimos 7 años para identificar la exposición al humo de leña y diagnosticar casos de EPOC durante ese período. A partir de hoy, se decide continuar el seguimiento prospectivo durante cinco años más, con visitas anuales y mediciones estandarizadas de función pulmonar, para registrar nuevos casos de EPOC.

En resumen, al leer un artículo de cohorte es indispensable identificar el tipo de diseño. Una cohorte prospectiva suele ser más válida pero costosa; una retrospectiva es más rápida y barata, pero depende de la calidad de los registros; y una ambispectiva busca un equilibrio entre ambos enfoques. Reconocer estas diferencias permite al lector crítico juzgar de inmediato las fortalezas y debilidades del estudio que tiene frente a sí, evitando interpretaciones simplistas o erróneas.

Tabla 1. Comparación entre los tipos de estudios de cohorte

Característica	Cohorte prospectiva	Cohorte retrospectiva	Cohorte ambispectiva
Momento de inicio del estudio	El estudio comienza en el presente y el seguimiento ocurre hacia el futuro.	El estudio se realiza en el presente, pero la exposición y el seguimiento ocurrieron en el pasado.	Combina datos del pasado y seguimiento prospectivo desde el presente.
Medición de la exposición	Se mide antes de que ocurra el desenlace.	Se reconstruye a partir de registros existentes.	Parte retrospectiva reconstruida + medición prospectiva posterior.
Temporalidad exposición–desenlace	Muy clara y bien establecida.	Puede ser menos clara si los registros son incompletos.	Clara en la fase prospectiva; depende de registros en la fase retrospectiva.
Control sobre calidad de datos	Alto (mediciones estandarizadas desde el inicio).	Variable (depende de la calidad de los registros históricos).	Moderado (hereda limitaciones del pasado, mejora en fase prospectiva).
Tiempo para obtener resultados	Largo (requiere años de seguimiento).	Corto (los eventos ya ocurrieron).	Intermedio.
Costo	Elevado.	Bajo o moderado.	Intermedio.
Riesgo de sesgo de información	Bajo si hay buen control metodológico.	Mayor (dependencia de registros previos).	Variable (alto en fase retrospectiva, menor en fase prospectiva).
Riesgo de pérdidas de seguimiento	Alto si el seguimiento es prolongado.	Bajo (los eventos ya ocurrieron).	Presente en fase prospectiva.

Característica	Cohorte prospectiva	Cohorte retrospectiva	Cohorte ambispectiva
Ejemplo aplicado al caso clínico (EPOC y humo de leña)	Reclutar hoy 1,200 adultos sin EPOC y seguirlos 12 años con espirometría anual.	Revisar registros 2010–2020 para identificar exposición y diagnóstico de EPOC.	Revisar registros 2015–2022 y continuar seguimiento 5 años más con mediciones estandarizadas.
Fortaleza principal	Mayor validez interna y mejor control de medición.	Rapidez y menor costo.	Equilibrio entre eficiencia y control metodológico.
Limitación principal	Costosa y susceptible a pérdidas diferenciales.	Dependencia absoluta de registros previos.	Combina limitaciones de ambos diseños.
Fuente: Elaboración propia			

4.2 SELECCIÓN DE LA POBLACIÓN

La validez de un estudio de cohorte depende en gran medida de cómo se selecciona la población. Para el lector crítico es fundamental preguntarse si los participantes incluidos realmente representan a la población de interés y si la muestra inicial estaba libre del desenlace en estudio.

En primer lugar, la población de estudio debe estar libre de la enfermedad al inicio del seguimiento. De lo contrario, lo que se estaría midiendo no es la incidencia de la enfermedad, sino su prevalencia, lo cual distorsionaría los resultados. En el caso, si se investiga la relación entre humo de leña y riesgo de EPOC, ningún participante debería tener un diagnóstico previo de EPOC al momento de ser reclutado.

Además, los autores deben describir claramente los criterios de inclusión y exclusión. Un buen artículo especifica quiénes fueron elegibles y quiénes quedaron fuera del estudio. En nuestro caso, podrían incluirse adultos mayores de 25 años residentes en comunidades rurales que utilicen leña o gas como fuente de energía. En cambio, podrían excluirse personas con enfermedades pulmonares intersticiales, cáncer de pulmón o con antecedentes de exposición ocupacional intensa a polvos industriales, ya que estas condiciones confundirían la relación entre humo de leña y EPOC.

Un aspecto clave es la representatividad de la muestra. El lector debe preguntarse: ¿los participantes incluidos reflejan a la población general a la que se quieren aplicar los resultados? Una cohorte que reclute únicamente pacientes de un hospital puede no ser representativa de toda la comunidad rural, porque quienes llegan a la atención médica suelen diferir en características socioeconómicas o en gravedad de síntomas. Por eso, en el caso clínico, la validez sería mayor si los investigadores reclutan participantes directamente en las comunidades, visitando hogares y aplicando cuestionarios estandarizados.

Otro criterio esencial es la comparabilidad entre grupos. En estudios de cohorte, los expuestos y los no expuestos deben ser similares en características importantes como edad, sexo, nivel socioeconómico, tabaquismo y comorbilidades. Si un grupo está compuesto mayoritariamente por personas mayores y el otro por jóvenes, la diferencia de riesgo podría deberse a la edad y no a la exposición estudiada. Aquí entra en juego el concepto de confusión, que se analizará en detalle más adelante, pero el lector debe reconocer desde la metodología si existe un desequilibrio inicial entre los grupos.

Finalmente, es importante considerar la tasa de participación. Un artículo debería indicar qué porcentaje de las personas elegibles aceptó participar. Una tasa muy baja podría introducir sesgo de selección, ya que quienes aceptan participar pueden diferir en motivaciones, salud o conductas respecto de quienes rechazan hacerlo.

En síntesis, al evaluar la selección de la población en un estudio de cohorte, el lector debe asegurarse de que los participantes estuvieran libres de la enfermedad al inicio, que los criterios de inclusión y exclusión estén bien definidos, que la muestra sea representativa de la población objetivo, que los grupos expuestos y no expuestos sean comparables, y que la tasa de participación sea adecuada. Solo así podrá confiar en que los resultados reflejan la realidad del problema en estudio y no una distorsión generada por una muestra sesgada.

4.3 DEFINICIÓN Y MEDICIÓN DE LA EXPOSICIÓN

La forma en que se define y mide la exposición es determinante para la validez de los resultados, el lector crítico debe prestar atención a cómo los investigadores conceptualizaron la exposición y con qué herramientas la evaluaron. Una exposición mal definida, evaluada de

manera imprecisa o heterogénea entre los grupos, puede generar sesgos graves que conduzcan a conclusiones equivocadas.

La definición operacional de la exposición es el primer aspecto a revisar. Es necesario que este definido su concepto general con claridad sobre lo que se entiende por exposición, así puede ser transformada en una variable concreta, medible y reproducible. No basta con mencionar que alguien estuvo “expuesto al humo de leña”; Por ejemplo, en nuestro caso clínico debería ser definido como “utilizar leña como principal combustible para cocinar al menos cuatro horas al día durante diez años consecutivos”.

En segundo lugar, es fundamental examinar los métodos de medición de la exposición. Estos pueden ser objetivos o subjetivos. Los métodos objetivos incluyen, por ejemplo, el uso de sensores ambientales para medir partículas contaminantes, biomarcadores de exposición (como niveles de monóxido de carbono en sangre o carboxihemoglobina), o registros oficiales sobre el tipo de combustible utilizado en los hogares. Los métodos subjetivos, en cambio, se basan en cuestionarios o autoinformes, donde los participantes declaran sus hábitos de uso de combustibles. Aunque más fáciles de aplicar, son susceptibles a errores de memoria o a respuestas socialmente deseables.

El lector crítico debe preguntarse si los investigadores utilizaron la misma técnica de medición para todos los participantes. La homogeneidad en la medición es crucial: si en un grupo se utilizaron cuestionarios y en otros biomarcadores, las diferencias encontradas podrían reflejar más el método de medición que la exposición en sí.

Otro punto a considerar es la clasificación de la exposición. Esta puede ser dicotómica (expuesto/no expuesto) o categórica, estableciendo niveles de intensidad. La estratificación no solo permite analizar si existe una relación dosis-respuesta, sino que fortalece la interpretación causal de los hallazgos. Por ejemplo, en el caso, se clasifica a los participantes como de exposición baja (menos de dos horas al día), moderada (dos a cuatro horas) o alta (más de cuatro horas diarias), así a mayor intensidad de exposición aumenta el riesgo de EPOC y la asociación gana consistencia.

Un aspecto avanzado que cada vez más estudios incluyen es la exposición dinámica. En cohortes largas, es habitual que los hábitos cambien con el tiempo. Los mejores estudios actualizan la exposición durante el seguimiento y la analizan como una variable dependiente del tiempo. Por ejemplo, una familia que inicialmente usaba leña puede adoptar gas o electricidad en el transcurso del estudio. Si el artículo clasifica a los participantes únicamente en función de la exposición inicial, se corre el riesgo de cometer errores de clasificación que subestimen o sobreestimen el efecto.

En síntesis, cuando se revisa un estudio de cohorte, el lector debe asegurarse de que la exposición haya sido definida de manera precisa, medida con métodos confiables, evaluada con el mismo procedimiento en todos los participantes y, cuando sea posible, actualizada a lo largo del seguimiento. Si la definición es vaga, la medición subjetiva o la clasificación inadecuada, las conclusiones del estudio pierden fuerza y su aplicabilidad clínica se ve comprometida.

4.4 DEFINICIÓN Y MEDICIÓN DE LOS DESENLACES

El desenlace es la variable de mayor relevancia, pues representa el evento clínico cuya relación con la exposición se busca establecer. La claridad en su definición y la precisión en su medición determinan la calidad de las conclusiones del estudio. Para el lector crítico, esta sección debe ser revisada con especial atención, ya que los errores en la definición del desenlace son una de las fuentes más frecuentes de sesgo.

Lo primero que debe verificarse es si el desenlace principal (resultado primario) está definido de manera explícita y objetiva. La definición debe basarse en criterios diagnósticos reconocidos internacionalmente. Si un estudio no aclara con qué criterio se diagnosticó el desenlace principal, o si utiliza definiciones vagas como “síntomas respiratorios crónicos”, su validez científica queda seriamente limitada. En nuestro caso clínico, el desenlace de interés es el diagnóstico de enfermedad pulmonar obstructiva crónica (EPOC), como el establecido por la Iniciativa Global para la EPOC (GOLD): un cociente FEV1/FVC menor al 70% confirmado mediante espirometría post-broncodilatador.

Además del desenlace principal, muchos estudios de cohorte evalúan desenlaces secundarios (resultados secundarios), que enriquecen el análisis y permiten una visión más amplia de la carga de enfermedad. Estos desenlaces secundarios ofrecen información complementaria sobre la evolución clínica y la severidad de la enfermedad. En el caso, los investigadores

registraron exacerbaciones graves de EPOC, hospitalizaciones, mortalidad por causa respiratoria o disminución progresiva del volumen espiratorio forzado (FEV1).

Otro aspecto esencial es la homogeneidad en la medición. Todos los participantes deben ser evaluados con los mismos métodos diagnósticos, en las mismas condiciones y con protocolos estandarizados. El lector debe comprobar que el artículo especifique cómo, cuándo y con qué instrumentos se midieron los desenlaces. En nuestro caso si algunos participantes son diagnosticados mediante espirometría y otros solo a partir de síntomas clínicos, el riesgo de sesgo de información aumenta considerablemente.

Los estudios también pueden incluir desenlaces compuestos, es decir, una combinación de varios eventos clínicos, como EPOC, hospitalización y mortalidad respiratoria. Aunque esto aumenta la potencia estadística, el lector crítico debe valorar si todos los componentes del desenlace compuesto son de igual relevancia clínica. Si se combinan eventos leves (tos crónica) con eventos graves (muerte por insuficiencia respiratoria), la interpretación puede resultar engañosa.

Finalmente, en cohortes largas, es fundamental que los desenlaces sean evaluados en intervalos regulares y definidos. El lector debe identificar si el artículo reporta estos intervalos y si hubo consistencia en la recolección de datos. En el caso, sería razonable realizar controles anuales con espirometría, complementados con visitas clínicas en caso de síntomas o exacerbaciones.

En conclusión, para que los resultados de una cohorte sean válidos, los desenlaces deben estar claramente definidos, medidos con métodos estandarizados, aplicados de forma homogénea en todos los participantes y evaluados en intervalos regulares durante el seguimiento. De no cumplirse estas condiciones, la confianza en los hallazgos se ve comprometida, sin importar cuán sofisticado sea el análisis estadístico posterior.

