MANUAL DE INVESTIGACIÓN

CUANTITATIVA EN MEDICINA:

Fundamentos, Metodología y Herramientas de Inteligencia Artificial

Yamil Liscano Martínez



VIGILADA MINEDUCACIÓN
USC UNIVERSIDAD SANTIAGO DE CALI
EDITORIAL

USC
UNIVERSIDAD SANTIAGO DE CALI
EDITORIAL

Cita este libro:

Liscano Martínez Y. (2026). *Manual de investigación cuantitativa en medicina: fundamentos, metodología y herramientas de inteligencia artificial*. Cali, Colombia: Editorial Universidad Santiago de Cali.

Palabras Claves / *Keywords*:

Investigación cuantitativa; metodología de la investigación; medicina basada en la evidencia; bioestadística; epidemiología; diseños de estudio; inferencia estadística; lectura crítica; redacción científica; ciencia de datos en salud; inteligencia artificial; educación médica.

Iuantitative research; research methodology; evidence-based medicine; biostatistics; epidemiology; study designs; statistical inference; critical appraisal; scientific writing; health data science; artificial intelligence; medical education.

Contenido relacionado:
https://investigaciones.usc.edu.co/

MANUAL DE INVESTIGACIÓN

CUANTITATIVA EN MEDICINA:

Fundamentos, Metodología y Herramientas de Inteligencia Artificial

Autor

Yamil Liscano Martínez

USC
UNIVERSIDAD SANTIAGO DE CALI
EDITORIAL

Manual de Investigación Cuantitativa en Medicina: Fundamentos, Metodología y Herramientas de Inteligencia Artificial / Yamil Liscano Martínez. = [Autores] - Cali, Universidad Santiago de Cali, 1° Edición, Sello Editorial – 2026

568 páginas; Figuras; Fotos; Imágenes, Graficas, Etc.; 24 cm

Incluye Índice; Referencias Bibliográficas

ISBN Impreso: 978-628-7914-26-1 **ISBN Digital:** 978-628-7914-27-8

Introducción, Parte I: Fundamentos Teóricos y Epistemológicos en Investigación Médica. Parte II: Manejo y Análisis de Datos en Investigación Médica. Parte III. Epidemiología básica. Parte IV. El Viaje de los Datos en Salud: De la Recolección al Conocimiento. Parte V. Inteligencia Artificial en Investigación. / 1. Yamil Liscano Martínez. Medicina. Universidad Santiago de Cali. Facultad de Ciencias de la Salud.

CDD 610.72 / Y19

Co-CaUSC
JJBV / 2026

USC
UNIVERSIDAD SANTIAGO DE CALI
EDITORIAL

Manual de Investigación Cuantitativa en Medicina:
Fundamentos, Metodología y Herramientas de Inteligencia Artificial

© Universidad Santiago de Cali

©Autor: Yamil Liscano Martínez

Edición 100 ejemplares
Cali, Colombia-2026

Fondo Editorial
University Press Team
Carlos Andrés Pérez Galindo
Rector
Anisbed Naranjo Rojas
Directora General de Investigaciones

Comité Editorial Universidad Santiago de Cali
Editorial Committee Universidad Santiago de Cali

Odín Ávila Rojas
Doris Lilia Andrade
Edgar Francisco Arcos
Florencio Arias Coronel
Jonathan Pelegrín
Héctor Cuevas Arenas
Liseth Suarez Osorio

Proceso de arbitraje doble ciego
"Double blind" peer-review

Recepción/Submission
Abril (April) de 2025

Evaluación de contenidos/
Peer-review outcome
Octubre (October) de 2025

Correcciones de autor/
Improved version submission
Noviembre (November) de 2025

Aprobación/Acceptance
Noviembre (November) de 2025

La editorial de la Universidad Santiago de Cali se adhiere a la filosofía de acceso abierto. Este libro está licenciado bajo los términos de la Atribución 4.0 de Creative Commons (http://creativecommons.org/licenses/by/4.0/), que permite el uso, el intercambio, adaptación, distribución y reproducción en cualquier medio o formato, siempre y cuando se dé crédito al autor o autores originales y a la fuente https://creativecommons.org/licenses/by/4.0/

Dedicatoria

Esta obra no habría encontrado su camino sin el faro constante de mi familia. A ustedes, mi más profundo agradecimiento. Su apoyo incondicional, los valores que sembraron en mí y la paciencia infinita que me brindaron, no fueron solo un soporte, sino los cimientos mismos sobre los que construí la fortaleza necesaria para volcar estas ideas en palabras. Sin su amor y guía, esta travesía intelectual y personal no habría sido posible.

Agradecimientos

Agradezco a la Universidad Santiago de Cali, y en especial a la Dirección General de Investigaciones y a su Editorial, por creer en este proyecto y acompañarlo hasta su publicación. Su compromiso con el acceso abierto permite que esta obra llegue a quienes más la necesitan.

Doy las gracias a los pares evaluadores, cuya lectura rigurosa y generosa mejoró la claridad y la solidez de cada capítulo. Agradezco a mis colegas de la Facultad de Ciencias de la Salud y a los integrantes del grupo de investigación GISI, con quienes he compartido discusiones, proyectos y aprendizajes que de una u otra forma están presentes en estas páginas.

Mi gratitud especial es para los estudiantes y residentes que, con sus preguntas, sus dudas y sus ganas de aprender, me mostraron qué faltaba explicar y cómo hacerlo mejor. Este libro nació en las aulas y en las asesorías de tesis, y es tan suyo como mío.

Por último, agradezco a mi familia, que me sostuvo en cada etapa de esta obra.

CONTENIDO

Content

Introducción 9

Parte I: Fundamentos Teóricos y Epistemológicos en Investigación Médica

Capítulo 1.
Introducción a la investigación en medicina 13

Capítulo 2.
Medicina basada en la evidencia y fundamentos epistemológicos..... 25

Parte II: Manejo y Análisis de Datos en Investigación Médica

Capítulo 3.
Fundamentos: determinantes de la salud, variables y confusores.....67

Capítulo 4.
Descripción y visualización de datos clínicos.....83

Capítulo 5.
Tablas y figuras: ¿cómo darle forma visible a los datos clínicos? 99

Capítulo 6.
Cuando dos ejes ya no bastan: visualizaciones para tres variables.....113

Capítulo 7.
Probabilidad y distribuciones teóricas en ciencias de la salud 129

Capítulo 8.
Inferencia y estimación: muestreo, intervalos y formulación de hipótesis.....151

Capítulo 9.
Inferencia estadística en investigación clínica. Pruebas de hipótesis, errores y estimación de la incertidumbre..... 163

Capítulo 10.
Selección y aplicación de pruebas estadísticas: de lo paramétrico a lo no paramétrico 181

Capítulo 11.
Asociación y predicción: correlaciones y regresiones 191

Parte III. Epidemiología básica

Capítulo 12.
Medidas epidemiológicas 221

Capítulo 13.
Medidas de asociación 229

Capítulo 14.
Medidas de impacto 237

Capítulo 15.
Diseños de estudios epidemiológicos 243

Capítulo 16.
Interpretación crítica de la literatura científica 385

Capítulo 17.
Estructura y redacción del artículo científico 419

Parte IV. El Viaje de los Datos en Salud: De la Recolección al Conocimiento

Capítulo 18.
Fundamentos del ciclo de vida de los datos 441

Capítulo 19.
Recolección y diseño de datos en salud 451

Capítulo 20.
Pretratamiento y preparación de datos 463

Capítulo 21.
De los datos al conocimiento aplicable 477

Parte V. Inteligencia Artificial en Investigación

Capítulo 22.
Historia y fundamentos de la IA..... 495

Capítulo 23.
Herramientas clave y sus nichos 503

Capítulo 24.
Desde la idea al protocolo..... 509

Capítulo 25.
Análisis y visualización asistidos por IA.....521

Capítulo 26.
Búsqueda y gestión de referencias..... 527

Capítulo 27.
Redacción y maquetación automatizada 533

Capítulo 28.
Presentación de resultados..... 539

Capítulo 29.
Divulgación científica con IA..... 545

Referencias 551

Acerca del autor 565

Pares evaluadores..... 567

INTRODUCCIÓN

Introduction

Desde los albores de la civilización, la humanidad ha mirado el mundo, y a sí misma, con una curiosidad insaciable. Esa chispa, ese deseo de comprender por qué suceden las cosas, de ir más allá de la simple observación para desentrañar los mecanismos subyacentes, es la fuerza que impulsa el progreso. En el vasto y complejo universo de la medicina, esta curiosidad se traduce en investigación, el motor incansable que nos permite avanzar, transformar la práctica clínica y, en última instancia, mejorar la salud y el bienestar de las personas.

La investigación en medicina no es un fin en sí misma, sino el cimiento sobre el cual se construyen las innovaciones que definen el cuidado del paciente en el siglo XXI. Nos permite pasar de la observación empírica, fundamental en los albores de la medicina, a la aplicación de métodos sistemáticos y experimentales que validan nuevas hipótesis, como ocurrió con hitos históricos que van desde el pensamiento racional de la Grecia clásica hasta la consolidación del método experimental en el Renacimiento y más allá. Este viaje histórico, repleto de descubrimientos revolucionarios, nos recuerda que el conocimiento médico nunca está quieto, sino que evoluciona constantemente, impulsado por la búsqueda de una comprensión más profunda.

Sin embargo, para muchos profesionales de la salud, especialmente aquellos en la primera línea de la atención clínica, como los residentes de medicina interna o los estudiantes de pregrado, el mundo de la investigación puede parecer un laberinto complejo y ajeno. Enfocados, comprensiblemente, en el arte y la ciencia del diagnóstico y tratamiento directo del paciente, a menudo sienten que carecen de las herramientas o el conocimiento necesario para navegar las aguas de la metodología científica, la estadística o la interpretación crítica de la literatura. Surge entonces una brecha entre la práctica clínica diaria y la necesidad imperante de comprender, evaluar y contribuir a la evidencia que da forma a esa misma práctica.

Esta guía nace precisamente de esa necesidad. Reconocemos que el médico clínico, inmerso en la atención directa, a menudo se enfrenta a la

exigencia de abordar la investigación, ya sea para completar una tesis, para mantenerse actualizado con los últimos avances o, fundamentalmente, para aplicar el pensamiento crítico que le permita discernir la mejor evidencia para sus pacientes. No se trata de convertir a cada clínico en un investigador de laboratorio, sino de equiparlo con una brújula y un mapa conceptual que le permitan entender por qué la investigación es vital, cómo se genera el conocimiento fiable y cómo puede utilizarlo para tomar decisiones más informadas.

A lo largo de estas páginas, exploraremos los fundamentos teóricos y epistemológicos que subyacen a la investigación médica, desmitificando conceptos y presentando la metodología como una ruta lógica y estructurada hacia el descubrimiento. Nos adentraremos en el manejo y análisis de datos, esas piezas fundamentales que, una vez limpias y organizadas, revelan patrones y tendencias ocultas. Abordaremos la epidemiología, la disciplina que nos permite comprender la salud y la enfermedad a nivel poblacional, y exploraremos cómo los diseños de estudio, desde los observacionales hasta los ECA, actúan como diferentes lentes para observar la realidad. Finalmente, nos asomaremos al horizonte de la Inteligencia Artificial, una herramienta emergente que, utilizada con juicio y conocimiento de los fundamentos, promete potenciar la investigación de maneras antes inimaginables.

Este libro es una invitación a ver la investigación no como una tarea árida y separada de la clínica, sino como una extensión natural de la curiosidad que llevó a muchos a elegir la medicina. Es una herramienta para construir puentes entre el consultorio y la publicación científica, entre la pregunta clínica y la respuesta basada en evidencia. Al embarcarnos en este recorrido, descubriremos que comprender la investigación no solo es posible, sino que es una habilidad esencial para el médico del presente y del futuro, capacitándolo no solo para tratar, sino también para innovar y contribuir activamente al avance de la ciencia médica.

Yamil Liscano Martínez
Universidad Santiago de Cali, Facultad de Ciencias de la Salud
Docente investigador y autor

"El verdadero conocimiento no reside en la vehemencia de la discusión, sino en la quietud de la escucha humilde; es en el silencio atento donde el saber echa raíces profundas, nutriéndose de la perspectiva ajena sin necesidad de imponer la propia"

—Yamil Liscano

Parte I:
Fundamentos Teóricos y Epistemológicos en Investigación Médica

INTRODUCCIÓN A LA INVESTIGACIÓN EN MEDICINA

Introduction to Research in Medicine

Resumen

La investigación en medicina es la base para innovar en la práctica clínica, posibilitando avances en diagnóstico y tratamiento, y adquiere gran relevancia al integrarse en la formación médica, donde la combinación de teoría y práctica fomenta habilidades analíticas, formulación de hipótesis y uso de herramientas estadísticas y tecnológicas. Su evolución parte de la observación empírica de las primeras civilizaciones, pasando de prácticas basadas en la religión y la superstición a métodos sistemáticos y experimentales evidenciados en hitos históricos como el pensamiento racional de la Grecia clásica, el renacimiento con figuras como Vesalio y Harvey, y la consolidación del método experimental mediante sociedades científicas. La revolución digital y la biología molecular han impulsado la transición hacia la medicina basada en la evidencia, integrando análisis estadístico y tecnologías avanzadas como *big data*, ómicas y *machine learning*. El pensamiento investigativo se caracteriza por una combinación de curiosidad, sistematicidad, lógica y creatividad, utilizando métodos deductivos e inductivos que refuerzan el análisis crítico y la validación de hipótesis, lo cual mejora la calidad en la toma de decisiones clínicas.

Palabras clave: investigación, medicina, evidencia, pensamiento crítico, historia, tecnología

Abstract

Research in medicine is the foundation for innovation in clinical practice, enabling advancements in diagnosis and treatment. It gains significant relevance when integrated into medical education, where the combination of theory and practice fosters analytical skills, hypothesis formulation, and the use of statistical and technological tools. Its evolution began with the empirical observation

of early civilizations, transitioning from practices based on religion and superstition to systematic and experimental methods. These are evidenced in historical milestones such as the rational thought of classical Greece, the Renaissance with figures like Vesalius and Harvey, and the consolidation of the experimental method through scientific societies. The digital revolution and molecular biology have driven the transition towards evidence-based medicine, integrating statistical analysis and advanced technologies like big data, omics, and machine learning. Investigative thinking is characterized by a combination of curiosity, systematicity, logic, and creativity. It utilizes deductive and inductive methods that strengthen critical analysis and hypothesis validation, thereby improving the quality of clinical decision-making.

Keywords: *research, medicine, evidence, critical thinking, history, technology*

1.1 Definición y Relevancia de la Investigación para la Práctica Clínica

La investigación en medicina es el cimiento sobre el cual se construyen las innovaciones que transforman la práctica clínica, permitiendo a los profesionales avanzar en el diagnóstico y tratamiento de enfermedades. Desde sus inicios, la curiosidad y la observación han impulsado el descubrimiento de nuevos conocimientos que, al ser sistemáticamente examinados y validados, se convierten en herramientas esenciales para la mejora de la atención al paciente. Este proceso, que va desde la formulación de preguntas hasta la aplicación de nuevos diagnósticos, se enriquece con métodos tradicionales y modernos, integrando tanto la experiencia clínica como las innovaciones tecnológicas.

La formación en investigación se ha convertido en un componente vital de los programas de educación médica, donde la exposición temprana a proyectos de investigación estimula el pensamiento crítico y la creatividad. Estudios como el de "Teaching Medical Research to Medical Students" y el curso en línea *Research* 101 demuestran que, al combinar la teoría con actividades prácticas, los estudiantes pueden aprender a formular hipótesis, diseñar estudios y analizar datos de manera efectiva. Por ejemplo, a través de módulos interactivos, los alumnos aprenden a utilizar herramientas de análisis y a interpretar resultados que luego aplican en proyectos reales, lo cual sienta las bases para una carrera robusta en investigación y práctica clínica basada en evidencia.

La evolución histórica de la investigación médica es un recorrido fascinante que comienza en la antigüedad, cuando la medicina se basaba en la observación empírica, y se transforma en la era moderna con el desarrollo de ensayos clínicos y metodologías estadísticas rigurosas. Desde la inoculación de la vacuna contra la viruela, inspirada en las primeras observaciones empíricas, hasta el descubrimiento del *Helicobacter pylori* en úlcera estomacales, cada avance ha sido posible gracias a la integración de conocimientos y la aplicación de nuevos métodos. Estos hitos históricos demuestran que el progreso en medicina depende de una constante revisión y actualización de los modelos investigativos.

El pensamiento investigativo se caracteriza por la combinación de curiosidad, sistematicidad, análisis crítico y colaboración; cualidades que se cultivan a través de una formación rigurosa y práctica. Los estudiantes, al enfrentar problemas complejos, deben aprender a desglosar preguntas amplias en ítems específicos y a utilizar herramientas analíticas, como se resalta en el libro *Introduction to research methods and data analysis in the health sciences*. Un buen ejemplo es el uso de técnicas estadísticas y software especializado, lo que permite a los futuros clínicos evaluar la evidencia y tomar decisiones basadas en datos confiables, reforzando así su capacidad para innovar en el tratamiento de diversas patologías.

La integración de la tecnología en la investigación, especialmente mediante el uso de ómicas, *big data* y *machine learning*, ha revolucionado la forma de abordar los desafíos médicos contemporáneos. La capacidad para analizar datos masivos y complejos mediante algoritmos avanzados no solo acelera el descubrimiento de biomarcadores y nuevas terapias, sino que también optimiza la precisión diagnóstica. Por ejemplo, estudios basados en técnicas de ómicas han permitido identificar variantes genéticas y perfiles metabólicos asociados con enfermedades multifactoriales, abriendo la puerta a un enfoque más personalizado en la medicina. Este avance, apoyado en gran medida por desarrollos computacionales, representa una convergencia entre la biología, la estadística y la informática que redefine la frontera del conocimiento clínico.

En el contexto de los países en desarrollo, la investigación biomédica enfrenta desafíos particulares que no solo limitan la producción científica, sino que también afectan la capacidad de resolver problemas de salud loca-

les. En países en desarrollo, la escasez de recursos, la falta de capacitación especializada y las limitadas infraestructuras son barreras que, sin embargo, se pueden transformar en oportunidades mediante alianzas internacionales y el fortalecimiento de la capacidad investigativa. La colaboración entre instituciones de países desarrollados y en desarrollo permite compartir conocimientos, formar redes de mentoría y adaptar metodologías según las necesidades específicas de cada contexto, lo que en última instancia contribuye a una medicina basada en evidencia que beneficia a toda la comunidad.

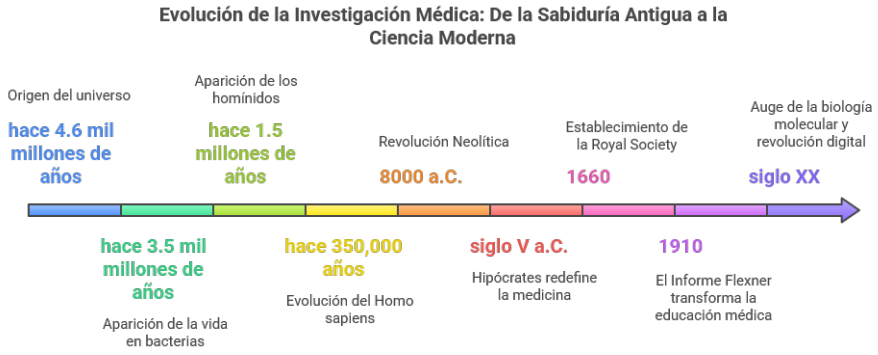
1.2. Historia y Evolución de la Investigación Médica

La historia de la investigación médica se fundamenta en la capacidad innata del ser humano para indagar en su entorno y buscar explicaciones a lo inexplicable. Se puede remontar, en términos muy amplios, a los orígenes del universo, hace aproximadamente 4.6 mil millones de años, y a la aparición de la vida, que se estima ocurrió hace alrededor de 3.5 mil millones de años en forma de bacterias. Con el paso del tiempo, nuestros ancestros evolucionaron y, aproximadamente hace 1.5 millones de años, surgieron los primeros homínidos con características humanas, evolucionando posteriormente hacia el *Homo sapiens* hace alrededor de 350.000 años, lo que marca el inicio de una historia larga de descubrimientos y cuestionamientos que han forjado el terreno de la medicina moderna.

Con el advenimiento de la agricultura hacia el 8000 a. n. e., los humanos experimentaron un cambio radical en su estilo de vida. Los cazadores-recolectores, al enfrentarse a la presión poblacional y a la necesidad de asegurar una fuente constante de alimento, comenzaron a domesticar plantas y animales, dando lugar a la revolución neolítica. Este cambio no solo transformó la forma de vida, sino que también impulsó la curiosidad por observar y registrar fenómenos naturales, lo que sentó las bases para el desarrollo de métodos empíricos y la eventual sistematización del conocimiento.

La investigación en medicina, en sus albores, surgió de la necesidad de comprender y combatir enfermedades mediante la observación directa (ver figura 1).

Figura 1. *Línea de tiempo de la evolución de la investigación médica. Elaborado con napkin: https://www.napkin.ai/es/*



En las antiguas civilizaciones, donde el saber se fusionaba con la religión, los sanadores eran tanto sacerdotes como médicos. Estas primeras prácticas, basadas en la superstición y en rituales, utilizaron remedios naturales, como el opio, y diversos minerales, para expulsar lo que se creía eran fuerzas malignas responsables de la enfermedad. Este enfoque, aunque rudimentario, representó el primer intento sistemático de intervenir en la salud humana a partir de la experiencia y la observación empírica.

La Grecia clásica trajo consigo una transformación radical en el pensamiento médico. Durante el siglo V a.C., figuras como Hipócrates redefinieron la medicina al promover una práctica basada en la observación y el razonamiento crítico, alejándose de explicaciones místicas. Se introdujo la teoría de los humores, que postulaba que la salud dependía del equilibrio entre la sangre, la flema, la bilis amarilla y la bilis negra, y se incentivó la adopción de dietas y hábitos higiénicos que anticipaban los conceptos modernos de medicina preventiva. Esta época marcó la transición de la medicina religiosa a un enfoque más racional y metódico. Durante la Edad Media, aunque se produjo un estancamiento en el progreso científico en Europa, otras culturas continuaron desarrollando conocimientos. Mientras que en Occidente la influencia de la dogmática autoridad religiosa limitó la libertad de exploración, en regiones como la India y el mundo islámico se preservaron y expandieron algunos conocimientos científicos. Sin embargo, fue durante el renacimiento cuando la medicina experimentó una verdadera revolución,

gracias a la reintroducción de los textos clásicos y al surgimiento de nuevas técnicas de disección. Pioneros como Andreas Vesalio y William Harvey introdujeron métodos experimentales que permitieron una comprensión más profunda de la anatomía y de la circulación sanguínea, respectivamente.

El establecimiento de sociedades científicas, como la *Royal Society* de Londres en 1660, y la publicación de la *Philosophical Transactions* en 1665, consolidaron el método experimental y la comunicación científica. Estos foros permitieron que los científicos compartieran sus hallazgos, replicaran experimentos y, lo más importante, cuestionaran y refinaran sus métodos. La producción de informes detallados evolucionó gradualmente, dando origen a los formatos modernos que estructuran la presentación de los estudios de investigación médica.

El siglo XIX vio el florecimiento de la medicina moderna, impulsado por avances tecnológicos y teóricos. Eventos como la inoculación con la vacuna contra la viruela, descubierta a partir de observaciones empíricas, y la sistematización de la anatomía y la fisiología mediante la disección y el experimento, marcaron hitos fundamentales. El Informe Flexner, publicado en 1910, transformó la educación médica en Estados Unidos al eliminar escuelas carentes de preparación científica y consolidar una formación rigurosa y basada en evidencia, lo que sentó las bases para el avance incesante del conocimiento médico.

En el siglo XX, el auge de la industrialización y la revolución digital aceleraron los descubrimientos y la difusión del conocimiento. La aparición de la biología molecular, los avances en la estadística y la incorporación de nuevas tecnologías, como las técnicas de ómicas y el análisis de *big data*, han permitido a los investigadores abordar problemas complejos con una precisión sin precedentes. Con el apoyo de potentes computadoras y métodos analíticos avanzados, se han identificado biomarcadores y se han desarrollado terapias personalizadas que han revolucionado la práctica clínica.

Simultáneamente, la evolución de la investigación médica no solo se reflejó en los avances tecnológicos, sino también en la transformación del pensamiento y la cultura científica. Desde la cautela de los antiguos griegos, que introdujeron el uso de "hedges" en sus escritos para expresar modestia y cautela en sus afirmaciones, hasta la adopción de un método más riguroso

y reproducible en tiempos modernos, la investigación se ha moldeado para responder a las demandas de precisión y verificación que exige la práctica clínica actual. Esta evolución metodológica ha permitido superar dogmas y abrir paso a nuevos paradigmas que integran la innovación con la ética y la responsabilidad social.

No obstante, la carga de la transformación recae sobre los hombros de los profesionales de la salud y de los pacientes, quienes deben ser los impulsores del cambio. La educación médica actual requiere formar investigadores que no solo dominen las técnicas tradicionales, sino que también se adapten a las metodologías emergentes en el análisis de datos y en la integración de diversas disciplinas. Tal como lo evidencia el curso en línea *Research 101* (*https://libguides.depaul.edu/research101*), la combinación de teoría con actividades prácticas fortalece el pensamiento crítico y prepara a los futuros médicos para liderar innovaciones en un sistema de salud en constante cambio.

La transición desde la medicina empírica hasta la medicina basada en la evidencia (MBE) ha sido un proceso largo y complejo, impulsado por el deseo de mejorar la atención al paciente a través de la ciencia. Desde las primeras observaciones que permitieron la elaboración de remedios rudimentarios hasta los sofisticados ensayos clínicos actuales, la investigación médica ha demostrado que la búsqueda del conocimiento es un viaje inacabable. Cada descubrimiento, desde la identificación de los agentes infecciosos hasta los modelos predictivos actuales basados en inteligencia artificial, ha sido un peldaño más hacia una atención médica más precisa y humana.

Con todo este recorrido histórico, se vislumbra cómo el legado de antiguos pensadores como Hipócrates, Galeno, Vesalio, y Harvey se fusiona con las tecnologías del siglo XXI para formar un campo en constante evolución. La investigación médica ha emergido como la piedra angular para mejorar e innovar en la atención sanitaria, y su historia nos ofrece lecciones fundamentales sobre la importancia de la observación, la experimentación y el rigor científico.

Esta rica trayectoria histórica, repleta de avances revolucionarios y de momentos de profunda transformación, prepara el terreno para abordar con mayor profundidad las características del pensamiento investigativo. El conocimiento acumulado a lo largo de los siglos nos invita a reflexionar so-

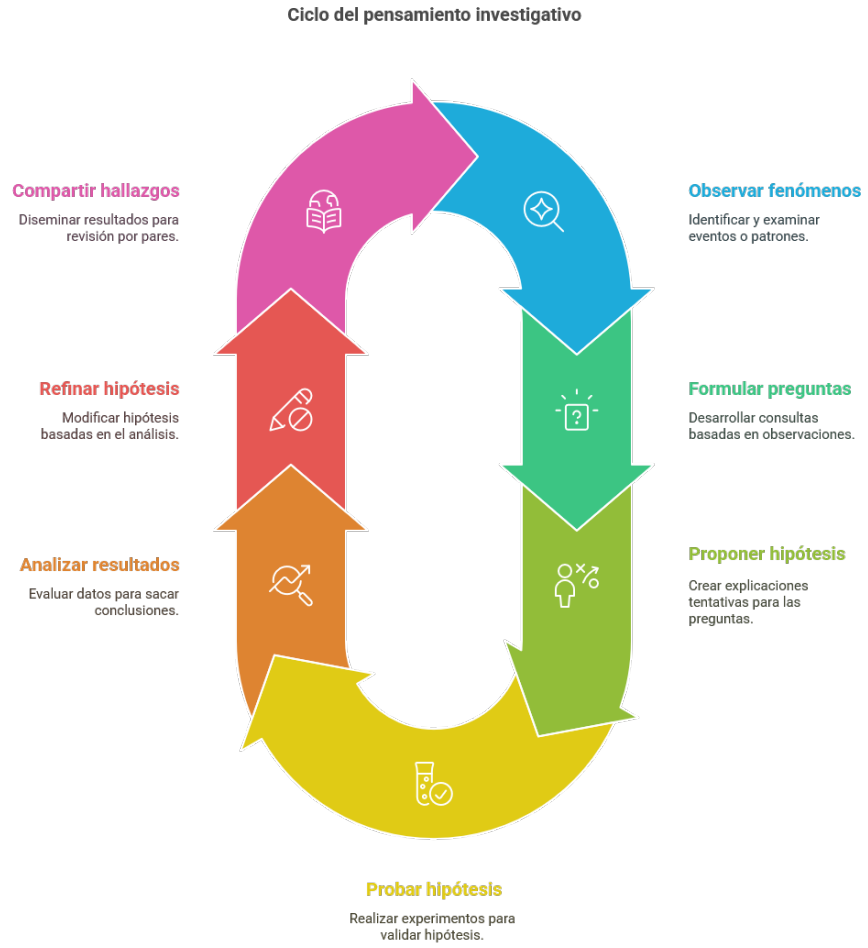
bre cómo las ideas y descubrimientos del pasado continúan inspirando y definiendo las estrategias de investigación que hoy transforman la práctica clínica, abriendo la puerta a explorar nuevos modelos de análisis y colaboración que serán fundamentales para el futuro de la medicina.

1.3. Características del Pensamiento Investigativo

El pensamiento investigativo se erige como la chispa que ha impulsado al ser humano a explorar lo desconocido, a cuestionar lo establecido y a buscar, de manera sistemática, las explicaciones que den sentido a fenómenos complejos. Esta capacidad se ha forjado a lo largo de la historia y se sustenta en la curiosidad, la lógica y la creatividad, elementos que se combinan para formar un proceso metódico en el que cada interrogante abre la puerta a nuevos descubrimientos en el ámbito clínico y científico.

Una de las características esenciales de este pensamiento es la sistematicidad, la cual obliga al investigador a estructurar su trabajo en fases claras que van desde la formulación de preguntas hasta la aplicación práctica de los conocimientos obtenidos (ver figura 2).

Figura 2. *Ciclo del proceso investigativo. Elaborado con napkin: https://www.napkin.ai/es/*



En este proceso, cada paso se revisa y valida, lo que asegura que los hallazgos puedan ser replicados y utilizados para mejorar los diagnósticos y tratamientos en la práctica médica. Así, la integración de métodos tanto tradicionales como modernos permite que la experiencia clínica se enriquezca con innovaciones tecnológicas que responden a la necesidad de evidencias robustas.

El uso de la lógica en la investigación se manifiesta en dos grandes métodos: la deducción y la inducción. La deducción parte de premisas generales, basadas en teorías o leyes ya establecidas, para llegar a conclusiones específicas. Por ejemplo, si se acepta que "todos los medicamentos basados en *beta bloqueadores* reducen la presión arterial" y se sabe que un paciente ha sido tratado con un medicamento de este grupo, se deduce que la presión arterial del paciente debería disminuir. Este método, heredado de la antigua tradición griega, ha sido fundamental para el desarrollo de sistemas de razonamiento riguroso en diversas disciplinas, aunque en ocasiones se queda corto a la hora de generar nuevo conocimiento. Otro ejemplo, si se conoce el principio general de que la exposición prolongada a sustancias tóxicas altera la función pulmonar, y se sabe que el tabaquismo implica la inhalación de múltiples sustancias dañinas, se puede deducir que un paciente fumador tendrá mayor probabilidad de sufrir alteraciones en su función pulmonar. Por tanto, observamos que este razonamiento permite aplicar conocimientos generales para evaluar situaciones concretas.

Por otro lado, la inducción se basa en la observación de casos específicos para formular generalizaciones que pueden llegar a convertirse en teorías o leyes. Imaginemos que tras analizar varios estudios donde se evidencia que la exposición a determinados biomarcadores en la sangre se correlaciona con un mayor riesgo de enfermedad cardiovascular, el investigador puede inducir que dichos biomarcadores son predictivos del riesgo en una población más amplia. Esta aproximación "de lo particular a lo general" ha ganado aceptación especialmente desde que figuras como Francis Bacon la promovieron hace aproximadamente 450 años, sentando las bases para el método científico moderno. Un ejemplo claro de pensamiento inductivo se presenta cuando un grupo de investigadores observa que en varios pacientes con antecedentes de tabaquismo se detecta una mayor incidencia de enfermedad pulmonar obstructiva crónica. A partir de estas observaciones específicas, se formula la hipótesis de que el tabaquismo podría ser un factor de riesgo fundamental para esta patología. Esta inferencia, aunque basada en casos particulares, invita a ampliar la investigación a otros contextos y a corroborar la hipótesis mediante estudios controlados.

La combinación de la inducción y la deducción se traduce en el método hipotético-deductivo, una estrategia que permite a los investigadores proponer

hipótesis basadas en observaciones iniciales y luego someterlas a pruebas rigurosas. En la práctica, esto implica que, tras observar un fenómeno y generar una hipótesis a partir de esos datos específicos, se deducen consecuencias que deben poder verificarse en nuevos experimentos o aplicaciones clínicas. De esta manera, las observaciones se contrastan y se refinan continuamente, lo que fortalece el avance del conocimiento científico y el pensamiento crítico.

El pensamiento crítico se distingue por la habilidad de cuestionar la información, no aceptándola sin antes analizar su fuente, su contexto y la metodología empleada para obtenerla. Un ejemplo de pensamiento crítico en medicina es el análisis detallado de un estudio que relaciona la dieta con la reducción de la presión arterial: el investigador no se limita a aceptar la conclusión, sino que revisa el tamaño de la muestra, la duración del estudio y las posibles variables de confusión. Esta actitud evita conclusiones precipitadas y fomenta una evaluación más completa. Por el contrario, el pensamiento no crítico se evidencia cuando se adopta una postura de aceptación inmediata de cualquier afirmación, sin analizarla en profundidad. Un ejemplo de este enfoque se da cuando un profesional de la salud recomienda un tratamiento únicamente basado en la experiencia anecdótica de otros colegas, sin cuestionar la validez de los datos o la existencia de estudios controlados que respalden la eficacia de dicho tratamiento. Esta falta de análisis puede derivar en prácticas no fundamentadas y, en el peor de los casos, en riesgos para los pacientes.