4.5 SEGUIMIENTO DE LA COHORTE

El seguimiento constituye la columna vertebral de un estudio de cohorte. Es en esta fase donde se observa si los participantes desarrollan o no el desenlace en cuestión, y de la calidad del seguimiento dependerá que las medidas de incidencia y asociación reflejen la realidad. Para el lector crítico, evaluar cómo se llevó a cabo este proceso es tan importante como revisar la definición de exposición o desenlace.

El primer aspecto a considerar es la duración del seguimiento. Éste debe ser lo suficientemente larga como para que aparezca el desenlace de interés. Si un estudio presenta un tiempo de seguimiento demasiado corto, sus conclusiones sobre causalidad deben ser interpretadas con cautela. En el caso, un seguimiento breve de dos o tres años sería insuficiente para detectar la aparición de EPOC, una enfermedad de evolución crónica y lenta. En cambio, un período de 10 a 12 años, ofrece mayor probabilidad de registrar nuevos diagnósticos y de establecer diferencias claras entre expuestos y no expuestos.

En segundo lugar, es esencial examinar los intervalos de evaluación. Un estudio serio no se limita a observar a los participantes al inicio y al final, sino que programa controles periódicos para recolectar datos de forma continua. Estos pueden ser anuales, semestrales o adaptados al desenlace de interés. Cuando los intervalos son irregulares o no se explicitan en el artículo, el riesgo de datos incompletos o de diagnósticos tardíos aumenta. En el caso, los controles anuales con espirometría y cuestionarios respiratorios serían un estándar apropiado.

Un tercer aspecto crucial es el problema de las pérdidas de seguimiento. En cualquier cohorte es esperable que algunos participantes abandonen el estudio, se muden, fallezcan por causas no relacionadas o simplemente dejen de acudir a los controles. El lector debe fijarse en cuántas personas se perdieron y en qué grupo ocurrieron esas pérdidas. Una pérdida global inferior al 10% se considera aceptable; cuando supera el 20%, la validez de los resultados queda seriamente comprometida. Más grave aún es cuando las pérdidas son diferenciales: si, por ejemplo, abandonan más participantes del grupo expuesto porque desarrollaron síntomas respiratorios severos, los resultados podrían subestimar la verdadera incidencia del desenlace. Los autores deben reportar no solo el número de pérdidas, sino también las razones, y describir qué métodos utilizaron para minimizar su impacto (contacto telefónico, visitas domiciliarias, análisis de sensibilidad).

Otro elemento cada vez más importante es el manejo de la exposición dinámica. En estudios prolongados, es frecuente que los hábitos de los participantes cambien con el tiempo como se mencionó previamente. Los mejores artículos actualizan la información de exposición en cada visita de seguimiento y utilizan modelos estadísticos que permiten tratar la exposición como una variable dependiente del tiempo. En nuestro caso, algunas familias inicialmente expuestas al

humo de leña pueden migrar al uso de gas natural o electricidad, mientras que otras, por razones económicas, pueden volver a usar leña tras un período sin hacerlo.

Finalmente, el lector debe considerar la calidad de los registros durante el seguimiento. Es común que algunos estudios mencionen el tiempo de observación, pero no detallen cómo se documentaron los eventos. Un buen artículo debe especificar si los desenlaces fueron confirmados mediante pruebas objetivas (como espirometrías), si hubo un comité independiente que validó los diagnósticos y si se utilizaron fuentes múltiples de información (historias clínicas, entrevistas, registros de defunción).

En conclusión, el seguimiento en una cohorte debe ser suficiente en duración, realizado en intervalos regulares, con pérdidas mínimas y bien documentadas, considerando los cambios en la exposición y asegurando la calidad en la recolección de datos. Para el lector crítico, un estudio con seguimiento incompleto o deficiente es como una película mal editada: puede tener un inicio prometedor, pero carece de un final confiable que permita interpretar la historia en su totalidad.

4.6 CONTROL DE LA CONFUSIÓN

Uno de los grandes retos de los estudios de cohorte es el manejo de la confusión. El lector crítico debe estar atento a cómo el artículo identifica, mide y controla los posibles confusores, ya que de ello depende la validez de la asociación observada.

Se entiende por confusión a la distorsión de la relación entre exposición y desenlace producida por una variable (externa). Las variables confusoras deben tener 3 condiciones, estar asociadas con la exposición investigada; estar asociadas con el resultado investigado; no deben estar en el camino causal entre exposición y resultado. Variables a considerar como confusores son: la edad, el sexo, el nivel socioeconómico o las condiciones ocupacionales. Todas ellas deben estar claramente registradas en la cohorte y reportadas en la metodología. En el caso clínico de humo de leña y riesgo de EPOC, el tabaquismo es un confusor evidente: los fumadores tienen más probabilidad de estar expuestos a humo en el hogar (por factores culturales o socioeconómicos) y, al mismo tiempo, el tabaco es un factor de riesgo reconocido para desarrollar EPOC. Si no se mide y ajusta adecuadamente, se podría concluir erróneamente que el humo de leña tiene un efecto mayor al real.

La confusión podría dar lugar a una distorsión de los efectos; puede conducir a sobreestimación o subestimación de un efecto, o incluso invertir la dirección de un efecto/resultado.

El lector debe analizar si el artículo aplicó estrategias de control de confusión. Estas pueden plantearse en dos niveles:

- Durante el diseño del estudio. Algunos autores emplean técnicas como la restricción (incluir solo no fumadores para eliminar la influencia del tabaco), el emparejamiento (hacer que los grupos expuestos y no expuestos sean comparables en edad, sexo u otras variables) o la selección de una muestra homogénea. Estas medidas mejoran la comparabilidad desde el inicio, aunque reducen la generalización de los resultados.
- Durante el análisis estadístico. Aquí se aplican métodos más sofisticados que buscan ajustar los resultados a los confusores medidos. Los más frecuentes son la regresión multivariable (incluyendo edad, sexo, tabaquismo, ocupación), la estratificación por categorías (calcular riesgos separados en fumadores y no fumadores) o el uso de puntuaciones de propensión (propensity scores) que equilibran los grupos mediante técnicas estadísticas. En cohortes con análisis de supervivencia, la regresión de Cox ajustada es una de las herramientas más utilizadas.

Un aspecto importante que el lector crítico debe valorar es cómo los autores seleccionaron las variables de ajuste. Si eligieron ajustar únicamente aquellas que resultaron “significativas” en un análisis preliminar, existe riesgo de sobreajuste o de dejar fuera variables realmente importantes. La mejor práctica consiste en justificar la selección de confusores con base en la literatura científica y en modelos conceptuales, como los diagramas causales dirigidos (DAGs).

Por último, es esencial recordar que ningún estudio observacional está completamente libre de confusión. Aun con los mejores métodos estadísticos, siempre puede quedar un grado de confusión residual debido a factores no medidos o mal clasificados. Por eso, un artículo sólido debe reconocer esta limitación en su discusión.

En resumen, al leer un estudio de cohorte, el lector crítico debe preguntarse: ¿los autores identificaron los confusores más relevantes? ¿Los midieron adecuadamente? ¿Aplicaron técnicas de ajuste apropiadas? Y, sobre todo, ¿reconocieron la posibilidad de confusión residual? Si la respuesta es afirmativa, la confianza en los hallazgos aumenta significativamente.

4.7 ANÁLISIS ESTADÍSTICO

El análisis estadístico tiene como propósito comparar la frecuencia de los desenlaces entre los grupos expuestos y no expuestos. Para ello se utilizan medidas de ocurrencia y asociación que permiten al lector cuantificar el riesgo y entender la magnitud de la relación observada.

El punto de partida suelen ser las medidas de incidencia, que informan cuántos casos nuevos de la enfermedad aparecen en un periodo determinado. A partir de estas incidencias, se calculan las medidas de asociación más utilizadas en cohortes: el riesgo relativo y la tasa de incidencia. En estudios donde interesa el tiempo hasta que ocurre el evento, se emplean análisis de supervivencia, como las curvas de Kaplan–Meier, y modelos más complejos como la regresión de Cox, que generan el hazard ratio.

Al evaluar un artículo, el lector debe fijarse en que estas medidas estén acompañadas de intervalos de confianza (IC95%). Un valor *p* aislado nunca es suficiente. Además, es recomendable que los autores presenten no solo medidas relativas (RR o HR), sino también medidas absolutas, como la diferencia de riesgos, que aportan mayor relevancia clínica.

En este apartado basta con reconocer que las cohortes requieren análisis que consideren la temporalidad y el tiempo de seguimiento. Más adelante, en el capítulo de interpretación estadística, se abordará en detalle cómo calcular estas medidas, cómo interpretarlas en la práctica clínica y qué errores comunes deben evitarse.

5. EVALUACIÓN DE LA VALIDEZ Y RELEVANCIA

Una vez comprendida la metodología, el siguiente paso para el lector crítico es evaluar si el estudio tiene validez interna (es decir, si los resultados reflejan la realidad dentro de la población estudiada) y validez externa (si esos resultados pueden aplicarse a otros contextos, incluidos los pacientes propios del lector). Al mismo tiempo, debe juzgar si los hallazgos son clínicamente relevantes, más allá de la significancia estadística.

5.1 VALIDEZ INTERNA

La validez interna depende de que el estudio haya controlado adecuadamente los sesgos y la confusión. Un artículo de cohorte es sólido cuando:

- Los grupos comparados estaban libres del desenlace al inicio y fueron definidos de manera clara.
- La exposición fue medida de forma precisa y homogénea.
- Los desenlaces fueron definidos con criterios diagnósticos objetivos y evaluados de manera uniforme.
- El seguimiento fue suficientemente largo, con pérdidas mínimas y explicadas en detalle.
- Los confusores más relevantes fueron identificados, medidos y ajustados en el análisis.

5.2 VALIDEZ EXTERNA

La validez externa hace referencia a la generalización de los hallazgos. El lector debe preguntarse si los participantes de la cohorte se parecen a los pacientes de su práctica clínica. Un estudio realizado en un hospital especializado de Estados Unidos puede tener resultados difíciles de aplicar en una comunidad rural de Bolivia, y viceversa. En nuestro caso clínico, los hallazgos de una cohorte rural sobre humo de leña y EPOC tendrían una validez externa muy alta para poblaciones latinoamericanas.

5.3 RELEVANCIA CLÍNICA

No todo hallazgo estadísticamente significativo es clínicamente importante. El lector debe fijarse en la magnitud del efecto y en las medidas absolutas de riesgo. Un riesgo relativo de 2.0 puede sonar impactante, pero si la incidencia absoluta es muy baja, el impacto real en salud pública puede ser limitado. Por el contrario, una diferencia de riesgos de 15% entre expuestos y no expuestos al humo de leña tendría implicaciones directas en prevención y políticas sanitarias, incluso si el intervalo de confianza es relativamente amplio, esto lo veremos a detalle en el acápite

de interpretación estadística.

5.4 INTERPRETACIÓN EQUILIBRADA

Un buen lector crítico no se deja llevar únicamente por la significancia estadística. Analiza los intervalos de confianza (¿son estrechos o demasiado amplios?), evalúa si los resultados son consistentes con estudios previos, y considera la plausibilidad biológica del hallazgo. También se pregunta si los autores reconocieron las limitaciones de su diseño y si realizaron análisis de sensibilidad para explorar la robustez de sus resultados.

5.5 EN LA PRÁCTICA CLÍNICA

Finalmente, la evaluación de validez y relevancia se traduce en una pregunta práctica: ¿los resultados de este estudio cambian la manera en que debo aconsejar, diagnosticar o tratar a mis pacientes? En el caso de nuestro ejemplo, si la cohorte muestra que el humo de leña triplica el riesgo de EPOC y que la fracción atribuible poblacional es alta, el médico debería incorporar este conocimiento a la educación de sus pacientes y apoyar programas de transición energética en comunidades rurales.

En conclusión, un estudio de cohorte solo será útil si es válido y relevante. La validez asegura que los resultados no son producto del azar ni de sesgos metodológicos; la relevancia garantiza que esos resultados tienen sentido en la práctica clínica y en la salud pública.

6. INTERPRETACIÓN ESTADÍSTICA

Se debe comprender dos ideas fundamentales:

- Riesgo (incidencia acumulada): es la proporción de personas que enferman durante un periodo.
- Tasa de incidencia (persona-tiempo): es la frecuencia a la que ocurren casos por unidad de tiempo en riesgo (casos por persona-años).

Ambas miden “aparición de enfermedad”, pero no son iguales: el riesgo ignora el detalle fino del tiempo (solo mira el período completo), mientras la tasa lo incorpora (si alguien se va antes, aporta menos tiempo).

También necesitaremos tres ideas más:

- Censura: alguien deja de ser observado antes de presentar el evento. No es “fallo”; es tiempo conocido sin evento y luego desconocido.
- Supervivencia libre de evento: probabilidad de no haber presentado el desenlace a lo largo del tiempo (se ve con Kaplan–Meier).
- Riesgo instantáneo (hazard): velocidad a la que ocurre el evento en un instante; comparar hazards da el hazard ratio (HR).

Tabla 2. Incidencia acumulada, tasa de incidencia y razón de tasas de EPOC según exposición al humo de leña durante 12 años de seguimiento

Exposición	N al inicio	Casos nuevos de EPOC	Persona-años	Incidencia acumulada (%)	Tasa de incidencia (por 100 persona-años)	Razón de tasas (IC95%)
Expuestos	600	132	5,28	22.0	2.5	3.0 (2.1–4.3)
No expuestos	600	48	5,76	8.0	0.83	Ref.

Fuente: Elaboración propia

Para fijar ideas, usaremos la tabla 2 asociado al caso, a partir de esta podremos calcular las medidas fundamentales.

6.2 INCIDENCIA ACUMULADA (RIESGO)

La incidencia acumulada es la medida fundamental en los estudios de cohorte. Se define

como la proporción de individuos inicialmente libres de enfermedad que desarrollan el desenlace durante un período determinado de seguimiento.

A diferencia de la prevalencia, que mide cuántos tienen la enfermedad en un momento dado, la incidencia acumulada refleja el riesgo de enfermar en el futuro. Por eso, es la medida más intuitiva para hablar de probabilidad y para establecer relaciones temporales entre una exposición y un desenlace.

Su fórmula básica es:

$$\text{Incidencia acumulada} = \frac{\text{Número de casos nuevos en el período}}{\text{Número de personas en riesgo al inicio}}$$

Sus características principales son:

- Período definido: siempre se calcula en relación a un lapso concreto (un año, cinco años, diez años, etc.).
- Población en riesgo: el denominador debe estar libre del desenlace al inicio.
- Medida de riesgo absoluto: expresa directamente la probabilidad de que un individuo desarrolle la enfermedad en el tiempo estudiado.
- Fácil comunicación: más comprensibles, responde a la pregunta: “De cada 100 personas, ¿cuántas enfermarán en un tiempo dado?”.

La principal limitación de la incidencia acumulada es que no toma en cuenta las diferencias en el tiempo de seguimiento individual. Se asume que todos los participantes han sido observados durante todo el periodo, aunque en la realidad muchos pueden haberse perdido en el seguimiento, haber cambiado de exposición o haber fallecido por otras causas. En estos casos, la incidencia acumulada puede dar una idea parcial de la realidad. Cuando estas variaciones son importantes, es más recomendable utilizar la tasa de incidencia, que incorpora el tiempo exacto de observación de cada persona.

En nuestro caso clínico, siguiendo la tabla 2:

- La columna N al inicio indica el número de personas libres de enfermedad al comenzar el seguimiento; este es el denominador de la incidencia acumulada.
- La columna de casos nuevos representa el numerador tanto de la incidencia acumulada como de la tasa de incidencia.
- La incidencia acumulada expresa qué proporción de la cohorte enfermó al final de los 12 años
- Incidencia acumulada en expuestos = $132 \div 600 = 0.22 \rightarrow 22\%$.
- Incidencia acumulada en no expuestos = $48 \div 600 = 0.08 \rightarrow 8\%$.
- Interpretación clínica: durante 12 años de seguimiento, 22 de cada 100 personas expuestas al humo de leña desarrollaron EPOC, mientras que entre quienes no estaban expuestos solo 8 de cada 100 enfermaron.

Este resultado ya permite al lector crítico detectar una diferencia clara entre los grupos. Pero aún no dice nada sobre cuán grande es esa diferencia en términos relativos o absolutos, ni sobre su velocidad en el tiempo. Para eso necesitamos avanzar hacia otras medidas.

6.3 TASA DE INCIDENCIA (PERSONA-TIEMPO) O DENSIDAD DE INCIDENCIA

La tasa de incidencia, también llamada densidad de incidencia, mide la frecuencia con la que ocurren nuevos casos de una enfermedad en relación al tiempo total de observación aportado por la población en riesgo. Su fórmula es:

$$\text{Tasa de incidencia} = \frac{\text{Número de casos nuevos}}{\text{Suma del tiempo en riesgo de cada persona}}$$

A diferencia de la incidencia acumulada, que trata a todos los participantes como si hubieran sido seguidos exactamente el mismo periodo, la tasa de incidencia reconoce que en los estudios longitudinales no todos permanecen en observación el mismo tiempo. Algunas personas pueden abandonar el estudio, cambiar de exposición o morir por otras causas antes de presentar el desenlace. La tasa ajusta el denominador sumando únicamente el tiempo efectivamente observado.

El denominador, entonces, no es “número de personas”, sino persona-tiempo: por ejemplo, 100 personas seguidas en 1 año equivalen a 100 persona-años. Si una persona se pierde al cabo de 6 meses, aporta solo 0.5 persona-año.

Sus características principales son:

- Mide velocidad de ocurrencia: refleja qué tan rápido aparecen los eventos en la población.
- Incorpora pérdidas y censura: es especialmente útil cuando hay tiempos de seguimiento heterogéneos.
- Permite comparaciones válidas: cuando los grupos no tuvieron exactamente el mismo tiempo de observación, la tasa es más justa que el riesgo acumulado.
- Expresada en unidades de tiempo: casos por 100 persona-años, casos por 1,000 persona-meses, etc.

Su principal limitación es que es menos intuitiva que la incidencia acumulada. Decir que hay “2.5 casos por 100 persona-años” no se comprende tan fácilmente como “22 de cada 100 enfermaron en 12 años”. Por eso, en la práctica clínica suele complementarse ambas medidas: el riesgo absoluto (incidencia acumulada) comunica probabilidad, mientras que la tasa de incidencia informa sobre velocidad.

En nuestro caso clínico retomando la tabla 2:

- El dato de persona-años refleja la suma del tiempo que cada participante estuvo realmente en riesgo, es decir, el período de observación efectivo aportado por cada individuo hasta enfermar, abandonar el estudio o concluir el seguimiento.
- La tasa de incidencia nos muestra la frecuencia de nuevos casos ajustada por tiempo de seguimiento.
- La razón de tasas nos muestra cuántas veces más rápido ocurre la enfermedad en un grupo respecto al otro, considerando el tiempo de observación aportado por cada participante.

La etiqueta “**Tasa de incidencia (por 100 persona-años)**” indica cuántos casos nuevos aparecen, en promedio, cuando sumas 100 años de observación entre las personas del grupo (por ejemplo: 100 personas seguidas 1 año, o 50 personas seguidas 2 años, etc.). No es una probabilidad, sino una velocidad de aparición del evento.

Analogía clínica para reforzar la comprensión: Piensa en la tasa de incidencia como en la velocidad de un auto:

- “60 km/h” no significa que el auto recorra 60 km cada hora siempre.
- Significa que, si mantiene esa velocidad promedio, tras 1 hora recorrerá unos 60 km; tras 2 horas, 120 km; y así sucesivamente.
- De la misma forma, “2.5 por 100 persona-años” es la velocidad promedio de aparición de casos en la cohorte. Nos dice que aparecerán 2.5 casos por cada 100 personas seguidas en 1 año. Bajo esto, nuestros resultados se interpretarían de esta forma:
- **Expuestos:** 2.5 por 100 persona-años. Si se acumulan 100 años de observación en personas expuestas, se esperarían en promedio entre 2 y 3 casos nuevos de EPOC en ese período total, independientemente de cómo se distribuyan esos años entre los individuos.
- **No expuestos:** 0.83 por 100 persona-años. Si se acumulan 100 años de observación en personas no expuestas, se esperarían en promedio ≈ 0.8 casos (menos de 1).
- **Comparación directa (razón de tasas).** Dividir 2.5 entre 0.83 da ≈ 3.0 . Eso significa que la enfermedad aparece tres veces más rápida entre los expuestos que entre los no expuestos, teniendo en cuenta el tiempo efectivo que cada persona estuvo en riesgo.

Por qué es útil esta lectura: La tasa integra pérdidas de seguimiento y diferentes duraciones de observación (cada quien aporta solo el tiempo real que estuvo “en riesgo”). Por eso es una comparación más justa cuando los tiempos no son iguales entre grupos.

Precauciones al interpretar:

- Una tasa no es una probabilidad acumulada: no debes “multiplicarla por años” para obtener riesgo a 12 años; para eso se usan curvas de Kaplan–Meier o modelos de supervivencia.
- En un artículo completo, busca los intervalos de confianza de estas tasas o de la razón de tasas (idealmente estimados con métodos de Poisson u otros), y confirma cómo se construyó el denominador de persona-años (quién se perdió, cuándo enfermó, si hubo cambios de exposición).
- Recuerda que esta comparación aún es cruda; más adelante debes ver si el estudio ajustó por confusores clave (p. ej., tabaquismo, edad, sexo, ocupación) y si la conclusión se mantiene.

Finalmente concluyendo la lectura crítica correcta de la tabla 2 es:

- El riesgo absoluto de EPOC a 12 años fue del 22% en los expuestos frente al 8% en los no expuestos.

- La tasa de incidencia muestra con claridad que la enfermedad aparece con mayor velocidad en quienes usan leña.

- Ambas medidas —riesgo e incidencia— son consistentes: la exposición al humo de leña prácticamente triplica la carga de enfermedad.

No obstante, el lector crítico debe ir más allá de la lectura numérica. Una pregunta obligada es si los autores ajustaron o no por posibles confusores. Esta información no suele aparecer en la tabla básica, sino en los análisis multivariantes y en la discusión del artículo. Aun así, la tabla ya ofrece una orientación clara sobre el efecto crudo de la exposición.

Finalmente, conviene destacar que la tasa de incidencia es particularmente útil en estudios con tiempos de observación variables o con pérdidas durante el seguimiento, porque ofrece una comparación más justa que la incidencia acumulada. Sin embargo, para comunicar estos hallazgos a pacientes o a la comunidad, sigue siendo preferible traducirlos a probabilidades acumuladas o a cifras absolutas, que resultan más fáciles de comprender y de aplicar en la práctica clínica.

6.4 RIESGO RELATIVO (RR)

En estudios de cohorte, una de las medidas más utilizadas para comparar la ocurrencia de un evento entre expuestos y no expuestos es el RR, también llamado Razón de Riesgos.

El RR expresa cuántas veces más (o menos) probable es que ocurra el evento en los expuestos en comparación con los no expuestos. Se calcula dividiendo la incidencia acumulada del grupo expuesto entre la incidencia acumulada del grupo no expuesto:

$$RR = \frac{\text{Riesgo en expuestos}}{\text{Riesgo en no expuestos}}$$

El valor resultante se interpreta de la siguiente manera:

- RR = 1: no existe diferencia en el riesgo entre expuestos y no expuestos (la exposición no parece asociarse con la enfermedad).
- RR > 1: la exposición se asocia con un mayor riesgo de enfermedad. Por ejemplo, un RR = 2 indica que los expuestos tienen el doble de probabilidad de enfermar que los no expuestos.
- RR < 1: la exposición parece tener un efecto protector. Por ejemplo, un RR = 0.5 significa que los expuestos tienen la mitad del riesgo de enfermar que los no expuestos.

El RR se expresa casi siempre acompañado de un intervalo de confianza (IC95%), que nos indica la precisión de la estimación y su p valor que indica si la asociación podría deberse al azar. Es especialmente útil porque se basa en proporciones absolutas (incidencia acumulada) y permite comunicar el riesgo de forma sencilla:

- En investigación clínica, ayuda a cuantificar la fuerza de la asociación entre una exposición y un desenlace.
- En salud pública, permite estimar el impacto que tendría eliminar una exposición nociva o promover una intervención protectora.
- En práctica clínica, traduce la evidencia en probabilidades fáciles de explicar a pacientes y comunidades.

Sus principales ventajas radican en que es intuitivo: responde a la pregunta “¿cuántas veces más riesgo tienen los expuestos?”. Es fácil de entender y comunicar. Sin embargo, sus limitaciones son: no tiene en cuenta el tiempo de seguimiento (a diferencia de la tasa de incidencia). En presencia de pérdidas de seguimiento o censura importante, puede subestimar o sobrestimar el efecto. Por todo ello siempre debe interpretarse junto con medidas absolutas (como la diferencia de riesgos) para tener una visión completa del impacto real.

En nuestro caso, se presenta la siguiente tabla:

Tabla 3. Riesgo acumulado y riesgo relativo (RR) de EPOC según exposición al humo de leña durante 12 años de seguimiento

Exposición	N al inicio	Casos nuevos	Incidencia acumulada (%)	Riesgo Relativo (IC95%)
Expuestos	600	132	22.0	2.75 (2.0–3.7)
No expuestos	600	48	8.0	Ref.

Nota: El intervalo de confianza (IC95%) es ilustrativo; en un artículo real se obtiene con métodos estadísticos (generalmente regresión binomial o modelos de Poisson con varianza robusta).