La dialéctica entre pensamiento crítico y no crítico es especialmente relevante en un mundo saturado de información, donde la capacidad de discernir entre datos confiables y conclusiones erróneas puede marcar la diferencia en la toma de decisiones médicas. Los profesionales que ejercen un pensamiento crítico sistemático están mejor preparados para identificar sesgos, errores metodológicos o interpretaciones simplistas que pueden comprometer el bienestar del paciente. Así, la combinación de observación, formulación de hipótesis y deducción en un marco de pensamiento crítico e investigativo sienta las bases para enfoques modernos como la MBE. Esta corriente integra la experiencia clínica con estudios rigurosos, permitiendo que cada diagnóstico y tratamiento se fundamente en información precisa, contrastada y actualizada, lo que favorece el avance constante de la salud y la calidad de la atención en nuestros centros médicos.

Conclusiones

La investigación en medicina es el motor que impulsa la evolución de la práctica clínica, combinando la tradición empírica con las innovaciones tecnológicas contemporáneas. Desde los orígenes en la observación rudimentaria hasta la aplicación de sofisticados métodos estadísticos y computacionales, este campo ha experimentado una transformación constante que refleja el avance del conocimiento científico. La integración de metodologías tradicionales e innovadoras permite no solo la formulación y prueba de hipótesis, sino también la adaptación de la práctica clínica a los desafíos emergentes de un entorno cada vez más complejo.

Además, la formación en investigación y el desarrollo del pensamiento crítico se posicionan como pilares fundamentales para la consolidación de una MBE. La capacidad de analizar de manera sistemática, evaluar críticamente los datos y aplicar herramientas modernas para la toma de decisiones, no solo mejora la calidad de la atención sanitaria, sino que también abre nuevas posibilidades para la innovación en el diagnóstico y tratamiento de enfermedades. Este enfoque integral y colaborativo es esencial para enfrentar los retos actuales y futuros en el ámbito de la salud.

CAPÍTULO 2

MEDICINA BASADA EN LA EVIDENCIA Y FUNDAMENTOS EPISTEMOLÓGICOS

Evidence-Based Medicine and Epistemological Foundations

Resumen

La MBE ha evolucionado para integrar la experiencia clínica, los valores del paciente y la solidez de los datos científicos, adaptándose a la complejidad del individuo. Este paradigma surge como respuesta a las limitaciones de los ensayos controlados aleatorizados, cuyos resultados poblacionales pueden no reflejar la heterogeneidad de los casos individuales. Para abordar estas deficiencias, se introduce el concepto de MBE, que se orienta a perfilar de forma detallada la historia biológica, clínica, psicológica y social del paciente índice, permitiendo comparar su evolución con la de individuos con perfiles similares y orientar decisiones terapéuticas más personalizadas. Además, se enfatiza la importancia de contar con una "biblioteca de experiencia clínica" que combine información diagnóstica, de laboratorio y biográfica, superando así la rigidez de los modelos tradicionales y ofreciendo un enfoque adaptativo y dinámico en la práctica médica.

Palabras clave: medicina basada en la evidencia, personalización; ensayos controlados, paciente índice, investigación clínica, datos científicos

Abstract

Evidence-Based Medicine (EBM) has evolved to integrate clinical expertise, patient values, and the best available scientific evidence, adapting to the complexity of the individual. This paradigm emerged as a response to the limitations of randomized controlled trials (RCTs), whose population-level results may not fully capture the heterogeneity of individual cases. To address these shortcomings, the concept of EBM is presented, aimed at profiling in detail the biological, clinical, psychological, and social history of the index patient. This detailed profile allows comparing their evolution with that of individuals sharing similar characteristics and guiding more personalized therapeutic decisions. Furthermore,

the importance of developing a "clinical experience library" that integrates diagnostic, laboratory, and biographical information is emphasized. This approach overcomes the rigidity of traditional models, offering a more adaptive and dynamic practice of medicine.

Keywords: *evidence-based medicine, personalization, randomized trials, index patient, clinical research, scientific data*

2.1. ¿Qué es la Medicina Basada en la Evidencia?

La MBE ha evolucionado para integrar la experiencia clínica, los valores del paciente y los datos científicos más sólidos, adaptándose a la complejidad del individuo. Este enfoque no solo supera las limitaciones de los ensayos controlados aleatorizados, cuyos promedios pueden no ajustarse a casos particulares, sino que también ofrece estrategias para una atención clínica más personalizada y efectiva.

Históricamente, la MBE se consolidó como respuesta a los métodos observacionales mal diseñados y a la dependencia de la experiencia individual para la toma de decisiones. Los ensayos controlados aleatorizados y los metaanálisis se convirtieron en las principales herramientas para establecer tratamientos eficaces a nivel poblacional, permitiendo la verificación rigurosa de intervenciones y el desarrollo de guías clínicas. Sin embargo, estas metodologías, al centrarse en resultados promedio, presentan desafíos a la hora de abordar la heterogeneidad de la población clínica.

Las limitaciones de este enfoque se hacen evidentes cuando se observa que criterios estrictos de inclusión en los ensayos controlados suelen excluir a muchos de los pacientes que finalmente utilizarán el tratamiento, y cuando la dependencia exclusiva de "puntos finales" puede obviar resultados igualmente relevantes, como el bienestar psicológico o social. Además, la insistencia en análisis de subgrupos basados en una o pocas variables a menudo no captura la complejidad de la situación clínica de un paciente.

Para dar respuesta a estas deficiencias, surge el concepto de MBE, el cual se orienta a perfilar de forma detallada la historia biológica, clínica, psicológica y social de un paciente índice. Se entiende por paciente índice al individuo representativo cuya información detallada es utilizada para orientar

decisiones terapéuticas que se adaptan a las particularidades de cada caso. Este enfoque propone comparar la evolución y respuesta al tratamiento del paciente en cuestión con la de pacientes que presenten perfiles similares, lo que permite obtener evidencia de manera personalizada y ajustada a las necesidades individuales de cada caso. En la práctica, este paradigma implica la creación de una base de datos que recopile historias clínicas completas, donde se registren y analicen los cambios en el curso de la enfermedad y las intervenciones realizadas a lo largo del tiempo. Por ejemplo, se podría realizar un estudio longitudinal con un paciente con hipertensión crónica, comparando su evolución clínica detallada con la de otros pacientes que hayan recibido tratamientos diferentes. Este análisis permitiría identificar cuál de los regímenes terapéuticos ofrece mejores resultados para las características y necesidades específicas de ese paciente.

La metodología de MBE enfatiza la importancia de contar con una "biblioteca de experiencia clínica", donde cada registro no solo contenga datos diagnósticos o de laboratorio, sino también información sobre eventos biográficos, cambios en el entorno social y otras variables que pueden influir en la respuesta al tratamiento. De este modo, se supera la rigidez del modelo tradicional y se ofrece una herramienta dinámica que se adapta a la evolución del paciente.

Asimismo, mientras que la MBE se fundamenta en estudios que generan resultados promedios a partir de grandes muestras, MBE permite al médico responder a la pregunta crucial: "¿Cómo responderá mi paciente en particular a un tratamiento específico?". Esta aproximación busca cerrar la brecha entre los resultados de la investigación y la experiencia clínica individual, ofreciendo una visión más fina y personalizada.

El proceso para aplicar este nuevo paradigma abarca desde la formulación de preguntas clínicas específicas, usando herramientas como el formato PICO (Paciente, Intervención, Comparación, Resultado), hasta la búsqueda sistemática de evidencia en bases de datos especializadas, y la evaluación crítica de la misma para incorporar no solo los datos de ensayos controlados, sino también de estudios observacionales, registros clínicos y experiencias acumuladas. Esta integración fortalece la toma de decisiones basadas en evidencia al tiempo que respeta la singularidad de cada paciente. Entre

los desafíos de la MBE se destacan la necesidad de desarrollar algoritmos que permitan el emparejamiento aproximado de pacientes, la recolección exhaustiva de datos longitudinales y la capacitación de los profesionales de la salud para la interpretación y aplicación de esta información personalizada. Además, la adaptación de sistemas informáticos y la integración de tecnologías emergentes, como la inteligencia artificial, son elementos clave para que este enfoque sea viable en el entorno clínico actual.

La transformación hacia MBE representa un avance significativo en la forma en que entendemos y aplicamos la evidencia en medicina. Al fusionar rigurosamente los datos científicos con la experiencia clínica y la individualidad del paciente, se forja un modelo que no solo mejora la calidad y seguridad de la atención, sino que también sienta las bases para una medicina verdaderamente personalizada y adaptativa, abriendo la puerta a futuras innovaciones en la práctica clínica.

2.2. Formulación de Preguntas Clínicas y Establecimiento de Objetivos

La formulación de preguntas clínicas y el establecimiento de objetivos constituyen etapas esenciales en el proceso de investigación y en la práctica clínica basada en evidencia. Estos procesos inician con la generación de ideas, tanto a partir del conocimiento previo como de la exploración de aspectos no abordados anteriormente. En este sentido, se pueden emplear diversas metodologías para estimular la creatividad y construir preguntas significativas. Por ejemplo, la lluvia de ideas (*brainstorming*) es una técnica grupal en la que los participantes comparten todas las ideas sin críticas iniciales; imaginemos a un equipo de atención primaria que, al enfrentarse a pacientes con síntomas inespecíficos de fatiga crónica, generan ideas sobre posibles factores implicados, desde causas metabólicas hasta ambientales o psicológicas. Otro método es el pensamiento lateral, que invita a salir de las rutas convencionales y considerar aspectos poco evidentes, como explorar la relación entre el entorno social del paciente y la persistencia de síntomas, abriendo el camino a nuevas preguntas que normalmente podrían pasarse por alto.

Una vez generada la idea, es crucial revisar la información disponible en internet y en la literatura científica, utilizando criterios de calidad y relevancia. En este contexto, los principios FINER (factibilidad, interesante, novedoso, ético y relevante) se presentan como un marco evaluativo para determinar la solidez de una pregunta clínica. Por ejemplo, un investigador que formula la pregunta "¿Cómo influye la calidad del sueño en la eficacia del tratamiento antihipertensivo en adultos mayores?" puede evaluar su factibilidad mediante la disponibilidad de recursos para medir parámetros del sueño, su interés al abordar una problemática común, la novedad si existen pocos estudios que relacionen estos dos aspectos, su ética y su relevancia en mejorar la calidad de vida del paciente.

Para profundizar en estos criterios, consideremos cada uno de ellos en detalle. La factibilidad evalúa si es posible llevar a cabo la investigación con los recursos técnicos, financieros y humanos disponibles, por ejemplo, disponiendo de equipos para monitorear la calidad del sueño y acceso a una población adecuada para el estudio. El criterio de interesante se refiere al potencial del estudio para captar la atención de la comunidad médica y científica, abordando problemáticas que resuenen con desafíos actuales en la práctica clínica. La novedad se analiza en función de si la pregunta amplía el conocimiento existente o responde a interrogantes poco explorados; en nuestro ejemplo, la relación específica entre calidad del sueño y respuesta al tratamiento antihipertensivo puede ser un área emergente con estudios limitados. Por otro lado, el aspecto ético garantiza que la investigación se realice respetando los derechos y la seguridad de los participantes, y cumpliendo con los estándares éticos y normativos vigentes. Finalmente, la relevancia valora la importancia del estudio en términos de impacto clínico, es decir, si la respuesta a la pregunta puede contribuir a mejorar la calidad de vida de los pacientes o a optimizar la práctica médica. Estos criterios, aplicados de forma sistemática, permiten al investigador determinar si la pregunta formulada tiene el potencial de generar evidencia sólida y útil para la toma de decisiones clínicas.

Los escenarios derivados de la búsqueda de literatura en internet pueden variar ampliamente. En un primer escenario, al buscar información sobre la pregunta formulada, se puede encontrar una gran cantidad de estudios, revisiones sistemáticas y metaanálisis que respalden la idea; en este caso, se

tiene un amplio soporte empírico para realizar una revisión sistemática o incluso un ensayo controlado aleatorizado. Por ejemplo, si existen numerosos estudios sobre la relación entre calidad del sueño y control de la presión arterial, el diseño ideal podría implicar un metaanálisis que sintetice estos hallazgos, respetando los criterios FINER de manera robusta.

En otro escenario, el tema podría ser demasiado novedoso y la literatura existente ser muy limitada, lo cual llevaría a considerar diseños exploratorios o estudios de casos, en donde la pregunta se formulase inicialmente como una investigación preliminar que sirva para generar hipótesis futuras, reconociendo la necesidad de ampliar el conocimiento antes de lanzar ensayos a gran escala.

Consideremos, por ejemplo, el caso ideal en el que la pregunta clínica cumple con todos los criterios FINER: supongamos que se plantea investigar "¿Cuál es el impacto de la intervención de mindfulness en pacientes con dolor lumbar crónico refractario a tratamientos convencionales?". Aquí, la factibilidad se garantiza con herramientas disponibles para evaluar el dolor y la atención psicológica, el tema resulta interesante y novedoso en la medida en que solo hay estudios aislados, la intervención es ética y relevante para un problema clínico frecuente. Este escenario ideal promueve el establecimiento de objetivos claros, como evaluar la reducción del dolor en un período determinado, y justificar la realización de un ensayo clínico controlado.

Paralelamente, el establecimiento de objetivos en la formulación de preguntas clínicas se beneficia del uso de la taxonomía de Bloom, que organiza los niveles de aprendizaje y análisis en la educación. Esta taxonomía clasifica las competencias cognitivas en niveles jerárquicos, desde el conocimiento básico y la comprensión hasta el análisis, la síntesis y la evaluación (ver figura 3). En la práctica clínica, al formular objetivos de investigación o de aprendizaje para el equipo médico, se puede diseñar cada objetivo según estos niveles; por ejemplo, se puede requerir que el profesional "identifique" (nivel de conocimiento) los factores asociados al insomnio en pacientes hipertensos, que "analice" (nivel de análisis) la interrelación entre dichos factores y la adherencia terapéutica, y que "evalúe" (nivel de evaluación) la eficacia de una intervención basada en mindfulness en la mejora del sueño.

Figura 3. *Pirámide de la taxonomía de Bloom.*



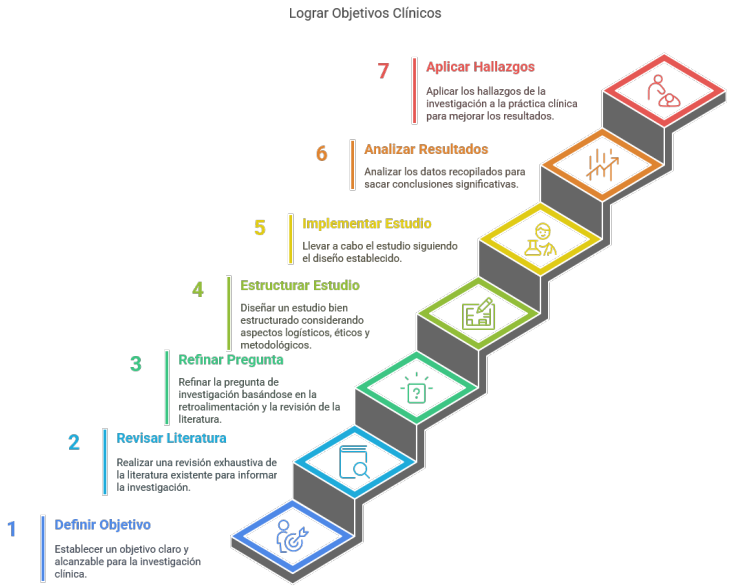
En la aplicación práctica de la taxonomía de Bloom, un objetivo a nivel de "aplicación" podría ser: "Implementar un programa piloto de mindfulness para mejorar la calidad del sueño en pacientes con hipertensión, documentando los cambios clínicos a lo largo de tres meses". Este objetivo se deriva de un conocimiento previo y de la generación de nuevas preguntas, integrando tanto la experiencia clínica como la evidencia científica disponible. Asimismo, la utilización de un objetivo en este nivel permite organizar la intervención en fases, establecer indicadores de éxito y facilitar la transferencia de resultados a la práctica clínica diaria.

La formulación de preguntas clínicas también implica un análisis crítico y sistemático de la literatura previa. Esto abarca la búsqueda estructurada de evidencias mediante bases de datos como PubMed, Cochrane y otras fuentes especializadas, utilizando criterios de inclusión y exclusión rigurosos. Por ejemplo, al investigar la relación entre calidad del sueño y eficacia del tratamiento antihipertensivo, se deben definir variables claras, identificar estudios relevantes y evaluar su consistencia, confiabilidad y aplicabilidad, todo ello en un proceso ordenado que respalde la pregunta inicial.

El proceso metodológico implica, además, la consideración de estrategias de retroalimentación entre la generación de ideas y la revisión de la literatura. Así, una idea inicial derivada de la lluvia de ideas o el pensamiento lateral puede ser refinada tras la búsqueda en internet, donde se encuentran estudios que confirman, contradicen o amplían el conocimiento previo. Este proceso iterativo es fundamental para alcanzar preguntas clínicas sólidas y objetivos específicos, los cuales permitan diseñar investigaciones con bases robustas y orientadas a resultados prácticos.

Una vez que la pregunta clínica se ha definido y se han establecido los objetivos específicos, la estructura del estudio debe considerar aspectos logísticos, éticos y metodológicos. Esto incluye la definición de la población de estudio, la intervención a evaluar, los métodos de recolección de datos y los criterios de análisis estadístico, siempre enmarcados en los principios FINER y en los niveles de la taxonomía de Bloom. Este enfoque integrado no solo garantiza la relevancia de la investigación, sino que además sienta las bases para la mejora continua en la práctica clínica (ver figura 4).

Figura 4. *Escalera hacia los Objetivos Clínicos. Elaborado con napkin: https://www.napkin.ai/es/*



Nota. Etapas del Proceso de Investigación Médica Aplicada. La figura representa de forma escalonada las siete etapas clave para lograr objetivos clínicos mediante investigación aplicada. Desde la definición clara del objetivo hasta la implementación de los hallazgos en la práctica clínica, este esquema visual destaca el carácter secuencial y progresivo del proceso investigativo. Cada peldaño enfatiza aspectos metodológicos fundamentales, promoviendo un enfoque sistemático que vincula la generación de conocimiento con la mejora de los resultados en salud.

2.3. Redactando la pregunta problema

Comprender la relación de causa y efecto es fundamental para la investigación clínica y científica, ya que permite distinguir entre relaciones verdaderas y meras coincidencias. En este contexto, identificar correctamente la variable dependiente (Y, la respuesta, desenlace o el fenómeno que queremos explicar) y la variable independiente (X, la causa o factor que se sospecha influya sobre Y) resulta esencial para establecer relaciones causales. Esto implica no solo observar que dos variables están asociadas, sino también determinar si una realmente provoca cambios en la otra.

Por ejemplo, supongamos que se investiga el efecto de una intervención de ejercicio físico (variable X) sobre la presión arterial (variable Y) en pacientes hipertensos. Si se observa que, en promedio, la presión arterial disminuye cuando los pacientes realizan ejercicio de manera regular, se puede plantear la hipótesis de que el ejercicio tiene un efecto causal sobre la reducción de la presión arterial. Sin embargo, es necesario investigar y descartar otros factores, ya que la mera correlación no implica causalidad; puede tratarse de una coincidencia o de la influencia de variables de confusión.

La distinción entre causalidad y casualidad es vital para formular una pregunta problema precisa. Una relación causal implica que un cambio en la variable X produce un cambio en la variable Y, mientras que una relación de casualidad es simplemente una asociación sin fundamento en un mecanismo que explique la modificación de Y a partir de X. Por ejemplo, observar que la incidencia de gripe aumenta en invierno (Y) y al mismo tiempo se incrementa el consumo de chocolate caliente (X) no significa que el chocolate caliente cause la gripe, sino que ambos pueden estar relacionados con un tercer factor, como las bajas temperaturas.

Partiendo de este entendimiento, al plantear una pregunta problema se debe tener claro qué variable se desea explicar (Y) y cuál es el factor que se pretende investigar como posible causa (X). Esto es la base para definir con precisión la hipótesis de la investigación. Conocer el fenómeno de causalidad permite estructurar una pregunta que realmente aborde la relación directa entre ambos fenómenos, eliminando la confusión con relaciones espurias o fortuitas.

Un ejemplo para ilustrar este concepto podría ser el siguiente: un investigador quiere saber si la cantidad de horas de sueño (X) afecta el rendimiento académico (Y) en estudiantes universitarios. Si, al analizar los datos, se observa que los estudiantes que duermen al menos ocho horas tienen mejores calificaciones que aquellos que duermen menos, es posible sugerir una relación causal, siempre y cuando se hayan controlado otras variables confusoras, como el estrés o la carga de estudios. Este ejemplo ayuda a visualizar cómo se debe plantear la pregunta problema basándose en la relación entre X e Y.

La fórmula sintáctica básica para construir una pregunta problema que aborde una relación causal es la siguiente:

¿Cuál es el efecto de [variable X*]*
sobre [variable Y*] en [contexto o población] y [temporalidad]*?

Siguiendo este formato, el ejemplo anterior se formularía de la siguiente manera: "¿Cuál es el efecto de la cantidad de horas de sueño sobre el rendimiento académico en estudiantes universitarios de medicina desde 2024 a 2025?" Esta estructura permite que la pregunta sea clara, específica y orientada a identificar o medir una relación de causa y efecto, lo cual es indispensable para el desarrollo de estudios rigurosos y de alta calidad.

2.4. De la Idea de Investigación a los Objetivos

Para comprender cómo formular Hipótesis nula y alternativa a partir de una idea de investigación, conviene primero aclarar las diferencias entre conceptos como hipótesis, teoría, ley y hecho, y luego ver cómo, a partir de la pregunta de investigación, podemos definir nuestros objetivos empleando los verbos adecuados de la taxonomía de Bloom. A continuación, se presenta un paso a paso que integra estos elementos de manera pedagógica:

1. Conceptos fundamentales

- **Hipótesis**: es una suposición fundada en alguna evidencia u observación previa que busca explicar un fenómeno. Está *en proceso de validación* y puede terminar siendo confirmada o refutada.
- **Teoría**: se consolida cuando una hipótesis, a lo largo del tiempo, acumula suficiente respaldo empírico y logra explicar un amplio conjunto de observaciones. Aun así, en ciencia, las teorías siempre están abiertas a ajustes si nuevas evidencias las contradicen o las perfeccionan.
- **Ley**: es un enunciado que ha sido validado de manera muy robusta, y que suele describir comportamientos reproducibles con alto grado de certeza (por ejemplo, las leyes de Mendel en genética). Aun así, ninguna ley es inmutable; simplemente es lo más cercano a la verdad disponible según la evidencia actual.
- **Hecho**: alude a una observación empírica ampliamente comprobada y aceptada, pero que en ciencia también puede revisarse o ser reinterpretada si emergen nuevos datos.

2. De la idea de investigación a la pregunta de investigación

Comenzamos con una *idea de investigación*, que puede ser fruto de la observación clínica, de una revisión de la literatura científica o incluso de la identificación de un vacío de conocimiento. Por ejemplo, supongamos que un equipo de salud sospecha que la *práctica de ejercicio físico regular* (variable X) influye positivamente en la *calidad del sueño* (variable Y) en adultos con insomnio crónico.

La *pregunta de investigación* podría formularse siguiendo la estructura de causa-efecto:

"*¿Cuál es el efecto de la práctica regular de ejercicio físico sobre la calidad del sueño en adultos con insomnio crónico?*"

3. Formulación de la hipótesis nula (H_0) y la hipótesis alternativa (H_a) Una vez definida la pregunta de investigación, se redactan dos hipótesis opuestas:

- **H_0**: establece que *no* existe relación o diferencia significativa entre las variables. Siguiendo el ejemplo, la **H_0** sería:

• "La práctica de ejercicio físico regular *no* influye en la calidad del sueño en adultos con insomnio crónico."

- **H_a**: plantea que *sí* existe relación o diferencia significativa. En el ejemplo, la hipótesis alternativa sería:

"La práctica de ejercicio físico regular *mejora* la calidad del sueño en adultos con insomnio crónico."

4. El contraste entre H_0 y H_a conducirá al investigador a diseñar un estudio (observacional o experimental) para recopilar datos y evaluar la validez de las hipótesis.

5. Rol de la hipótesis y el camino hacia teorías o leyes

- Si la H encuentra suficiente respaldo empírico (en distintos estudios, con metodologías rigurosas, reproducibles y con evidencia estadística sólida), y consigue explicar un fenómeno de forma coherente, se consolidará como una teoría.

- Una teoría de muy amplia aceptación y fuertemente validada podría en el futuro, tras años de confirmación, adquirir el estatus de ley; aunque en muchas disciplinas, especialmente en biomedicina y en ciencias sociales, tal salto es inusual debido a la complejidad y variabilidad de los fenómenos estudiados.

6. Delimitación de objetivos empleando la Taxonomía de Bloom
Una vez que se define la hipótesis y la pregunta de investigación, el siguiente paso es formular objetivos claros que orienten la recolección y el análisis de datos. Aquí resulta útil la Taxonomía de Bloom, la cual propone *niveles cognitivos* que van desde lo más básico (recordar, entender) hasta niveles superiores (analizar, evaluar, crear). Cada nivel se asocia con verbos de acción específicos, útiles para redactar objetivos de aprendizaje o de investigación. Por ejemplo:

- **Recordar (nivel básico)**
  - Objetivo ejemplo: "*Recordar* los principales criterios de diagnóstico de insomnio crónico."
  - Verbo clave: Recordar, enumerar, describir.
- **Entender (nivel intermedio)**
  - Objetivo ejemplo: "*Explicar* de qué manera los factores fisiológicos y conductuales pueden afectar la calidad del sueño."
  - Verbos clave: explicar, interpretar, resumir.
- **Aplicar (nivel intermedio)**
  - Objetivo ejemplo: "*Aplicar* un protocolo de ejercicio físico durante cuatro semanas en un grupo de adultos con insomnio para recopilar datos sobre su calidad del sueño."
  - Verbos clave: aplicar, implementar, emplear.
- **Analizar (nivel avanzado)**
  - Objetivo ejemplo: "*Analizar* estadísticamente los resultados obtenidos en el grupo de intervención frente al grupo control."
  - Verbos clave: analizar, comparar, diferenciar.
- **Evaluar (nivel avanzado)**
  - Objetivo ejemplo: "*Evaluar* la efectividad del programa de ejercicio físico en la mejora de la calidad del sueño, considerando variables como edad y comorbilidades."
  - Verbos clave: evaluar, juzgar, criticar.
- **Crear (nivel más elevado)**
  - Objetivo ejemplo: "*Diseñar* una nueva guía de recomendaciones de salud para el manejo integral del insomnio que incluya rutinas de ejercicio basadas en los hallazgos del estudio."
  - Verbos clave: crear, diseñar, proponer, formular.

Ejemplo de paso a paso

- **Idea de investigación**: Se sospecha que el ejercicio físico regular mejora la calidad del sueño en adultos con insomnio crónico.
- **Pregunta de investigación**: "¿Cuál es el efecto del ejercicio físico regular en la calidad del sueño de adultos con insomnio crónico?"
- **H_0**: "El ejercicio físico regular no influye en la calidad del sueño."
- **H_a**: "El ejercicio físico regular mejora la calidad del sueño."
- **Objetivos (Taxonomía de Bloom)**:

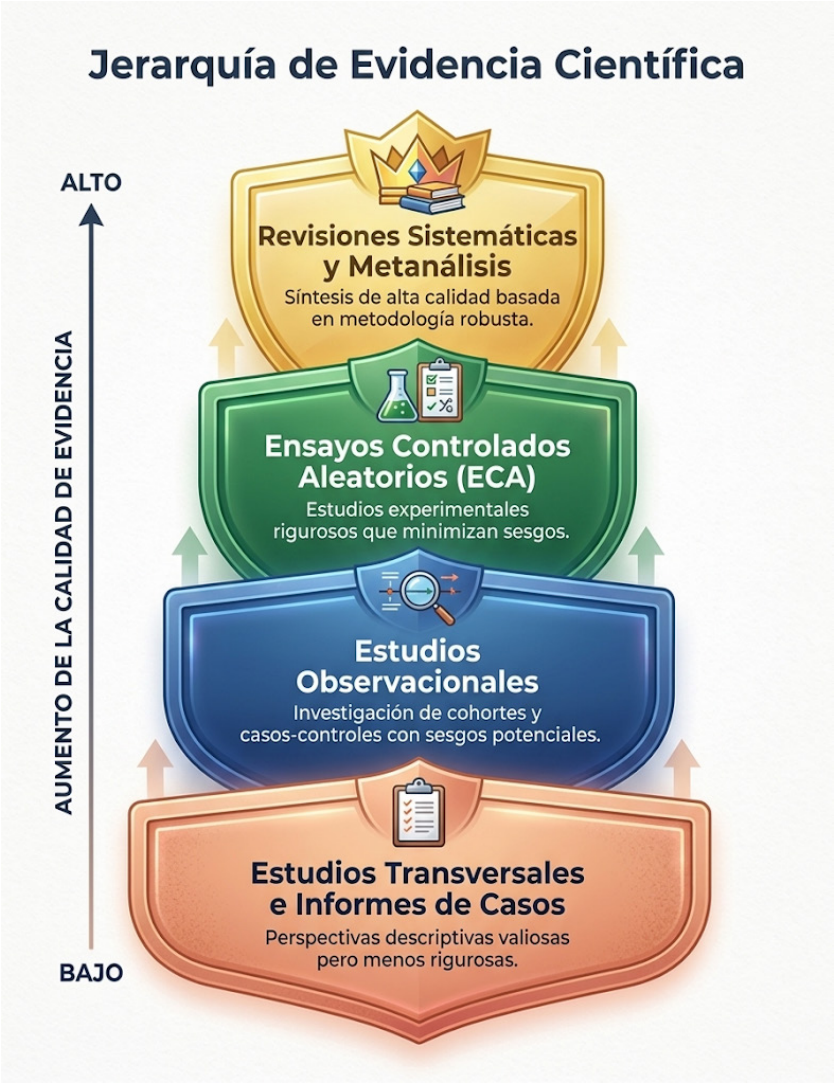
- *Recordar* los criterios diagnósticos de insomnio crónico (nivel "recordar").
- *Aplicar* un protocolo de ejercicio a un grupo de pacientes con insomnio (nivel "aplicar").
- *Analizar* los resultados pre y postintervención mediante estadística descriptiva e inferencial (nivel "analizar").
- *Evaluar* la eficacia del programa comparando los datos de sueño en el grupo experimental y el control (nivel "evaluar").
- *Diseñar* recomendaciones para la práctica clínica en base a los hallazgos (nivel "crear").

2.5. La Pirámide de Evidencia en la Investigación Clínica: del Laboratorio a la Síntesis Sistemática

La pirámide de evidencia es una representación jerárquica de los distintos tipos de estudios de investigación, ordenados de menor a mayor rigor metodológico y capacidad para establecer relaciones causales. Va desde investigaciones *in vitro* en la base, hasta revisiones sistemáticas y metanálisis en la cúspide, o incluso en su versión más actualizada en estudios *umbrella* (ver figura 5). Los estudios *umbrella*, también conocidos como revisiones sistemáticas de revisiones sistemáticas o metanálisis, representan el nivel más alto en esta jerarquía. Su propósito es sintetizar y comparar los resultados de múltiples revisiones sistemáticas sobre un mismo tema o conjunto de

intervenciones, permitiendo una visión panorámica de la evidencia disponible. Esta metodología es especialmente útil en campos como la medicina preventiva, la farmacología o la salud pública, donde las decisiones clínicas requieren integrar una gran cantidad de información proveniente de distintas fuentes de alta calidad.

Figura 5. *Niveles de evidencia científica*. *Elaborado con Gemini PRO.*



A continuación, vamos a describir cada nivel de la pirámide con un ejemplo clínico sencillo, utilizando la misma idea de investigación (variable X y variable Y) pero aplicado a distintos diseños, y mostrando preguntas progresivas según la Taxonomía de Bloom.

Investigaciones in vitro

En la base de la pirámide se ubican los estudios *in vitro*, es decir, experimentos realizados en laboratorio a nivel celular o molecular, fuera del organismo vivo.

- **Ejemplo de la idea X–Y**: *Efecto de un extracto natural* (X) *sobre células endoteliales* (Y) para evaluar si podría mejorar la función vascular.
- **Tipo de pregunta (nivel "recordar" o "comprender" en Bloom)**:
  - "¿Qué cambios observamos en las células endoteliales cuando se exponen al extracto natural?" (Se busca describir reacciones básicas, por ejemplo, producción de óxido nítrico.)

Investigaciones con animales

Siguen los estudios en modelos animales, que permiten ensayos preclínicos para explorar efectos fisiológicos y obtener una aproximación inicial a la seguridad y eficacia de una intervención.

- **Ejemplo de la idea X–Y**: *Efecto de un suplemento alimenticio* (X) *sobre la presión arterial* (Y) *en ratas hipertensas*.
- **Tipo de pregunta (nivel "comprender" o "aplicar" en Bloom)**:
  - "¿Cómo *varía la presión arterial en ratas hipertensas que reciben el suplemento durante 8 semanas*?" (Se inicia la aplicación de la intervención, pero aún no en humanos.)

Informes de Casos (Caso Clínico)

Describen en detalle la evolución de uno o unos pocos pacientes con un mismo diagnóstico o tratamiento, sin incluir grupo comparativo formal.

- **Ejemplo de la idea X–Y**: *Informe del caso de un paciente de 60 años con hipertensión descontrolada que mejora tras iniciar un programa de ejercicio físico* (X), y se observa *reducción de la presión arterial* (Y) en pocos meses.

- **Tipo de pregunta (nivel "aplicar" o "analizar" en Bloom)**:

  - "*¿Cómo se comportó la presión arterial de un paciente específico al aplicar un programa personalizado de ejercicio?*"

(Ofrece indicios o hipótesis iniciales, pero no permite generalizar ni confirmar causalidad.)

Estudios *Transversales*

Observan la prevalencia o asociación entre variables en una población en un momento puntual, sin seguir a los participantes en el tiempo.

- **Ejemplo de la idea X–Y**: *Relación entre el consumo diario de sodio* (X) *y la presión arterial* (Y) *en adultos de un centro de salud*. Se aplican encuestas y medición de presión arterial en un único corte temporal.

- **Tipo de pregunta (nivel "analizar" en Bloom)**:

  - "*¿Cuál es la proporción de adultos con ingesta elevada de sodio que presentan presión arterial alta en este momento?*" (Permite ver asociación, pero no confirma quién causa a quién.)

Estudios *de Casos y Controles*

Seleccionan un grupo con la condición de interés (casos) y otro sin la condición (controles), y se investiga retrospectivamente la exposición a la variable X.

- **Ejemplo de la idea X–Y**: *Comparar adultos hipertensos (casos) y no hipertensos (controles) para evaluar si practicaban ejercicio regular* (X).

- **Tipo de pregunta (nivel "analizar" o "evaluar" en Bloom)**:

  - "*¿Existen diferencias significativas en la frecuencia de ejercicio físico entre las personas con hipertensión y las que no la tienen?*" (Es un estudio retrospectivo que sugiere asociación y da pistas sobre causalidad, pero no la asegura al 100 %).

Estudios de Cohorte

Siguen a un grupo de personas a lo largo del tiempo para ver si desarrollan el evento de interés (Y), comparando la exposición (X) entre quienes sí y no desarrollan dicho evento.

- **Ejemplo de la idea X–Y:** *Seguir durante 5 años a adultos sin hipertensión (Y) y ver si la aparición de hipertensión se asocia con sus niveles de actividad física* (X).

- **Tipo de pregunta (nivel "evaluar" en Bloom)**:

  - "*¿Cuál es la incidencia de hipertensión en personas que practican ejercicio regular frente a quienes no lo hacen?*" (Diseño prospectivo que permite aproximarse más a la causalidad.)

Ensayo Controlado Aleatorizado (ECA)

Consiste en asignar al azar a los participantes a un grupo que recibe la intervención (X) y a un grupo control, para posteriormente comparar resultados (Y).

- **Ejemplo de la idea X–Y**: "*¿Cuál es el efecto de un programa de ejercicio supervisado de 12 semanas sobre la presión arterial comparado con cuidados habituales en adultos hipertensos?*"

- **Tipo de pregunta (nivel "evaluar" o incluso "crear" en Bloom)**:

  - Se pueden plantear objetivos más complejos y definitivos:

1. *Evaluar* la efectividad del programa midiendo la reducción de la presión arterial sistólica y diastólica.
2. *Diseñar* posibles guías clínicas para incorporar la intervención al cuidado estándar, si los resultados son positivos.

Revisión Sistemática y Metanálisis

Son análisis que recopilan y sintetizan la evidencia de múltiples estudios primarios, siguiendo métodos rigurosos y transparentes.

- **Ejemplo de la idea X–Y**: "*Revisión sistemática y metanálisis de ensayos controlados aleatorizados que evalúen el efecto del ejercicio físico sobre la presión arterial*".

- **Tipo de pregunta (nivel "crear" en Bloom)**:

  - "*¿Cuál es la evidencia acumulada acerca del impacto del ejercicio en la presión arterial y cuán consistente es en distintas poblaciones?*" (Se busca integrar hallazgos de numerosos estudios para emitir conclusiones robustas y guiar la práctica clínica.)

2.6. Perspectivas Epistemológicas: Empirismo, Racionalismo, Falsabilidad y el Principio de Parsimonia

La filosofía y la ciencia se entrelazan en un diálogo que busca desvelar los misterios del universo y, en particular, comprender la complejidad del cuerpo humano y de sus enfermedades. Desde tiempos antiguos se han desarrollado diversas perspectivas epistemológicas que fundamentan el método científico, permitiendo que la búsqueda del conocimiento se erija sobre cimientos sólidos y críticos.

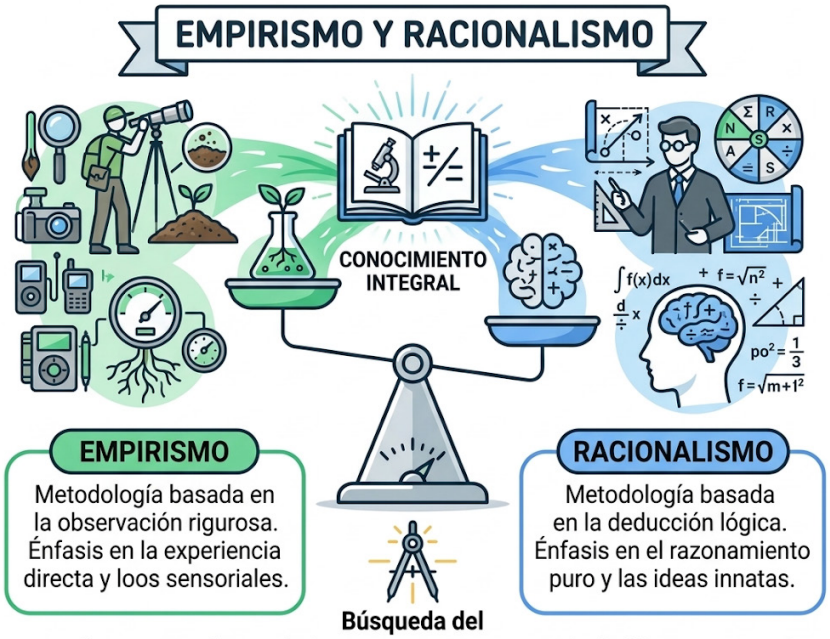
La epistemología, cuyo término proviene del griego *epistêmê*, es la disciplina encargada de estudiar la naturaleza, el origen y los límites del conocimiento. Esta rama no solo se cuestiona cómo llegamos a saber lo que sabemos, sino que examina la validez de nuestras creencias y teorías. Para el investigador, comprender estos fundamentos es vital, ya que orienta la elección de métodos, la interpretación de los datos y la formulación de hipótesis, asegurando que las conclusiones sean sólidas y reproducibles.

El conocimiento se define como la suma de datos verificados y la experiencia acumulada a lo largo del tiempo. Este conjunto abarca hechos empíricamente comprobados y teorías que, gracias a su capacidad predictiva, aportan comprensión sobre fenómenos complejos. Un investigador se enfrenta constantemente al desafío de garantizar la fiabilidad y exactitud de la información, sometiéndola a un escrutinio riguroso que diferencie lo comprobado de lo que aún requiere validación.

Dentro de la epistemología se distinguen dos corrientes fundamentales: el empirismo y el racionalismo. El empirismo enfatiza la importancia de la experiencia directa y la observación, base de experimentos y estudios clínicos que, por ejemplo, miden la variación de la presión arterial en respuesta a un tratamiento. En contraposición, el racionalismo confía en la deducción y

el razonamiento lógico para descubrir relaciones causales; así, modelos matemáticos basados en principios fisiológicos pueden predecir el impacto de cambios en la dieta sobre el metabolismo (ver figura 6).

Figura 6. *Empirismo y Racionalismo. Elaborado con Gemini PRO.*



El escepticismo, otro componente esencial, impulsa al investigador a cuestionar cualquier afirmación, sometiéndola a pruebas que puedan refutarla. Inspirado en la tradición griega, este enfoque garantiza que solo las hipótesis robustas sobrepasen el riguroso proceso de validación.

La falsabilidad, acuñada por Karl Popper, complementa el escepticismo, al exigir que toda teoría pueda ser refutada mediante la evidencia; de esta manera, una hipótesis que afirme que "todos los pacientes mejorarán con un tratamiento" es científica solo si existe la posibilidad de demostrar lo contrario en determinadas condiciones.

El concepto de pseudociencia surge en este contexto, distinguiendo aquellas afirmaciones que utilizan lenguaje científico sin basarse en evidencia empírica ni adherirse a métodos verificables. Un ejemplo en medicina es la homeopatía, cuyos principios no cumplen con el requisito de falsabilidad, lo que la desacredita frente a tratamientos probados mediante ensayos clínicos controlados.

Para abordar fenómenos complejos, se emplean enfoques complementarios: el reduccionismo y el holismo. El reduccionismo, que consiste en descomponer un sistema en sus componentes básicos, se ejemplifica cuando se estudia la función del corazón a nivel celular, analizando cada señal y reacción bioquímica. Sin embargo, para comprender completamente el impacto de un tratamiento en un paciente, es necesario adoptar un enfoque holístico, que integre también factores psicológicos y sociales.

La noción de paradigma y sus cambios, introducida por Thomas Kuhn, explica cómo la ciencia progresa a través de revoluciones que desafían las concepciones establecidas. Cuando surgen anomalías que el marco actual no puede explicar, se produce un cambio paradigmático, similar a la transición de la teoría geocéntrica a la heliocéntrica, o más recientemente, la incorporación de la terapia génica en el tratamiento de enfermedades hereditarias, que reconfigura el enfoque terapéutico vigente.

Por otro lado, las ideas del postmodernismo han cuestionado la objetividad y la universalidad del conocimiento, sugiriendo que incluso los "hechos" y las "leyes" son construcciones culturales. La hipótesis de *Gaia* de James Lovelock, por ejemplo, propone ver a la Tierra como un organismo vivo, una idea revolucionaria pero objeto de críticas por su falta de precisión empírica. Estas corrientes invitan a reflexionar sobre la naturaleza del conocimiento y a ser conscientes de que la verdad en la ciencia está siempre sujeta a revisión.

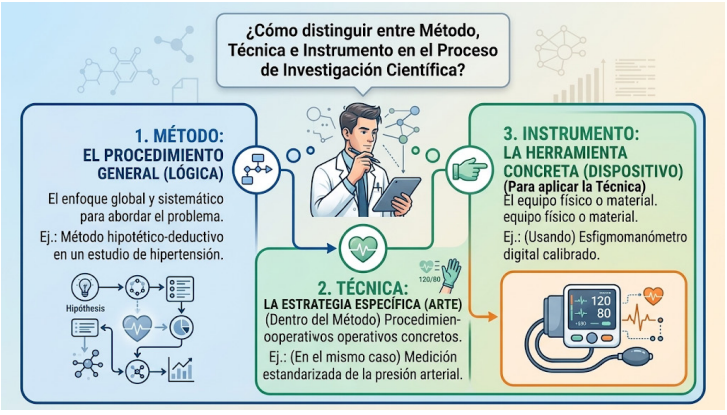
Esta amalgama de enfoques, desde el empirismo y el racionalismo hasta el escepticismo, la falsabilidad, el reduccionismo y el holismo, ofrece al investigador una perspectiva multidimensional que fortalece tanto el proceso de generación como el de evaluación del conocimiento. Al integrar estas ideas, la ciencia, especialmente en el ámbito de la medicina, se mantiene en constante evolución, permitiendo que las prácticas clínicas se actualicen y se adapten a los nuevos descubrimientos de manera crítica y fundamentada.

2.7 El Método Científico: la Ruta hacia el Conocimiento

El método científico se erige como el pilar fundamental en la búsqueda del saber, un conjunto de pasos estructurados que transforman la observación casual en conocimiento riguroso y verificable. En la medicina, este método permite que cada hallazgo se someta a un proceso sistemático de formulación, experimentación y análisis, lo que reduce la incertidumbre y aporta solidez a las decisiones terapéuticas. Los principios del método, basados en la integración de enfoques inductivos y deductivos, han permitido desarrollar protocolos clínicos que se actualizan de acuerdo con nuevas evidencias empíricas.

Dentro de este proceso, es esencial distinguir entre método, técnica e instrumento, conceptos que, aunque están interrelacionados, cumplen funciones distintas. El método es el camino o procedimiento general que se sigue para investigar un fenómeno, por ejemplo, el método hipotético-deductivo que se utiliza para formular y contrastar hipótesis en estudios clínicos. En cambio, la técnica se refiere a las estrategias o procedimientos específicos empleados para llevar a cabo cada paso del método; por ejemplo, la medición de la presión arterial mediante esfigmomanómetro es una técnica utilizada en estudios de hipertensión. Por último, el instrumento es la herramienta concreta que facilita la aplicación de la técnica, como el propio esfigmomanómetro o, en investigaciones más sofisticadas, un monitor de presión arterial digital conectado a un sistema de registro de datos (ver figura 7).

Figura 7. *Diferencias entre Método, Técnica e Instrumento en el Proceso de Investigación Científica. Elaborado con Gemini PRO.*



La clasificación de la investigación, a su vez, se organiza en función del enfoque y la finalidad del estudio. En medicina, los estudios pueden clasificarse en investigación básica, que explora los mecanismos fundamentales de los procesos biológicos; en investigación clínica, que se centra en la evaluación de tratamientos y diagnósticos en pacientes; y en investigación traslacional, que busca integrar los hallazgos de la investigación básica en la práctica clínica. Por ejemplo, un estudio básico puede investigar a nivel molecular cómo interactúa un fármaco con sus receptores, mientras que uno clínico evaluará la eficacia de dicho fármaco en mejorar la salud de los pacientes, y uno traslacional integrará ambos conocimientos para desarrollar nuevas guías terapéuticas.

La observación es el primer paso en cualquier investigación, y en el campo de la medicina se realiza a través de la recopilación sistemática de datos en el entorno clínico y de laboratorio. Imaginemos que en un centro de salud se observa que numerosos pacientes con hipertensión presentan, además, alteraciones en su estilo de vida. Este patrón conduce al investigador a formular la hipótesis de que una intervención combinada, que incluya un fármaco y cambios en la dieta, puede modular eficazmente la presión arterial.

Una vez formulada la hipótesis, se diseñan los estudios experimentales con grupos control y experimental, utilizando técnicas específicas que permitan la medición de variables de interés. En este caso, se podrían comparar dos grupos de pacientes, uno recibiendo el tratamiento convencional y otro con la intervención combinada, usando técnicas como la medición automatizada de la presión arterial y encuestas sobre hábitos alimenticios. Aquí, tanto el método como las técnicas se apoyan en instrumentos precisos que garantizan la reproducibilidad de los resultados.

El análisis de los datos recogidos se realiza mediante técnicas estadísticas robustas, que permiten determinar si las diferencias observadas entre grupos son significativas. Este paso requiere no solo del dominio de las técnicas estadísticas, sino también del uso de instrumentos informáticos y software especializado que facilitan la interpretación de los resultados. Estos recursos, integrados en un plan metodológico coherente, refuerzan la validez del estudio y aseguran que las conclusiones sean fiables y aplicables a la práctica clínica.

El proceso de producción del conocimiento en la investigación médica se nutre además de una rigurosa evaluación crítica de la evidencia. Esto incluye la revisión sistemática de estudios previos, los metaanálisis, y la elaboración de guías clínicas, que permiten sintetizar la información disponible. Un investigador, por ejemplo, puede recurrir a revisiones sistemáticas publicadas en la *Cochrane Library* para comparar la efectividad de dos intervenciones en pacientes con insuficiencia cardíaca, garantizando que la mejor evidencia respalde sus conclusiones.

Los avances metodológicos y las técnicas de análisis también requieren que los instrumentos sean calibrados y validados. En la investigación médica, el uso de dispositivos de última generación, como los monitores de ritmo cardíaco conectados a sistemas informáticos, es crucial para obtener datos precisos, lo que a su vez impacta en la calidad de la evidencia acumulada y en la elaboración de protocolos clínicos efectivos.

En el mundo de la investigación, las diferencias entre método, técnica e instrumento son fundamentales, ya que cada uno contribuye en forma única al proceso investigativo. Este enfoque permite una mayor comprensión de los mecanismos de la enfermedad y, en última instancia, el desarrollo de intervenciones más efectivas y personalizadas en medicina.

La clasificación de la investigación en básica, clínica y traslacional permite que se aborde la complejidad de la salud desde distintos ángulos, enriqueciendo el cuerpo de conocimientos disponibles. Cada tipo de investigación aporta datos valiosos: desde el nivel molecular hasta la aplicación en un entorno real, lo que facilita que los descubrimientos se conviertan en mejoras palpables en el cuidado del paciente.

A lo largo de este proceso, la integración del método científico y la MBE se destaca como un camino práctico y riguroso para resolver problemas clínicos. La observación inicial se transforma en una hipótesis cuidadosamente elaborada, que es puesta a prueba mediante experimentos controlados, y finalmente se somete a un análisis crítico que orienta la toma de decisiones en el ámbito médico.

2.8. Formulación de Preguntas Clínicas y Estrategias de Búsqueda en Medicina Basada en la Evidencia

La formulación de preguntas clínicas es el primer paso para transformar la experiencia del consultorio en investigación aplicada. Un enfoque clásico es el marco PICO, que divide la interrogante en cuatro elementos: Paciente o Problema, Intervención, Comparación y Resultado. Por ejemplo, en el caso de evaluar el beneficio de administrar carvedilol en niños con insuficiencia cardíaca, la pregunta se estructura de la siguiente manera: ¿En niños con insuficiencia cardíaca (P), mejora la administración de carvedilol (I) la función cardíaca (O) en comparación con la ausencia de dicho tratamiento (C)? Esta estructura permite que el investigador delimite con claridad qué se espera y cuáles serán los indicadores de éxito o fracaso de la intervención.

Cuando el marco incluye el factor tiempo, se utiliza PICOT, agregando una dimensión temporal que resulta esencial en muchos estudios clínicos. Por ejemplo, una pregunta formulada en este modelo podría ser: ¿En mujeres embarazadas con riesgo de preeclampsia (P), ¿reduce la administración de aspirina a dosis baja (I), en comparación con el placebo (C), la incidencia de preeclampsia (O) durante el tercer trimestre (T)? Esta formulación no solo define quiénes, qué y cómo, sino también cuándo se medirá el efecto, lo que es particularmente útil para estudios longitudinales.

La búsqueda de la evidencia pertinente comienza identificando fuentes de información que se dividen en tres categorías: primarias, secundarias y terciarias. Las fuentes primarias son los estudios originales, como un ensayo clínico que evalúa la eficacia de carvedilol en la mejora de la función cardíaca en niños, publicado en una revista indexada. Las fuentes secundarias incluyen revisiones sistemáticas y metaanálisis, que sintetizan datos de múltiples estudios primarios, mientras que las fuentes terciarias, como libros de texto y enciclopedias médicas, proporcionan resúmenes generales y se usan para contextualizar la evidencia. Cada tipo de fuente tiene su propio valor y se selecciona dependiendo del grado de detalle y especificidad que se requiera (ver figura 8).

Figura 8. *Proceso de identificación de fuentes de evidencia. Elaborado con napkin: https://www.napkin.ai/es/*



Extraer palabras clave de la pregunta clínica es fundamental para orientar la estrategia de búsqueda. Se identifican los términos esenciales de cada componente del modelo PICO, y se complementan con sinónimos, acrónimos y variantes ortográficas. Por ejemplo, de la pregunta "¿En niños con insuficiencia cardíaca, mejora carvedilol la función cardíaca?", se extraerían términos como "carvedilol", "insuficiencia cardíaca", "función cardíaca", "niños" y sus equivalentes, como "heart failure" y "pediatric" en inglés. Esta extracción no sólo facilita la búsqueda, sino que también asegura que se abarque la mayor cantidad de literatura relevante posible.

El uso de términos controlados, como los *MeSH* (*Medical Subject Headings*) en PubMed y *DeCS* (Descriptores en Ciencias de la Salud) en la BVS, permite afinar la búsqueda con mayor precisión. Al introducir "Heart Failure" en la base de datos de MeSH, por ejemplo, el sistema "explota" el término para incluir todas sus subcategorías relacionadas. De igual manera, en DeCS se utilizará "Insuficiencia Cardíaca" para recuperar todos los estudios indexados bajo ese descriptor, garantizando así una búsqueda sistemática y coherente que abarque la totalidad de la evidencia disponible.

En la tabla 1 se comparan diferentes bases de datos en la búsqueda de literatura científica. Se abordan aspectos esenciales como el contenido que indexa cada base de datos, el tipo de acceso que ofrece (gratuito o mediante suscripción) y las estrategias de búsqueda recomendadas, que incluyen el uso de términos controlados (como MeSH) y búsquedas en texto libre. Además, se explican operadores booleanos y herramientas específicas (por ejemplo, truncamiento, operadores de proximidad y filtros de búsqueda) que permiten personalizar y mejorar los resultados.

Cada base de datos tiene características particulares: PubMed se destaca por su amplia cobertura y acceso gratuito, mientras que Embase y Scopus ofrecen una mayor amplitud en ciertos campos y cobertura de publicaciones europeas o no inglesas. La Cochrane Library se centra en la obtención de evidencia para revisiones sistemáticas, y Web of Science proporciona una perspectiva amplia en múltiples áreas del conocimiento. Con estos ejemplos y definiciones, el investigador puede elegir la estrategia de búsqueda que mejor se adapte a sus necesidades y asegurarse de que su revisión de literatura sea sistemática, completa y precisa.

Tabla 1. *Comparativa de Bases de Datos para la Búsqueda de Literatura en Medicina Basada en la Evidencia*

Base de Datos	Contenido	Acceso	Estrategia de Búsqueda	Ejemplos/ Comentarios
PubMed	Indexa artículos internacionales en medicina, enfermería, odontología, veterinaria, ciencias de la salud y preclínicas; incluye más de 23 millones de citas desde 1946.	Gratuito a través de www.pubmed.gov, administrado por la Biblioteca Nacional de Medicina de EE. UU.	Utiliza Encabezados (MeSH) para identificar términos controlados y búsqueda en texto libre; dispone de opciones avanzadas para combinar búsquedas con operadores booleanos y aplicar filtros (fecha, tipo de publicación, grupo etario).	Ejemplo: ("Insuficiencia Cardíaca"[MeSH] AND carvedilol[tiab]) AND (niño*[tiab] OR pediatr*[tiab]) para recuperar artículos sobre el efecto de carvedilol en niños con insuficiencia cardíaca.
Medline (EBSCOhost)	Es el componente central de PubMed; abarca estudios originales de alta calidad en diversas áreas de la medicina.	Acceso por suscripción a través de proveedores como EBSCO.	Búsqueda con términos MeSH y texto libre; permite el uso de truncamientos, comodines y operadores booleanos para refinar resultados	Ejemplo: ("Infarto de Miocardio" OR "Cardiac Infarction") AND (enfermer*[tiab]) para identificar estudios sobre el papel de la enfermería en el manejo del infarto de miocardio.
Embase (Ovid)	Base de datos biomédica y farmacológica que indexa artículos y resúmenes de congresos; contiene más de 6 millones de registros (desde 1974) y, adicionalmente, registros clásicos.	Acceso mediante suscripción a través de proveedores como Ovid (Wolters Kluwer).	Utiliza términos de Emtree (equivalentes a MeSH) y permite búsquedas en texto libre; ofrece funciones de truncamiento, comodines y operadores de proximidad (ADJn) para definir relaciones precisas entre términos.	Ejemplo: (child* ADJ3 "insuficiencia cardíaca") para buscar artículos donde "child" y "insuficiencia cardíaca" aparecen cercanos.

The Cochrane Library	Colección de seis bases de datos que incluyen revisiones sistemáticas (Cochrane Database of Systematic Reviews – CDSR), resúmenes (DARE), ensayos controlados (CENTRAL) y otros.	Gratuito en el Reino Unido a través de www.thecochranelibrary.com.	Permite búsquedas utilizando términos MeSH y búsquedas en texto libre en el Search Manager; admite la combinación de búsquedas y la aplicación de limitaciones por fecha u otros criterios.	Ejemplo: Uso de MeSH para "Cáncer de Mama" y combinar búsquedas con operadores, como (#1 OR #2) AND #3, donde cada número representa una búsqueda previa en el Search Manager.
Scopus	Indexa revistas revisadas por pares en ciencia, tecnología, medicina y ciencias sociales, además de algunos libros y actas de congresos; cuenta con aproximadamente 32 millones de registros.	Acceso mediante suscripción a través de Elsevier.	Búsqueda en texto libre en títulos, resúmenes y palabras clave; permite el uso de truncamiento (*) y operadores de proximidad (W/n) para establecer relaciones entre términos; ofrece la opción de combinar búsquedas mediante números en el historial.	Ejemplo: {salud mental} W/2 política para hallar artículos en los que "salud mental" y "política" se encuentren cercanos en el texto.
Web of Science	Incluye artículos de revistas, actas de congresos y capítulos de libros en ciencias, ciencias sociales, artes y humanidades; abarca más de 55 millones de registros desde 1900.	Acceso mediante suscripción a través de Thomson Reuters.	Búsqueda en texto libre; utiliza el truncamiento (*) y comodines (por ejemplo, $ para variaciones), operadores de proximidad (NEAR/n) y la función de combinación de búsquedas mediante el Search History; permite filtrar por año o idioma.	Ejemplo: "déficit de atención" NEAR/2 "hiperactividad" para encontrar documentos en los que ambas frases aparezcan cerca una de la otra.

La tabla 2 presenta una serie de operadores booleanos y símbolos que resultan esenciales para elaborar algoritmos de búsqueda precisos en bases de datos científicas. Cada término se acompaña de una definición clara y un ejemplo práctico, lo que permite comprender cómo combinarlos para optimizar la recuperación de información relevante.

Tabla 2. *Operadores Booleanos y Símbolos de Búsqueda en Bases de Datos*

Término/ Operador	Definición	Ejemplo
AND	Se utiliza para unir palabras de conceptos diferentes, reduciendo y afinando la búsqueda al exigir que ambos términos estén presentes.	*Salbutamol* AND *asma*: busca artículos que contengan tanto "salbutamol" como "asma".
OR	Recupera documentos que contengan al menos uno de los términos conectados, ampliando la búsqueda.	*Salbutamol* OR *Ventolín*: incluye artículos que mencionen cualquiera de los nombres, ya que un fármaco puede referirse por su nombre genérico o de marca.
NOT	Excluye documentos que contengan el término que sigue, restringiendo la búsqueda al eliminar resultados indeseados.	*Asma* NOT *ejercicio*: excluye resultados que contengan "ejercicio", aun si se menciona "asma".
( )	Los paréntesis se usan para agrupar términos y establecer el orden lógico de la búsqueda, similar a las operaciones matemáticas.	(*Salbutamol* OR *Ventolín*) AND *asma*: recupera artículos que contengan "asma" y, además, "salbutamol" o "Ventolín".
NEAR/adj#	Requiere que el término siguiente aparezca dentro de una proximidad definida respecto al término precedente; reduce resultados pero aumenta la relevancia.	"*corazón* NEAR/3 *insuficiencia*": encontrará artículos donde "corazón" e "insuficiencia" aparezcan a no más de tres palabras de distancia.
" "	Las comillas se utilizan para buscar una frase exacta, garantizando que las palabras aparezcan en el orden especificado.	"*insuficiencia cardíaca*": solo recupera documentos que contienen exactamente esa frase.

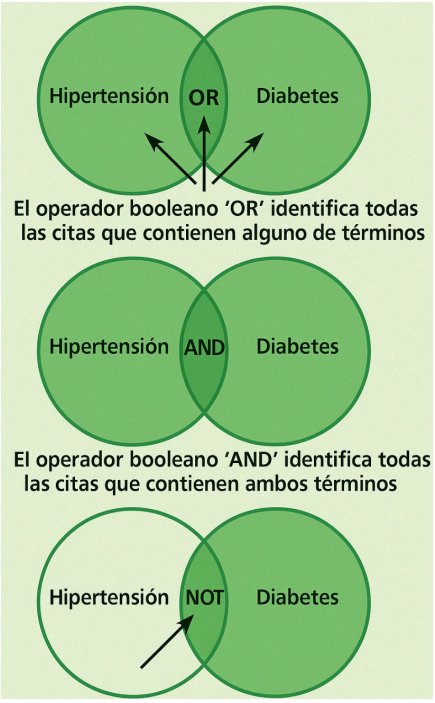
Término/ Operador	Definición	Ejemplo
*	El asterisco actúa como truncamiento o comodín, permitiendo buscar palabras con diferentes terminaciones a partir de una raíz.	*niñ**: recupera "niño", "niños" y "niñez".
XOR	Menos común en búsquedas clínicas; exclusivo OR recupera documentos que contengan uno u otro término, pero no ambos al mismo tiempo.	*Salbutamol* XOR *Ventolín*: encuentra artículos que mencionen únicamente "salbutamol" o solo "Ventolín", según la configuración de la base de datos.
ADJ/adj#	Indica proximidad específica; "adj#" asegura que dos términos aparezcan juntos, separados por un número máximo de palabras definido por "#".	(*ventricular adj3 disfunción*): encuentra documentos donde "ventricular" y "disfunción" aparezcan próximos, no separados por más de tres palabras.
TI	Operador que limita la búsqueda a los títulos de los artículos, focalizando en aquellos documentos cuyo título contenga el término indicado.	TI("*insuficiencia cardíaca*"): retorna documentos donde "insuficiencia cardíaca" aparece en el título.
AB	Se utiliza para buscar términos en el resumen (abstract) de los documentos, ampliando la búsqueda a esta sección.	AB(*carvedilol*): identifica artículos que mencionen "carvedilol" en el abstract.
EXP	Utilizado en bases de datos como PubMed para "explotar" un término MeSH, incluyendo todas sus subcategorías.	EXP "*Insuficiencia Cardíaca*": recupera artículos indexados bajo "Insuficiencia Cardíaca" y todas sus subcategorías.
?	Comodín de carácter único que reemplaza a una sola letra o dígito, útil para buscar variaciones ortográficas.	*p?ediatr**: puede recuperar "pediátrico" o "paediátrico", según la variante utilizada en la literatura.

La construcción de un algoritmo de búsqueda robusto es una habilidad esencial en la práctica de la MBE. Con operadores booleanos como AND, OR y NOT se combinan los términos extraídos de la pregunta. Por ejemplo, un algoritmo sencillo podría ser:

("Heart Failure"[MeSH] OR "cardiac failure"[tiab]) AND ("Carvedilol"[MeSH] OR carvedilol[tiab]) AND (child[MeSH] OR pediatric*[tiab] OR paediatric*[tiab])

Aquí, "[MeSH]" indica el uso del vocabulario controlado, "tiab" se refiere a la búsqueda en título y resumen, y los operadores AND y OR ayudan a incluir sinónimos y a filtrar la información. Además, se pueden incorporar operadores de proximidad, como "adj3" en Ovid MEDLINE, para buscar frases cercanas y unir conceptos con mayor precisión (ver figura 9).

Figura 9. *Uso de operadores booleanos en búsquedas bibliográficas sobre hipertensión y diabetes*. Elaborado con GeminiPRO.



Nota. La figura muestra el funcionamiento de los operadores booleanos OR, AND y NOT aplicados a los términos hipertensión y diabetes en búsquedas científicas. El operador OR amplía la búsqueda incluyendo resultados que contengan cualquiera de los términos; AND restringe la búsqueda a citas que incluyan ambos términos; mientras que NOT excluye resultados que contengan el segundo término, mostrando únicamente aquellos que mencionan el primero.

A continuación, se detalla una tabla que muestra algunos de los términos clave utilizados en la indexación de artículos mediante MeSH y otros componentes asociados, junto con sus definiciones y ejemplos en el ámbito de la medicina (Tabla 3):

Tabla 3. *Organización de términos MeSH*

Categoría	Definición	Ejemplos
Encabezados (descriptores)	Términos principales que representan el concepto central de un tema; son la base para clasificar el contenido de los artículos en las bases de datos.	Peso corporal, insuficiencia cardíaca, asma
Subencabezados (calificadores)	Términos que especifican o describen un aspecto particular del concepto principal, permitiendo una clasificación más detallada.	Diagnóstico, tratamiento, efectos adversos
Registros de Concepto Suplementario	Una lista de más de 200,000 términos que se actualiza semanalmente y que se utiliza para indexar sustancias, protocolos o enfermedades poco comunes, de manera complementaria a los encabezados.	Atorvastatina, protocolo MOPP, mixoma atrial, familiar
Características de la Publicación	Aspectos relativos al tipo de publicación que se indexa, los cuales ayudan a identificar el formato y la naturaleza del documento.	Carta, revisión, ensayo controlado aleatorizado (ECA)

Esta tabla 3 está diseñada para proporcionar una visión práctica y clara de cómo se organizan y se utilizan los términos en bases de datos médicas, especialmente en el contexto de MeSH. Los **encabezados** o descriptores constituyen los términos principales que identifican el concepto central de un artículo, tales como "insuficiencia cardíaca" o "asma". Los **subencabezados** o calificadores, por su parte, permiten especificar aspectos particulares del tema, como por ejemplo "diagnóstico" o "tratamiento".

Los **registros de concepto suplementario** se diferencian porque ofrecen términos adicionales, generalmente de sustancias, protocolos o enfermedades raras, que no están incluidos en los encabezados clásicos, y su actualización semanal garantiza que se incluya la literatura más reciente. Finalmente, las **características de la publicación** se refieren al formato en el que se presenta la información, lo cual es fundamental para clasificar los documentos según su tipo (como revisiones o ensayos clínicos). Esta estructura permite a los investigadores y clínicos realizar búsquedas precisas y sistemáticas, facilitando la identificación de la evidencia más relevante para la práctica médica y la investigación clínica.

Estos pasos, que van desde la formulación de la pregunta utilizando PICO o PICOT hasta la construcción del algoritmo de búsqueda en bases de datos especializadas, constituyen la columna vertebral de la MBE. Al delimitar con exactitud la población, intervención, comparación, resultados y tiempo, el investigador se asegura de orientar la búsqueda hacia estudios relevantes que podrán ser evaluados críticamente y, posteriormente, aplicados en la práctica clínica.

Cada búsqueda debe estar respaldada por la revisión de fuentes primarias, secundarias y terciarias, lo que permite validar los hallazgos a través de múltiples niveles de evidencia. Por ejemplo, los estudios primarios pueden identificar una posible mejora en la función cardíaca en pacientes tratados con carvedilol, mientras que las revisiones sistemáticas consolidan estos hallazgos y los libros de texto contextualizan su relevancia en la práctica diaria.