Para su Interpretación: Un RR de 2.75 significa que, en el periodo de 12 años, los expuestos al humo de leña tuvieron casi tres veces más riesgo de desarrollar EPOC que quienes no estaban expuestos, además su IC95% (2.0–3.7) no incluye el 1, por lo que podríamos concluir que la

asociación no solo es fuerte, sino también estadísticamente significativa.

Como lector crítico, este resultado nos orienta de forma clara hacia la existencia de un riesgo aumentado, pero siempre debemos revisar si el análisis fue ajustado por factores de confusión (como tabaquismo, edad, nivel socioeconómico), ya que podrían explicar parte del efecto observado.

6.5 DIFERENCIA ABSOLUTA DE RIESGOS (RIESGO ATRIBUIBLE ABSOLUTO, RAR)

La Diferencia Absoluta de Riesgos (DAR), también llamada Riesgo Atribuible Absoluto (RAR), mide el exceso de riesgo que se observa en el grupo expuesto en comparación con el grupo no expuesto.

Se calcula restando el riesgo acumulado de los no expuestos al de los expuestos:

$$RAR = \text{Riesgo en expuestos} - \text{Riesgo en no expuestos}$$

Su interpretación general es de la siguiente forma:

- Si $RAR > 0$, la exposición añade riesgo: existe un exceso de casos atribuibles a la exposición.
- Si $RAR < 0$, la exposición reduce riesgo: refleja un efecto protector.
- Si $RAR = 0$, no hay diferencia entre los grupos.

Esta medida es muy valiosa porque expresa el impacto absoluto de la exposición. Mientras que el RR indica la “fuerza de la asociación” (por ejemplo, triplica el riesgo), la RAR muestra cuántos casos adicionales ocurren realmente por efecto de la exposición.

La RAR también se utiliza para calcular el Número Necesario a Tratar (NNT) o el Número Necesario para Dañar (NNH), conceptos muy útiles en la práctica clínica:

$$NNT/NNH = \frac{1}{RAR}$$

Estos valores responden a preguntas como: ¿cuántas personas debo tratar (o cuántas estarán expuestas) para que aparezca o se evite un caso adicional de la enfermedad?

Las ventajas de la RAR son:

- Muestra el impacto absoluto de la exposición (número de casos adicionales).
- Es muy útil para decisiones en salud pública, porque permite dimensionar el problema a nivel poblacional.
- Facilita la comunicación con pacientes: es más comprensible decir “14 casos adicionales por cada 100 adultos” que “un riesgo 2.75 veces mayor”.

Entre sus limitaciones:

- Depende de la magnitud del riesgo basal: el mismo RR puede dar diferentes RAR en poblaciones con distinta incidencia.
- Puede variar según el tiempo de seguimiento; si la cohorte fuera de 20 años, la RAR probablemente sería mayor.
- En presencia de censura importante o pérdida de seguimiento, puede estar sesgada si no se ajusta adecuadamente.

Para que lo apliquemos en nuestra cohorte de ejemplo, se presenta la siguiente tabla:

Tabla 4. Riesgo acumulado, diferencia absoluta de riesgos (RAR) y número necesario a dañar (NNH) de EPOC según exposición al humo de leña durante 12 años de seguimiento

Exposición	N al inicio	Casos nuevos	Incidencia acumulada (%)	RAR (%)	NNH
Expuestos	600	132	22.0	+14.0	7
No expuestos	600	48	8.0	Ref.	–

Fuente: Elaboración propia

Para su interpretación de RAR: 14% significa que, en 12 años de seguimiento, 14 de cada 100 adultos expuestos al humo de leña desarrollaron EPOC como exceso atribuible a la exposición, en comparación con los no expuestos.

Este dato se puede traducir a la práctica clínica de la siguiente forma:

- Por cada 100 adultos expuestos al humo de leña, habrá 14 casos de EPOC que no se hubieran

producido si hubieran tenido condiciones de no exposición.

- Para su interpretación de NNH: por cada 7 adultos expuestos al humo de leña, 1 desarrollará EPOC atribuible a la exposición en 12 años.

Ahora si Imaginamos un ensayo clínico hipotético en el que se implementa ventilación mejorada en cocinas para reducir la exposición al humo de leña. Tras 12 años nos reporta un Numero Necesario a tratar (NNT) de 8, esto se interpreta que por cada 8 mujeres que reciben ventilación mejorada, se previene 1 caso de EPOC en 12 años.

Reflexión para el lector crítico:

- El NNH comunica el daño atribuible a una exposición en términos absolutos y prácticos.
- El NNT comunica el beneficio de una intervención de la misma manera.
- Ambas medidas son muy poderosas porque permiten a médicos, pacientes y autoridades traducir los números de un artículo a decisiones concretas.

6.6 HAZARD Y HAZARD RATIO (HR)

HAZARD

El hazard (también llamado tasa de riesgo instantánea) es la probabilidad de que ocurra el evento de interés en un instante de tiempo determinado, entre quienes todavía no lo han presentado.

- A diferencia de la incidencia acumulada (que mira la proporción de personas que enferman al final de un periodo) o de la tasa de incidencia (que mide la velocidad promedio de aparición de casos por unidad de tiempo), el hazard se enfoca en el momento exacto en que ocurre el evento.
- Puede pensarse como la “intensidad instantánea del riesgo”: ¿qué tan probable es que una persona que hasta ahora está sana enferme justo en el siguiente momento?

Matemáticamente, el hazard se expresa como una tasa por unidad de tiempo, pero en la práctica casi nunca lo calculamos manualmente, sino que se estima a través de modelos de supervivencia, como el modelo de Cox.

HAZARD RATIO (HR)

El HR compara el hazard entre dos grupos (expuestos vs. no expuestos) y se lo interpreta de esta forma:

- HR = 1: No hay diferencia en el riesgo instantáneo entre grupos.
- HR > 1: El evento ocurre más rápido (o con mayor intensidad) en el grupo expuesto.
- HR < 1: El evento ocurre más lentamente en el grupo expuesto (efecto protector).

Por todo esto, debemos diferenciarlo con RR y tasa de incidencia

- RR (Riesgo Relativo): compara proporciones acumuladas de enfermos al final de un seguimiento.
- Tasa de incidencia: compara la velocidad promedio de aparición de casos.
- HR: compara la velocidad *instantánea* de ocurrencia a lo largo del tiempo.

Por eso, el HR es más adecuado cuando:

- El tiempo de seguimiento es largo.
- Hay censura (pérdida de seguimiento, muertes por otras causas).
- El riesgo no se distribuye uniformemente en el tiempo (por ejemplo, mayor riesgo en los primeros años).

Tabla 5. Análisis de supervivencia mediante modelo de Cox para riesgo de EPOC según exposición al humo de leña durante 12 años de seguimiento

Variable	HR	IC95%	p-valor
Exposición al humo de leña	2.8	2.0 – 3.9	<0.001
Edad (>50 años)	1.5	1.1 – 2.1	0.02
Tabaquismo	2.2	1.6 – 3.1	<0.001

Fuente: Elaboración propia

En nuestra cohorte: Supongamos que al aplicar un análisis de supervivencia con modelo de

Cox encontramos los resultados de la tabla 5, su interpretación clínica será:

- En cada momento del seguimiento, los pacientes expuestos al humo de leña tienen 2.8 veces más probabilidad de desarrollar EPOC que los no expuestos.
- A diferencia del RR (que solo nos dice que al final del seguimiento las expuestas tienen casi 3 veces más riesgo), el HR nos informa sobre la dinámica temporal: no solo enferman más, sino que además lo hacen antes.

Reflexión crítica para el lector

- El HR es una de las medidas más poderosas porque integra la información del tiempo y permite análisis ajustados por múltiples variables.
- Sin embargo, su interpretación requiere entender el supuesto del modelo de Cox y Kaplan – Meier.
- En lectura crítica, si un artículo reporta un HR, conviene verificar que los autores hayan evaluado la proporcionalidad de riesgos; de lo contrario, la estimación podría ser engañosa.

6.7 KAPLAN–MEIER Y REGRESIÓN DE COX EN ESTUDIOS DE COHORTE

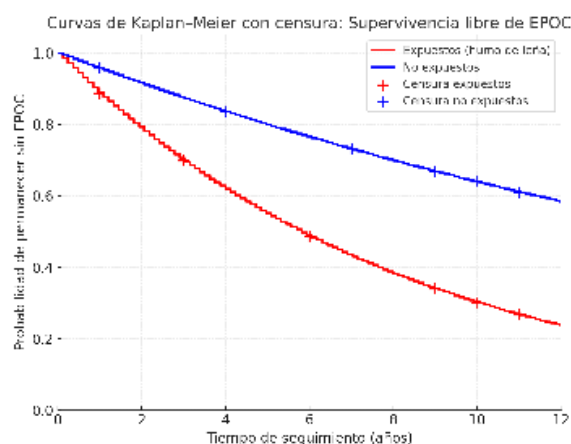
En los estudios de cohorte, cuando el interés se centra en analizar no solo cuántos participantes enferman, sino cuándo lo hacen, entran en juego dos herramientas estadísticas fundamentales: las curvas de Kaplan–Meier y el modelo de regresión de Cox. Ambas se utilizan de manera complementaria y permiten al lector crítico comprender la dinámica temporal del evento.

Las curvas de Kaplan–Meier son un método descriptivo que estima la probabilidad de permanecer libre del evento a lo largo del tiempo. Cada vez que ocurre un caso, la curva desciende en forma de escalón, mostrando el momento en que alguien desarrolla la enfermedad. Si un paciente se pierde en el seguimiento o termina el estudio sin enfermar, se registra como “censura” y se marca en la curva, generalmente con un pequeño símbolo. Esta técnica tiene la virtud de aprovechar al máximo la información del tiempo hasta el evento y de incluir a quienes no completan el seguimiento, algo muy común en cohortes largas. Cuando se comparan dos grupos, por ejemplo, expuestos y no expuestos, se construyen dos curvas y la diferencia global entre ellas se evalúa mediante la prueba de log-rank, que responde a la pregunta: ¿la supervivencia libre de enfermedad es distinta entre ambos grupos a lo largo del tiempo?

La regresión de Cox, también llamada modelo de riesgos proporcionales, va un paso más allá. No se limita a mostrar curvas descriptivas, sino que estima de manera cuantitativa la magnitud de la diferencia entre los grupos. Su resultado central es el HR, que indica cuántas veces más rápido ocurre el evento en un grupo comparado con otro en cualquier momento del seguimiento. Además, permite ajustar por múltiples variables de confusión (como edad, tabaquismo o nivel socioeconómico), lo cual le otorga gran poder analítico. Sin embargo, su validez depende de un supuesto fundamental: la proporcionalidad de riesgos, es decir, que la relación entre los hazards de los grupos se mantenga relativamente constante a lo largo del tiempo. Si este supuesto se viola, la interpretación del HR puede ser engañosa.

Veamos cómo estas herramientas se aplican a nuestro caso:

Figura 1. Curvas de Kaplan–Meier con censura para supervivencia libre de EPOC según exposición al humo de leña durante 12 años de seguimiento



Al observar una curva de Kaplan–Meier, lo primero es identificar el **eje X**, que representa el tiempo de seguimiento (en nuestro caso, los años del estudio), y el **eje Y**, que indica la probabilidad de permanecer libre del evento (EPOC). La curva comienza siempre en 1.0, es decir, el 100% de los participantes sin enfermedad al inicio.

Cada vez que un participante desarrolla el evento, la curva desciende en un escalón. La magnitud del descenso depende de cuántos enferman en ese momento en relación con el número de personas que aún estaban en riesgo. Por ejemplo, un gran salto hacia abajo indica que varios individuos presentaron EPOC en ese intervalo.

Los símbolos de censura, representados aquí como cruces, marcan el momento en que un individuo deja de aportar información al estudio sin haber desarrollado el evento. Esto puede deberse a que abandonó el seguimiento, falleció por otra causa o simplemente terminó el periodo de observación sin enfermar. El lector debe notar que la presencia de censuras significa que esos datos se interrumpieron en ese instante.

La comparación entre curvas de diferentes grupos es clave. En nuestro caso, la curva de los pacientes expuestos al humo de leña desciende más rápidamente y se separa visiblemente de la curva de las no expuestos desde los primeros años. Esto significa que la probabilidad de permanecer libre de EPOC es consistentemente menor en los expuestos. El espacio entre curvas se mantiene e incluso se amplía hacia el final, lo que indica un efecto sostenido de la exposición.

El análisis estadístico complementario suele ser el log-rank test, que evalúa si la diferencia observada entre curvas es estadísticamente significativa. Si el valor *p* es menor de 0.05, podemos afirmar que las curvas difieren más de lo que cabría esperar por azar.