El dominio de los términos clave, la capacidad para identificar sinónimos y la correcta utilización de las estrategias de búsqueda con operadores booleanos no solo fortalecen la calidad de las evidencias encontradas, sino que también garantizan que cada revisión de literatura sea exhaustiva y reproducible. Este enfoque metodológico aumenta la precisión y reduce la posibilidad de pasar por alto estudios importantes, creando una base sólida para la toma de decisiones en clínica.

Integrar todo este conocimiento en la formulación de preguntas clínicas y la estrategia de búsqueda es, en última instancia, el primer paso hacia la implementación de intervenciones basadas en evidencia. Este proceso meticuloso asegura que las decisiones médicas estén respaldadas por la mejor

información disponible, permitiendo que los clínicos entreguen un cuidado óptimo a sus pacientes y que, a través de la evaluación continua, se logren mejoras significativas en la práctica clínica a lo largo del tiempo.

2.9. Algunos Ejemplos de Algoritmo Booleanos

En caso de que no haya quedado claro la construcción de los algoritmos booleanos, a continuación, veremos algunos ejemplos.

Ejemplo 1: Efecto de la terapia con estatinas en pacientes con hipertensión

Pregunta problema:
"¿Cómo influye la terapia con estatinas en el control de la presión arterial en pacientes hipertensos?"
Pasos:

1. **Identificar términos clave y sinónimos:**

- Término principal: "hipertensión" (se puede ampliar a "alta presión arterial")
- Intervención: "estatinas"
- Resultado: "control de la presión arterial", "efecto" o "impacto".

2. **Construir el algoritmo booleano:**
   Se combinan los términos utilizando operadores lógicos:

("hipertensión" OR "alta presión arterial") AND "estatinas" AND ("control de la presión arterial" OR "efecto" OR "impacto")

Ejemplo 2: Influencia del ejercicio físico en el manejo de la diabetes tipo 2

Pregunta problema:
"¿Cómo afecta el ejercicio físico al manejo de la diabetes tipo 2?"
Pasos:

1. **Identificar términos clave y sinónimos:**

- Condición: "diabetes tipo 2" o "diabetes mellitus tipo 2"

- Intervención: “ejercicio físico” (se puede incluir “actividad física”)
- Resultado: “manejo”, “control” o “tratamiento” de la enfermedad.

2. Construir el algoritmo booleano:

Se estructura la búsqueda con operadores lógicos:

(“diabetes mellitus tipo 2” OR “diabetes tipo 2”) AND (“ejercicio físico” OR “actividad física”) AND (“manejo” OR “control” OR “tratamiento”)

Conclusiones

La transformación hacia una MBE personalizada representa el eje central de este capítulo, donde se explora cómo la integración de la experiencia clínica, los valores del paciente y los datos científicos robustos permite abordar la complejidad y heterogeneidad de cada caso individual. Se evidencia la necesidad de trascender los enfoques tradicionales basados en resultados promediados de ensayos controlados, abriendo paso a estrategias que detallan la historia biológica, clínica, psicológica y social del paciente índice. Además, se subraya la importancia de construir una base sólida a partir de métodos y algoritmos de búsqueda que optimicen la recopilación y análisis de información relevante, configurando así un modelo que promueve una atención médica más precisa, adaptada y sensible a las particularidades de cada situación. Este paradigma, respaldado por una integración crítica y dinámica de la evidencia, sienta las bases para la innovación continua en la práctica clínica y fomenta el desarrollo de intervenciones que respondan de manera efectiva a las necesidades reales de los pacientes.

Parte II:
Manejo y Análisis de Datos en Investigación Médica

Manejo de Datos: Variables y Distribuciones

En el vasto océano de la investigación médica, la estadística es al mismo tiempo brújula y carta de navegación, pues transforma datos dispersos en patrones discernibles que guían decisiones críticas. Sin ella, el avance clínico sería como navegar en la niebla, una sucesión de intuiciones sin coordenadas. Durante la pandemia de COVID-19, por ejemplo, el análisis estadístico comparó tasas de infección entre grupos vacunados y no vacunados, demostró la efectividad de las vacunas y ayudó a salvar millones de vidas.

La estadística, entendida como la ciencia de recolectar, organizar, analizar, interpretar y presentar datos, se divide en dos ramas que se complementan. La descriptiva resume la información, como calcular la media de la presión arterial en un estudio o trazar la distribución de edades en un histograma; la inferencial utiliza muestras para extraer conclusiones acerca de poblaciones más amplias, por ejemplo, al evaluar si un nuevo fármaco es verdaderamente eficaz mediante un ensayo clínico aleatorizado. Sin ese filtro, la práctica médica correría el riesgo de apoyarse en conjeturas no verificadas.

Cuando estos métodos se aplican a la salud surge la bioestadística, que adapta las herramientas estadísticas a fenómenos biológicos y sanitarios. Sus usos van desde diseñar estudios epidemiológicos para examinar la relación entre contaminación y asma, hasta analizar datos genómicos en busca de mutaciones relacionadas con cáncer. En la era del *big data*, la disciplina distingue con destreza la señal del ruido, optimiza políticas de vacunación y mejora la asignación de camas de cuidados intensivos. El punto de partida de cualquier análisis es la variable, entendida como una característica que puede asumir distintos valores entre individuos, lugares o momentos. Edad, tipo de sangre, dosis administrada de un medicamento, presencia de un síntoma o tensión arterial son ejemplos cotidianos. Definirlas con precisión evita confundir unidades y garantiza que los hallazgos tengan significado clínico.

Según su naturaleza, las variables se clasifican en categóricas, continuas u ordinales. Las categóricas agrupan datos en clases, como los grupos sanguíneos A, B, AB u O; las continuas abarcan escalas potencialmente infinitas, por ejemplo, la concentración de glucosa en sangre; las ordinales ordenan la realidad en niveles de intensidad sin asumir distancias idénticas, como dolor leve, moderado o severo. Clasificar correctamente estas variables de-

termina si conviene utilizar mediana o media, o si corresponde un modelo de regresión logística o lineal. Con las variables identificadas, la estadística descriptiva despliega su arsenal. Un histograma revela la distribución del peso al nacer, destacando *outliers* que pueden indicar errores o patologías; un diagrama de caja resume la variabilidad de la presión intraocular en pacientes con glaucoma; media, mediana y moda señalan el centro de los datos, mientras que rango y desviación estándar (DS) describen su dispersión.

La medicina, sin embargo, no se conforma con describir; aspira a explicar por qué suceden las cosas. El concepto de causalidad distingue correlación de efecto real y se vuelve indispensable para establecer relaciones fármacoevento o exposiciónenfermedad. Inicialmente, se creyó que el consumo de café aumentaba el riesgo cardiovascular. Sin embargo, análisis que controlaron factores como el tabaquismo demostraron que esta asociación se debilitaba considerablemente. En contraste, la relación entre el tabaco y el cáncer de pulmón fue confirmada mediante estudios longitudinales y experimentales que identificaron al humo como causa directa. Herramientas estadísticas como la regresión multivariante permiten aislar influencias y minimizar sesgos con la precisión de un cirujano al separar tejidos.

Por estas razones, la MBE descansa sobre cimientos estadísticos: cuantificar magnitudes de efecto, estimar intervalos de confianza y calcular números necesarios para tratar convierte hallazgos aislados en conocimiento colectivo. Cada variable correctamente clasificada, cada gráfico claro y cada análisis causal riguroso construyen el puente entre la observación y la acción clínica, abriendo paso a la discusión detallada sobre cómo definir, medir y relacionar variables, tema que exploraremos a continuación.

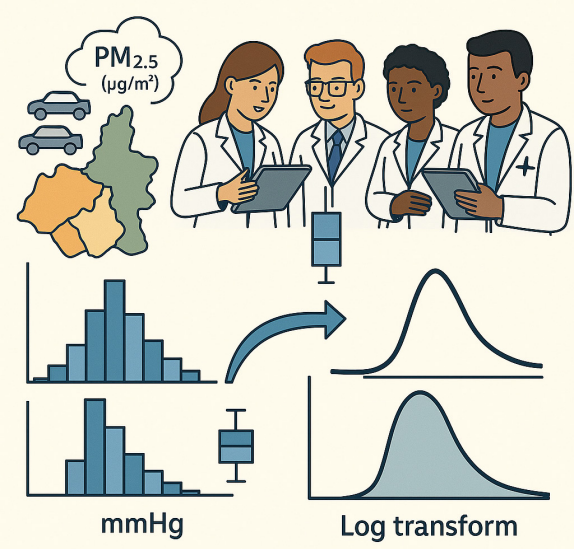
Antes de sumergirnos en la mecánica de clasificar variables, conviene ilustrar cómo todos los siguientes conceptos se enlazan en un caso real. Imaginaremos una investigación sencilla, pero completa, que muestre paso a paso para qué sirve cada herramienta estadística y por qué resulta crucial entender términos como “determinante de salud”, “variable confusora” o “poder estadístico”. El objetivo es que, aun sin experiencia previa, puedas seguir el razonamiento y ver cómo las piezas encajan en la práctica.

Ahora, imagina que un equipo de jóvenes médicos quiere investigar por qué en ciertos barrios de Cali la presión arterial media es mayor que en otros.

El determinante de salud inicial es la exposición crónica a material particulado proveniente del tráfico vehicular, un factor ambiental que, según los registros locales de calidad del aire, varía sustancialmente entre comunas. El grupo formula la variable independiente (X) como concentración mensual promedio, medida en microgramos por metro cúbico, y la variable dependiente (Y) como presión arterial sistólica en milímetros de mercurio registrada en adultos. Para no perder de vista la complejidad clínica, añaden variables demográficas, edad y sexo, y conductuales, consumo de sal y tabaquismo, que podrían actuar como confusoras.

Primero describen sus datos. Calculan media, mediana y DS de la presión arterial, construyen un histograma que revela una distribución ligeramente desplazada a la derecha y emplean un diagrama de caja para detectar un *outlier* o dato anómalo en el conjunto de datos: un paciente con hipertensión mal controlada. Al observar la asimetría positiva de la distribución de los datos, es decir, cola a la derecha y una curtosis pronunciada (picos pronunciados en la distribución de los datos), medidas de forma y apuntalamiento, deciden, antes de comparar grupos, verificar si una transformación logarítmica mejora la aproximación a la distribución normal.

Figura 10. *Distribución sesgada de la presión arterial y corrección estadística en relación con la exposición a* $PM_{2.5}$ *en barrios de Cali*. Elaborado con GeminiPRO.



Nota. $PM_{2.5}$ =material particulado con diámetro aerodinámico ≤ 2,5 μm, capaz de penetrar profundamente en las vías respiratorias en barrios de Cali. Infografía que resume el abordaje de un equipo de jóvenes médicos: se muestra el mapa de Cali con la concentración promedio mensual de $PM_{2.5}$ como variable independiente, un histograma y boxplot de la presión arterial sistólica con asimetría positiva (cola a la derecha) y alta curtosis, y la transformación logarítmica que aproxima los datos a la distribución normal para permitir comparaciones estadísticas adecuadas.

Antes de aplicar cualquier prueba, los estadísticos miran la forma de sus datos igual que un sastre observa la tela antes de cortar: si la tela es "recta" y uniforme, la famosa curva normal, que parece una campana, puede usar patrones clásicos que ahorran tiempo y desperdicio. Esa campana es un modelo que resume cómo se comportan la mayoría de las mediciones cuando intervienen muchos factores pequeños y azarosos, y permite calcular con regla de oro cuántos casos caen cerca del promedio (por ejemplo, casi todas las personas están a uno o dos pasos de la altura media). Si los datos se desvían, con cola larga o picos raros, esos patrones fallan: las mangas quedan cortas o la chaqueta torcida, es decir, los intervalos de confianza engañan, los errores se multiplican y las conclusiones pierden fiabilidad. Para enderezar la tela hay tres caminos: plancharla con una transformación (como el logaritmo, que aplana los extremos), elegir otro tipo de costura más flexible, pruebas no paramétricas, o directamente usar un patrón distinto, diseñado para otras formas (modelos binomiales, de Poisson, etc.). Ignorar este ajuste es como medir una superficie curva con una regla rígida: obtendrás un número, pero no la medida real.

A continuación, revisan la teoría de probabilidad y las distribuciones; comprueban mediante la prueba de Shapiro Wilk que, tras la transformación, los valores se ajustan razonablemente a una distribución normal, lo que habilita el uso de pruebas paramétricas. Con esa base pasan a la estadística inferencial: toman una muestra aleatoria de 200 residentes en zonas de alta exposición y 200 en zonas de baja exposición. Calculan el error estándar de la media y construyen un intervalo de confianza del 95 %, comprendiendo que esta franja expresa la imprecisión debida al muestreo.

Formulan entonces sus hipótesis: nula, "no existe diferencia en la presión arterial media entre zonas", y alternativa, "la presión arterial en áreas

con $PM_{2.5}$ elevado es al menos 5 mmHg mayor que en áreas limpias". Fijan $\alpha = 0,05$ y, antes de iniciar, realizan un cálculo de poder para asegurarse de que 400 sujetos ofrecen 80 % de probabilidades de detectar ese efecto. La decisión les alerta sobre el riesgo de errores tipo I y tipo II: si α es muy estricto podrían pasar por alto un efecto real, y si el tamaño muestral es pequeño podrían no alcanzar poder suficiente.

Ejecutan una prueba t de Student entre los dos grupos. El valor p obtenido es 0,018, estadísticamente significativo; sin embargo, revisan el intervalo de confianza (2,1 a 7,6 mmHg) para juzgar la relevancia clínica. Paralelamente, calculan la correlación de Pearson entre $PM_{2.5}$ y presión arterial ($r = 0,32$) y montan una regresión lineal ajustada por edad, sexo, consumo de sal y tabaquismo. La pendiente se reduce pero sigue siendo positiva, evidenciando que las confusoras explicaban parte, aunque no toda, de la asociación inicial.

Para diferenciar causalidad de simple casualidad aplican los criterios de Bradford Hill: la fuerza de asociación es moderada, existe gradiente biológico, mayor exposición, mayor presión, la relación es coherente con estudios in vivo sobre estrés oxidativo vascular y se observa plausibilidad fisiológica. Aunque no pueden realizar un ensayo controlado, proponen un estudio de intervención, colocar barreras verdes en avenidas, para fortalecer la evidencia.

Con este recorrido, incluso quien comienza desde cero observa cómo cada concepto estadístico y epidemiológico entra en juego de forma ordenada: desde definir las variables hasta interpretar intervalos, *p* y poder estadístico, pasando por la detección de confusores y la discusión sobre causalidad. Así se hace patente por qué dominar estos fundamentos resulta imprescindible antes de emprender cualquier investigación clínica.

No te preocupes si algunos términos aún resultan confusos, los explicaremos paso a paso con ejemplos sencillos y ejercicios guiados. Cuando concluyas este capítulo, te invito a releer el caso de la contaminación y la presión arterial; será como volver a un mapa después de aprender el significado de cada símbolo, descubrirás conexiones que antes pasaban inadvertidas y la historia cobrará pleno sentido.

CAPÍTULO 3

FUNDAMENTOS: DETERMINANTES DE LA SALUD, VARIABLES Y CONFUSORES

Fundamentals: Health Determinants, Variables, and Confounders

Resumen

Este capítulo introduce los fundamentos de la investigación en medicina poblacional, comenzando por los determinantes de la salud, entendidos como un ecosistema de factores biológicos, ambientales, sociales, económicos y culturales que influyen en la salud. Se distinguen los determinantes estructurales, como la distribución de la riqueza, de los intermedios, como la vivienda o los estilos de vida, destacando que rara vez actúan aisladamente y que su estudio permite priorizar intervenciones eficaces. Luego, el texto aborda las variables, las etiquetas que usamos para medir características cambiantes, y sus escalas: nominal, ordinal, intervalo y razón. Se explica cómo esta clasificación determina los análisis estadísticos apropiados, ilustrado con ejemplos clínicos como tipo sanguíneo, nivel de dolor, temperatura y presión arterial. Se detalla la conversión unidireccional entre escalas, de la más informativa a la menos, y la distinción entre variables dicotómicas y politómicas. Finalmente, el capítulo conecta estos conceptos al introducir las relaciones entre variables, diferenciando las independientes (factores) de las dependientes (desenlaces), usando la analogía del diagrama de Ishikawa. Se explica cómo la escala de la variable influye en el análisis de estas relaciones, desde riesgos relativos para variables nominales hasta regresión lineal para continuas. Se concluye presentando las variables confusoras, aquellas que distorsionan la relación aparente entre exposición y desenlace, y técnicas básicas para controlarlas como la aleatorización, restricción, emparejamiento, estratificación y modelos multivariables, subrayando el riesgo de ignorarlas.

Palabras clave: Determinantes de la salud, Variables, Escalas de medición, Variables confusoras, Epidemiología, Análisis estadístico.

Abstract

This chapter introduces the fundamentals of research in population medicine, starting with the determinants of health, understood as an ecosystem of biological, environmental, social, economic, and cultural factors influencing health. It distinguishes structural determinants, like wealth distribution, from intermediate ones, like housing or lifestyles, emphasizing that they rarely act in isolation and that their study allows prioritizing effective interventions. Next, the text discusses variables, the labels we use to measure changing characteristics, and their scales: nominal, ordinal, interval, and ratio. It explains how this classification determines appropriate statistical analyses, illustrated with clinical examples such as blood type, pain level, temperature, and blood pressure. The unidirectional conversion between scales, from most to least informative, and the distinction between dichotomous and polytomous variables are detailed. Finally, the chapter connects these concepts by introducing the relationships between variables, differentiating independent variables (factors) from dependent ones (outcomes), using the Ishikawa diagram analogy. It explains how variable scale influences the analysis of these relationships, from relative risks for nominal variables to linear regression for continuous ones. It concludes by presenting confounding variables, those that distort the apparent relationship between exposure and outcome, and basic control techniques like randomization, restriction, matching, stratification, and multivariable models, highlighting the risk of ignoring them.

Keywords: *Health determinants, Variables, Measurement scales, Confounding variables, Epidemiology, Statistical analysis.*

1.1 Determinantes de la Salud y sus Factores

En medicina poblacional la salud se asemeja a un ecosistema, ningún árbol crece aislado, cada hoja recibe luz, agua y nutrientes en proporciones variables y el suelo cambia con las estaciones. Ese ecosistema está formado por los **determinantes de la salud**, un conjunto de circunstancias biológicas, ambientales, sociales, económicas y culturales que, al entrelazarse, inclinan la balanza hacia la enfermedad o el bienestar. La altitud de una ciudad modifica la presión atmosférica y obliga al organismo a producir más hemoglobina, lo que reduce el riesgo de anemia en Bogotá en comparación con Car-

tagena. Del mismo modo, el nivel educativo promedio de un barrio influye en los ingresos futuros de quienes allí nacen y, a la larga, en la calidad de su dieta y en la posibilidad de practicar ejercicio de forma segura.

Para entender su funcionamiento conviene distinguir dos grandes categorías, **determinantes estructurales** y **determinantes intermedios**. Los estructurales son raíces profundas, la distribución de la riqueza, las leyes laborales y la infraestructura sanitaria, factores que definen a qué recursos accede una persona durante su vida. Los intermedios actúan como ramas próximas a la experiencia cotidiana, incluyen la vivienda, la exposición a contaminantes como el $PM_{2.5}$, los estilos de vida y el acceso a vacunas. Un joven que habita junto a una avenida con tráfico denso respira partículas finas a diario y experimenta inflamación pulmonar, mientras que disponer de un parque arbolado cercano y programas de actividad física gratuitos puede amortiguar ese efecto nocivo.

Las grandes series epidemiológicas, desde los mapas de cólera de John Snow hasta la cohorte de Framingham, muestran que los determinantes rara vez actúan de forma aislada. La obesidad, por ejemplo, no depende solo del exceso de calorías, influyen la disponibilidad de ultraprocesados, la publicidad dirigida a niños, la falta de espacios seguros para caminar, la genética familiar y la calidad del sueño; cada elemento suma o resta como piezas de ajedrez que se mueven en el mismo tablero. En la era del *big data* medimos muchos de esos desplazamientos en tiempo real, pero la interpretación exige una mirada integradora capaz de conectar bioquímica, sociología y política pública.

Estudiar los determinantes no implica resignarse al destino que marca el código postal, al contrario, permite priorizar intervenciones eficaces. Mejorar el saneamiento básico reduce las enfermedades diarreicas más que cualquier antibiótico y las políticas de espacios sin humo disminuyen las hospitalizaciones por crisis asmática en cuestión de meses. Reconocer cuáles hilos son más gruesos y tensos en cada comunidad guía la inversión de tiempo y recursos, y prepara el terreno para traducir estos determinantes en **variables** concretas, definir con precisión los **factores de riesgo** y los **desenlaces** que mediremos, tema que abordaremos a continuación para convertir el panorama general en preguntas de investigación específicas y cuantificables.

1.2. Tipos y Escalas de Variables

Cuando el clínico registra datos, en realidad coloca etiquetas en el mundo para que la ciencia pueda leerlo. Cada etiqueta es una **variable**, una característica que cambia entre personas, lugares o momentos. Algunas variables se comportan como interruptores de luz, solo encendido o apagado, otras ofrecen un regulador con infinitas posiciones y unas más se ordenan como los peldaños de una escalera. La estadística clasifica estas diferencias en **escalas** que determinan qué operaciones matemáticas son válidas y qué pruebas analíticas resultan apropiadas.

Imaginemos una sala de urgencias. Al ingreso, la enfermera anota el **tipo sanguíneo**, un dato que solo identifica grupos, sin orden implícito. Más tarde, pregunta al paciente por el **nivel de dolor** usando la clásica escala de 0 a 10. Después, mide la **temperatura corporal** en grados Celsius y registra la **presión arterial** en milímetros de mercurio. Cada uno de estos datos reside en una escalera distinta, conocer la escalera correcta evita dar pasos en falso cuando interpretamos cifras o comparamos grupos.

La conversión entre tipos de variables estadísticas se puede realizar de manera unidireccional, desde el tipo de variable más informativo hasta el menos informativo, Razón → Intervalo → Ordinal → Nominal (ver tabla 4). Por ejemplo, una concentración de glucosa (razón) puede reagruparse en rangos de ayuno alterado o diabetes (ordinal) o, aún más, en la categoría binaria "normal" versus "elevada" (nominal). El camino inverso no es posible, ya que la información perdida no se recupera. Además de distinguir nominal, ordinal, intervalo y razón, la práctica diaria exige reconocer si una variable **nominal** es **dicotómica** o **politómica**. Una variable dicotómica funciona como un interruptor de dos posiciones, por ejemplo, "fuma, no fuma". Las politómicas, en cambio, ofrecen varios estados, tal como "soltero, casado, conviviente, divorciado, viudo". Esta clasificación importa porque algunas pruebas estadísticas, por ejemplo, la regresión logística binaria, aceptan solo respuestas dicotómicas, mientras que la logística multinomial maneja categorías múltiples.

Tabla 4. *Tipos de variables y escalas*

Variable o medición	Tipo general	Escala específica*	Propiedades clave	Ejemplo clínico
Tipo sanguíneo (A, B, AB, O)	Cualitativa	Categórica nominal	Clases sin orden ni distancias	Selección de hemoderivados
Estado funcional (leve, moderado, severo)	Cualitativa	Categórica ordinal	Orden presente, intervalos desconocidos	Clasificación de insuficiencia cardíaca
Temperatura axilar (°C)	Cuantitativa	Intervalo	Distancias iguales, cero arbitrario	Fiebre de 38 °C no duplica el "calor" de 19 °C
Presión arterial sistólica (mmHg)	Cuantitativa	Razón	Distancias iguales, cero absoluto	120 mmHg es el doble de 60 mmHg
Número de crisis asmáticas al mes	Cuantitativa	Discreta	Valores enteros contables	0, 1, 2, 3…
Concentración de glucosa (mg/dL)	Cuantitativa	Continua	Infinitos valores dentro de un rango	87,3 mg/dL

*La categoría "intervalo" y la "razón" se agrupan dentro de las variables numéricas; la diferencia principal es que en la escala de razón existe un cero real y admite comparaciones de multiplicación, mientras que en la escala de intervalo el cero es arbitrario.

Para ilustrar por qué importa esta distinción, pensemos en la **temperatura** y la **presión arterial**. Si la temperatura sube de 36 °C a 40 °C, no podemos decir que el cuerpo está "el doble de caliente", porque 0 °C no representa ausencia absoluta de energía térmica; en cambio, si la presión arterial aumen-

ta de 60 mmHg a 120 mmHg, sí es correcto afirmar que se duplicó, ya que 0 mmHg indicaría ausencia de presión. Esta diferencia guía la elección de análisis: promediar tipos sanguíneos carece de sentido, pero calcular la media de la presión arterial es útil y compararla con un intervalo de confianza resulta válido.

Las variables **ordinales** exigen cautela adicional. Entre "dolor moderado" y "dolor severo" hay una jerarquía clara, aunque no sabemos cuánta distancia subjetiva separa cada peldaño. Por ello, pruebas paramétricas que asumen intervalos iguales, como la t de Student, no son apropiadas; conviene utilizar alternativas como la prueba de MannWhitney. En cambio, las variables **nominales** se analizan con tablas de contingencia y pruebas de χ^2 (chi-cuadrado), mientras que las variables **numéricas** suelen ajustarse a modelos de regresión cuando cumplen supuestos de normalidad.

Dominar los tipos y escalas de variables es como leer el código de colores en una resistencia electrónica, revela qué operaciones soporta cada dato y cuáles provocarían cortocircuitos lógicos. Con estas escalas claras, podemos transformar los determinantes de la salud en **factores de riesgo** bien definidos y, más adelante, relacionarlos con un **desenlace** medible, abriendo así la puerta al estudio de confusión y causalidad que abordaremos en la sección siguiente.

1.3. Conversión de Variables

En estadística, transformar una variable equivale a disminuir la resolución de una fotografía: se obtiene una imagen más simple, aunque con menos detalles finos, en esta sección ampliaremos el concepto de conversión de variables visto anteriormente. Ese proceso de transformación solo fluye en un sentido, desde el formato **más informativo**, que conserva valores exactos, hasta el **menos informativo**, que agrupa o nombra sin medir. El camino típico es

Razón → Intervalo → Ordinal → Nominal

De Razón a Intervalo

Un monitor registra la **presión arterial sistólica** con precisión de un milímetro de mercurio, 0 mmHg indica ausencia total de presión, por eso es una

variable de **razón**. Si el investigador decide expresar cada lectura como "diferencia respecto a 100 mmHg" (p. ej., +20, +15, –10), el cero deja de ser real y se convierte en arbitrario; la escala pasa a **intervalo**. Ahora podemos sumar y restar esas diferencias, pero ya no afirmar que 120 mmHg duplica a 60 mmHg.

De Intervalo a Ordinal

La **temperatura corporal** en grados Celsius es de intervalo, porque 0 °C no significa "sin calor". Para un boletín epidemiológico se agrupa en tres rangos: "hipotermia" (< 35 °C), "normotermia" (3537,5 °C) y "fiebre" (> 37,5 °C). Con ese recorte la variable se vuelve **ordinal**: mantiene un orden lógico, pero desconocemos cuántos grados separan de forma exacta cada categoría.

De Ordinal a Nominal

Una escala de **dolor** de 0 a 10 se simplifica para triage de urgencias en "dolor leve" (03) versus "dolor moderadosevero" (410). Perdemos la gradación fina y, aunque persiste cierto orden implícito, la nueva variable se analiza como **nominal dicotómica**, apta para tablas de 2 × 2 y pruebas χ^2.

Reducir la escala simplifica análisis y facilita la comunicación, sin embargo sacrifica precisión y limita las pruebas disponibles. Un estudio de factores de riesgo cardiovascular podría necesitar la presión arterial en forma numérica para calcular promedios y desviaciones estándar, mientras que una campaña de salud pública quizá solo requiera saber cuántos pacientes superan el umbral de 140 mmHg.

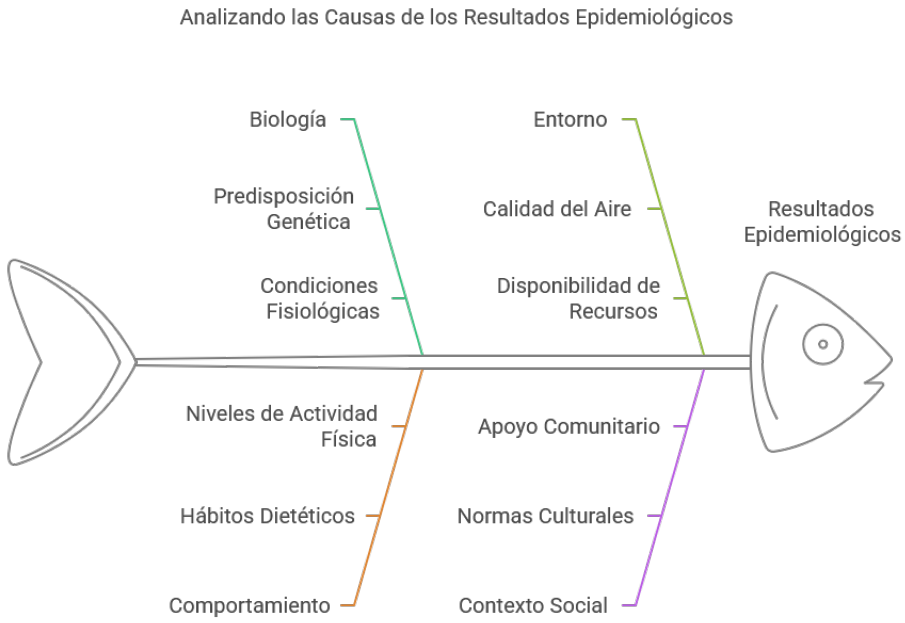
Por eso, antes de convertir conviene preguntarse: ¿Qué preguntas quiero responder? y ¿Cuánta información puedo permitirme perder? Una vez la fotografía se guarda con menos píxeles, no existe filtro que restaure los detalles originales.

1.4. Del "qué" al "por qué": conectar variables independientes y dependientes

Comprender los tipos y escalas de variables es solo la mitad del camino; el paso siguiente consiste en descubrir **cómo** se relacionan esas variables entre sí. En epidemiología trazamos la ruta que va del *factor* al *desenlace* igual que un ingeniero utiliza un diagrama de Ishikawa, el clásico "hueso de pescado",

para desmenuzar las causas de una avería. En lugar de "materiales, métodos o maquinaria", nuestro esqueleto suele llevar etiquetas como **biología**, **comportamiento**, **entorno**, **sistema de salud** y **contexto social**. Cada espina representa un conjunto de variables que empujan, con más o menos fuerza, hacia el problema estudiado: incidencia de diabetes, mortalidad por COVID19 o tasas de hipertensión en barrios contaminados (ver figura 11).

Figura 11. *Diagrama Ishikawa de los principales determinantes de los resultados epidemiológicos. Elaborado con napkin: https://www.napkin.ai/es/*

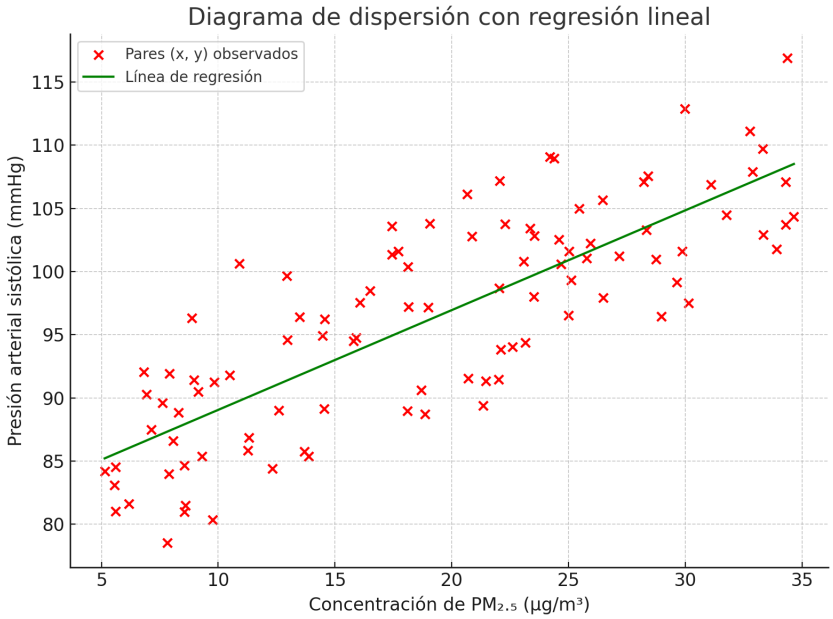


Nota. Diagrama de espina de pescado que esquematiza las cuatro categorías de causas, Biología (predisposición genética, condiciones fisiológicas), Comportamiento (hábitos dietéticos, niveles de actividad física), Entorno (calidad del aire, disponibilidad de recursos) y Contexto social (normas culturales, apoyo comunitario), y sus factores específicos que interactúan para determinar los resultados epidemiológicos de una población.

1.5. De "qué" a "por qué": variables independientes y dependientes

En ese armazón causal necesitamos diferenciar las **variables independientes**, los supuestos motores del cambio, de las **variables dependientes**, los resultados que observamos. Matemáticamente se resume en la notación $y = f(x)$, donde x simboliza la exposición o el factor de riesgo y y el desenlace. Si dibujamos los datos en un plano, cada punto rojo del diagrama de dispersión indica un par (x,y), y colocar una línea verde, la regresión, equivale a estimar la función f. Cuando los puntos forman una nube ascendente, podemos afirmar que a mayor x tiende a haber mayor y; por ejemplo, más microgramos de $PM_{2}\square_{5}$ al día, mayor presión arterial sistólica. Si la nube es plana o difusa, la relación tal vez sea débil o inexistente (ver figura 12).

Figura 12. *Diagrama de dispersión de variables independientes y dependientes con línea de regresión*. Elaborado con GeminiPRO.



Nota. La nube ascendente de puntos indica una relación positiva: a mayores niveles de $PM_{2.5}$ suele observarse mayor presión arterial sistólica, mientras que una nube plana o muy dispersa sugeriría una asociación débil o nula.

1.6. De la escala a la Relación

La escala de la variable determina el tipo de análisis que podemos aplicar para cuantificar esa flecha causal:

- Si "x" es **nominal dicotómica**, vacunado sí o no, y "y" es **nominal dicotómica**, infectado sí o no, calculamos **riesgos relativos** o **odds ratios** y usamos regresión logística.
- Si "x" es **ordinal**, baja, media, alta exposición solar, y "y" es **continua**, nivel sérico de vitamina D, podemos utilizar análisis de varianza (ANOVA) o modelos de regresión lineal con variables indicadoras.
- Si ambos, "x" y "y", son **continuos de razón**, edad gestacional y peso al nacer, aplicamos correlación de Pearson y regresión lineal para estimar cuántos gramos gana el recién nacido por cada semana adicional.

Cuando reducimos la resolución de x o y, siguiendo la secuencia razón, intervalo, ordinal, nominal, también cambiamos la lupa con la que buscamos patrones: un modelo de regresión lineal exige números continuos, mientras que una prueba χ^2 se nutre de categorías. Así, la decisión de convertir una variable no es solo estética, define el idioma en que podremos hablar sobre la relación causaefecto.

1.6.1. *Estrategias de Análisis de Causas*

Una vez que hemos formulado hipótesis sobre las posibles causas de un fenómeno, es crucial aplicar métodos analíticos rigurosos para evaluar estas relaciones. No basta con observar una simple correlación; debemos investigar a fondo para comprender la naturaleza y la fuerza de la conexión, distinguiendo las asociaciones genuinas del ruido aleatorio o de la influencia de otros factores. A continuación, se describen las estrategias clave para llevar a cabo este análisis causal:

1. **Exploración gráfica**: diagramas de dispersión para continuas, diagramas de cajas para comparar categorías, mapas de calor cuando hay más de dos dimensiones. Ver antes de calcular evita confundir ruido con señal.
2. **Modelos univariados**: estimar la asociación bruta, por ejemplo, la diferencia media de presión arterial entre zonas contaminadas y limpias.

3. **Modelos multivariables**: añadir confusoras conocidas, edad, sexo, tabaco, para aislar el efecto de interés, igual que ajustar los espejos retrovisores antes de arrancar.

4. **Evaluación de interacción**: explorar si una variable modifica la fuerza de la otra, por ejemplo, si el impacto de la contaminación sobre la presión es mayor en adultos mayores que en jóvenes.

1.6.2. *De la Asociación a la Acción*

Cuando los datos sostienen la relación $y=f(x)$ y los criterios de plausibilidad son razonables, avanzamos hacia la intervención: reducir x (exposición) o reforzar un mediador protector para mejorar y (desenlace). El ejemplo de las barreras verdes para filtrar $PM_{2.5}$ ilustra cómo una idea estadística desemboca en decisión de salud pública.

En síntesis, identificar la naturaleza de cada variable clarifica el tipo de lupa estadística, pero identificar su rol, independiente o dependiente, y las trayectorias que las conectan nos dice dónde colocar la lupa. Con este hilo conductor, estamos listos para examinar técnicas que miden la fuerza de esas relaciones y, sobre todo, para discernir cuándo una correlación deja de ser una simple coincidencia y se convierte en evidencia de causalidad. Sin embargo, debemos conocer que existen un tipo de variables que son llamadas confusoras que pueden distorsionar la asociación entre la exposición y el desenlace al relacionarse tanto con la variable independiente como con la dependiente, de modo que, si no las identificamos y controlamos adecuadamente, corremos el riesgo de extraer conclusiones causales equivocadas. A continuación las trataremos con más detalle.

1.7. Variables Confusoras y Técnicas básicas de Control

En la estadística aplicada a la salud cada relación entre causa y efecto se asemeja a una fotografía tomada desde un globo aerostático, las formas generales son visibles pero la atmósfera distorsiona los colores. Esa "atmósfera" la crean las **variables confusoras**, factores que se relacionan tanto con la exposición como con el desenlace y que, si no se corrigen, tiñen la imagen con tonos engañosos.

Un ejemplo clásico lo ilustra el canto del gallo y la salida del sol: cualquiera que observe que el gallo canta unos minutos antes del alba podría concluir que el ave provoca el amanecer, sin embargo, ambos fenómenos están sincronizados por una tercera variable, la rotación de la Tierra, que actúa como confusora. Si ignoramos ese reloj astronómico parecería que basta callar al gallo para retrasar la aurora. Identificar y controlar estas neblinas estadísticas nos permite revelar los contornos reales del paisaje y tomar decisiones basadas en causas auténticas, no en coincidencias matutinas (ver figura 13).

Figura 13. *Ejemplos de variable confusora en fenómenos naturales y epidemiológicos*. Elaborado con GeminiPRO.



1.7.1. ¿*Qué es una Confusora*?

Imagina que investigas si el café eleva el riesgo de infarto. Encuentras que quienes beben más café presentan más infartos y concluyes que la cafeína es peligrosa. Sin embargo, observas un detalle extra, la mayoría de los grandes cafeteros también fuman. El tabaquismo está conectado con el consumo de café y con el infarto, por lo tanto, actúa como confusor, emborrona la lente y hace que el efecto aparente del café sea mayor de lo real.

Otro ejemplo cotidiano: medir la relación entre número de horas de series *streaming* y obesidad infantil. Al principio emerge una correlación positiva,

pero si preguntamos por el tiempo de actividad física descubrimos que los niños que ven más pantalla, además, juegan menos en exteriores. La falta de ejercicio es el verdadero motor de la ganancia de peso y, al mismo tiempo, explica la afición a la pantalla, por tanto, confunde la fotografía.

1.7.2. ¿Cómo Reconocerlas?

Actuar frente a una confusora se parece al trabajo de un detective que busca al responsable de un robo en una habitación llena de huellas superpuestas. La tarea consiste en separar cada pisada para descubrir cuál pertenece al ladrón y cuál al dueño de la casa. La estadística ofrece varias "lámparas de luz ultravioleta" para aislar las huellas correctas.

A. Aleatorización

En un ensayo clínico se reparte al azar quién recibe la pastilla y quién el placebo, como si se lanzaran monedas. De esa manera las confusoras conocidas, edad, sexo, comorbilidades, y las desconocidas se distribuyen equitativamente, la nube de color que distorsiona la foto se difumina y la diferencia de resultados se atribuye con mayor confianza al fármaco (ver tabla 5).

Tabla 5. *Técnicas de control de confusión en estudios epidemiológicos: momento de aplicación, metáfora y ejemplo clínico*

Técnica de control	Momento de aplicación	Metáfora sencilla	Ejemplo clínico
Aleatorización	Diseño, antes de recolectar datos	Barajar y repartir cartas al azar, cualquier carta puede ir a cualquier jugador	Ensayo de fármaco antihipertensivo, la edad media queda equilibrada en ambos grupos
Restricción	Diseño	Un filtro en la puerta, solo entra gente de cierto tipo	Estudio de tabaquismo en embarazadas, se incluye solo a mujeres no fumadoras

Técnica de control	Momento de aplicación	Metáfora sencilla	Ejemplo clínico
Emparejamiento (*matching*)	Diseño	Parejas de baile con misma altura, se comparan con gracia	Casocontrol de cáncer de mama, cada caso se empareja con un control de misma edad
Estratificación	Análisis descriptivo	Dividir un mazo en montones ordenados y revisar uno a uno	Asociación colesterolinfarto examinada por rangos de edad
Modelos multivariables	Análisis inferencial	Balanza de varios platillos que compensa cada peso y revela la carga real	Regresión logística que ajusta simultáneamente por tabaquismo, dieta y actividad física

B. Restricción

Supón que quieres saber si trabajar de noche altera la secreción de melatonina y temes que el consumo de alcohol ensucie la relación. Puedes decidir reclutar solo participantes que no beban, así reduces el efecto de ese factor, aunque sacrificas representatividad.

C. Emparejamiento

En un estudio casocontrol sobre enfermedad pulmonar ocupacional emparejas cada caso con un control del mismo barrio y edad. Como gemelos virtuales, los pares ayudan a comparar exposiciones específicas manteniendo constantes otros rasgos.

D. Estratificación

En un análisis de colesterol, separas a los participantes por décadas de edad, comparas dentro de cada estrato y, si la relación desaparece en todos ellos, compruebas que la edad era la responsable de la asociación global. Es como observar una pintura a través de recortes para ver qué color domina en cada cuadrante.

E. Modelos multivariables

Cuando las confusoras son numerosas o continuas, recurrimos a la regresión, una balanza con múltiples platos que calcula cuánta carga sostiene cada variable manteniendo quietas las demás. Así descubrimos, por ejemplo, que la contaminación atmosférica incrementa la presión arterial en 3 mmHg tras ajustar por edad, sexo, tabaquismo y consumo de sal, cifra más modesta pero mucho más veraz que la estimada en el análisis crudo.

1.7.3. *Consecuencias de Ignorar la Confusión*

Analizar sin controlar confusores es como juzgar un partido de fútbol filmado con niebla en el lente, las camisetas se mezclan y no distinguimos qué jugador marcó el gol. El riesgo es tomar decisiones clínicas basadas en ilusiones: prohibir el café cuando el culpable es el cigarrillo o recomendar suplementos caros de vitamina D cuando bastaba con fomentar el deporte al aire libre.

Al dominar estas técnicas, el investigador convierte la cámara nebulosa en un instrumento de alta definición. Así, los determinantes descubiertos, las variables bien clasificadas y las relaciones causa-efecto se alinean para formar una historia coherente, lista para someterse a las pruebas de significancia, potencia y predicción que exploraremos en los capítulos que siguen.

Conclusiones

Comprender los determinantes que tejen el tapiz de la salud poblacional y aprender a etiquetar sus hilos con la precisión de las variables y sus escalas es el primer paso para descifrar el complejo lenguaje de la enfermedad y el bienestar. Hemos visto cómo la elección de la escala correcta define las preguntas que podemos hacerle a los datos y cómo la identificación de variables independientes y dependientes nos permite trazar las flechas de una posible causalidad. Sin embargo, este paisaje no está exento de neblinas; las variables confusoras, como sombras que distorsionan la luz, nos recuerdan la necesidad de una mirada crítica y herramientas analíticas sofisticadas para revelar las verdaderas relaciones. Con estos cimientos sólidos, esta comprensión de los bloques básicos de construcción, estamos preparados para adentrarnos en las técnicas que nos permitirán no solo medir la fuerza

de estas asociaciones, sino también discernir la diferencia crucial entre una simple coincidencia y la evidencia que apunta hacia una causa real.

CAPÍTULO 4

DESCRIPCIÓN Y VISUALIZACIÓN DE DATOS CLÍNICOS

Description and Display of Clinical Data

Resumen

Este capítulo introduce la estadística descriptiva como el primer paso fundamental en el análisis de datos clínicos, comparándola con la presentación de personajes y escenarios en una novela. Se enfoca en resumir y organizar la información cruda para revelar patrones y detectar valores atípicos. Se detallan las medidas de tendencia central (media aritmética, mediana, moda), explicando sus conceptos, usos clínicos, ventajas, desventajas y la importancia de elegir la métrica adecuada según la simetría de la distribución y la presencia de *outliers*. Posteriormente, se abordan las medidas de dispersión (rango, rango intercuartílico, varianza, DS, coeficiente de variación (CV), mediana absoluta de las desviaciones), que cuantifican la variabilidad de los datos y son cruciales para entender la amplitud del conjunto de valores. Finalmente, se explican las medidas de forma (asimetría) y apuntalamiento (curtosis), que describen la silueta de la distribución de los datos, indicando inclinaciones o picos que pueden influir en la elección de métodos de análisis posteriores. El capítulo enfatiza la selección práctica de cada medida y la necesidad de complementar las cifras con visualizaciones gráficas para una comprensión completa antes de proceder a la estadística inferencial.

Palabras Clave: Estadística descriptiva, datos clínicos, medidas de tendencia central, medidas de dispersión, asimetría, curtosis.

Abstract

This chapter introduces descriptive statistics as the essential first step in analyzing clinical data, likening it to introducing characters and settings in a novel. It focuses on summarizing and organizing raw information to reveal patterns and detect outliers. Measures of central tendency (arithmetic mean, median, mode) are detailed, explaining their concepts, clinical applications, advantages,

disadvantages, and the importance of selecting the appropriate metric based on distribution symmetry and the presence of outliers. Subsequently, measures of dispersion (range, interquartile range, variance, standard deviation, coefficient of variation, median absolute deviation) are discussed, quantifying data variability and crucial for understanding the spread of values. Finally, measures of shape (skewness) and peakedness (kurtosis) are explained, describing the distribution's silhouette, indicating skews or peaks that can influence the choice of subsequent analytical methods. The chapter emphasizes the practical selection of each measure and the need to complement numerical summaries with graphical visualizations for a comprehensive understanding before moving to inferential statistics.

Keywords: *Descriptive statistics, clinical data, measures of central tendency, measures of dispersion, skewness, kurtosis.*

Clasificación elemental de la estadística

> *Si la investigación médica fuera una novela, la estadística descriptiva sería la primera página, nos presenta a los personajes, edad, presión arterial, glucemia, y nos muestra la escena antes de que inicie la acción inferencial. Su misión consiste en resumir la información cruda, revelar patrones evidentes y señalar valores atípicos que podrían ser errores de digitación o pistas de nuevas hipótesis.*
>
> — *Yamil Liscano*

1. Descriptiva, recoge y organiza datos tal como un fotógrafo encuadra, ajusta la luz y entrega la imagen.

2. Inferencial, extrae conclusiones sobre la población a partir de la muestra, equivalente al crítico que lee la foto para adivinar la historia completa.

2.1. Medidas de Tendencia Central

Antes de profundizar en cada medida, es importante reconocer que las medidas de tendencia central resumen un conjunto de datos en un único valor representativo, facilitando la comparación y la comunicación de resultados clínicos. Su elección depende del tipo de distribución y de la presencia de va-

lores extremos: la media aritmética es ideal cuando los datos son simétricos y sin *outliers* pronunciados; la mediana resulta más robusta en distribuciones sesgadas o con valores atípicos; y la moda aporta información clave en variables categóricas. A continuación, se describen sus características, ventajas y limitaciones en contextos clínicos (ver tabla 6):

2.1.1 *Media Aritmética*

En un estudio de glucemia en ayunas, la media indica el "punto de balance" de todos los valores. Si un solo paciente presenta 400 mg/dL, la media sube como un columpio inclinado por un niño más pesado, de ahí la necesidad de revisar *outliers* antes de citársela a un colega.

Ventaja clínica: se enlaza con la DS y permite calcular intervalos de confianza bajo supuestos normales.

Desventaja: un único error de digitación, escribir 1000 en lugar de 100 mg/dL, puede alterar todo el análisis.

2.1.2 *Mediana*

En emergencias, la mediana del tiempo de espera refleja mejor la realidad que la media, puesto que algunos pacientes críticos pasan de inmediato y alargan la media de manera artificial.

Ventaja: resiste extremos y describe distribuciones sesgadas.

Desventaja: no utiliza la distancia entre valores, por lo que se pierde precisión comparada con la media en datos simétricos.

2.1.3 *Moda*

En un brote infeccioso la moda de los síntomas, tos, fiebre, diarrea, identifica rápidamente la presentación más frecuente y orienta el diagnóstico sindrómico.

Ventaja: útil cuando los datos son puramente categóricos.

Desventaja: si las frecuencias se reparten de forma similar entre varias categorías, la moda deja de ser informativa.

Tabla 6. *Resumen de medidas de tendencia central*

Métrica	Concepto narrado	Fórmula básica*	Uso habitual	Ventajas	Desventajas	Variables adecuadas	Variables problemáticas
Media aritmética	El "centro de gravedad" de los valores, cada dato empuja con fuerza proporcional a su magnitud	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$	Valores continuos de distribución simétrica, por ejemplo, presión arterial	Sencilla, se combina bien con la desviación estándar y métodos paramétricos	Sensible a *outliers*, distorsiona la imagen si hay cola larga	Numéricas de razón o intervalo sin sesgo extremo	Categóricas, ordinales, distribuciones muy asimétricas
Mediana	El punto que divide la fila al cien por ciento, tantos datos por encima como por debajo	Ordenar datos, tomar el (n+1)/2 ésimo valor	Salarios de residentes, recuentos de linfocitos con valores extremos	Resistente a *outliers*, describe bien distribuciones sesgadas	No aprovecha toda la información, menos eficiente si la distribución es normal	Numéricas o ordinales con asimetría o valores extremos	Variables nominales, muestras muy pequeñas con muchos empates
Moda	La categoría o valor que aparece con mayor frecuencia, el "trending topic" de la muestra	Valor con frecuencia máxima	Tipo de sangre más común, presentación de síntomas	Única medida para variables cualitativas, fácil de interpretar	Puede haber más de una, muy inestable en muestras pequeñas	Nominales y ordinales, discretas con repeticiones	Variables continuas con muchos valores únicos

Elección Práctica

- Media, cuando los datos son numéricos, de distribución aproximadamente normal y sin valores aberrantes.
- Mediana, cuando hay cola larga, pequeños tamaños muestrales o se quiere proteger la privacidad de valores extremos.
- Moda, cuando se trabaja con variables nominales o se desea señalar la categoría más común.

Visualizar las métricas con histogramas o diagramas de caja refuerza la narrativa y permite detectar discrepancias entre las medidas. En los siguientes apartados combinaremos estas cifras con herramientas gráficas para destacar patrones y prepararnos para el salto hacia la estadística inferencial.

2.2 Medidas de Dispersión

Con las medidas de tendencia central hemos localizado el punto en que la balanza de los datos se equilibra; ahora necesitamos saber cuán anchas son las alas de esa balanza. Las **medidas de dispersión** describen la amplitud del abanico numérico, informan si los valores se arremolinan como sardinas en un banco o se esparcen como gaviotas sobre el mar. Sin esta información la media, la mediana o la moda podrían engañarnos: dos barrios pueden compartir la misma presión arterial media y, sin embargo, uno albergar la mayoría de sus lecturas cerca de 120 mmHg mientras que el otro oscila entre 90 y 180 mmHg.

2.2.1 *Rango e* IQR:

En un estudio de tiempos de espera quirúrgica, el rango nos dice que el más rápido fue atendido en tres días y el más lento en ciento veinte. El IQR afina la imagen: la mitad de los pacientes cayó entre veinte y cuarenta y cinco días, señal de que los extremos quizá reflejan casos urgentes o demoras administrativas excepcionales.

Ventaja del IQR: sobrevive a retrasos exagerados, sirve al comparar servicios con colas asimétricas.

Desventaja: no informa cuán larga es la cola, puede enmascarar eventos críticos cuando se toman decisiones de gestión.

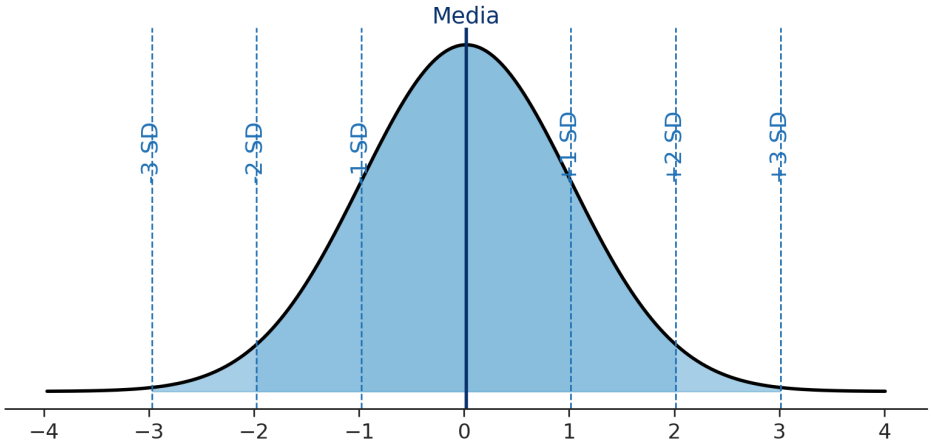
2.2.2 *Varianza y Desviación Estándar*

Supongamos que medimos la glicemia de cien pacientes en ayunas; la DS nos indica cuántos miligramos por decilitro se aleja, en promedio, cada glucemia de la media. En una curva normal el 68 % de los valores caen dentro de una DS y el 95 % dentro de dos DS, regla nemotécnica que facilita interpretar "hiperglucemias" o "hipoglucemias" sin memorizar rangos exactos (ver figura 14).

Ventaja: se integra en pruebas t, ANOVA, intervalos de confianza paramétricos.

Desventaja: un error de digitación que registre 1200 mg/dL eleva la DS y ensancha todos los intervalos, nublando la clasificación de riesgo.

Figura 14. *Campana de Gauss con límites de desviación estándar. Elaborado con Gemini Pro.*



Nota. Curva de densidad de la distribución normal estándar (μ = 0, σ = 1), con la media marcada por la línea vertical continua. Las líneas punteadas indican las posiciones de −3 σ, −2 σ, −1 σ, +1 σ, +2 σ y +3 σ; aproximadamente el 68 % de los datos cae dentro de ±1 σ, el 95 % dentro de ±2 σ y el 99.7 % dentro de ±3 σ de la media.

2.2.3 *Coeficiente de Variación*

Comparar la dispersión del peso (kilogramos) y de la troponina ultrasensible (nanogramos por litro) es tan inútil como cotejar grados Celsius con dólares. El CV "divide" la DS por la media y convierte ambos en porcentajes: un CV del 5 % en troponina indica un proceso analítico muy preciso, mientras que un 25 % en recuento plaquetario alerta de variabilidad biológica o técnica considerable.

2.2.4 *Mediana Absoluta de las Desviaciones* (MAD)

En microbiota intestinal, los recuentos de *Escherichia coli* pueden oscilar mil veces de un sujeto a otro. La MAD se basa en la mediana, no en la media, por lo que un valor extremo, una infección aguda con diez millones de bacterias, apenas altera esta medida, manteniendo una foto fiel para la mayoría (ver tabla 7).

Tabla 7. *Resumen de medidas de dispersión*

Métrica	Concepto narrado	Fórmula básica*	Uso habitual	Ventajas	Desventajas	Variables adecuadas	Variables problemáticas
Rango (R)	Distancia entre los extremos, como puntas de una cometa	$R = x_{\text{máx}} - x_{\text{mín}}$	Vistazo rápido a la amplitud de peso al nacer	Fácil de calcular y explicar	Muy sensible a *outliers*, ignora valores intermedios	Cualquier variable numérica en informes descriptivos	Muestras pequeñas con valores atípicos
Rango intercuartílico (IQR)	Anchura del "cinturón central" que envuelve al 50 % de los datos	$IQR = Q_3 - Q_1$	Distribuciones sesgadas, niveles de ferritina	Robusto frente a extremos, combina bien con la mediana	No refleja cola completa, menos intuitivo para público general	Variables ordinales largas o continuas, datos asimétricos	Datos bimodales donde el centro está vacío
Varianza (s^2)	Promedio de los cuadrados de las desviaciones, energía cinética de la nube de puntos	$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$	Paso intermedio para pruebas paramétricas	Fundamento de la desviación estándar, propiedades algebraicas útiles	Difícil de interpretar en unidades cuadradas	Variables continuas simétricas	Categóricas, ordinales puras

Métrica	Concepto narrado	Fórmula básica*	Uso habitual	Ventajas	Desventajas	Variables adecuadas	Variables problemáticas
Desviación estándar (DS)	Raíz cuadrada de la varianza, radio de la "zona de confort" alrededor de la media	$SD = \sqrt{s^2}$	Bioquímica sérica, tensión arterial, talla	Se expresa en mismas unidades que la media, directa en distribuciones normales	Distorsionada por *outliers* y asimetrías pronunciadas	Continuas de razón o intervalo con forma aproximadamente normal	Distribuciones muy sesgadas, datos acotados inferiormente como tiempo a evento
Coeficiente de variación (CV)	DS relativa al tamaño de la media, porcentaje de variabilidad	$CV = \frac{SD}{\bar{x}} \times 100\%$	Comparar variabilidad entre marcadores con unidades diferentes	Escala libre, útil para hormonas versus metabolitos	Inestable si la media se acerca a cero	Variables de razón siempre positivas	Variables con valores cercanos a cero o que puedan tomar negativos
Mediana absoluta de las desviaciones (MAD)	Desviación típica robusta alrededor de la mediana	$\text{MAD} = \text{mediana}(\lvert x_i - \text{mediana}(x) \rvert)$	Evaluación de dispersión en datos con outliers frecuentes (p. ej. conteos de citoquinas)	Muy resistente a valores extremos	Menos eficiente con distribuciones normales; poco difundida	Variables continuas con valores atípicos frecuentes	Muestras muy pequeñas o datos puramente categóricos

¿Cuál medida elegir?

- **Datos simétricos** y sin valores extremos: media + DS pintan el cuadro con colores realistas.
- **Datos sesgados** o con cola larga: mediana + IQR o mediana + MAD protegen frente a distorsiones.
- **Comparar dispersiones** entre variables de distinta escala: CV iguala el terreno numérico.

Visualizar SD con barras de error, o IQR dentro de diagramas de caja, proporciona al lector un mapa intuitivo de la variabilidad. En el próximo segmento combinaremos estas herramientas con medidas de forma y visualización gráfica para seguir revelando la historia oculta en el mar de números clínicos.

2.3. Medidas de Forma y Apuntalamiento

Tras ubicar el centro del grupo y medir cuán lejos se dispersan los datos, conviene mirar la **silueta** de la distribución. Igual que un pastel horneado puede crecer uniforme o arqueado en un lado, un conjunto de valores puede inclinarse hacia la derecha o hacia la izquierda y mostrar picos pronunciados o cimas achatadas. Las **medidas de forma** (asimetría) y **apuntalamiento** (curtosis) capturan esas sutilezas; detectar una inclinación acusada o un pico afilado ayuda a decidir si conviene transformar los datos o emplear pruebas no paramétricas.

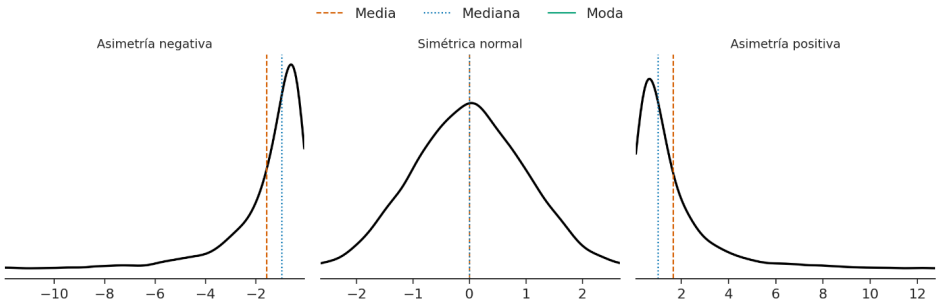
2.3.1 *Coeficiente de Asimetría*

En pacientes con triglicéridos, la mayoría se agrupa por debajo de 200 mg/dL, pero unos pocos superan 800. La distribución tiene **asimetría positiva** y el Sk puede rondar +2 o +3. Como un columpio con más peso en un extremo, la media se desplaza; por eso utilizamos la mediana y quizás aplicamos logaritmo antes de ajustar una regresión (ver figura 15).

Ventaja: alerta sobre la necesidad de transformaciones o de pruebas robustas.

Desventaja: un error de digitación en un solo valor puede cambiar drásticamente Sk.

Figura 15. *Posición relativa de media, mediana y moda en distribuciones con distinta asimetría. Elabora con Gemini Pro.*

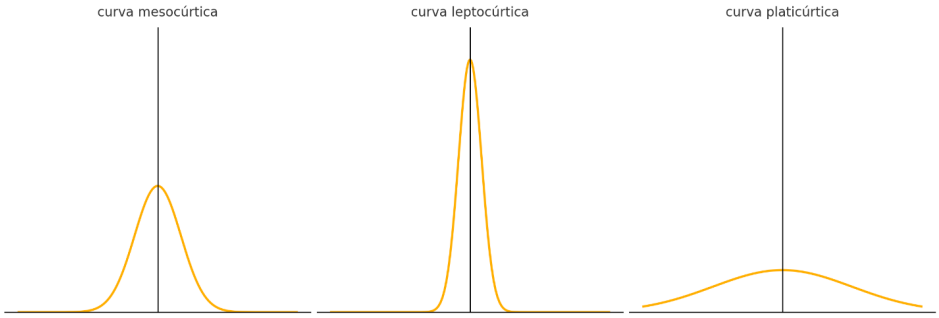


Nota. Densidad estimada de tres distribuciones: a la izquierda, asimetría negativa, donde la cola se extiende hacia la izquierda y la media (línea naranja) queda a la izquierda de la mediana (línea azul); en el centro, distribución simétrica normal con coincidencia de media, mediana y moda; y a la derecha, asimetría positiva, donde la cola se alarga hacia la derecha y la media se sitúa a la derecha de la mediana. Las líneas verde, azul y naranja representan la moda, la mediana y la media, respectivamente.

2.3.2 Coeficiente de Curtosis: "La Cúpula Afilada o el Panqueque"

Antes de entrar en la definición matemática, conviene imaginar la curtosis como la forma de la "cúpula": ¿aplanada como un panqueque o puntiaguda como un campanario? El coeficiente de curtosis (Ku) mide precisamente ese apuntalamiento y el peso de las colas. Por ejemplo, en población sana los valores de PCR ultrasensible generan una distribución amplia y baja (curtosis negativa o platicúrtica), como un pastel aplanado, mientras que los títulos de anticuerpos tras vacunación suelen formar un pico pronunciado con colas gruesas (curtosis positiva o leptocúrtica), parecido a una aguja. Un Ku de +5 sugiere colas tan pesadas que los intervalos construidos bajo supuestos normales podrían subestimar la probabilidad de extremos, por lo que es recomendable recurrir a métodos no paramétricos o *bootstrap* para estimar con seguridad los límites de confianza (ver figura 16).

Figura 16. *Comparación de curvas mesocúrtica, leptocúrtica y platicúrtica.* Elaborado con GeminiPRO.



Nota. Las tres densidades normales estándar muestran distintos grados de apuntalamiento (curtosis).

Ventaja: señala la presencia de colas que inflan la DS y pueden aumentar error tipo I si se usa t de Student.

Desventaja: en muestras pequeñas (n < 50) Ku varía mucho y puede llevar a conclusiones exageradas.

Tabla 8. *Resumen de medidas de forma y apuntalamiento.*

Métrica	Concepto narrado	Fórmula muestral*	Uso habitual	Ventajas	Desventajas	Variables adecuadas	Variables problemáticas
Coeficiente de asimetría (Sk)	Indica si la cola se alarga más a la izquierda (negativa) o a la derecha (positiva); la "palanca" que inclina el gráfico	$g_1 = \frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i - \bar{x}}{s}\right)^3$	Evaluar normalidad, decidir transformación logarítmica	Sensible a la dirección de la cola, fácil de estandarizar	Muy sensible a *outliers*, difícil de interpretar en valores absolutos	Continuas de razón o intervalo	Distribuciones multimodales donde la media no representa bien
Coeficiente de curtosis (Ku)	Mide el grado de pico o achatamiento respecto a una normal; "cuán puntiagudo" es el domo	$g_2 = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{x_i - \bar{x}}{s}\right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$	Detección de colas pesadas, robustez de varianza	Identifica riesgo de *outliers*, útil en bioquímica con picos extremos	Compleja de explicar a clínicos, puede ser inestable en n pequeño	Continuas con posible cola pesada, datos de laboratorio	Distribuciones discretas con pocos valores distintos

Cómo interpretar Forma y Apuntalamiento

1. **Sk ≈ 0 y Ku ≈ 0** sugiere forma casi normal, media ± DS e inferencia paramétrica son apropiadas.

2. **Sk > 1 o Sk < –1** indica asimetría marcada; preferir mediana + IQR, considerar logtransformación.

3. **Ku > 1** Denota colas más pesadas que la normal (picudos y extremos frecuentes). Los intervalos basados en percentiles o técnicas bootstrap resultan más fiables que los basados en DS bajo supuestos normales; además, el error tipo I de las pruebas t puede aumentar si no se corrige.

4. **Ku < –1** Refleja una distribución aplanada (colas ligeras, pico bajo). En estos casos, las estimaciones de DS infraestiman la variabilidad real de la cola y también es conveniente usar intervalos de percentiles o métodos no paramétricos.

Variables y Decisiones

- **Variables continuas en laboratorio**: Sk y Ku guían transformaciones antes de aplicar ANOVA.
- **Variables discretas con conteos bajos**: mejor usar tablas de frecuencia y evitar interpretar Sk/Ku que pueden ser erráticos.
- **Datos multimodales**: combinaciones de subpoblaciones producen Sk ≈ 0 pero múltiples picos; por eso siempre complementamos números con histogramas.

Con el retrato completo, centro, dispersión y forma, el investigador dispone de un altímetro, un velocímetro y un giroscopio para pilotar el análisis descriptivo. En el capítulo siguiente pasaremos del terreno de la descripción al de la inferencia, donde estos indicadores sirven de semáforo para escoger la pista de despegue adecuada.

Conclusiones

1. **Fundamento Descriptivo:** La estadística descriptiva es esencial para resumir, organizar y comprender las características básicas de los datos

clínicos antes de realizar análisis más complejos. Proporciona la "primera imagen" de la información.

2. **Elección de Métricas Clave:** La selección adecuada de medidas de tendencia central (media, mediana, moda), dispersión (DE, IQR, CV, etc.) y forma (asimetría, curtosis) es crucial y depende de la naturaleza de las variables (numéricas, categóricas), la forma de la distribución (simétrica, sesgada) y la presencia de valores atípicos.

3. **Guía Práctica:**

- Para datos simétricos sin *outliers*: Usar Media y DS.
- Para datos sesgados o con *outliers*: Preferir Mediana y (IQR) o MAD por su robustez.
- Para datos categóricos: La Moda es la medida de tendencia central indicada.
- Para comparar variabilidad entre variables con distintas unidades: Utilizar el CV.