Como lector crítico, es importante preguntarse: ¿cuántas censuras hay y en qué momentos ocurren? Si la mayoría de los participantes fueron censurados en los últimos años, la curva en ese tramo tiene menos precisión y las conclusiones sobre el largo plazo son menos seguras. También conviene revisar si los autores reportan un HR ajustado por regresión de Cox, como mencionaremos a continuación.

Tabla 6. Modelo de Cox de riesgos proporcionales para tiempo a EPOC durante 12 años de seguimiento

Variable	Categoría (referencia)	HR ajustado	IC95%	p
Exposición al humo de leña	Sí (vs No)	2.70	2.00–3.65	<0.001
Edad	≥50 años (vs <50)	1.55	1.12–2.14	0.008
Tabaquismo	Actual/exfumador (vs nunca)	2.15	1.54–3.00	<0.001
Exposición ocupacional a polvos	Sí (vs No)	1.38	1.01–1.90	0.045
Sexo	Mujer (vs Varón)	1.20	0.86–1.68	0.27
Nivel socioeconómico	Bajo (vs Medio/alto)	1.22	0.90–1.66	0.19

Notas del modelo: *n* = 1,200; eventos = 180; mediana de seguimiento ≈ 11 años; prueba de log-rank global entre expuestos vs no expuestos *p* < 0.001. Supuesto de proporcionalidad de riesgos evaluado con residuos de Schoenfeld (global *p* = 0.21), sin violaciones relevantes.

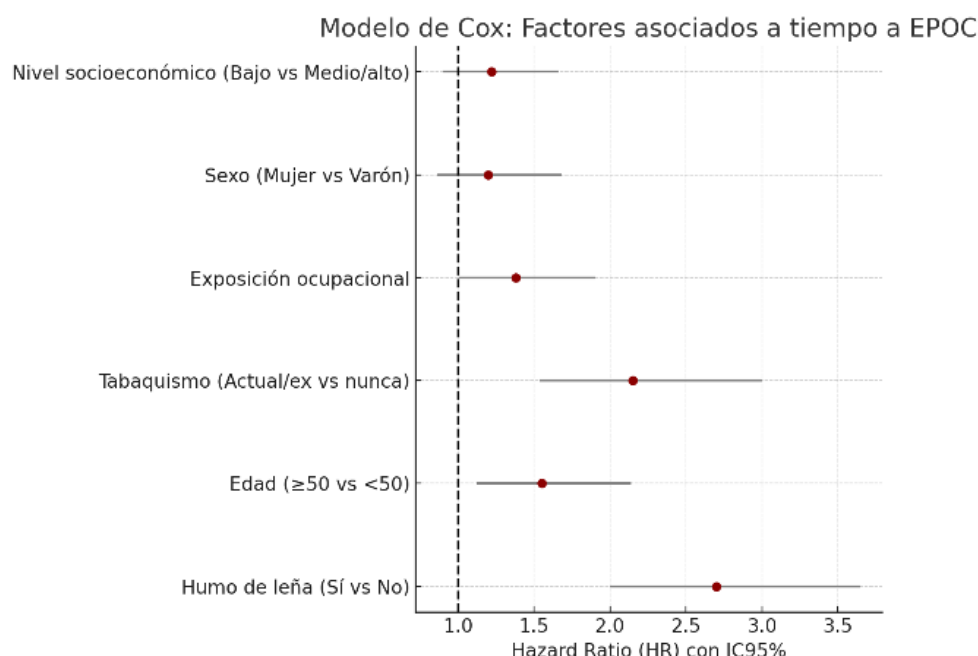
La interpretación comienza por la fila de la exposición principal. Un HR ajustado de 2.70 indica que, en cualquier momento del seguimiento, las personas expuestas al humo de leña presentan una velocidad de ocurrencia de EPOC aproximadamente 2.7 veces mayor que las no expuestas, tras controlar por edad, tabaquismo, exposición ocupacional, sexo y nivel socioeconómico. Dicho de forma clínica, no solo hay más casos entre los expuestos; además, los casos aparecen antes. Esta conclusión no depende de un único momento del estudio, sino que resume todo el proceso temporal, que ya pudimos visualizar con las curvas de Kaplan–Meier.

La edad muestra un efecto independiente: los participantes de 50 años o más tienen una aceleración del riesgo del 1.55 veces con respecto a los menores de 50, coherente con la historia natural de EPOC. El tabaquismo, como era esperable, duplica el hazard de EPOC; que el efecto del humo de leña se mantenga elevado aun ajustando por tabaco refuerza la plausibilidad etiológica de la exposición ambiental del hogar. La exposición ocupacional a polvos añade un incremento modesto pero significativo del hazard, lo que sugiere efectos aditivos en vías respiratorias cuando confluyen fuentes de partículas. Sexo y nivel socioeconómico, en este ejemplo, no alcanzan significación estadística; eso no implica que carezcan de relevancia clínica (sus intervalos de confianza IC cuentan con el 1), sino que, con estos datos y covariables, su contribución independiente al hazard no se demostró con precisión suficiente.

Es crucial recordar que un HR no es un riesgo relativo acumulado: cuantifica la relación entre tasas instantáneas de ocurrencia. Por ello, su interpretación siempre debe ir acompañada de la verificación del supuesto de proporcionalidad de riesgos. Aquí, la comprobación con residuos de Schoenfeld no sugiere violaciones; de haberlas, correspondería informar modelos

estratificados, términos dependientes del tiempo o alternativas como tiempos medios de supervivencia restringidos (RMST). También es buena práctica que el artículo especifique cómo se codificaron las variables (por ejemplo, tabaquismo agrupado como nunca vs actual/ex), cómo se manejaron empates en el tiempo (método de Efron o Breslow) y si hubo estratificación por centro o comunidad para controlar confusión no medida a nivel contextual.

Figura 2. Modelo de Cox: factores asociados al tiempo a EPOC con estimaciones de hazard ratio (HR) e intervalos de confianza al 95% durante 12 años de seguimiento



El gráfico de bosque o *Forest plot* es una representación visual de los resultados de la regresión de Cox. En el eje vertical aparecen las variables analizadas, mientras que en el eje horizontal se ubica la escala de los HR. Cada punto corresponde al valor puntual del HR ajustado, y las barras horizontales representan su IC95%.

La línea vertical en 1.0 marca el valor de “no efecto”. Si el intervalo de confianza cruza esa línea, significa que el efecto estimado podría ser nulo y, por lo tanto, el resultado no es estadísticamente significativo.

En nuestro ejemplo, la variable exposición al humo de leña presenta un HR de 2.70, con IC95% entre 2.00 y 3.65, como la barra de confianza no toca la línea de 1, este hallazgo es estadísticamente sólido. La variable sexo (HR 1.20, IC95% 0.86–1.68) y nivel socioeconómico bajo (HR 1.22, IC95% 0.90–1.66) no alcanzan significación, pues sus intervalos cruzan la línea de 1. Eso no implica que sean irrelevantes, sino que en este estudio los datos disponibles no permiten concluir con certeza sobre su efecto independiente.

Lo más importante para el lector crítico es entender que este gráfico resume en una sola mirada la dirección, magnitud y precisión de cada asociación. Mientras la tabla muestra los números, el Forest plot permite comparar rápidamente cuáles variables son relevantes, cuáles tienen gran incertidumbre y cuáles no muestran un efecto claro.

Finalmente, conviene recordar que estos HR son ajustados por las demás variables incluidas en el modelo. Esto significa que la asociación entre humo de leña y EPOC que se observa aquí no se debe a la edad, al tabaquismo u otros factores considerados, sino que refleja un efecto independiente de la exposición.

6.8 RAZÓN DE TASAS VS. RIESGO RELATIVO: ¿CUÁL USO?

Tabla 7. Comparación entre riesgo relativo (RR) y razón de tasas de incidencia (IRR) en estudios de cohorte

Medida	¿Qué compara?	Ventajas	Limitaciones	Ejemplo aplicado al caso clínico
Riesgo Relativo (RR)	Probabilidad acumulada de enfermar en un periodo fijo (incidencia acumulada)	- Fácil de entender y comunicar a pacientes y responsables de salud. - Se traduce directamente en cifras absolutas y en medidas como NNT/NNH. - Útil en cohortes cerradas con seguimiento uniforme.	- No incorpora diferencias en el tiempo de seguimiento. - Sesgado si hay muchas pérdidas o censuras. - Menos útil si la población en riesgo varía mucho en el tiempo.	Tras 12 años: 22% expuestas vs 8% no expuestas → RR ≈ 2.75
Razón de Tasas de Incidencia (IRR)	Velocidad de aparición de nuevos casos, ajustada por persona-tiempo (persona-años)	- Más precisa cuando los seguimientos son desiguales o hay censura. - Permite comparar "la rapidez" del evento. - Útil en cohortes dinámicas o con tiempos de observación variables.	- Menos intuitiva para pacientes/comunidad. - Requiere cálculo detallado de persona-tiempo. - No siempre se reporta en artículos clínicos.	Tasa en expuestas: 2.5/100 persona-años vs 0.83/100 en no expuestas → IRR ≈ 3.0

Fuente: Elaboración propia

Aunque los valores son parecidos, no siempre ocurre así. En estudios con alta censura o tiempos de observación desiguales, las diferencias entre RR e IRR pueden ser notables. Como lector crítico, conviene identificar qué medida se reporta, si corresponde al diseño y, sobre todo, cuál es más útil para interpretar la aplicabilidad clínica.

En la práctica clínica, para conversar con un paciente o la comunidad, el RR (o medidas derivadas como el NNT/NNH) suele ser más comprensible. En cambio, para evaluar políticas de salud, comparaciones internacionales o estudios con tiempos de exposición muy variables, la IRR aporta una visión más ajustada.

6.9 ATRIBUCIÓN DE RIESGO: FAE Y FAP

Además de estimar cuánto aumenta el riesgo con una exposición, podemos calcular qué proporción de casos se debe realmente a esa exposición. Esto se conoce como atribución del riesgo, y es una herramienta poderosa tanto para la práctica clínica como para la salud pública. Existen dos formas principales de expresarlo:

a) Fracción Atribuible en Expuestos (FAE): indica qué fracción de la enfermedad observada en los expuestos puede atribuirse directamente a la exposición, suponiendo causalidad. En otras palabras, responde: *"¿Si elimináramos la exposición, qué porcentaje de casos en el grupo expuesto no ocurrirían?"*

Se calcula con la fórmula:

$$FAE = \frac{RR - 1}{RR} \times 100$$

b) Fracción Atribuible Poblacional (FAP): amplía la mirada a toda la población, considerando tanto expuestos como no expuestos. Responde: *"¿Si elimináramos la exposición en la población, qué proporción total de casos se evitarían?"*

Se calcula con:

$$FAP = \frac{Pe \times (RR - 1)}{Pe \times (RR - 1) + 1} \times 100$$

Donde Pe es la proporción de población expuesta.

La FAP depende no solo de la fuerza de asociación (RR), sino también de cuán frecuente es la exposición en la comunidad. Así, una exposición moderadamente riesgosa pero muy frecuente puede tener mayor impacto poblacional que una muy peligrosa pero rara.

En nuestro caso se presenta la tabla de esta forma:

Tabla 8. Riesgo relativo, fracción atribuible en expuestos (FAE) y fracción atribuible poblacional (FAP) para EPOC según exposición al humo de leña

Medida	Estimación	IC95%
Riesgo relativo (RR)	2.75	1.90 – 3.98
Fracción atribuible en expuestos (FAE)	64%	–
Fracción atribuible poblacional (FAP)	47%	–

Fuente: Elaboración propia

Para interpretar la tabla: Durante el seguimiento de 12 años la FAE fue del 64%, indicando que casi dos tercios de los casos de EPOC entre los adultos expuestos podrían atribuirse al humo de leña.

A nivel poblacional, considerando que el 50% de la cohorte estuvo expuesta, la FAP fue del 47%. Esto implica que la eliminación de la exposición podría prevenir aproximadamente la mitad de los casos de EPOC en esta población.

Nota: En la mayoría de artículos, la FAE y la FAP rara vez aparecen con todos los detalles; normalmente se presentan directamente los porcentajes, a veces en tablas secundarias. El lector crítico debe saber de dónde vienen estos números (RR y prevalencia de la exposición) para verificar si están correctamente estimados y si su interpretación es válida.

6.10 RIESGOS COMPETITIVOS: PORQUE IMPORTAN

En la práctica clínica los pacientes pueden experimentar otros eventos que impidan o modifiquen la probabilidad de que ocurra el desenlace principal. Estos son los llamados riesgos competitivos. Un riesgo competitivo es cualquier evento que:

- Previene la ocurrencia del desenlace principal (ejemplo: un paciente fallece por infarto antes de desarrollar EPOC).
- Modifica la probabilidad de que ocurra el desenlace (ejemplo: un paciente es diagnosticado de otra enfermedad pulmonar crónica y cambia de ambiente o de exposición, reduciendo la probabilidad de llegar a EPOC clínico).