4. **Importancia de la Forma:** La asimetría y la curtosis informan sobre la desviación de la normalidad, lo cual es vital para decidir si se necesitan transformaciones de datos o si se deben emplear métodos no paramétricos en la inferencia estadística. Valores altos de curtosis, por ejemplo, alertan sobre colas pesadas y mayor probabilidad de valores extremos.
5. **Complementariedad con Visualización:** Las medidas numéricas deben complementarse siempre con herramientas gráficas (como histogramas o diagramas de caja) para obtener una comprensión completa de la distribución de los datos y detectar patrones o anomalías que los números por sí solos podrían ocultar.
6. **Preparación para la Inferencia:** Un análisis descriptivo riguroso sienta las bases y guía la selección de las técnicas estadísticas inferenciales apropiadas para extraer conclusiones válidas sobre la población a partir de la muestra.

CAPÍTULO 5

TABLAS Y FIGURAS: ¿CÓMO DARLE FORMA VISIBLE A LOS DATOS CLÍNICOS?

Tables and Figures: How to Give Visible Form to Clinical Data?

Resumen

Este capítulo se enfoca en cómo las tablas y figuras dan forma visible a los datos clínicos, transformando cifras en imágenes comprensibles. Detalla la construcción y utilidad de las tablas de frecuencia, incluyendo el cálculo de frecuencias relativas y acumuladas y el uso de reglas como la de Sturges para agrupar datos continuos. Explica las tablas de contingencia como herramientas para analizar la asociación entre dos o más variables cualitativas, ejemplificando con estudios de exposición-resultado y su extensión a análisis estratificados. Posteriormente, describe diversas herramientas gráficas: histogramas para visualizar la distribución de frecuencias de variables continuas, diagramas de caja para resumir la mediana, rangos y detectar atípicos, diagramas de dispersión para explorar relaciones entre variables numéricas, gráficos de barras (simples y apiladas) para comparar categorías o composiciones relativas, gráficos de sectores para proporciones con pocas categorías, y polígonos de frecuencia para mostrar la forma de la distribución de manera continua. Finaliza con una tabla guía para seleccionar la visualización más adecuada según el objetivo y tipo de variable.

Palabras clave: Visualización de datos, tablas de frecuencia, tablas de contingencia, histogramas, diagramas de caja, gráficos estadísticos.

Abstract

This chapter focuses on how tables and figures give visible shape to clinical data, transforming numbers into understandable images. It details the construction and utility of frequency tables, including the calculation of relative and cumulative frequencies and the use of rules like Sturges' for grouping continuous data. It explains contingency tables as tools for analyzing the association between two

or more qualitative variables, exemplified by exposure-outcome studies and their extension to stratified analyses. Subsequently, it describes various graphical tools: histograms for visualizing the frequency distribution of continuous variables, box plots for summarizing the median, ranges, and detecting outliers, scatter plots for exploring relationships between numerical variables, bar charts (simple and stacked) for comparing categories or relative compositions, pie charts for proportions with few categories, and frequency polygons for showing the distribution's shape continuously. It concludes with a guide table for selecting the most appropriate visualization based on the objective and variable type.

Keywords: Data visualization, frequency tables, contingency tables, histograms, box plots, statistical graphics.

3.1. Tablas de Frecuencia

La estadística descriptiva cobra vida cuando pasa de cifras a imágenes, igual que un ecógrafo revela el latido fetal oculto tras la pared abdominal. A continuación, presentamos las herramientas gráficas y tabulares más usadas en investigación médica, su lógica interna, la fórmula necesaria cuando procede y un ejemplo práctico que conecte con la experiencia clínica.

—Yamil Liscano

Listados que cuentan cuántas veces aparece cada valor o categoría. Se acompañan de *frecuencia relativa* (fi/n) y *frecuencia acumulada* para convertir conteos en porcentajes útiles.

Intervalos de Clase

Al agrupar una variable continua conviene elegir el ancho de las casillas, w. Una regla clásica es la de Sturges:

$$k = 1 + 3{,}322\ \log_{10}(n), \qquad w = \frac{\max(x) - \min(x)}{k}$$

donde k es el número aproximado de clases.

Ejemplo

En una cohorte de 120 neonatos, la bilirrubina total oscila entre 0,5 y 18 mg/

dL. Sturges sugiere k≈8; cada intervalo será de (18–0,58)/8 ≈ 2,2 mg/dL. La tabla resultante muestra cuántos recién nacidos caen en 0–2,2, 2,2–4,4, y así sucesivamente, junto con los porcentajes y frecuencias acumuladas (ver tabla 9).

Paso a Paso

1. **Reunir los datos crudos**
   – Tienes 120 valores de bilirrubina neonatal (en mg/dL)

2. **Redondear (opcional)**
   – Para simplificar, redondeamos a 1 decimal.

3. **Construir la tabla de frecuencias crudas** (ver tabla 9).

Tabla 9. *Frecuencias crudas*.

Valor (mg/dL)	*fi*	*fi/nf*	Fi	Fi/n × 100
0.7	1	0.0083	1	0.83 %
0.8	2	0.0167	3	2.50 %
0.9	3	0.0250	6	5.00 %
1.0	4	0.0333	10	8.33 %
1.1	5	0.0417	15	12.50 %
…	…	…	…	…

– Contar cuántas veces aparece cada valor x_i → frecuencia absoluta f_i.

–Calcular frecuencia relativa: f_i/n.

– Calcular frecuencia acumulada

$$F_i = \sum_{j \le i} f_j \text{ y porcentaje acumulado } F_i/n \times 100.$$

4. **Aplicar la Regla de Sturges para Intervalos**

$$k = 1 + 3{,}322\ \log_{10}(n) \approx 1 + 3{,}322 \cdot \log_{10}(120) \approx 8 \quad \text{clases}$$

$$w = \frac{\max(x) - \min(x)}{k} \approx \frac{18 - 0{,}5}{8} \approx 2{,}2 \text{ mg/dL}$$

5. Definir los intervalos de clase

– De 0,5 a 2,7; 2,7 a 4,9; ... cada uno de ancho ≈ 2,2 mg/dL.

6. Construir la tabla de frecuencias por clase (ver tabla 10).

– Asignar cada valor de bilirrubina a su intervalo.

– Contar f_i, calcular f_i/n, Fi y Fi/n ×100

Tabla 10. *Frecuencias por intervalos de clase* (n=120)

Intervalo (mg/dL)	f_i	f_i/nf	**Fi**	$\mathbf{F_i/n \times 100}$
[0.5 – 2.7)	22	0.1833	22	18.33 %
[2.7 – 4.9)	34	0.2833	56	46.67 %
[4.9 – 7.1)	24	0.2000	80	66.67 %
[7.1 – 9.3)	23	0.1917	103	85.83 %
[9.3 – 11.5)	10	0.0833	113	94.17 %
[11.5 – 13.7)	4	0.0333	117	97.50 %
[13.7 – 15.9)	2	0.0167	119	99.17 %
[15.9 – 18.1)	1	0.0083	120	100.00 %

- $\mathbf{f_i}$ (frecuencia absoluta): número de observaciones que caen exactamente en el valor xi (en la tabla cruda) o dentro del intervalo de clase i.
- **fi/n** (frecuencia relativa): proporción de la muestra que corresponde a esa frecuencia, calculada como fi dividido por el tamaño total n.
- $\mathbf{F_i}$ o **Fac** (frecuencia acumulada): suma sucesiva de las frecuencias absolutas desde la primera categoría o clase hasta la i-ésima; sirve para saber cuántos casos quedan por debajo del límite superior de cada fila.
- $\mathbf{F_i/n}$ **o** $\mathbf{F_i/n \times 100}$ (frecuencia relativa acumulada o porcentaje acumulado): proporción o porcentaje de la muestra que corresponde a todas las observaciones hasta la i-ésima categoría o clase.

3.2 Tablas de Contingencia

Las tablas de contingencia surgen de la necesidad de entender, de manera sencilla y visual, cómo se relacionan dos variables cualitativas en un conjunto de datos clínicos o epidemiológicos, pues cada celda de la tabla refleja la correspondencia entre categorías, convirtiendo números crudos en historias reveladoras. En medicina, por ejemplo, imaginemos un estudio sobre el consumo de tabaco y la incidencia de hipertensión arterial: la tabla 2 × 2 nos permite ver cuántos pacientes fumadores versus no fumadores presentan o no presentan hipertensión, y a partir de esa disposición geométrica de frecuencias absolutas y relativas, podemos calcular medidas como el riesgo relativo o la odds ratio, que cuantifican la fuerza de la asociación y orientan la toma de decisiones en salud pública, sin perder de vista que cada celda puede esconder variaciones según la edad o el sexo, lo que invita a estratificar las tablas para afinar el análisis.

Cuando elevamos la mirada, las tablas de contingencia pueden ampliarse para incluir tres o más variables, formando cubos o tablas de mayor dimensión, y en estos casos se recurre a pruebas como el chicuadrado de independencia o exactas de Fisher, las cuales, pese a su aparente complejidad estadística, descansan en el mismo principio de comparar lo observado con lo esperado bajo la hipótesis de independencia; así, en un ensayo clínico de un nuevo antibiótico frente a *Staphylococcus aureus* resistente, podríamos cruzar el tratamiento (*antibiótico* A versus *placebo*) con la resolución de la infección y la aparición de efectos adversos, construyendo una tabla 2 × 2 × 2 que nos permita diferenciar no solo eficacia, sino también seguridad, y con ello diseñar intervenciones más precisas y personalizadas.

Al trasladar estos conceptos al laboratorio, una tabla de contingencia facilita explorar la relación entre la presencia de anticuerpos y la severidad de la enfermedad, por ejemplo categorizando niveles de inmunoglobulina G en rangos bajos, medios y altos, frente a eventos clínicos leves, moderados o graves; de ese modo, no solo cuantificamos tendencias, sino que también descubrimos patrones inesperados, como que un subtipo de paciente con respuesta intermedia presente mayor riesgo de complicaciones, una pista

que invita a explorar otros factores biológicos o de comportamiento. Dominar la construcción e interpretación de tablas de contingencia, por tanto, supone dotar al investigador de un altímetro estadístico capaz de medir asociaciones cualitativas, y así preparar el terreno para técnicas más sofisticadas de modelado multivariable o para reproducir hallazgos en metanálisis que integren evidencias de distintas poblaciones.

Ejemplo Práctico de Tabla de Contingencia

Imaginemos un estudio clínico con 100 pacientes en el que queremos analizar la asociación entre el **consumo de tabaco** (variable cualitativa dicotómica) y la **presencia de hipertensión arterial** (también dicotómica). Para organizar la tabla (ver tabla 11):

- **Filas**: variable explicativa o "exposición" (aquí, consumo de tabaco).
- **Columnas**: variable de resultado u "outcome" (aquí, hipertensión).

Tabla 11. *Tabla de contingencia para hipertensión y fumadores*.

	Hipertensión (Sí)	**Hipertensión (No)**	**Total filas**
Fumadores	30	20	50
No fumadores	10	40	50
Total columnas	40	60	100

Interpretación de celdas:

- La celda (fila "Fumadores", columna "Sí") contiene 30, es decir, 30 de los 50 fumadores tienen hipertensión.
- La celda (fila "No fumadores", columna "No") contiene 40, es decir, 40 de los 50 no fumadores no tienen hipertensión.

3.3. Histogramas

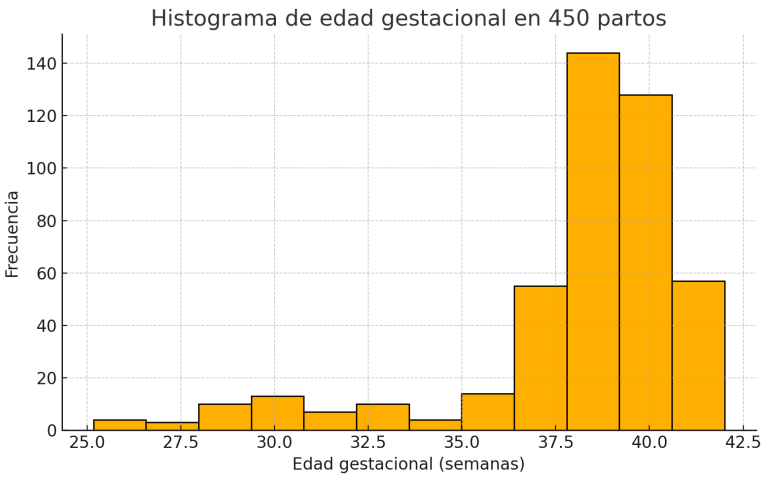
¿Qué muestran?

La frecuencia de una variable continua en intervalos adyacentes. El área de cada barra es proporcional a la cantidad de casos en ese rango.

Ejemplo

La distribución de la edad gestacional en 450 partos. Un histograma permite confirmar la simetría (o no) y detectar una posible cola izquierda de prematuros extremos.

Figura 17. *Histograma de edad gestacional.* Elaborado con GeminiPRO.



Nota. Observa cómo la distribución central alrededor de 38–40 semanas forma un "montículo" claro, mientras que la barra más baja a la izquierda ilustra la cola de partos prematuros extremos. Esto permite confirmar visualmente tanto la simetría general como la presencia de valores atípicos a la izquierda, guiando la elección de estadísticas descriptivas y cualquier transformación necesaria antes de análisis posteriores.

3.4. Diagramas de Caja (Box and Whisker)

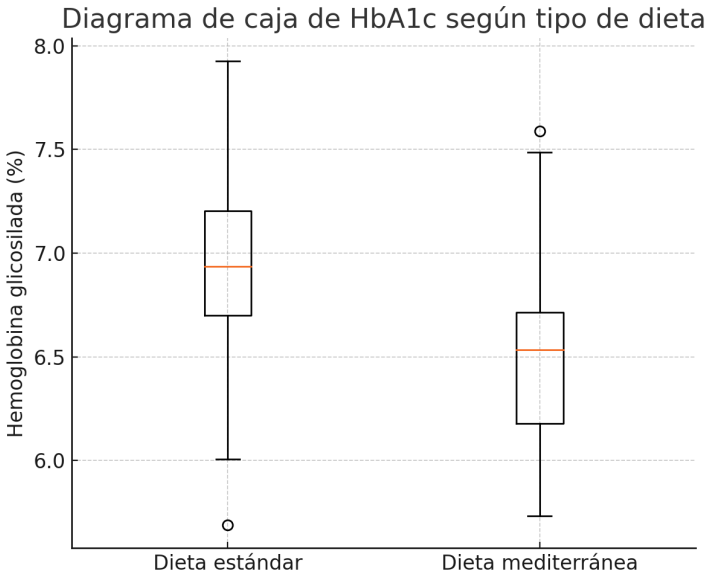
¿Qué Muestran?

Mediana, IQR, valores mínimos y máximos razonables, además de puntos atípicos.

Ejemplo

Comparar la hemoglobina glicosilada en pacientes con dieta estándar frente a dieta mediterránea. Un solo diagrama de cajas por grupo ilustra diferencias centrales y dispersión sin saturar al lector con números (ver figura 18).

Figura 18. *Diagrama de caja y bigotes de hemoglobina glicosilada según tipo de dieta*. Elaborado con GeminiPRO.



Nota. La **línea naranja** dentro de cada caja señala la mediana de HbA1c (%) en cada grupo. La **caja** abarca el rango intercuartílico (Q1–Q3), mostrando la dispersión central. Los **bigotes** se extienden hasta 1.5 × IQR o hasta el valor mínimo/máximo razonable, ilustrando el rango de datos esperable. Los **círculos** representan valores atípicos externos al rango de los bigotes.

Se aprecia que el grupo con **dieta mediterránea** tiene una mediana más baja y una caja más compacta, lo cual sugiere una menor variabilidad y mejores niveles de HbA1c, en comparación con la **dieta estándar**. Esta visualización resume eficazmente diferencias centrales y de dispersión sin saturar al lector con estadísticas numéricas detalladas, definiendo claramente dónde reside la mayor concentración de resultados y cómo emergen los valores extremos.

3.5. Diagramas de Dispersión (Scatter Plots)

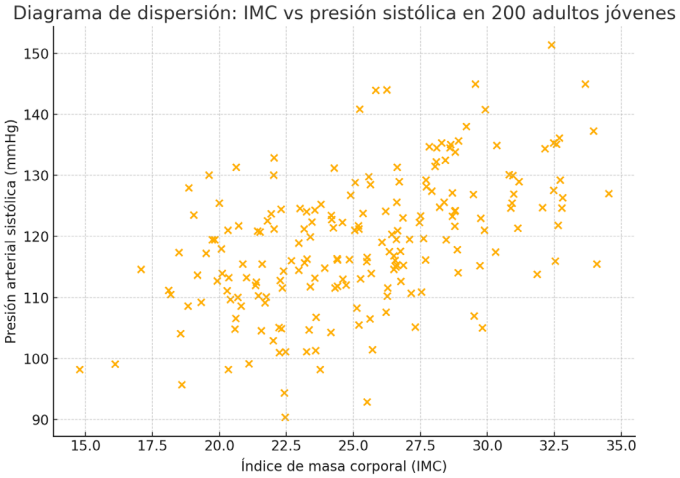
¿Qué muestran?

Cada punto es un individuo con dos coordenadas numéricas, x y y. Revelan correlaciones, nubes de puntos, patrones no lineales.

Ejemplo

Relación entre índice de masa corporal y presión sistólica en 200 adultos jóvenes; la nube ascendente motiva luego un modelo de regresión lineal (ver figura 19).

Figura 19. *Diagrama de dispersión IMC vs presión sistólica*. Elaborado con GeminiPRO.



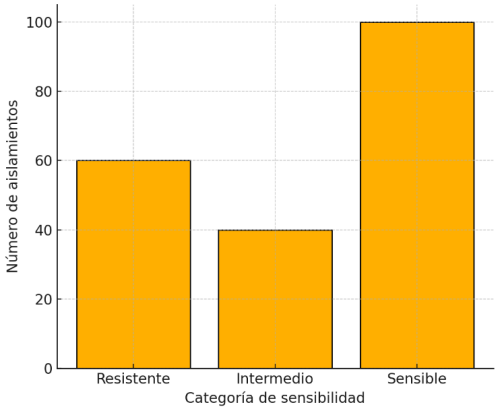
El diagrama de dispersión muestra 200 puntos, cada uno representando un adulto joven con su **IMC** en el eje horizontal y su **presión arterial sistólica** en el vertical, lo que permite visualizar una nube ascendente que sugiere correlación positiva; este patrón fundamenta luego un modelo de regresión lineal para cuantificar cómo varía la presión con el aumento del IMC, y ayuda a detectar posibles valores atípicos que podrían merecer estudio individual o explicación adicional.

3.6. Gráficos de Barras

¿Qué muestran?

Barras verticales o horizontales que representan el conteo o porcentaje de categorías. Conviene aplicar a variables nominales u ordinales, por ejemplo, prevalencia de tipos de Pseudomonas resistente, sensible, intermedio (ver figura 20).

Figura 20. *Gráfico de barras de prevalencia de sensibilidad a Pseudomonas.* Elaborado con GeminiPRO.



3.7. Barras Apiladas

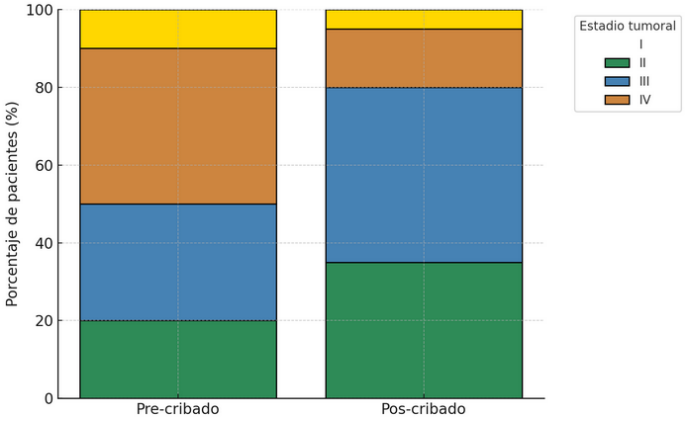
¿Qué muestran?

Cada barra suma 100 % y se divide en segmentos coloreados. Útiles para comparar la composición relativa entre grupos.

Ejemplo

Distribución de estadios tumorales (I, II, III, IV) en cáncer de colon para pacientes pre y poscribado; se visualiza si la introducción de la prueba fecal desplaza diagnósticos a estadios tempranos (ver figura 21).

Figura 21. *Barras apiladas de distribución de estadios tumorales en cáncer de colon antes y después del cribado fecal.* Elaborado con GeminiPRO.



Nota. Cada barra, correspondiente a pacientes antes y después del cribado, suma 100 % y está segmentada por estadio tumoral (I, II, III, IV). Se observa que el cribado desplaza diagnósticos hacia estadios tempranos (porción verde de estadio I aumenta de 20 % a 35 %), lo cual sugiere un beneficio potencial en la detección precoz y una mejora en la proporción de pacientes diagnosticados en etapas tratables.

3.8. Gráficos de Sectores (Pie Charts)

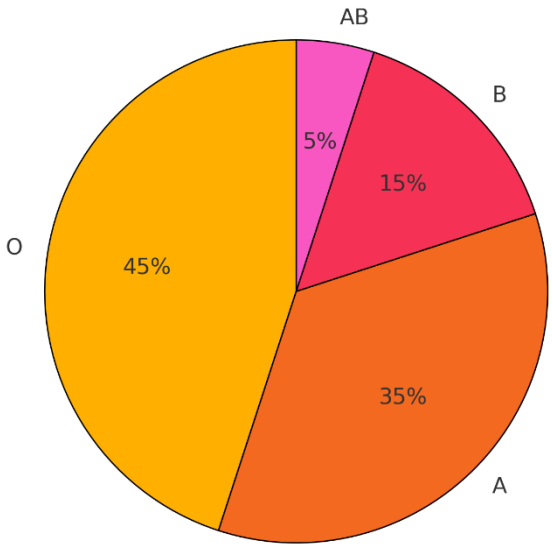
Regla Práctica

Útiles solo cuando hay pocas categorías (no más de cuatro rebanadas) y la suma debe ser 100 %.

Ejemplo

Proporción de grupos sanguíneos en un banco de donantes: O, A, B, AB. Más categorías diluirían el mensaje; convendrá un gráfico de barras (ver figura 22).

Figura 22. Distribución de grupos sanguíneos en un banco de donantes. Elaborado con GeminiPRO.



3.9. Polígonos de Frecuencia

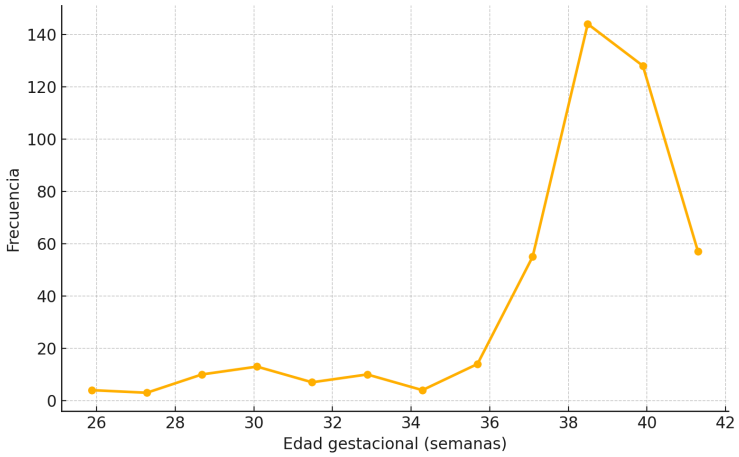
¿Qué Muestran?

Un polígono de frecuencia traza la frecuencia de cada clase en el eje vertical contra el punto medio de cada intervalo en el horizontal, y luego conecta esos puntos con líneas. De esta forma, aporta una visión continua de la forma de la distribución, facilitando la comparación entre varias muestras o la detección de picos y valles en la densidad de los datos.

Ejemplo

Volvamos al ejemplo de la edad gestacional en 450 partos. Tras construir el histograma con 12 intervalos de 1.5 semanas cada uno, podemos calcular el punto medio de cada clase, por ejemplo, 25.75, 27.25, 28.75 semanas, etc., y representar la frecuencia absoluta en cada uno de esos puntos. Al unirlos con segmentos de línea, obtenemos el polígono de frecuencia, que resalta con claridad la "colina" central de nacimientos a término y la pendiente descendente hacia prematuros extremos, sin los bloques de un histograma tradicional (ver figura 23).

Figura 23. *Polígono de frecuencia de edad gestacional en 450 partos*. Elaborado con GeminiPRO.



Nota. El trazado muestra los puntos medios de cada intervalo de edad gestacional y une sus frecuencias con líneas continuas. Este gráfico enfatiza la

continuidad de la distribución, permitiendo identificar de un vistazo el rango más frecuente, la simetría o asimetría de la nube de datos y cualquier meseta o valle que pudiera indicar subgrupos poblacionales, todo ello con un diseño ligero que favorece la comparación con otras cohortes o con series temporales de evolución perinatal.

Para que revises nuevamente que gráfico usar la Tabla 12 resume de manera práctica qué tipo de visualización se adapta mejor a cada objetivo estadístico y a la naturaleza de las variables estudiadas. Por ejemplo, cuando el propósito es describir frecuencias en datos cualitativos se recurre a una tabla de frecuencia o a un gráfico de barras; para proporciones con pocas categorías, el gráfico de sectores resulta más intuitivo; en cambio, las variables continuas requieren histogramas o, para enfatizar la continuidad de la distribución, polígonos de frecuencias. Asimismo, la mediana y el IQR se sintetizan en diagramas de caja, mientras que las relaciones entre dos variables numéricas se exploran con diagramas de dispersión, y las asociaciones entre dos variables cualitativas con tablas de contingencia y barras apiladas. Esta guía facilita la toma de decisiones al diseñar reportes estadísticos claros y coherentes.

Tabla 12. *Selección de gráficos según el objetivo de análisis y tipo de variable*

Objetivo	Tipo de variable	Figura sugerida
Describir frecuencias	Nominal / Ordinal	Tabla de frecuencia, gráfico de barras
Visualizar proporciones	Nominal (≤ 4 categorías)	Gráfico de sectores
Mostrar distribución	Continua	Histograma
Resumir rango y mediana	Continua o Numérica vs Categoría	Diagrama de caja
Analizar forma continua	Continua	Polígono de frecuencias
Explorar relación lineal	Continua vs Continua	Scatter plot
Relación entre dos categorías	Nominal vs Nominal	Tabla de contingencia, barras apiladas

Conclusiones

1. **Visualización Esencial:** Las representaciones gráficas y tabulares son fundamentales para dar vida y comprensión a la estadística descriptiva, permitiendo revelar patrones y relaciones en los datos clínicos que las cifras por sí solas no muestran.

2. **Tablas para Frecuencias y Asociaciones:** Las tablas de frecuencia organizan y resumen el conteo de datos categóricos o continuos agrupados, mientras que las tablas de contingencia son clave para explorar y cuantificar asociaciones entre variables cualitativas.

3. **Diversidad de Gráficos:** Existe una variedad de herramientas gráficas, cada una con un propósito específico:

- **Histogramas y Polígonos de Frecuencia:** Muestran la forma de la distribución de variables continuas.

- **Diagramas de Caja:** Resumen eficazmente la tendencia central (mediana), dispersión (IQR) y valores atípicos.

- **Diagramas de Dispersión:** Ideales para visualizar la relación y posible correlación entre dos variables continuas.

- **Gráficos de Barras/Sectores:** Adecuados para representar frecuencias o proporciones de variables categóricas (nominales/ordinales), siendo los de sectores mejores para pocas categorías.

- **Barras Apiladas:** Útiles para comparar la composición proporcional entre diferentes grupos.

4. **Selección Adecuada:** La elección de la tabla o gráfico correcto depende del objetivo del análisis y del tipo de variables que se estén manejando (cualitativas/cuantitativas, continuas/discretas). La Tabla 12 del capítulo proporciona una guía práctica para esta selección.

5. **Interpretación Contextual:** Más allá de la construcción, es crucial interpretar correctamente lo que cada tabla y figura muestra en el contexto del problema clínico o de investigación.

CAPÍTULO 6

CUANDO DOS EJES YA NO BASTAN: VISUALIZACIONES PARA TRES VARIABLES

When Two Axes Are No Longer Enough: Visualizations for Three Variables

Resumen

Este capítulo aborda la necesidad de visualizar datos clínicos cuando intervienen tres o más variables relevantes, superando las limitaciones de los gráficos bidimensionales. Presenta diversas estrategias gráficas avanzadas para incorporar dimensiones adicionales. Se describen técnicas como el gráfico de burbujas y la dispersión con color, que añaden una tercera variable numérica a un diagrama de dispersión mediante el tamaño o el color de los puntos. Las facetas (paneles múltiples) permiten comparar relaciones bivariadas a través de los niveles de una tercera variable categórica usando paneles separados con escalas idénticas. Los mapas de calor visualizan una tercera variable mediante intensidad de color en una matriz definida por dos variables discretizadas. Se introducen también los gráficos de contorno y superficies 3D para modelar respuestas continuas, los gráficos ternarios para visualizar composiciones de tres partes que suman 100%, las coordenadas paralelas para mostrar perfiles multivariados, y los diagramas de Sankey, *alluvial* y *chord* para representar flujos, transiciones y relaciones entre múltiples categorías o estados. El capítulo concluye con una tabla resumen que guía la selección de la técnica más apropiada.

Palabras Clave: Visualización multivariada, tres variables, gráficos avanzados, gráfico de burbujas, facetas, mapa de calor.

Abstract

This chapter addresses the need to visualize clinical data when three or more relevant variables are involved, overcoming the limitations of two-dimensional graphs. It presents various advanced graphical strategies to incorporate additional dimensions. Techniques described include the bubble plot and

color-mapped scatter, which add a third numerical variable to a scatter plot through point size or color. Facets (small multiples) allow comparison of bivariate relationships across levels of a third categorical variable using separate panels with identical scales. Heatmaps visualize a third variable through color intensity on a matrix defined by two discretized variables. Also introduced are contour plots and 3D surfaces for modeling continuous responses, ternary plots for visualizing three-part compositions summing to 100%, parallel coordinates for displaying multivariate profiles, and Sankey, alluvial, and chord diagrams for representing flows, transitions, and relationships among multiple categories or states. The chapter concludes with a summary table guiding the selection of the most appropriate technique.

Keywords: *Multivariate visualization, three variables, advanced graphics, bubble plot, facets, heatmap*

Las historias clínicas rara vez contienen solo dos columnas relevantes; con frecuencia queremos explorar la relación entre una exposición, un desenlace **y** un modulador, dosis, sexo, edad o marcador biológico, que pueda alterar la trama. A continuación, se presentan las estrategias más eficaces (y fáciles de reproducir) para poner en la misma escena tres dimensiones de la información.

4.1 Gráfico de Burbujas

Las historias clínicas rara vez contienen solo dos columnas relevantes; con frecuencia queremos explorar la relación entre una exposición, un desenlace y un modulador, dosis, sexo, edad o marcador biológico, que pueda alterar la trama. A continuación, se presentan las estrategias más eficaces (y fáciles de reproducir) para poner en la misma escena tres dimensiones de la información.

—Yamil Liscano

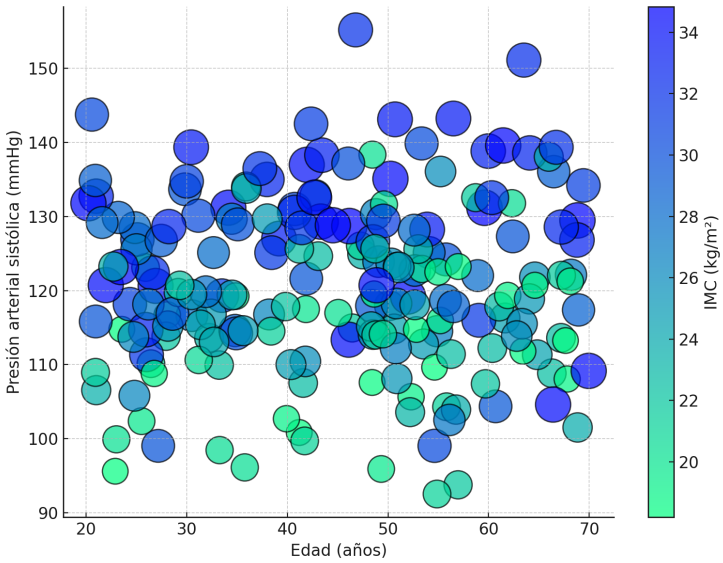
¿Qué Muestran?

Un ***buble Plot*** combina un diagrama de dispersión tradicional con un tercer eje codificado en el tamaño (o el color) de cada punto. Así, cada burbuja representa un caso, su posición en x e y refleja dos variables numéricas, y su diámetro (o tinte) aporta la magnitud de la tercera.

Ejemplo

En un estudio de hipertensión arterial, situamos la **edad** en el eje x y la **presión sistólica** en el eje y, y usamos el **IMC** para dimensionar las burbujas. Pacientes de mayor IMC aparecen como círculos más grandes, lo que permite ver de un vistazo si la obesidad agrava la presión a cualquier edad (ver figura 24).

Figura 24. *Bubble plot de edad vs presión sistólica (n=200), con el tamaño de la burbuja proporcional al* IMC (kg/m^2). Elaborado con GeminiPRO.



Nota. El gradiente de verdes a azules ayuda a distinguir IMC medios de altos, facilitando la identificación de subgrupos de riesgo.

4.2 Dispersión con Mapa de Color

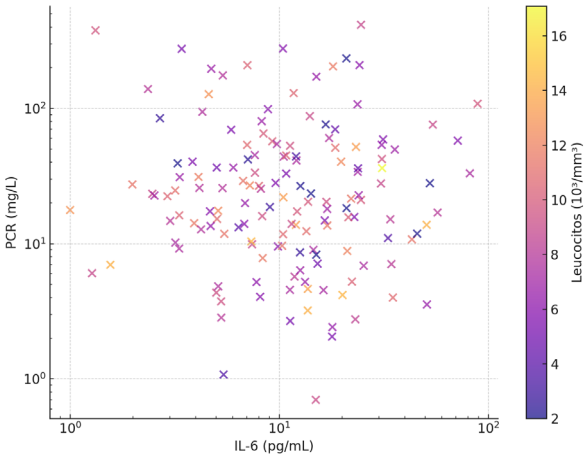
¿Qué Muestran?

Un **scatter** normal, donde cada punto tiene igual tamaño, pero se colorea mediante un degradado continuo según el valor de la tercera variable. Se adiciona así una capa cromática que actúa como termómetro cuantitativo.

Ejemplo

Para investigar inflamación sistémica, graficamos **IL6** en el eje x y **PCR** en el eje y, y coloreamos cada punto según el recuento de **leucocitos**. La escala de púrpuras a naranjas revela zonas de «tormenta citocínica», donde el recuento elevado acompaña a valores altos de ambos marcadores inflamatorios (ver figura 25).

Figura 25. *Diagram scatter de PCR (mg/L) vs IL6 (pg/mL) en escala logarítmica, coloreado por leucocitos ($10^3/mm^3$).* Elaborado con GeminiPRO.



Nota. El degradado continúa de púrpura (bajo) a amarillo (alto), destacando pacientes con respuesta inmune exacerbada.

4.3 Facetas

¿Qué Muestran?

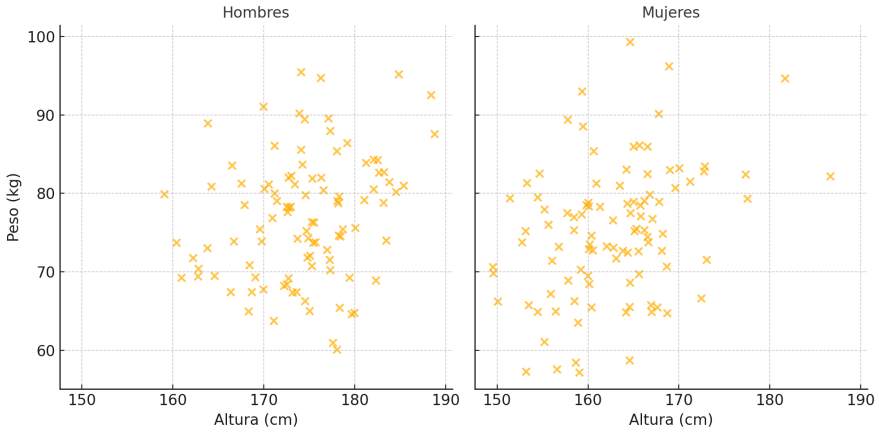
Varios **paneles** idénticos en escala, uno por nivel de la tercera variable cualitativa. Cada faceta despliega la misma relación x – y para facilitar comparaciones directas.

Ejemplo

Para explorar diferencias por **sexo**, trazamos **altura** en x y **peso** en y en dos paneles paralelos: uno para **hombres**, otro para **mujeres**. Así se aprecia cla-

ramente que la pendiente y dispersión son ligeramente distintas entre ambos sexos, sin cambiar escalas ni colores (ver figura 26).

Figura 26. *Gráficos por facetas de altura (cm) vs peso (kg) según sexo*. Elaborado con GeminiPRO.



Nota. Cada panel usa ejes idénticos, permitiendo comparar forma de nube y posibles valores atípicos en hombres y mujeres.

4.4 Mapa de Calor

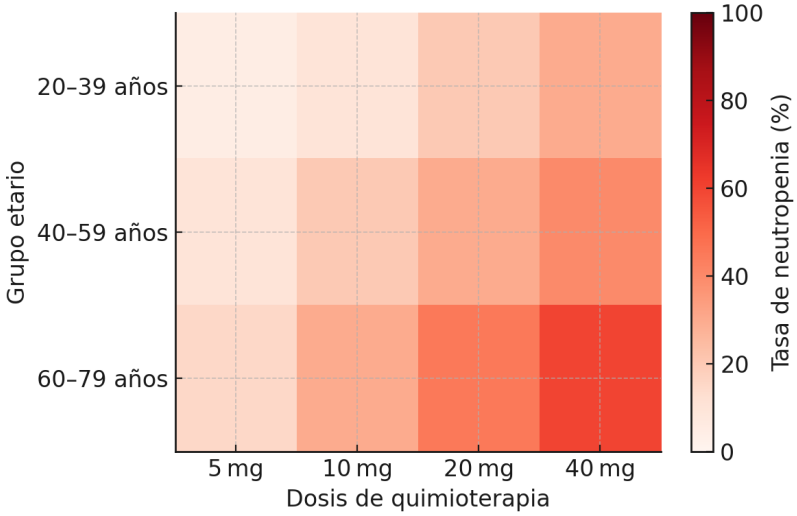
¿Qué Muestran?

Un **mapa de calor** agrupa en filas y columnas dos variables (pueden ser discretizaciones de continuas), y colorea cada celda según el valor medio o la tasa de una tercera variable. El resultado es un mosaico con gradientes que resaltan "puntos calientes".

Ejemplo

En un ensayo de quimioterapia, agrupamos dosis (**5 mg**, **10 mg**, **20 mg**, **40 mg**) en el eje x y rangos de edad (**20–39**, **40–59**, **60–79 años**) en el eje y. Cada casilla muestra la **tasa de neutropenia** (%), coloreada de blanco a rojo según la frecuencia (ver figura 27).

Figura 27. *Heatmap de tasa de neutropenia (%) según dosis y grupo etario*. Elaborado con GeminiPRO.



Nota. Las celdas más oscuras identifican combinaciones con mayor riesgo, orientando ajustes de protocolo en población mayor.

4.5 Gráfico de Contorno / Superficie 3D

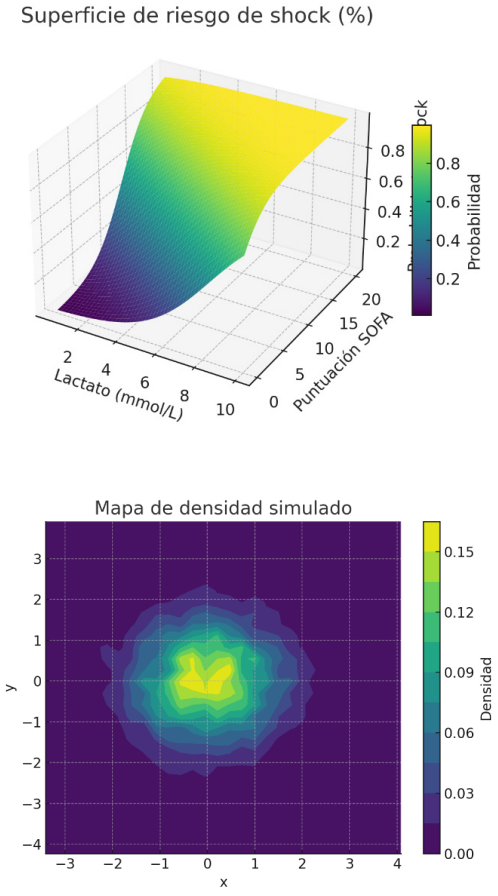
¿Qué Muestran?

Un ***contour plot*** dibuja curvas de nivel que conectan puntos con idéntico valor de la tercera variable, mientras que un **3D *surface*** levanta un modelo continuo en forma de malla. Ambos revelan mesetas, valles y crestas de probabilidad o magnitud.

Ejemplo

Modelamos la **probabilidad de shock séptico** como función del **lactato** y la puntuación **SOFA**. Un *contour map* muestra líneas que unen riesgos del 10 %, 20 %, 50 % y 80 %, delineando la frontera donde el riesgo se dispara y ayudando a definir umbrales clínicos críticos (ver figura 28).

Figura 28. *Contour plot de probabilidad de shock (%) según lactato (mmol/L) y* SOFA. Elaborado con GeminiPRO.



Nota. Las curvas de nivel guían la identificación de zonas de alto riesgo sin necesidad de paneles repetitivos.

4.6 Gráfico Ternario

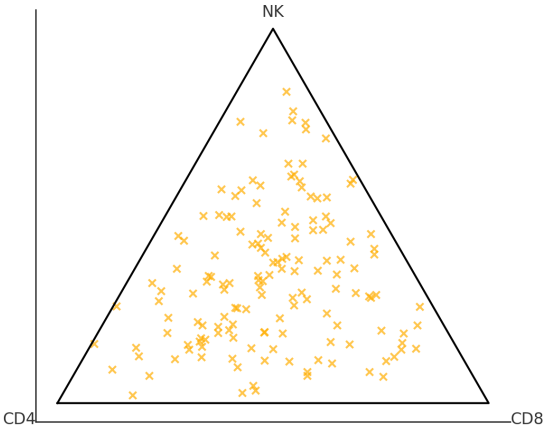
¿Qué Muestran?

Un **gráfico ternario** ubica tres proporciones, que suman 100 %, en un triángulo. Cada vértice representa el 100 % de un componente, y cualquier punto interior indica mezcla relativa de los tres.

Ejemplo

En estudios de inmunidad, los linajes **CD4**, **CD8** y **células NK** suman el 100 % de linfocitos. El *ternary plot* sitúa cada paciente según el balance entre estos tres subconjuntos, evidenciando perfiles inmunes dominantes o desvinculados (ver figura 29).

Figura 29. *Ternary plot de porcentajes de CD4, CD8 y NK en 150 pacientes*. Elaborado con GeminiPRO.



Nota. El triángulo facilita comparar perfiles celulares sin perder la unidad total.

4.7 Coordenadas Paralelas

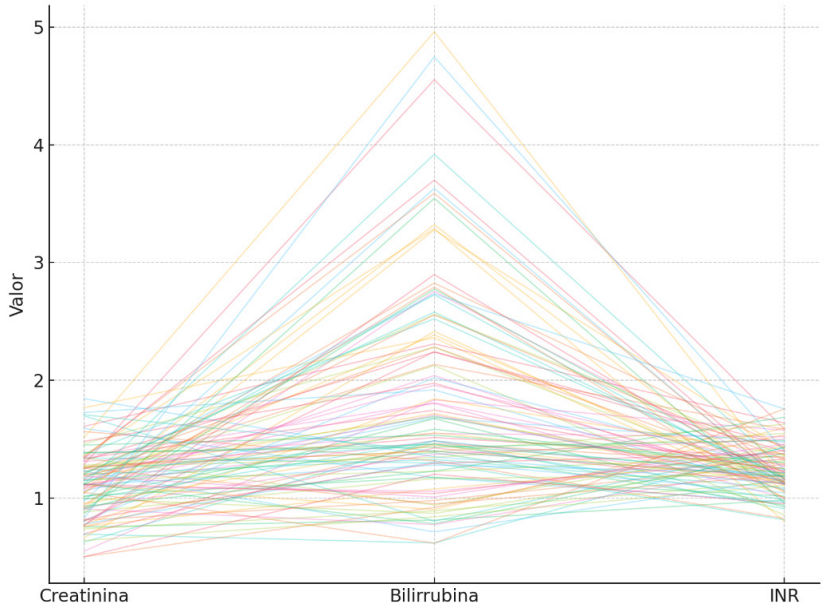
¿Qué Muestran?

Una serie de ejes verticales paralelos, uno por cada variable continua, y una **línea** que atraviesa todos describe un caso. Permiten visualizar trayectorias multivariantes en una sola trama.

Ejemplo

Para caracterizar la **gravedad de la cirrosis**, trazamos en ejes paralelos **función renal** (creatinina), **bilirrubina** e **INR**. Cada paciente aparece como una línea que sube o baja, mostrando patrones combinados de disfunción orgánica (ver figura 30).

Figura 30. *Parallel coordinates de función renal, bilirrubina e INR en pacientes con cirrosis (n=100).* Elaborado con GeminiPRO.



Nota. La superposición de líneas densas destaca subgrupos con insuficiencia múltiple.

4.8 *Sankey diagram*

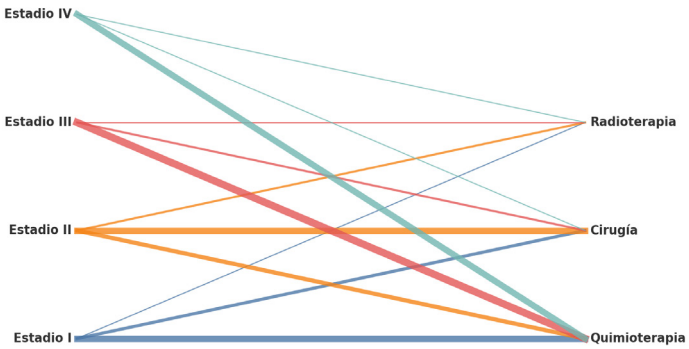
¿Qué Muestran?

Un diagrama de *Sankey* representa flujos entre dos o más estados o categorías, donde el grosor de cada "corriente" es proporcional al volumen o número de unidades que transitan de un nodo a otro.

Ejemplo

En un estudio de seguimiento oncológico, podemos ilustrar cómo 200 pacientes pasan de un estadio tumoral inicial (I–IV) a distintos tipos de tratamiento (quimioterapia, cirugía, radioterapia). Cada banda gruesa revela las transiciones más frecuentes, ayudando a identificar cuellos de botella donde muchos pacientes abandonan un protocolo o cambian de terapia (ver figura 31).

Figura 31. *Sankey diagram de flujo de pacientes entre estadio tumoral y línea de tratamiento (n=200).* Elaborado con GeminiPRO.



Nota. Las bandas más anchas marcan las trayectorias dominantes, facilitando el rediseño de rutas terapéuticas para reducir pérdidas en cada etapa.

4.9 *Alluvial Pslot*

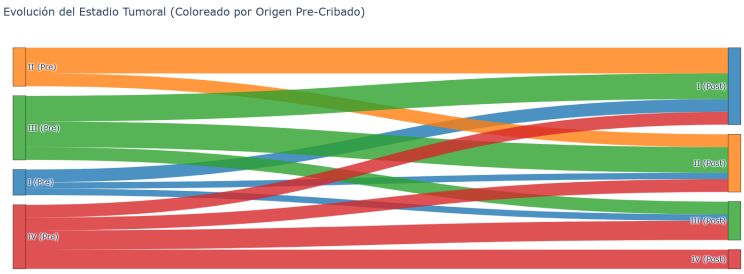
¿Qué Muestran?

Un *alluvial plot* conecta categorías a través de varias etapas o "bloques" mediante bandas fluídas, mostrando cómo se redistribuyen los elementos entre niveles sucesivos.

Ejemplo

En un ensayo de cáncer de colon pre y poscribado, un *alluvial plot* puede revelar cómo los pacientes migran entre estadios I–IV antes y después de implementar la prueba de sangre oculta en heces, destacando el aumento de flujos hacia estadios tempranos (ver figura 32).

Figura 32. *Alluvial plot de evolución de estadio tumoral en cohortes pre y poscribado (n=150).* Elaborado con GeminiPRO.



Nota. Las "riberas" muestran pérdidas y ganancias de pacientes en cada estadio, evidenciando el impacto del cribado en el diagnóstico precoz.

4.10 *Chord Diagram*

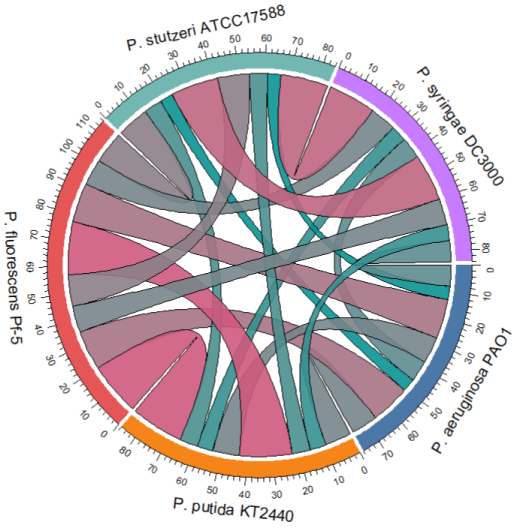
¿Qué Muestran?

Un *chord diagram* dibuja categorías alrededor de un círculo y conecta pares de ellas con arcos cuyo grosor indica la magnitud de la relación o intercambio entre cada par.

Ejemplo

Para mapear resistencias microbianas, un *chord diagram* de cinco cepas de Pseudomonas ilustra la frecuencia de genes de resistencia compartidos entre cada par de cepas, mostrando en un solo gráfico conexiones bidireccionales y facilitando la detección de cepas "puente" que transmiten resistencia a otras (ver figura 33).

Figura 33. *Chord diagram de resistencias cruzadas en cepas de Pseudomonas* (*n*=50). Elaborado con GeminiPRO.



Nota. Cada arco une dos cepas y su grosor refleja el porcentaje de genes de resistencia compartidos, resaltando posibles puntos de intervención en el control de infecciones.

Ahora bien, veremos un resumen de la información dada en la Tabla 13, ésta agrupa diversas técnicas de visualización multivariada que permiten trascender el diagrama de dispersión clásico y explorar hasta seis dimensiones en un solo gráfico, o bien comparar subgrupos de forma intuitiva. Por ejemplo, un *bubble plot* combina la orientación espacial de un *scatter plot* con el tamaño de las burbujas para incorporar el índice de masa corporal, mientras que un *heatmap* resume de un vistazo relaciones bidimensionales discretizadas, como dosis de fármaco y edad, mediante un gradiente de color. Para variables categóricas adicionales se prefieren las *facetas*, que reproducen varios paneles con la misma escala, evitando ambigüedades cromáticas. Cada estrategia describe el momento idóneo de uso, su principal fortaleza y qué evitar, de modo que el investigador pueda seleccionar la técnica más apropiada según la complejidad de sus datos y el mensaje que desee transmitir.

Tabla 13. *Estrategias gráficas avanzadas para visualizar múltiples dimensiones*

Estrategia	Idea visual	Cuando luce mejor	Ventaja principal	Precaución	Ejemplo clínico
Bubble plot (diagrama de burbujas)	Dispersión xy; el tamaño o el color del punto codifica la tercera variable	Tres métricas numéricas continuas	El lector ya conoce la lógica del scatter; añadir tamaño o color es intuitivo	Evitar solapamientos en grupos densos	Edad vs presión sistólica, burbuja proporcional al IMC
Colormapped scatter	Todos los puntos del mismo tamaño, degradado de color continuo	Variable tercera continua de amplio rango	Fácil de añadir escala de color con leyenda	Degradados mal impresos en escala de grises	PICC linfocitario vs IL6 con color para PCR
Small multiples / Facetas	Varios paneles (o símbolo diferente) para niveles de la variable categórica	Tercera variable cualitativa (sexo, grupo de riesgo)	Conserva la escala y evita mezclas de color	Multiplica paneles si hay muchas categorías	Relación tallapeso por sexo (hombres arriba, mujeres abajo)
Heatmap	Eje x y eje y discretizados; tono o intensidad representa la tercera variable	Dos variables continuas pero fácilmente agrupables; tercera como media/porcentaje	Patrón general salta a la vista (gradiente)	Se pierde detalle de caso individual	Frecuencia de neutropenia según dosis y edad
Contour plot / 3D surface	Líneas de igual valor (o superficie)	Modelos suaves de probabilidad	Muestra picos y valles de riesgo	Requiere interpolación fiable	Probabilidad de shock según lactato y SOFA

Estrategia	Idea visual	Cuando luce mejor	Ventaja principal	Precaución	Ejemplo clínico
Ternary plot	Triángulo equilátero, cada vértice = 100 % de un componente	Tres proporciones que suman 1	Representa mezclas (macronutrientes, genotipos)	No apto para variables independientes	Porcentajes CD4, CD8 y NK en pacientes
Sankey diagram	Flujo de unidades de una categoría a otra, grosor proporcional al volumen	Trayectorias o transiciones entre dos o más etapas	Revela pérdidas o ganancias en cada paso del flujo	Puede saturarse si hay demasiados nodos o ramas	Flujo de pacientes entre estadios de enfermedad y líneas de tratamiento
Alluvial plot	Barras multicategoría conectadas por "riberas" que muestran cambios de estado o grupo	Transiciones categóricas a través de varias etapas/tramos	Destaca patrones de migración o evolución	Legibilidad baja si hay muchas categorías o cambios	Evolución de estadio tumoral en cohortes pre y postratamiento
Chord diagram	Arcos que conectan categorías circulares, ancho proporcional al volumen de la relación	Relaciones simétricas entre múltiples grupos	Ilustra conexiones bidireccionales con un solo gráfico	Difícil de interpretar sin leyenda clara	Interacciones potenciales entre antimicrobial resistances en cepas
Parallel coordinates	Ejes paralelos verticales, cada línea = un caso	Tres o más métricas continuas	Permite > 3 dimensiones	Difícil leer densas	Función renal, bilirrubina y INR en cirrosis

Conclusiones

1. **Limitaciones de Dos Ejes:** El análisis de datos clínicos complejos frecuentemente requiere visualizar la interacción de más de dos variables (ej. exposición, desenlace y un modulador como edad, sexo, IMC).
2. **Incorporación de Tercera Dimensión:** Existen múltiples técnicas para añadir información de una tercera variable a gráficos bidimensionales:

- **Tamaño/Color:** *Bubble plots* y *color-mapped scatters* usan el tamaño o color del punto en un diagrama de dispersión para representar una tercera variable numérica.
- **Paneles Múltiples (Facetas):** Permiten comparar una relación X-Y entre diferentes subgrupos definidos por una tercera variable categórica, manteniendo la consistencia de las escalas.
- **Intensidad/Tono:** Los *heatmaps* usan color en una cuadrícula para mostrar el valor de una tercera variable en función de dos variables discretizadas.

3. **Visualizaciones Especializadas:** Se presentaron gráficos para relaciones específicas:

- **Superficies y Contornos:** *Contour plots* y superficies 3D son útiles para visualizar modelos de probabilidad o magnitud en función de dos predictores continuos.
- **Composiciones:** Los *ternary plots* son específicos para representar proporciones de tres componentes que suman 100%.
- **Flujos y Relaciones Múltiples:** Diagramas de Sankey, *alluvial* y *chord* visualizan flujos, transiciones entre estados o relaciones complejas entre múltiples entidades.
- **Perfiles Multivariados:** Los *parallel coordinates* permiten comparar perfiles individuales a través de múltiples variables continuas.

4. **Selección Guiada:** La elección de la técnica de visualización avanzada depende de la naturaleza de las variables (continuas, categóricas, proporciones), el número de dimensiones a representar y el objetivo especí-

fico del análisis (comparar grupos, mostrar relaciones, seguir flujos). La Tabla 13 ofrece una guía práctica para esta selección.

5. **Claridad y Precaución:** Aunque potentes, estas visualizaciones requieren cuidado en su diseño (evitar solapamientos, elegir colores adecuados, usar leyendas claras) para asegurar que comuniquen la información de forma efectiva y no generen confusión.

CAPÍTULO 7

PROBABILIDAD Y DISTRIBUCIONES TEÓRICAS EN CIENCIAS DE LA SALUD

Probability and Theoretical Distributions in Health Sciences

Resumen

Este capítulo introduce los conceptos fundamentales de probabilidad como base para manejar la incertidumbre en ciencias de la salud, definiendo experimento aleatorio, espacio muestral, suceso, axiomas de probabilidad, probabilidad condicional y el teorema de Bayes. Explica la idea de variable aleatoria y su distribución de probabilidad (masa o densidad). Posteriormente, detalla las características, propiedades y aplicaciones clínicas de las principales distribuciones teóricas: la Normal (Gaussiana) como modelo central para variables continuas; la t de Student para muestras pequeñas o varianza poblacional desconocida; la Chi-cuadrado (χ^2) para pruebas de bondad de ajuste y tablas de contingencia; la F para comparación de varianzas y ANOVA; la Binomial para contar éxitos en ensayos fijos; y la de Poisson para eventos raros en un intervalo. Finalmente, aborda las relaciones entre estas distribuciones, así como el uso de transformaciones de datos (log, raíz cuadrada, Box-Cox) para cumplir supuestos de normalidad y las aproximaciones entre distribuciones (Binomial a Poisson/Normal, Poisson/t a Normal) para simplificar cálculos en escenarios específicos.

Palabras clave: Probabilidad, distribuciones de probabilidad, distribución normal, distribución T de Student, distribución binomial, distribución de Poisson.

Abstract

This chapter introduces fundamental probability concepts as a basis for handling uncertainty in health sciences, defining random experiment, sample space, event, probability axioms, conditional probability, and Bayes' theorem. It explains the idea of a random variable and its probability distribution (mass or density). Subsequently, it details the characteristics, properties, and clinical

applications of the main theoretical distributions: the Normal (Gaussian) as a central model for continuous variables; the t-Student for small samples or unknown population variance; the Chi-squared (χ^2) for goodness-of-fit tests and contingency tables; the F for variance comparison and ANOVA; the Binomial for counting successes in fixed trials; and the Poisson for rare events in an interval. Finally, it addresses the relationships between these distributions, as well as the use of data transformations (log, square root, Box-Cox) to meet normality assumptions and approximations between distributions (Binomial to Poisson/Normal, Poisson/t to Normal) to simplify calculations in specific scenarios.

***Keywords**: Probability, probability distributions, normal distribution, t-distribution, binomial distribution, Poisson distribution*

5.1 Conceptos Elementales de Probabilidad

Tras haber recorrido el paisaje de la estadística descriptiva y aprendido a contar, agrupar y visualizar datos con histogramas, diagramas de caja o *scatter plots*, damos ahora un paso hacia el terreno de la incertidumbre, donde la probabilidad se convierte en brújula para orientar nuestras conclusiones. En su esencia, la probabilidad mide la plausibilidad de que ocurra un evento, ya sea que un paciente presente una reacción adversa tras un tratamiento, que un patógeno desarrolle resistencia o, en la vida cotidiana, que llueva mañana al salir del hospital. Para abordarla con rigor, partimos de tres pilares: la definición del espacio muestral, el conjunto de sucesos y la función que asigna a cada suceso un número entre cero y uno, donde cero significa imposibilidad y uno certeza absoluta.

Imaginemos una caja con cien cápsulas numeradas, en la que diez contienen un marcador genético de resistencia. El espacio muestral incluye todas las cápsulas, cada suceso es "extraer la cápsula 17" o "extraer una cápsula resistente", y la probabilidad de este último suceso es 10/100, es decir, 0,1. De modo parecido funciona el razonamiento en epidemiología cuando establecemos la frecuencia esperada de un evento, como un efecto secundario grave, y lo expresamos como proporción entre casos observados y total de pacientes, un enfoque que enlaza con la frecuencia relativa que ya vimos en tablas de frecuencias.

Para que este edificio teórico no se derrumbe ante la complejidad, formalizamos tres axiomas elementales, basados en la teoría de Kolmogórov: la probabilidad de cualquier suceso es no negativa, la suma de probabilidades de sucesos mutuamente excluyentes equivale a la probabilidad de su unión y la probabilidad del suceso seguro, es decir del espacio completo, es uno. Estos principios nos permiten derivar reglas indispensables, como la probabilidad complementaria, $P(\text{no ocurre}) = 1 - P(\text{ocurre})$, y las reglas de adición y multiplicación, que gobiernan desde simples combinaciones hasta complejas cadenas de eventos dependientes.

Cuando los sucesos interactúan, la noción se expande a la probabilidad condicional, $P(A|B)$, que cuantifica la probabilidad de que A ocurra sabiendo que B ya ha ocurrido. En prácticas clínicas, esto equivale a evaluar el riesgo de complicaciones dado un biomarcador positivo, una idea central en el cálculo de riesgos relativos o en la aplicación de la fórmula de Bayes, que integra información previa, la prevalencia, con evidencia nueva, la sensibilidad y especificidad de una prueba diagnóstica. Gracias a ella, podemos traducir los resultados de un test rápido de antígenos o de PCR en probabilidades posttest que informen decisiones terapéuticas individualizadas.

Al volver la mirada desde las tablas de frecuencia y los histogramas hacia la incertidumbre, descubrimos que la probabilidad no es un capricho matemático, sino la herramienta que traduce la variabilidad clínica en un mapa de rutas posibles. Para asentar buenos cimientos, repasemos cada concepto con metáforas que resuenen en el consultorio y ejemplos que nazcan del mismo latido del paciente.

Imagine por un momento que cada visita al consultorio es como lanzar un dado médico: aunque el protocolo sea idéntico, pequeñas variaciones en la respiración, el estrés o la colocación del manguito hacen que la **medición de la presión arterial** sea un *experimento aleatorio*, pues repetirlo bajo las mismas condiciones puede arrojar valores distintos. Este ejercicio cotidiano nos aproxima al primer concepto elemental de la probabilidad:

Experimento Aleatorio

Es cualquier procedimiento cuyos resultados no podemos predecir a ciencia cierta, aunque conozcamos las reglas. Medir la glucemia, evaluar un elec-

trocardiograma o lanzar un dado justo son ejemplos de experimentos aleatorios.

Una vez definido el experimento, debemos conocer el **espacio muestral (Ω)**, el "armario" de todas las lecturas posibles. Si categorizamos la presión sistólica en tres niveles, 'normal' (N), 'elevada' (E) o 'hipertensión grave' (G), entonces

$$\Omega = \{N, E, G\}$$

Antes de interpretar cualquier resultado, sabemos que debe provenir de este escaparate.

Un **suceso** o **evento** es simplemente un subconjunto de Ω. Por ejemplo,

$$A = \{E, G\}$$

podría describir el evento "lectura ≥ 140 mmHg". En el dado, el evento "salga un número par" es {2,4,6}.

A cada evento A le asignamos una **probabilidad** P(A) ∈ [0,1], de modo que P(Ω)=1 (algo tiene que ocurrir). Si A y B son mutuamente excluyentes, por ejemplo, "tensión < 90" y "tensión ≥ 140", entonces

$$P(A \cup B) = P(A) + P(B)$$

El **suceso complementario** A^c recoge las situaciones en que A no ocurre. Si A es "fiebre al ingreso", A^c será "ausencia de fiebre", y matemáticamente vale

$$P(A^c) = 1 - P(A)$$

Al combinar dos eventos, hablamos de **unión** ($A \cup B$: "fiebre o dolor de cabeza") o **intersección** ($A \cap B$: "fiebre y dolor de cabeza").

Cuando conocer que B ha ocurrido no altera la probabilidad de A, los llamamos **independientes**, con

$$P(A \cap B) = P(A)\,P(B)$$

Al lanzar dos dados, el resultado del primero no influye en el segundo; de modo similar, una mutación en un gen no relacionada no cambia la proba-

bilidad de hallar otra mutación en un gen distinto, si son verdaderamente independientes.

Más allá, la **probabilidad condicionada**

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

nos permite responder la pregunta clínicamente relevante: "¿Cuál es el riesgo de neumonía (A) dado que el paciente está intubado (B)?"

De aquí surge la **regla de la multiplicación**

$$P(A \cap B) = P(A)\,P(B \mid A) = P(B)\,P(A \mid B)$$

que en farmacocinética, por ejemplo, calcularía la probabilidad de un pico tóxico en dos tomas sucesivas condicionándose la segunda al nivel alcanzado tras la primera.

El **teorema de Bayes**,

$$P(A \mid B) = \frac{P(B \mid A)\,P(A)}{P(B)}$$

es el pilar de la MBE: convierte la sensibilidad y especificidad de un test en la probabilidad de enfermedad tras un resultado positivo o negativo, guiando decisiones tan críticas como iniciar antibióticos empíricos tras una PCR positiva.

Para formalizar cuantitativamente cada resultado, definimos una **variable aleatoria** X, que puede ser **discreta** (conteo de leucocitos) o **continua** (nivel de glucosa). Su **distribución de probabilidad** describe cómo se reparte la masa (discreta) o la densidad (continua) de probabilidad sobre los posibles valores de X.

Supongamos una consulta con **100 pacientes** adultos, en la que registramos dos condiciones frecuentes: **hipertensión (A)** y **diabetes (B)**. De esos 100:

- 40 son hipertensos,
- 30 son diabéticos,
- 15 tienen ambas enfermedades.

A partir de estos datos podemos ilustrar todos los conceptos:

1. **Probabilidad marginal**

   - P(A)= probabilidad de hipertensión = 40/100 = **0,40**.
   - P(B) = probabilidad de diabetes = 30/100 = **0,30**.

2. **Probabilidad conjunta**

   - P(A∩B) = probabilidad de ser hipertenso **y** diabético = 15/100 = **0,15**.

3. **Probabilidad de unión**

Como A y B no son mutuamente excluyentes (15 pacientes tienen ambas),

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0{,}40 + 0{,}30 - 0{,}15 = 0{,}55$$

Es decir, un **55 %** de la muestra tiene al menos una de las dos condiciones.

4. **Suceso complementario**

   - $P(A^c)$ = probabilidad de **no** ser hipertenso = 1 – 0,40 = **0,60**.
   - $P(B^c)$ = probabilidad de **no** ser diabético = 1 – 0,30 = **0,70**.

5. **Probabilidad condicional**

$$P(B \mid A) = \text{“diabetes dado que hay hipertensión”} = \frac{P(A \cap B)}{P(A)} = \frac{0{,}15}{0{,}40} = 0,375,$$

es decir, el **37,5** % de los hipertensos también son diabéticos.

   - De igual modo P(A|B) = 0,15/0,30 = 0,50, la **mitad** de los diabéticos son hipertensos.

6. Eventos independientes o dependientes

Comparamos $P(A \cap B)$ con $P(A) \times P(B)$:

$P(A)\ P(B) = 0{,}40 \times 0{,}30 = 0{,}12$, $P(A \cap B) = 0{,}15$

Dado que **$0{,}15 \neq 0{,}12$**, los eventos **no** son independientes, sino **dependientes**: conocer que un paciente es hipertenso **aumenta** del 30 % al 37,5 % la probabilidad de que sea diabético.

Conocer estos conceptos, experimento, espacio muestral, evento, probabilidad, complemento, unión, intersección, independencia, probabilidad condicionada, multiplicación, Bayes, variable aleatoria y distribución, es haber armado el andamiaje teórico que da sentido a cualquier análisis estadístico. En el siguiente tema, exploraremos las **distribuciones de probabilidad** concretas (binomial, Poisson, normal, t de Student, etc.), que modelan fenómenos clínicos específicos y convierten la teoría en predicción rigurosa.