En teoría estadística, los riesgos competitivos afectan la incidencia acumulada: si se ignoran, se asume que todos los pacientes “siguen en riesgo” hasta el final, lo que sobreestima la probabilidad de enfermar.

En nuestra cohorte de 1,200 pacientes:

- La incidencia acumulada de EPOC fue 22% en expuestos y 8% en no expuestos.
- Sin embargo, durante el seguimiento algunos pacientes fallecieron por otras causas (ej. cáncer, enfermedad cardiovascular) antes de desarrollar EPOC.
- Estas muertes actúan como riesgos competitivos: esos pacientes ya no podían enfermar de EPOC, por lo que el análisis que no los considere podría inflar la incidencia real atribuida a la exposición.
- En cohortes largas, con pacientes adultos o con comorbilidades, este problema es frecuente y relevante.

Como lector crítico de un artículo, hay que revisar si los autores reconocieron la existencia de riesgos competitivos y usaron métodos adecuados, como curvas de incidencia acumulada en presencia de riesgos competitivos o modelos de Fine & Gray, en lugar de limitarse a Kaplan–Meier o a incidencias acumuladas crudas.

Ignorar los riesgos competitivos puede llevar a:

- Sobreestimación del riesgo de EPOC.
- Distorsión del efecto real de la exposición.
- Poca aplicabilidad clínica, ya que en la práctica muchos pacientes no alcanzarán a desarrollar el evento de interés porque otro evento se interpuso antes.

En nuestro ejemplo: los expuestos al humo de leña siguen mostrando mayor riesgo de EPOC, pero si un 10% muere por enfermedad cardiovascular antes, el “impacto real” de la exposición es distinto al que se calcula ignorando esas muertes.

6.11 AJUSTE POR CONFUSIÓN: MEDIDAS CRUDAS VERSUS AJUSTADAS

Cuando se interpreta un estudio de cohorte, es fundamental diferenciar entre las medidas crudas (sin ajuste) y las medidas ajustadas (tras controlar por confusores).

Medidas crudas: son las primeras que reporta un estudio: riesgos relativos, tasas de incidencia o hazard ratios calculados sin considerar confusores. Estas medidas nos muestran el efecto “total” observado, pero pueden estar distorsionadas por variables externas.

Ejemplo con nuestros datos de tabla 3:

- RR crudo de EPOC en expuestos vs. no expuestos = 2.75 (IC95%: 2.0–3.7), indica casi 3 veces más riesgo de desarrollar EPOC.
- Sin embargo, todavía no sabemos si este efecto se debe exclusivamente al humo de leña o si parte de la asociación se explica por tabaquismo, edad, sexo u otros factores.

Medidas ajustadas: Para acercarse a la verdadera magnitud de la asociación, los investigadores usan modelos multivariantes (regresión de Cox, regresión de Poisson, regresión logística según el diseño). Estos modelos permiten ajustar por confusores y reportar un efecto “depurado”.

Ejemplo (hipotético, usando el modelo de Cox de nuestro caso en tabla 6):

- HR ajustado por edad, sexo y tabaquismo = 2.1 (IC95%: 2.0–3.65).
- Aquí vemos que el efecto del humo de leña sigue siendo significativo, pero menor que en la estimación cruda (de 2.75 baja a 2.10). Esto sugiere que parte del riesgo observado inicialmente estaba explicado por otros factores, en particular el tabaquismo.

Reflexión crítica: El lector crítico debe preguntarse siempre:

- ¿Los autores reportaron tanto medidas crudas como ajustadas?
- ¿Qué variables incluyeron en el ajuste? ¿Tienen base causal o fueron seleccionadas solo por significancia estadística?
- ¿Cambió mucho el estimador al ajustar? Un cambio importante indica confusión relevante.
- ¿Persistió la asociación tras el ajuste? Si se desvanece, es posible que el efecto crudo fuera explicado casi por completo por confusores.

6.12 MODIFICADORES DEL EFECTO

En un estudio de cohorte, no siempre la asociación entre exposición y desenlace es uniforme en todos los subgrupos de pacientes. A veces, la magnitud del efecto cambia según otra variable. En esos casos hablamos de modificación del efecto, también conocida como interacción.

Un modificador de efecto es una variable que altera la fuerza o dirección de la asociación entre exposición y desenlace.

- Si el efecto de la exposición es mayor en un grupo que en otro, decimos que hay interacción.
- A diferencia de un confusor, el modificador de efecto no distorsiona la relación: forma parte de la realidad clínica y es un hallazgo importante por sí mismo.

En términos prácticos:

- Confusión: un “artefacto” que debe eliminarse (ajuste estadístico).
- Modificación de efecto: una característica real de cómo actúa la exposición, que debe describirse y reportarse.

Identificación en estudios: La modificación de efecto se explora de dos maneras:

- Análisis estratificado: se calcula el RR o HR en distintos subgrupos (por ejemplo, hombres vs. mujeres, fumadores vs. no fumadores). Si los estimadores difieren sustancialmente, puede existir interacción.
- Modelos con términos de interacción: en regresión de Cox o Poisson se incluyen variables producto (ej. exposición × sexo). Si resultan significativas, indican modificación de efecto.

Tabla 9. Riesgo relativo (RR) de EPOC asociado a exposición al humo de leña: análisis global y por estratos

Estratificación	RR crudo	IC95%	p de interacción
Análisis global	2.75	1.90 – 3.98	–
Sexo			
Hombres	1.90	1.30 – 2.80	0.02
Mujeres	3.40	2.30 – 5.00	
Tabaquismo			
No fumadores	3.20	2.10 – 4.80	0.01
Fumadores	1.50	1.00 – 2.20	

Fuente: Elaboración propia

En nuestro caso el análisis global mencionaba un RR crudo = 2.75, tras la estratificación por sexo, se evidencia una modificación de RR en hombres y mujeres, su interpretación: el efecto del humo de leña sobre el riesgo de EPOC es más fuerte en mujeres que en hombres. Esto refleja una modificación real del efecto: probablemente porque en contextos rurales son las mujeres quienes permanecen más tiempo en la cocina, acumulando mayor exposición. Así mismo se hizo la estratificación por tabaquismo, en esto se evidencia como interpretación: el humo de leña tiene un mayor impacto en no fumadores, ya que en fumadores parte del riesgo ya está “ocupado” por el tabaco.

Reflexión crítica: El lector crítico debe distinguir entre confusión y modificación de efecto.

- Si el efecto desaparece al ajustar, probablemente era confusión.
- Si el efecto se mantiene, pero varía en magnitud según un subgrupo, hablamos de modificación del efecto.

Este concepto es de enorme importancia clínica y en salud pública, porque permite identificar poblaciones particularmente vulnerables. En nuestro ejemplo, reconocer que las mujeres y los no fumadores tienen mayor susceptibilidad al humo de leña orienta las estrategias preventivas: programas de sustitución de combustibles y ventilación deben priorizar estos grupos.

Tabla 10. Comparación entre confusión y modificación del efecto en estudios de cohorte

Aspecto	Confusión	Modificación del efecto (interacción)
Definición	Distorsión aparente de la asociación exposición– desenlace causado por una tercera variable relacionada con ambos.	Cambio real en la magnitud o dirección de la asociación según los niveles de otra variable.
Naturaleza	Es un artefacto metodológico. No refleja la realidad.	Es un fenómeno real de la relación causal.
Efecto sobre la asociación	Hace que la asociación observada sea más fuerte, más débil o incluso cambie de dirección respecto a la asociación verdadera.	La asociación existe en todos los grupos, pero su magnitud varía (ej.: más fuerte en mujeres que en hombres).
Cómo se detecta	Comparando medidas crudas vs. ajustadas.	Estratificación (comparar riesgos relativos en subgrupos).
	Si cambian sustancialmente, había confusión.	Inclusión de términos de interacción en modelos multivariantes.
Qué hacer	Debe controlarse en el análisis (ajuste estadístico, restricción, emparejamiento).	Debe reportarse y destacarse, pues refleja una característica real de la asociación.
Ejemplo en el caso clínico	Tabaquismo: si no se ajusta, el efecto del humo de leña sobre EPOC se sobrestima. Tras ajustar, el HR baja de 2.75 a 2.10.	Sexo: el RR del humo de leña es 3.4 en mujeres y 1.9 en hombres. El efecto existe en ambos, pero es más fuerte en mujeres.

Fuente: Elaboración propia

6.13: P-VALOR, INTERVALOS DE CONFIANZA Y RELEVANCIA CLÍNICA

En la lectura crítica de un estudio de cohorte, los números crudos (riesgos, tasas, RR o HR) no bastan por sí solos. Es indispensable interpretar p-valores e intervalos de confianza (IC95%), siempre en conjunto con la relevancia clínica de los hallazgos.

EL P-VALOR

Indica la probabilidad de observar los datos (o algo más extremo) si en realidad no hubiera asociación entre la exposición y el desenlace.

- Si $p < 0.05$, se considera que la asociación observada es poco compatible con el azar y se denomina “estadísticamente significativa”.
- Si $p \geq 0.05$, no hay evidencia suficiente para descartar el azar
Sin embargo, el p-valor tiene limitaciones claras:
 - No mide la magnitud del efecto.
 - No refleja la relevancia clínica.
 - Puede ser “significativo” incluso con efectos muy pequeños si el tamaño muestral es grande, o “no significativo” pese a efectos importantes si la muestra es pequeña.

Por eso, un lector crítico nunca debe basar su interpretación solo en el p-valor.

INTERVALOS DE CONFIANZA (IC95%)

Proporciona un rango de valores dentro del cual es razonable que se encuentre la verdadera magnitud de la asociación, considerando la variabilidad muestral.

- Un IC95% estrecho indica precisión en la estimación.
- Un IC95% amplio indica incertidumbre.
- Si el IC95% incluye el valor nulo ($RR=1$ o $HR=1$), la asociación puede no ser estadísticamente significativa.

El intervalo de confianza es más informativo que el p-valor porque muestra simultáneamente:

- La dirección del efecto (mayor, menor o nula).
- La magnitud de la asociación.
- La precisión de la estimación.

RELEVANCIA CLÍNICA

Más allá de la significación estadística, el lector debe preguntarse si el hallazgo tiene impacto clínico o de salud pública.

- Un RR de 1.2 con $p < 0.001$ puede ser estadísticamente sólido, pero clínicamente irrelevante si el exceso de riesgo es mínimo.
- En cambio, un RR de 2.5 con $p = 0.07$ puede ser clínicamente relevante, aunque no alcance la significación estadística por falta de tamaño muestral.

La combinación de magnitud del efecto, intervalos de confianza y plausibilidad biológica es lo que determina si el hallazgo es útil para la práctica clínica.

Reflexión crítica: El mensaje para el lector crítico es claro:

- El p-valor solo responde a la pregunta “¿es probable que esto sea azar?”.
- El IC95% responde a “¿qué tan grande y preciso es el efecto?”.
- La relevancia clínica responde a “¿este hallazgo importa en la práctica?”.

Solo integrando los tres niveles se logra una interpretación madura de los resultados de una cohorte.

Tabla 11. Comparación entre p-valor, intervalo de confianza e importancia clínica en estudios de cohorte

Aspecto	p-valor	Intervalo de confianza (IC95%)	Relevancia clínica
Definición	Probabilidad de observar los datos (o más extremos) si la hipótesis nula fuera cierta.	Rango de valores plausibles para la verdadera magnitud del efecto, considerando la variabilidad muestral.	Valor práctico o impacto real del hallazgo en la salud de pacientes y comunidades.
Pregunta que responde	¿Es probable que el hallazgo sea producto del azar?	¿Cuál es la magnitud y precisión del efecto observado?	¿Importa este hallazgo para la práctica clínica o la salud pública?
Valor nulo	$p < 0.05$ suele considerarse significativo (umbral convencional).	El efecto no es significativo si el IC incluye el valor nulo ($RR=1$, $HR=1$, diferencia=0).	No depende de umbrales fijos; requiere juicio clínico, epidemiológico y social.
Fortalezas	Fácil de calcular e interpretar en términos de significación estadística.	Aporta dirección, magnitud y precisión. Más informativo que el p-valor solo.	Permite contextualizar si un efecto “estadísticamente significativo” realmente cambia la práctica o las políticas.
Limitaciones	No informa magnitud ni relevancia clínica.	Puede ser amplio si la muestra es pequeña, reflejando incertidumbre.	Subjetiva: depende del contexto, del desenlace y de la factibilidad de intervenir.
	Puede ser significativo con efectos mínimos en muestras grandes, o no significativo con efectos importantes en muestras pequeñas.	Requiere interpretación conjunta con la clínica.	
Ejemplo aplicado al caso	$p < 0.001$ → muy improbable que la asociación entre humo de leña y EPOC sea por el azar.	$RR = 2.75$ (IC95%: 1.90–3.98) → el riesgo verdadero se encuentra entre 1.9 y 4.0 veces mayor en expuestos.	Un incremento de casi 3 veces en el riesgo de EPOC es clínicamente muy relevante, pues justifica intervenciones preventivas.

Fuente: Elaboración propia

6.14 ERRORES FRECUENTES Y COMO EVITARLOS

Al leer un estudio de cohorte, incluso médicos con experiencia pueden caer en errores

comunes de interpretación. Estos errores no siempre son técnicos: a veces provienen de trasladar de forma automática los números a la práctica clínica sin analizar el contexto. Reconocerlos es clave para evitar conclusiones engañosas.

1. Traducir el RR como si fuera un riesgo absoluto

Uno de los errores más frecuentes. Por ejemplo, decir que *“el humo de leña aumenta en 275% el riesgo de EPOC”* puede sonar alarmante, pero sin indicar la incidencia absoluta se pierde perspectiva. En nuestro caso, el riesgo pasó de 8% en no expuestos a 22% en expuestos: la diferencia real es del 14%. Lo correcto es buscar ambos: medida relativa y absoluta.

2. Confiar solo en el p-valor

Esto lleva a catalogar hallazgos como “significativos” o “no significativos” sin considerar su magnitud clínica. Un RR de 2.5 con $p=0.06$ en una cohorte pequeña puede ser muy relevante clínicamente, aunque no cruce el umbral estadístico. El lector crítico debe mirar también el intervalo de confianza y la plausibilidad clínica.

3. Ignorar el diseño y los tiempos de seguimiento

Muchos lectores pasan por alto si el seguimiento fue prospectivo o retrospectivo, o si hubo pérdidas importantes de pacientes. Estos detalles son fundamentales: un estudio con seguimiento desigual puede sesgar la incidencia acumulada y exagerar los riesgos. Siempre se debe preguntar: *¿cuánto tiempo estuvieron realmente los pacientes en riesgo?*

4. Ajustar por todo (o solo por lo “significativo”)

Un error frecuente es pensar que mientras más variables se ajusten, mejor. Sin embargo, ajustar por variables irrelevantes puede distorsionar el resultado. El ajuste debe basarse en criterios causales, no en caprichos estadísticos.

5. Confundir confusión con modificación del efecto

A diferencia de la confusión, de cuyos efectos queremos deshacernos en nuestro análisis, la modificación de efectos es un hallazgo interesante en sí mismo, y lo reportamos. Para verificar la modificación del efecto, realizar un análisis estratificado.

6. Olvidar la precisión de las estimaciones

Un RR elevado con un intervalo de confianza muy amplio (ej. 1.2–9.0) es mucho menos confiable que un RR de 2.5 con IC95% 2.0–3.0. Los lectores suelen fijarse solo en el valor puntual y olvidan que la amplitud del intervalo refleja la incertidumbre de la estimación.

7. Desestimar riesgos competitivos

En cohortes largas, algunos pacientes mueren por otras causas antes de desarrollar el evento de interés. Si se ignora este hecho, se sobreestima la incidencia. El lector debe revisar si se aplicaron métodos para riesgos competitivos o al menos si se reconoce esta limitación.

8. Generalizar sin evaluar validez externa

Un error clásico es extrapolar resultados de una cohorte específica a toda la población. Si el estudio se hizo en comunidades rurales expuestas a biomasa, no podemos generalizar sin cautela a pacientes urbanos que usan gas. El lector crítico debe siempre contextualizar: *¿se parece esta cohorte a mis pacientes?*

Reflexión crítica: Los errores de interpretación no son detalles menores: pueden cambiar la manera en que aplicamos la evidencia en la clínica. Por eso, el lector crítico debe adoptar una actitud sistemática: verificar si el estudio reporta medidas absolutas y relativas, revisar los intervalos de confianza, distinguir confusión de modificación del efecto, y valorar si los hallazgos son clínicamente relevantes y aplicables a su población.

7. SESGOS Y LIMITACIONES EN ESTUDIOS DE COHORTE

Ningún estudio está libre de imperfecciones. Incluso las cohortes más cuidadosamente diseñadas pueden presentar sesgos (errores sistemáticos que distorsionan la asociación real) y limitaciones (factores que restringen la interpretación de los resultados). El lector crítico debe reconocerlos, porque de ello depende decidir si los hallazgos son confiables y aplicables.

7.1. SESGO DE SELECCIÓN

Ocurre cuando los grupos de comparación (expuestos y no expuestos) no son realmente comparables desde el inicio, o cuando las pérdidas en el seguimiento no ocurren al azar.

- Ejemplo en nuestro caso: si los pacientes expuestos al humo de leña eran en promedio de mayor edad o con peor estado de salud inicial, el riesgo de EPOC se vería aumentado por

factores distintos a la exposición. Asimismo, si durante el seguimiento se pierde más del 20% de los expuestos porque migran a áreas urbanas, la estimación del riesgo puede ser sesgada.

7.2. SESGO DE INFORMACIÓN

Se produce cuando la exposición o el desenlace se miden de forma incorrecta. Puede ser diferencial (más error en un grupo que en otro) o no diferencial (igual error en ambos grupos).

- Ejemplo: si la exposición al humo de leña se evaluó solo con entrevistas y no con mediciones objetivas de partículas, puede haber clasificación errónea. Del mismo modo, si el diagnóstico de EPOC en zonas rurales se basó en síntomas sin espirometría, podrían clasificarse falsamente enfermos y sanos.

7.3. CONFUSIÓN

Ya hemos visto que ocurre cuando una tercera variable se relaciona tanto con la exposición como con el desenlace.

- Ejemplo: el tabaquismo es un confusor clave. Si no se mide y ajusta, parte del riesgo atribuido al humo de leña en realidad se debe al tabaco.

7.4. SESGO DE SUPERVIVENCIA O PREVALENCIA

En cohortes retrospectivas o ambispectivas, puede haber una “paradoja de supervivencia”: quienes ingresan al estudio ya sobrevivieron cierto tiempo a la exposición, lo que puede subestimar el efecto real.

- Ejemplo: incluir solo pacientes que aún viven en áreas rurales y llevan décadas expuestos al humo de leña puede excluir a los que murieron antes por EPOC severo.

7.5. RIESGOS COMPETITIVOS

Como vimos, los pacientes pueden fallecer por otras causas antes de presentar EPOC. Si se ignora este fenómeno, la incidencia acumulada puede sobreestimarse.

LIMITACIONES HABITUALES DE LAS COHORTES

- **Tiempo y costo:** Las cohortes prospectivas requieren años de seguimiento y gran inversión. Esto limita su factibilidad y tamaño de muestra.
- **Pérdidas de seguimiento:** Una limitación casi inevitable. Cuando son numerosas, la validez interna se debilita.
- **Exposiciones cambiantes:** Los pacientes pueden modificar hábitos durante el seguimiento (ej. cambiar de leña a gas), lo cual complica la clasificación.
- **Generalización restringida:** La validez externa depende de cuánto se parezca la cohorte a la población de interés. Cohortes rurales pueden no ser comparables a poblaciones urbanas.
- **Medición de variables:** No siempre es posible medir con precisión exposición, desenlace o confusores.

8. CHECKLIST

La lectura crítica de un estudio de cohorte requiere una estructura clara que guíe al lector paso a paso. Para elaborar este checklist hemos integrado las principales guías internacionales:

- STROBE (Strengthening the Reporting of Observational Studies in Epidemiology): que define los elementos mínimos que debe reportar un estudio observacional.
- CASP (Critical Appraisal Skills Programme): que propone preguntas prácticas para valorar validez, resultados y aplicabilidad.
- JAMA Users' Guides to the Medical Literature: una de las guías más influyentes de la medicina basada en evidencia, centrada en utilidad clínica.

- EQUATOR Network: iniciativa internacional que reúne estándares de calidad en publicación científica.

El resultado es una herramienta didáctica y práctica que permite, por un lado, revisar de manera exhaustiva todos los aspectos metodológicos de un estudio de cohorte, y por otro, contar con un esquema resumido para la consulta rápida en la práctica clínica o docente.

Tabla 13. Checklist completo para lectura crítica de estudios de cohorte

Dimensión	Preguntas clave
1. Pregunta de investigación	¿Está claramente definida y es clínicamente relevante?
	¿La estructura sigue un esquema PICO (Población, Exposición, Comparador, Desenlace)?
2. Selección de la cohorte	¿Se describen claramente los criterios de inclusión y exclusión?
	¿Los grupos de expuestos y no expuestos eran comparables al inicio?
	¿Se justifica el tamaño muestral y la potencia del estudio?
3. Medición de la exposición	¿La exposición se midió con métodos válidos y fiables (ej. cuestionario validado, biomarcadores, registros)?
	¿Se definieron claramente las categorías de exposición (sí/no, intensidad, duración)?
	¿Se consideraron cambios de exposición durante el seguimiento (ej. exposición dependiente del tiempo)?
4. Medición del desenlace	¿El desenlace está claramente definido y medido con criterios estándar (ej. diagnóstico de EPOC por espirometría)?
	¿La medición fue ciega o independiente respecto a la exposición?
	¿Los desenlaces son clínicamente relevantes?
5. Seguimiento	¿La duración del seguimiento fue adecuada para que aparezca el evento?
	¿Se reportó la proporción de pérdidas de seguimiento? ¿Fue <20%?
	¿Se analizaron las pérdidas para verificar si fueron diferenciales entre grupos?
	¿Se consideraron riesgos competitivos (ej. muerte por otra causa antes del desenlace)?
6. Análisis estadístico	¿Se presentaron medidas crudas (RR, HR, tasas) y ajustadas?
	¿Se ajustó por confusores clave (edad, sexo, tabaquismo, comorbilidades)?
	¿Se exploró la modificación de efecto (estratificación o interacción)?
	¿Se verificaron los supuestos de los modelos (ej. proporcionalidad de riesgos en Cox)?
7. Resultados	¿Se reportan p-valores e intervalos de confianza (IC95%)?
	¿Los resultados son clínicamente relevantes además de estadísticamente significativos?
	¿Se reportan medidas absolutas además de relativas (ej. diferencia de riesgos, NNT/NNH)?
8. Comparación y validez externa	¿Los resultados son consistentes con otros estudios?
	¿La población estudiada es comparable a la de mi práctica clínica?
	¿Las conclusiones pueden generalizarse con cautela?
9. Sesgos y limitaciones	¿El estudio reconoce sus limitaciones (sesgo de selección, información, confusión)?
	¿Los autores discuten cómo estas limitaciones pueden haber afectado los resultados?
Fuente: Elaboración propia	

Tabla 14. Checklist breve para lectura crítica de estudios de cohorte (versión aplicable)

✓	Pregunta clave
☐	¿Es válida la cohorte?
☐	¿La población fue bien definida, con exposición y desenlace medidos de forma confiable?
☐	¿Hubo buen seguimiento (tiempo suficiente, pérdidas <20%)?
☐	¿Se controlaron los confusores (edad, sexo, tabaquismo u otras variables clave)?
☐	¿Cuál fue la magnitud y precisión del efecto (RR/HR con IC95% estrecho y plausible)?
☐	¿Son aplicables los resultados a mis pacientes? ¿El efecto observado es clínicamente relevante?

Fuente: Elaboración propia

9. UTILIDAD DE LOS ESTUDIOS DE COHORTE

Los estudios de cohorte son una de las herramientas más valiosas en la investigación clínica y epidemiológica. Aunque no alcanzan el mismo nivel de evidencia que los ensayos clínicos aleatorizados para demostrar causalidad, poseen características únicas que los convierten en instrumentos indispensables tanto para comprender la historia natural de las enfermedades como para orientar políticas de salud.

Una primera utilidad es que permiten seguir a los pacientes a lo largo del tiempo y describir cómo surge y progresa una enfermedad desde la exposición hasta el desenlace. Esto aporta información esencial sobre la historia natural. El ejemplo clásico es el Estudio de Framingham, que identificó factores de riesgo cardiovascular como la hipertensión, el tabaquismo y la hipercolesterolemia. En nuestro caso, el seguimiento de pacientes expuestos al humo de leña permite observar cómo la exposición continua produce inflamación crónica de la vía aérea, deterioro progresivo de la función pulmonar y finalmente EPOC.

Además, las cohortes permiten identificar factores de riesgo y también factores protectores. Una sola cohorte puede arrojar información sobre múltiples desenlaces relacionados con una misma exposición. En el caso de la biomasa, un estudio no solo puede medir su impacto en la EPOC, sino también su relación con cáncer de pulmón, neumonía, exacerbaciones respiratorias y mortalidad general. Esta capacidad de explorar múltiples desenlaces a partir de una misma población hace que el diseño sea especialmente valioso en entornos de salud pública.

Otra fortaleza central es que las cohortes permiten calcular directamente medidas de ocurrencia y asociación, como la incidencia acumulada, la tasa de incidencia, el riesgo relativo y el hazard ratio. Estas medidas son la base de toda la epidemiología analítica porque aportan información tanto absoluta como relativa sobre la magnitud del problema. Gracias a esto, los médicos y los decisores de salud pueden estimar no solo cuánto riesgo tiene un paciente, sino también qué impacto tendría una intervención preventiva en la comunidad.

Un aspecto clave es que el diseño de cohorte asegura la temporalidad: la exposición se registra antes del desenlace, lo que refuerza la plausibilidad causal. Ningún otro estudio observacional aporta con tanta claridad esta dimensión temporal. En nuestro ejemplo, el hecho de demostrar que los pacientes expuestos al humo de leña enferman más y además lo hacen antes que los no expuestos fortalece la hipótesis causal y reduce la posibilidad de que la asociación observada sea espuria.

Los estudios de cohorte también tienen un enorme valor para la salud pública, ya que generan evidencia que fundamenta políticas y programas preventivos. Así como las cohortes sobre tabaquismo fueron determinantes para las leyes de espacios libres de humo, las cohortes sobre biomasa pueden sustentar intervenciones como la sustitución de combustibles y la ventilación adecuada en viviendas rurales. Cuando se demuestra que casi la mitad de los casos de EPOC en una comunidad pueden atribuirse al humo de leña, se establece una base sólida para la acción sanitaria.

Siguiendo a individuos durante periodos prolongados, los estudios de cohorte aclaran como diversos factores contribuyen a los resultados de salud con el tiempo. Proporcionan pruebas sólidas que informa la formulación de políticas más efectivas y específicas de intervenciones y estrategias preventivas de salud.

Las cohortes son insumos esenciales para revisiones sistemáticas y metaanálisis, especialmente en áreas donde un ensayo clínico sería poco ético o inviable. La acumulación de cohortes de diferentes contextos refuerza la consistencia de los hallazgos y permite establecer conclusiones de mayor alcance.

10. AUTOEVALUACIÓN

PREGUNTAS DE SELECCIÓN MÚLTIPLE

1. ¿Cuál es la principal característica que diferencia a los estudios de cohorte de otros diseños observacionales?

- a) Se realizan en un solo momento en el tiempo.
- b) Se parte de la exposición y se sigue a los individuos hasta observar el desenlace.
- c) Seleccionan participantes según su desenlace y analizan retrospectivamente la exposición.
- d) No requieren grupo comparador.

2. ¿Qué medida de asociación puede calcularse directamente en una cohorte prospectiva?

- a) Odds ratio (OR).
- b) Incidencia acumulada.
- c) Sensibilidad y especificidad.
- d) Prevalencia.

3. En una cohorte sobre humo de leña y EPOC se encontró un RR de 2.75 (IC95% 1.9–3.9; $p < 0.001$). ¿Cuál es la interpretación más adecuada?

- a) La asociación es estadísticamente significativa, pero clínicamente irrelevante.
- b) Los expuestos tienen 2.75 veces más riesgo de desarrollar EPOC que los no expuestos.
- c) Los expuestos tienen un 275% de riesgo absoluto de desarrollar EPOC.
- d) El humo de leña no es un factor de riesgo independiente.

4. ¿Qué diferencia principal existe entre confusión y modificación del efecto?

a) La confusión refleja una característica real de la asociación y la modificación es un error metodológico.

b) Ambas deben eliminarse mediante ajuste estadístico.

c) La confusión es un artefacto que debe controlarse; la modificación del efecto es real y debe reportarse.

d) No existe diferencia entre ambos conceptos.

5. ¿Cuál de los siguientes es un ejemplo de riesgo competitivo en una cohorte de EPOC?

- a) Un paciente cambia de exposición durante el seguimiento.
- b) Un paciente muere por infarto antes de desarrollar EPOC.
- c) Un paciente presenta exacerbaciones repetidas de EPOC.
- d) Una paciente se pierde en el seguimiento por migración.

PREGUNTAS ABIERTAS

6. Explique con sus palabras la diferencia entre incidencia acumulada, tasa de incidencia y hazard. ¿En qué situación es más útil cada medida?

7. En el caso clínico del humo de leña y EPOC, ¿por qué es importante reportar tanto el riesgo relativo como la diferencia absoluta de riesgos?

8. ¿Qué información adicional aportan las curvas de Kaplan–Meier frente a la incidencia acumulada simple en un estudio de cohorte?

9. Mencione tres sesgos frecuentes en estudios de cohorte y explique cómo podrían afectar los resultados.

10. Imagine que un artículo reporta un RR crudo de 3.0, pero tras ajustar por tabaquismo y edad el RR baja a 1.5. ¿Cómo interpretaría esta diferencia?

RESPUESTAS:

1: b) Se parte de la exposición y se sigue a los individuos hasta observar el desenlace.

La característica fundamental del estudio de cohorte es que los participantes se clasifican según su exposición inicial, y posteriormente se observan en el tiempo para identificar la aparición del desenlace. Esto garantiza la secuencia temporal exposición → evento, lo que permite calcular incidencia y riesgo relativo

2: b) Incidencia acumulada.

En una cohorte, los individuos están libres de enfermedad al inicio y se siguen en el tiempo. Por ello, es posible calcular directamente la incidencia acumulada, la tasa de incidencia y el riesgo relativo. El Odds Ratio es característico de estudios de casos y controles, no de cohortes.

3: b) Los expuestos tienen 2.75 veces más riesgo de desarrollar EPOC que los no expuestos.

Un RR de 2.75 significa que el riesgo en el grupo expuesto es 2.75 veces el riesgo en el grupo no expuesto. El intervalo de confianza no incluye el valor nulo (1), por lo que la asociación es estadísticamente significativa. No implica un 275% de riesgo absoluto, sino una comparación relativa entre grupos.

4: c) La confusión es un artefacto que debe controlarse; la modificación del efecto es real y debe reportarse.

La confusión distorsiona la asociación real y debe ajustarse mediante estratificación o modelos multivariantes. La modificación del efecto, en cambio, indica que la magnitud del efecto varía entre subgrupos (por ejemplo, mayor efecto en mujeres que en hombres), y debe informarse como hallazgo relevante.

5: b) Un paciente muere por infarto antes de desarrollar EPOC.

Un riesgo competitivo es un evento que impide la ocurrencia del desenlace principal. Si un paciente fallece antes de desarrollar EPOC, ya no puede experimentar ese evento. Esto modifica la probabilidad acumulada de aparición del desenlace y debe considerarse en el análisis.

6. La incidencia acumulada representa la proporción de individuos que desarrollan el evento durante un período determinado. Es una medida de riesgo absoluto y responde a la pregunta: ¿qué probabilidad tiene una persona de enfermar en un intervalo específico? Es más útil cuando el seguimiento es relativamente uniforme y completo.

La tasa de incidencia incorpora el tiempo de observación y se expresa en persona-tiempo (por ejemplo, casos por 100 persona-años). Es especialmente útil cuando los participantes tienen tiempos de seguimiento variables o existe censura, ya que ajusta la frecuencia del evento por el tiempo real en riesgo. El hazard describe la tasa instantánea de ocurrencia del evento en un momento específico, condicionado a haber permanecido libre del evento hasta ese punto. Se utiliza en análisis de supervivencia y modelos como la regresión de Cox, siendo particularmente útil cuando interesa evaluar cómo cambia el riesgo a lo largo del tiempo.

7. El RR cuantifica la magnitud comparativa del efecto de la exposición. Indica cuántas veces mayor es el riesgo en el grupo expuesto respecto al no expuesto. La diferencia absoluta de riesgos (RAR) muestra el aumento real del riesgo en términos absolutos. Es fundamental para estimar el impacto clínico y poblacional, y permite calcular medidas como el número necesario a dañar (NNH). En el caso del humo de leña y EPOC, el RR informa sobre la fuerza de la asociación, mientras que la RAR permite dimensionar el impacto real en la población. Ambas medidas son complementarias y necesarias para una interpretación completa.

8. La incidencia acumulada simple ofrece una estimación global al final del seguimiento, pero no informa cuándo ocurren los eventos. Las curvas de Kaplan-Meier permiten visualizar la evolución temporal del riesgo, comparar grupos a lo largo del tiempo, identificar la velocidad de aparición de los eventos y mostrar la distribución de censuras. Además, permiten aplicar pruebas estadísticas como el log-rank para evaluar diferencias entre grupos. Por tanto, aportan una perspectiva dinámica del riesgo que la incidencia acumulada aislada no proporciona.

9. Un sesgo frecuente es el sesgo de selección, que ocurre cuando los grupos no son comparables al inicio. Puede generar estimaciones exageradas o atenuadas del efecto. El sesgo de información se produce cuando la exposición o el desenlace se miden de forma inexacta o diferente entre grupos, lo que puede distorsionar la asociación. La confusión ocurre cuando una variable externa, como el tabaquismo, está asociada tanto con la exposición como con el desenlace. Si no se controla adecuadamente, puede generar asociaciones espurias o infladas. Estos sesgos comprometen la validez interna del estudio.

10. Si el RR crudo es 3.0 y, tras ajustar por edad y tabaquismo, disminuye a 1.5, esto indica que parte de la asociación inicial estaba influenciada por confusión. El ajuste elimina la influencia de variables que distorsionaban la relación entre exposición y desenlace. El RR ajustado representa una estimación más precisa del efecto independiente de la exposición. Esta diferencia sugiere que edad y tabaquismo eran factores importantes que explicaban parcialmente la asociación observada en el análisis crudo.

REFERENCIAS

- Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al; STROBE Initiative. STROBE explanation and elaboration. *Int J Surg*. 2014;12(12):1500–24. Available at: <https://doi.org/10.1016/j.ijsu.2014.07.014>.
- Critical Appraisal Skills Programme (CASP). CASP Cohort Study Checklist. Oxford: CASP; 2024 [cited 2025 Feb 8]. Available at: <https://casp-uk.net/casp-tools-checklists/cohort-study-checklist/>.
- Guyatt G, Rennie D, Meade MO, Cook DJ, editors. *Users' guides to the medical literature: A manual for evidence-based clinical practice*. 3rd ed. New York: McGraw-Hill Education; 2015.
- Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ, editors. *Modern epidemiology*. 4th ed. Philadelphia: Wolters Kluwer; 2020. ISBN: 9781451193282.
- Hernán MA, Robins JM. *Causal inference: What if*. Boca Raton: Chapman & Hall/CRC; 2020 [updated 2024 Jan 2; cited 2025 Feb 8]. Available at: <https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>.
- Zabert G, Patino CM, Ferreira JC. Prospective cohort studies and their contribution to public health and evidence-based medicine. *J Bras Pneumol*. 2025;51(4):e20250355. Available at: <https://doi.org/10.36416/1806-3756/e20250355>.
- Wang X, Kattan MW. Cohort studies: Design, analysis, and reporting. *Chest*. 2020 Jul;158(1 Suppl):S72–8. Available from: <https://doi.org/10.1016/j.chest.2020.03.014>.
- Dias S, Batalha MJ. Epidemiologic study designs: Cohort studies. *Lusiadas Sci J*. 2024;5(2). Available at: <https://doi.org/10.48687/ljsj.216>.
- Capili BA, Anastasi JK. Cohort studies. *Am J Nurs*. 2021 Dec 1;121(12):45–8. Available at: <https://doi.org/10.1097/01.NAJ.0000803196.49507.08>.
- 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>.
- Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*. 1999;94(446):496–509. Available at: <https://doi.org/10.1080/01621459.1999.10474144>.
- Royston P, Parmar MKB. Restricted mean survival time: An alternative to the hazard ratio for the analysis of time-to-event data. *BMC Med Res Methodol*. 2013;13:152. Available at: <https://doi.org/10.1186/1471-2288-13-152>.
- Pottegård A. Core concepts in pharmacoepidemiology: Fundamentals of the cohort and case-control study designs. *Pharmacoepidemiol Drug Saf*. 2022;31(8):817–26. Available at: <https://doi.org/10.1002/pds.5482>.
- Perez-Guerrero EE, Guillen-Medina MR, Marquez-Sandoval F, Vera-Cruz JM, Gallegos-Arreola MP, et al. Methodological and statistical considerations for cross-sectional, case-control, and cohort studies. *J Clin Med*. 2024;13(14):4005. Available at: <https://doi.org/10.3390/jcm13144005>.
- Xing Z, Zhang Y, Sun S, Cao W, Zhu J, Wang S, et al. Combined effect of ozone and household air pollution on COPD in young adults: A prospective cohort study. *Thorax*. 2024;79(5):449–57. Available at: <https://doi.org/10.1136/thorax-2023-220017>.