5.2 Distribuciones de probabilidad (normal, t, χ^2, F, Bsinomial, Poisson)

5.2.1. *Introducción a las distribuciones de probabilidad*

Tras haber sentado las bases de la probabilidad condicionada y las operaciones entre sucesos, necesitamos un marco más amplio que nos permita describir cómo se distribuyen los posibles resultados de cualquier experimento aleatorio. Entra entonces el concepto de **distribución de probabilidad**, un mapa que asocia a cada valor numérico de interés su "peso" o "densidad" de ocurrencia. Con esta introducción, podemos adentrarnos en tres ideas fundamentales: la definición de variable aleatoria y la distinción entre masa y densidad de probabilidad, la función de distribución acumulada y los momentos estadísticos.

Variable Aleatoria, Masa vs. Densidad de Probabilidad

Imaginemos que, en lugar de anotar simplemente "hipertenso" o "no hipertenso", asignamos a cada paciente un número real que refleja su presión sistólica exacta. Esa regla de asignación, convertir cada resultado del experimento clínico en un valor numérico, es lo que denominamos *variable aleatoria*. Cuando sus posibles valores son finitos o contables, hablamos de

una *función de masa de probabilidad*, donde asignamos a cada valor x_i la probabilidad $P(X= x_i)$. Por ejemplo, si registramos el número de crisis asmáticas mensuales de un paciente, X solo puede tomar 0, 1, 2, … ataques. En cambio, cuando la variable adopta un abanico infinito de valores, como la concentración de glucosa en sangre, describimos su comportamiento mediante una *función de densidad* $f_X(x)$, de modo que la probabilidad de hallar X entre *a* y *b* se calcula como

$$P(a \leq X \leq b) \;=\; \int_a^b f_X(x)\,\mathrm{d}x$$

Así, entender si trabajamos con masa o con densidad es el primer paso para elegir las herramientas adecuadas.

Función de Distribución Acumulada (FDA o CDF)

Tanta multitud de valores reclama un resumen progresivo: la *función de distribución acumulada*, $F_X(x)$, que en cualquier punto xxx nos dice la probabilidad de que la variable aleatoria sea menor o igual a ese valor,

$$F_X(x) = P(X \leq x)$$

Visualicemos pilas de fichas apilándose sobre la línea del número real, acumulándose hasta el nivel *x*. En un estudio de glucemias en ayunas, trazar $F_X(x)$ permite ver de un vistazo cuántos pacientes están por debajo de 100 mg/dL, cuántos entre 100 y 125, y así sucesivamente. Además, F_X es la llave para generar muestras simuladas y para invertir la probabilidad, fundamento de las pruebas de hipótesis y de la generación de números aleatorios con cualquier distribución.

Momentos: Esperanza, Varianza, Sesgo y Curtosis

Para caracterizar aún más la forma de una distribución, recurrimos a sus *momentos*. El **primer momento**, o *esperanza* E[X], es el centro de gravedad de la distribución, el valor promedio que esperaríamos si repitiésemos el experimento infinitas veces. El **segundo momento central**, o *varianza*

$$\mathrm{Var}(X) = E\big[(X - E[X])^2\big]$$

mide la dispersión, cuán alejados tienden a estar los valores del promedio. Más allá, el **tercer momento** nos habla del *sesgo* (o *skewness*), indicando si la cola de la distribución se extiende más hacia la derecha o hacia la izquierda, mientras que el **cuarto momento**, la *curtosis*, valora la concentración de probabilidad en el centro frente a las colas (picos agudos o planicies).

Por ejemplo, si estudiamos la duración de estancia hospitalaria tras una cirugía, una distribución con cola larga a la derecha revela que, aunque la mayoría de pacientes sale en pocos días, un grupo pequeño permanece mucho tiempo. El sesgo positivo y una curtosis elevada alertan de que una media aritmética podría no resumir bien la realidad, recomendando quizás mediana y percentiles para describir el dato.

Con estos conceptos bien asentados, variable aleatoria, función de masa o densidad, distribución acumulada y momentos, contamos ya con el vocabulario y las herramientas para adentrarnos en la **distribución normal**, esa "campana" omnipresente en la estadística médica, que en la siguiente sección expondremos en detalle como modelo paradigmático de fenómenos influenciados por multitud de pequeñas causas independientes.

5.2.2. *Distribución Normal (Gaussiana)*

La **distribución Normal**, también llamada *Gaussiana*, es la protagonista indiscutible de la estadística médica, pues muchos fenómenos biológicos, desde la talla al nacer hasta la presión arterial, se comportan como si obedecieran a una "campana" simétrica impuesta por la suma de infinidad de causas pequeñas e independientes. Su función de densidad se define así:

$$f(x) = \frac{1}{\sqrt{2\pi\,\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\,\sigma^2}\right)$$

donde μ es la **media**, centro de la campana, y σ^2 la **varianza**, que regula su anchura.

Propiedades clave

- **Simetría**: el gráfico es idéntico a ambos lados de μ; en consecuencia, **media = mediana = moda**.

- **Regla 68–95–99**: aproximadamente el 68 % de los valores cae en [μ − σ, μ + σ], el 95 % en [μ − 1,96 σ, μ + 1,96 σ] y el 99 % en [μ − 2,58 σ, μ + 2,58 σ]. Estas franjas, llamadas *intervalos de referencia*, permiten definir rangos "normales" y detectar valores atípicos con fundamento teórico.
- **Transformación a la normal estándar**: al restar μ y dividir por σ, se obtiene la variable z, con distribución N (0,1). Con ello, podemos emplear tablas o funciones integradas en software para calcular probabilidades y cuantiles sin volver a integrar cada vez.

Ejemplo Clínico

Imaginemos un estudio de presión arterial sistólica en adultos sanos de una población urbana. Al medir a 500 voluntarios, hallamos μ=120 mmHg y σ=10 mmHg. La regla 68–95–99 nos dice que:

- Alrededor del 68 % de los sujetos tendrá cifras entre 110 y 130 mmHg,
- el 95 % entre 100 y 140 mmHg,
- el 99 % entre 94 y 146 mmHg.

De este modo, definir "hipertensión" como presión ≥ 140 mmHg equivale a señalar el **percentil 97,5**, un umbral sólidamente anclado en la teoría de la Campana de Gauss.

Aplicaciones

La Normal es la base de gran número de pruebas *paramétricas*, como el **test z** y los **intervalos de confianza** de la media, pues cuando los datos se ajustan razonablemente a este modelo, estos métodos son más potentes que sus homólogos no paramétricos. Además, la *normalidad* de los errores es un requisito en regresiones y en el ANOVA, garantizando validez y reproducibilidad en ensayos clínicos y estudios epidemiológicos.

Con la distribución Normal bien asentada, estamos listos para explorar cómo varía este modelo cuando la muestra es pequeña y la varianza poblacional desconocida, lo que nos lleva de inmediato a la **distribución t de Student**, tema que abordaremos a continuación.

5.2.3. *Distribución t de Student*

Cuando el tamaño de la muestra es pequeño o la varianza poblacional es desconocida, la curva normal pierde precisión en las colas y corre el riesgo de subestimar la frecuencia de valores extremos. En esas situaciones entra en escena la **distribución t**, descubierta por William Gosset bajo el seudónimo "Student". Su función de densidad se define como

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

donde ν son los **grados de libertad** (generalmente, el tamaño de muestra menos uno) y Γ es la función Gamma.

Propiedades clave

- **Colas más pesadas**: comparada con la normal, la t asigna mayor área a los extremos, lo que reduce el riesgo de confundir valores poco frecuentes con pura "suerte".
- **Simetría**: como la normal, es perfectamente simétrica en torno a cero, y para $\nu > 1$ la media es 0; a partir de $\nu > 2$ la varianza existe y vale $\nu/(\nu-2)$.
- **Convergencia**: al crecer ν, la distribución t se aproxima rápidamente a la normal estándar; para $\nu \geq 30$ suelen considerarse casi indistinguibles.

Ejemplo clínico

Supongamos un ensayo piloto de 12 pacientes sometidos a una nueva técnica de rehabilitación cardíaca, donde medimos la mejoría en la distancia caminada en seis minutos. La media muestral fue de 50 m con DS de 8 m. Para construir un intervalo de confianza del 95 % sobre el verdadero aumento medio, usamos la t con $\nu = 11$:

Así, con 95 % de confianza, el efecto real de la intervención se halla entre 44,9 m y 55,1 m, un rango que refleja la imprecisión propia de muestras pequeñas.

Aplicaciones

El test t para una muestra, el t para dos muestras independientes o pareadas, y los **intervalos de confianza** de medias se basan en esta distribución. Su mayor robustez frente a colas pesadas garantiza conclusiones más conservadoras cuando no disponemos del tamaño muestral ideal ni de la varianza poblacional.

En el siguiente apartado, exploraremos la **distribución χ^2**, fruto de sumar cuadrados de variables normales, esencial para pruebas de bondad de ajuste y ANOVA en estudios clínicos.

5.2.4. *Distribución χ^2 (Ji-cuadrado)*

La **distribución χ^2** surge de sumar los cuadrados de varias variables que siguen una distribución normal estándar. Imagina que cada valor normal representa un "movimiento" en un eje, y al elevarlo al cuadrado y acumularlo, obtenemos una medida acumulativa de dispersión que solo toma valores no negativos, creando una curva asimétrica con cola derecha.

Propiedades clave

- **Asimetría positiva**: la chicuadrado está sesgada hacia la derecha, pues los valores al cuadrado nunca bajan de cero y los resultados muy grandes permanecen posibles aunque raros.
- **Dependencia de grados de libertad (df)**: cuantos más componentes sumemos, más cercana a una forma simétrica se vuelve la curva. Con pocos grados de libertad, la cola es muy pesada; con muchos, se aproxima a la normal.
- **Media y varianza**: dependen del número de grados de libertad, reflejando cómo crece y se extiende la curva conforme añadimos más cuadrados.

Ejemplo clínico

Supongamos que queremos comprobar si la frecuencia de grupos sanguíneos en una población de 500 donantes se ajusta a la proporción esperada (O: 42 %, A: 45 %, B: 10 %, AB: 3 %). Registramos los conteos observados y los comparamos con los esperados mediante un **test de bondad de ajuste χ^2.**

Si obtenemos un valor alto de χ^2, significa que las diferencias entre lo observado y lo esperado son demasiado grandes como para atribuirlas al azar, sugiriendo la necesidad de investigar factores demográficos o de selección que alteren la proporción.

Aplicaciones

La distribución χ^2 es la base de:

- **Pruebas de independencia** en tablas de contingencia, para decidir si dos variables categóricas, por ejemplo "fumar" y "enfermedad pulmonar", están relacionadas.
- **Pruebas de bondad de ajuste**, que contrastan si un conjunto de datos sigue una distribución teórica dada.
- **Estimación de varianza poblacional**, pues el cuadrado de la desviación muestral normalizado sigue una chicuadrado, muy útil para construir intervalos de confianza de la varianza.

Con la χ^2 bien comprendida, podemos avanzar a la **distribución F**, que compara varianzas y sirve como cimiento de los ANOVA en estudios clínicos y ensayos controlados.

5.2.5. *Distribución* F

La **distribución F** nace del cociente de dos medidas de varianza obtenidas de muestras independientes, y refleja cómo varía la dispersión entre dos grupos. Visualmente, es una curva asimétrica hacia la derecha, pues al dividir dos valores siempre positivos el resultado sólo puede tomar valores no negativos, concentrándose cerca de 1 y extendiéndose hacia la derecha tanto más cuanto más dispar estén las varianzas comparadas.

Propiedades clave

- **Asimetría positiva**: la densidad cae rápidamente tras su pico inicial, pero deja una cola larga que captura razones de varianzas muy elevadas.
- **Dependencia de dos grados de libertad**: uno para el numerador (varianza del primer grupo) y otro para el denominador (varianza del segundo), lo que modifica la forma exacta de la curva según el tamaño de ambas muestras.

- **Media > 1** cuando los grados de libertad del denominador son pequeños, tendiendo a 1 conforme ambas muestras crecen.

Ejemplo clínico

Imaginemos un ensayo que compara la variabilidad de la **presión arterial sistólica** en dos unidades de cuidados intensivos—una con protocolo estándar y otra con un nuevo sistema de ventilación mecánica. Calculamos la varianza de cada grupo y formamos su cociente, obteniendo un valor $F = 2,8$. Al consultar la tabla F con los grados de libertad de ambas muestras, vemos que este cociente excede el cuantíl crítico para $\alpha = 0,05$, lo que sugiere que la dispersión de presiones no es la misma en ambos protocolos, quizás indicando un efecto del nuevo sistema en la estabilidad hemodinámica.

Aplicaciones

- **ANOVA**: el estadístico F evalúa si las medias de más de dos grupos difieren significativamente, comparando la variabilidad entre grupos con la variabilidad dentro de ellos.
- **Prueba de igualdad de varianzas**: tradicionalmente conocida como test F, comprueba si dos poblaciones comparten la misma varianza, paso previo a varios análisis.
- **Comparación de modelos** en regresión: el cociente de los residuos de dos ajustes diferentes sigue una F, permitiendo decidir si incluir variables adicionales mejora la explicación del fenómeno.
- Armados con la Normal, la t, la χ^2 y la F, cubrimos las grandes distribuciones continuas de la estadística paramétrica, y podemos ahora pasar a las distribuciones discreta, la **binomial** y la **Poisson**, que modelan conteos y eventos raros en investigación biomédica.

5.2.6. Distribución Binomial

La **distribución binomial** modela el número de "éxitos" en una secuencia de n ensayos independientes, cada uno con probabilidad fija de éxito p. Es como plantearse cuántos pacientes, de un grupo de n, responderán positivamente a un tratamiento que tiene una eficacia esperada del p.

Propiedades clave

- **Discreta y de dos salidas**: cada ensayo sólo puede concluir en "éxito" o "fracaso".
- **Media y dispersión**: el número medio de éxitos es np, y su variabilidad crece con np(1–p).
- **Asimetría según *p***: si p = 0,5 la distribución es casi simétrica; si p < 0,5 se sesga hacia la izquierda, y si p > 0,5 hacia la derecha.

Ejemplo clínico

Supongamos un estudio con 20 pacientes ingresados por neumonía, y que la literatura sugiere que la probabilidad de presentar una complicación grave es 0,15. La distribución binomial nos da la probabilidad de ver exactamente k complicaciones entre los 20 casos, por ejemplo, 2, 3 o más. Esto nos ayuda a planificar camas de cuidados intensivos y recursos de enfermería con anticipación.

Aplicaciones

- **Ensayos clínicos** donde contamos cuántos pacientes en un grupo alcanzan un umbral de respuesta.
- **Cálculo de tamaños muestrales** al diseñar estudios dicotómicos: cuántos sujetos necesitamos para esperar un cierto número de éxitos.
- **Pruebas de hipótesis** de proporciones, como el test de proporciones z, donde la binomial subyacente fundamenta el cálculo.

5.2.7. Distribución de Poisson

La **distribución de Poisson** describe el conteo de eventos que ocurren de forma aleatoria e independiente a una **tasa constante** por unidad de tiempo, espacio o población. Imagina cuántas urgencias llegan a un servicio por hora, asumiendo que cada caso es un evento raro y que la probabilidad de dos llegadas exactas al mismo instante es prácticamente nula.

Propiedades clave

- **Discreta y sin límite superior**: el número de eventos puede crecer sin tope, aunque valores muy altos sean cada vez más improbables.
- **Media = varianza**: el parámetro λ que define la tasa de ocurrencia también fija la dispersión, de modo que media y varianza coinciden en λ.
- **Aproximación de la binomial**: cuando los ensayos son muchos (n grande) y la probabilidad de éxito p pequeña, la binomial tiende a la Poisson con $\lambda = np$.

Ejemplo clínico

Si en promedio llegan 3 pacientes con fracturas de muñeca al servicio de urgencias cada día, la Poisson con $\lambda = 3$ permite calcular la probabilidad de que lleguen exactamente 0, 1, 2 o más de 5 fracturas un día dado. Estos cálculos son esenciales para asignar turnos de traumatología y planificar quirófanos.

Aplicaciones

- **Estudios de frecuencia** de eventos poco frecuentes: complicaciones postoperatorias, reingresos hospitalarios, reacciones adversas a vacunas.
- **Control de calidad en salud**: conteo de errores de medicación por cantidad de dosis administradas.
- **Modelos de regresión de Poisson** para relacionar tasas de eventos con covariables (edad, comorbilidades, tratamiento).

Con la binomial y la Poisson dominadas, obtenemos dos herramientas discretas que, junto con las continuas (Normal, t, χ^2, F), completan el abanico de distribuciones esenciales para modelar conteos, proporciones y mediciones clínicas. El siguiente paso nos llevará a explorar sus **relaciones y aplicaciones prácticas**, identificando cuándo y cómo recurrir a cada una en la investigación biomédica.

5.2.8. *Relaciones y Aplicaciones Prácticas*

Hasta aquí hemos explorado cada distribución como si fueran herramientas sueltas en un banco de trabajo; ahora veamos cómo encajan entre sí y cómo

escoger la adecuada para cada escenario clínico. Por ejemplo, la **binomial** y la **Poisson** son primas cercanas: cuando el número de ensayos es grande y la probabilidad de "éxito" pequeña, una binomial con parámetros (n,p) puede aproximarse elegantemente por una Poisson con $\lambda = np$. Esto resulta útil al estimar la incidencia de efectos adversos raros tras vacunar a miles de personas, pues sustituir muchos ensayos por un solo conteo Poisson simplifica los cálculos sin perder precisión.

Del mismo modo, las familias continuas se entrelazan: al sumar cuadrados de desviaciones normalizadas nacen la χ^2 y, al dividir dos χ^2 normalizadas y ajustar por grados de libertad, emerge la **F**, la piedra angular del ANOVA que compara grupos múltiples. Cuando el tamaño muestral es pequeño y desconocemos la varianza poblacional, la **t** cobra protagonismo, pero conforme crecen los grados de libertad, la t vuelve a fundirse con la **normal**, esa campana omnipresente que subyace al teorema del límite central.

En la práctica clínica, esta red de relaciones guía decisiones diarias: para evaluar si la media de un biomarcador difiere de un umbral, usarás test t o z según conozcas o no la DS poblacional; para comparar varianzas de dos lotes de reactivos antígeno, recurrirás al test F; ante un 2×2 de sensibilidad y especificidad, desplegarás la χ^2; y al modelar el número de recaídas por mes, elegirás la Poisson o su extensión binomial negativa. El objetivo es siempre casar el patrón de tus datos, continuo, discreto, rareza, tamaño muestral, con la curva estadística que mejor represente la incertidumbre inherente y maximice la potencia de tus pruebas.

5.3 Transformaciones y Aproximaciones

Hasta ahora hemos visto distribuciones ideales, pero los datos reales a menudo se desvían de los supuestos teóricos o implican conteos raros y muestras grandes. Para salvar esa brecha, disponemos de dos estrategias: **transformar** los datos para que encajen mejor en un modelo, o **aproximar** unas distribuciones por otras más manejables.

Cuando la variable presenta asimetrías marcadas o colas pesadas, por ejemplo, los niveles de PCR o de un marcador tumoral que crecen exponencialmente, una simple transformación logarítmica o de raíz cuadrada puede "enderezar" la distribución, acercándola a la normal. Esto permite luego aplicar intervalos de confianza y test paramétricos con mayor validez.

- *Ejemplo clínico*: si los valores de PCR en una cohorte están muy sesgados hacia la derecha, al graficar su histograma observamos una larga cola de pacientes con niveles altísimos. Al trazar log(PCR), la nube de puntos se compacta en forma de campana, y al calcular la media y la DS sobre la escala logarítmica obtenemos intervalos de confianza que, al volver a exponerlos en la escala original, describen la variabilidad con precisión.

Otras transformaciones como **Box–Cox** permiten optimizar automáticamente el exponente de la potencia para maximizar la normalidad de los residuos en regresiones, un recurso muy útil en modelización de biomarcadores o tiempos de supervivencia.

Aproximaciones entre Distribuciones

Al diseñar estudios, a veces trabajamos con cantidades tan grandes o tan pequeñas de eventos que resulta ventajoso sustituir una distribución por otra:

- **Binomial → Poisson**

Cuando el número de ensayos n es grande y la probabilidad de "éxito" p es pequeña ($np < 5$), la distribución binomial se aproxima muy bien a una Poisson con $\lambda = np$. Así, contar casos de un efecto adverso raro en miles de dosis administradas pasa de una fórmula combinatoria a un cómputo directo de Poisson, agilizando cálculos de probabilidad.

- **Binomial → Normal**

Si n es grande y p no está demasiado cerca de 0 ni de 1 (por ejemplo, $np > 5$ y $n(1-p) > 5$, la binomial puede aproximarse a la normal con media np y varianza $np(1-p)$. Esto facilita el uso de test z en comparaciones de proporciones sin recurrir a tablas binomiales.

- **Poisson → Normal**

Para λ suficientemente grande ($\lambda > 10$), la Poisson se comporta casi como una normal con media y varianza λ. Es útil al analizar tasas de incidencia diarias cuando los conteos promedio superan la decena.

- **t → Normal**

Cuando los grados de libertad crecen ($df > 30$), la t de Student converge a la normal estándar, de modo que en muestras moderadamente grandes podemos aplicar test z en lugar de t sin pérdidas apreciables de precisión.

¿Por qué y cuándo usar estas Aproximaciones?

1. **Simplicidad computacional**: trabajar con la normal evita cálculos de factoriales o sumas largas.
2. **Disponibilidad de software**: muchas funciones estadísticas nativas suponen normalidad o Poisson.
3. **Potencia estadística**: en ciertos rangos de parámetros, las aproximaciones «suavizan» la discreción y ofrecen mayor robustez.

Ejemplo integrado

En un ensayo de vacunación con 10 000 voluntarios, la incidencia histórica de una reacción adversa grave es p=0,001p=0{,}001p=0,001. Calcular exactamente la probabilidad de ver k = 5 casos mediante la binomial requerirían sumas combinatorias. En cambio, aproximamos a Poisson con $\lambda = 10$ y usamos

$$P(X = 5) \approx e^{-10}\,\frac{10^5}{5!}$$

o incluso, si preferimos z, con la normal $N(10,10)$ escribimos

$$P(4.5 < X < 5.5) \approx \Phi\Big(\frac{5.5 - 10}{\sqrt{10}}\Big) - \Phi\Big(\frac{4.5 - 10}{\sqrt{10}}\Big)$$

Así agilizamos la inferencia, manteniendo la precisión suficiente para la toma de decisiones regulatorias y de seguridad.

Con las **transformaciones** y **aproximaciones** en la manga, podemos enfrentar datos clínicos reales que rara vez encajan a la perfección en los modelos ideales, garantizando que nuestras pruebas de hipótesis e intervalos de confianza sigan descansando sobre fundamentos válidos, y abriendo el camino hacia análisis más avanzados como la regresión y los modelos de supervivencia.

Conclusiones

1. **Fundamentos de Probabilidad:** La teoría de la probabilidad (incluyendo espacio muestral, eventos, axiomas, probabilidad condicional y Teorema de Bayes) es esencial para cuantificar la incertidumbre inherente a los fenómenos biológicos y clínicos.

2. **Variables Aleatorias y Distribuciones:** Las variables aleatorias asignan valores numéricos a los resultados de experimentos, y sus distribuciones de probabilidad (de masa para discretas, de densidad para continuas) describen cómo se reparte la probabilidad entre esos valores.

3. **Distribuciones Clave:** Varias distribuciones teóricas son fundamentales en ciencias de la salud:

- **Normal:** Modelo central para muchas mediciones continuas; base de pruebas paramétricas.

- **t de Student:** Crucial para inferencias sobre medias con muestras pequeñas o varianza desconocida.

- **Chi-cuadrado (χ^2):** Utilizada en pruebas de independencia (tablas de contingencia) y bondad de ajuste.

- **F:** Base del ANOVA para comparar medias de múltiples grupos y para comparar varianzas.

- **Binomial:** Modela el número de éxitos en un número fijo de ensayos independientes.

- **Poisson:** Modela el número de eventos raros que ocurren en un intervalo fijo de tiempo o espacio.

4. **Interconexión y Aplicación:** Estas distribuciones están relacionadas (ej., t converge a Normal, Binomial puede aproximarse por Poisson o Normal bajo ciertas condiciones), y la elección de cuál usar depende del tipo de datos y la pregunta de investigación.

5. **Manejo de Datos Reales:** Las transformaciones (logarítmica, raíz cuadrada, Box-Cox) pueden ayudar a que los datos se ajusten mejor a los supuestos de las distribuciones (especialmente la normalidad).

6. **Aproximaciones Útiles:** Las aproximaciones entre distribuciones (ej., Normal para Binomial o Poisson con N o λ grandes) simplifican los cálculos y permiten usar herramientas estadísticas estándar en ciertos contextos.

CAPÍTULO 8

INFERENCIA Y ESTIMACIÓN: MUESTREO, INTERVALOS Y FORMULACIÓN DE HIPÓTESIS

Inference and Estimation: Sampling, Confidence Intervals, and Hypothesis Formulation

Resumen

Este capítulo introduce la estadística inferencial como el proceso de extraer conclusiones sobre una población a partir de una muestra. Explica la importancia del muestreo, detallando métodos probabilísticos (aleatorio simple, sistemático, estratificado, por conglomerados) que permiten generalizar resultados, y métodos no probabilísticos (conveniencia, cuotas, intencional, bola de nieve) útiles en contextos específicos pero con limitaciones de representatividad. Define la estimación de parámetros poblacionales mediante estimadores muestrales y la construcción de intervalos de confianza para cuantificar la precisión de dichas estimaciones. Introduce el marco de la formulación y prueba de hipótesis, estableciendo la H_0 y la Ha, y el uso del p-valor para decidir si rechazar H_0. Discute los errores tipo I (falsa alarma) y tipo II (pasar por alto un efecto real). Finalmente, aborda el cálculo del tamaño de la muestra necesario para lograr una precisión deseada en la estimación o un poder estadístico adecuado en la comparación de grupos, incluyendo ajustes por posibles pérdidas de seguimiento.

Palabras clave: Inferencia estadística, muestreo, intervalo de confianza, prueba de hipótesis, tamaño de muestra, error tipo I y II.

Abstract

This chapter introduces statistical inference as the process of drawing conclusions about a population from a sample. It explains the importance of sampling, detailing probabilistic methods (simple random, systematic, stratified, cluster) that allow generalization of results, and non-probabilistic methods (convenience, quota, purposive, snowball) useful in specific contexts but with representativeness limitations. It defines the estimation of population parameters using

sample estimators and the construction of confidence intervals to quantify the precision of these estimates. It introduces the framework for formulating and testing hypotheses, establishing the null (H_0) and alternative (Ha) hypotheses, and using the p-value to decide whether to reject H_0. It discusses Type I (false alarm) and Type II (missing a real effect) errors. Finally, it addresses the calculation of the required sample size to achieve desired precision in estimation or adequate statistical power in group comparison, including adjustments for potential dropouts.

Keywords: *Statistical inference, sampling, confidence interval, hypothesis testing, sample size, type I & II error.*

Introducción

Al adentrarnos en el mundo de la estadística inferencial, dejamos atrás la mera descripción de los datos para emprender un viaje de descubrimiento, en el que extraemos conclusiones sobre toda una población a partir de una muestra limitada. Así como un cardiólogo no ausculta todo el torrente sanguíneo, sino que valora pulsos y signos vitales para inferir el estado global, el investigador utiliza técnicas de muestreo, estimación y prueba de hipótesis para transformar unos pocos datos en certezas razonables.

En primer lugar, el muestreo actúa como brújula, define cómo seleccionar pacientes o mediciones de manera que reflejen fielmente la diversidad del universo clínico. Un muestreo aleatorio simple, por ejemplo, es como barajar un mazo de cartas antes de repartir, asegurando que cada individuo tenga la misma probabilidad de ser elegido, mientras que un muestreo estratificado, semejante a dividir el mazo por palos, permite equilibrar grupos de interés, como edades o comorbilidades, y garantiza que todos los subgrupos estén representados.

A partir de esa muestra, construimos estimadores, como la media, la proporción o la varianza muestral, que funcionan como periscopios para asomar al vasto océano poblacional. Sin embargo, toda mirada tiene un margen de error, y por ello acompañamos nuestras estimaciones con intervalos de confianza, que, cual redes de seguridad, miden la precisión del estimador. Por ejemplo, un intervalo al 95 % alrededor de la media de presión arterial sistólica en un ensayo farmacológico nos dice que, si repitiésemos el estudio

cien veces, en noventa y cinco de ellas la verdadera media poblacional recaería dentro de ese margen, brindando un mapa de incertidumbre manejable.

Cuando necesitamos decidir si un efecto observado es real o fruto del azar, recurrimos a la formulación de hipótesis. Establecemos una H_0, por ejemplo "el nuevo antihipertensivo no reduce la presión más que el placebo", y una hipótesis alternativa, contraria a la nula, que declara la presencia de un efecto. A través de una prueba estadística calculamos un pvalor, que expresa la probabilidad de obtener resultados al menos tan extremos como los observados si la $\mathbf{H_0}$ fuera cierta. Si ese pvalor cae por debajo de un umbral predefinido, $\alpha = 0{,}05$, rechazamos la hipótesis nula, convencidos de que el efecto es poco compatible con el azar.

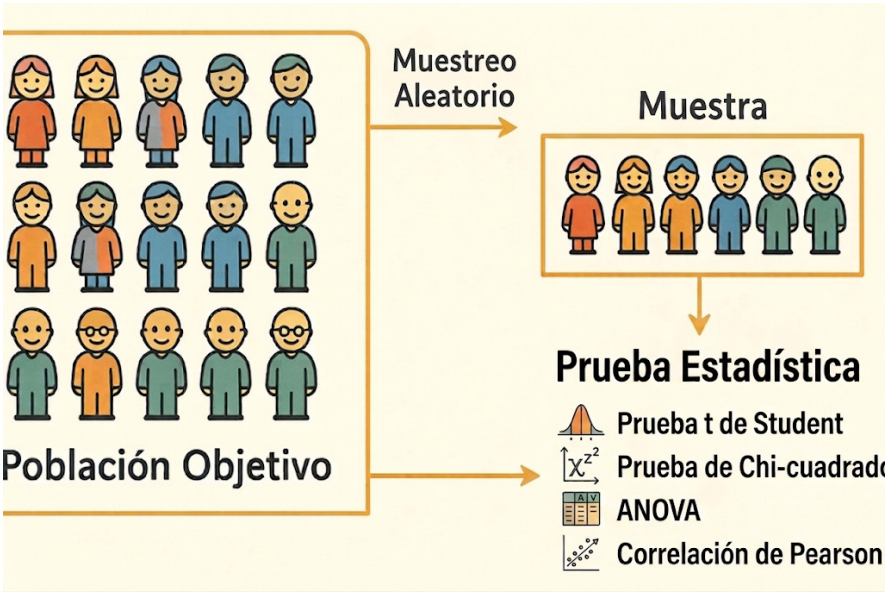
En todo este proceso conviven dos tipos de error, el error tipo I, equivalente a una falsa alarma, rechazar la nula cuando es verdadera, y el error tipo II, pasar por alto un efecto real. Diseñar el estudio implica equilibrar estos errores, ajustando el tamaño de la muestra o el nivel de significación, para no sacrificar la sensibilidad ni la especificidad de nuestras conclusiones.

Con este panorama general de la inferencia hemos visto cómo la estadística trasciende la simple suma de números y se convierte en un puente entre la observación limitada y las certezas médicas. A continuación, en la sección de **Distribuciones de probabilidad**, exploraremos los modelos que describen cómo se dispersan los datos teóricos, como la binomial para eventos dicotómicos o la normal para variables continuas, y que constituyen la base de la inferencia y la predicción en salud.

6.1. Muestreo en Estadística: de la Teoría a la Práctica Clínica

El muestreo es el arte de elegir una parte representativa de un universo mayor, de modo que, al estudiar unos pocos enfermos o mediciones, podamos inferir con solidez sobre toda la población. Si imaginamos un gran océano poblacional, el muestreo es como lanzar redes selectas, cada una con un diseño distinto, para capturar peces, datos que reflejen la diversidad del ecosistema; en epidemiología, esa "red" podría ser un protocolo de selección para pacientes con hipertensión, cáncer o infección respiratoria.

Figura 34. *Muestreo estadístico*. Elaborado con GeminiPRO.



Nota. Diagrama que representa la extracción de una muestra aleatoria de una población objetivo diversa (niños, mujeres adultas, hombres adultos y ancianos), ilustrada en la caja cerrada de la izquierda. La flecha "Muestreo Aleatorio" conduce a la muestra seleccionada, y la flecha "Prueba Estadística" muestra cómo los resultados obtenidos se proyectan de nuevo a la población para la inferencia.

6.1.1. *Tipos de Muestreo Probabilístico*

Los muestreos probabilísticos garantizan que cada individuo tiene una probabilidad conocida (y no nula) de ser seleccionado, lo cual sustenta la validez de los intervalos de confianza y las pruebas de hipótesis.

A. **Muestreo Aleatorio Simple**

- **Qué es**: cada miembro de la población aparece en bolitas numeradas y se extraen al azar, igual que un bombo de lotería.
- **Ejemplo**: en un hospital de 1000 pacientes con diabetes, seleccionamos 100 números al azar para estudiar el control glucémico.

- **Ventajas**: elimina sesgos sistemáticos, facilita el análisis estadístico.
- **Desventajas**: requiere lista completa de la población, no captura subgrupos raros.
- **Usos**: ensayos clínicos fase inicial, encuestas de satisfacción en toda la clínica.

B. **Muestreo Sistemático**

- **Qué es**: elegimos un punto de partida al azar y luego seleccionamos cada késimo individuo, como tomar cada décimo resultado en una máquina de laboratorio.
- **Ejemplo**: de una lista ordenada de ingresos hospitalarios, tomamos al paciente 5, luego al 15, al 25, y así sucesivamente.
- **Ventajas**: fácil de implementar, distribuye la muestra uniformemente.
- **Desventajas**: vulnerable a patrones periódicos en la lista (por ejemplo, horarios de ingreso).
- **Usos**: auditorías de historias clínicas, muestreo de fármacos dispensados.

C. **Muestreo Estratificado**

- **Qué es**: dividimos la población en estratos homogéneos, como rangos etarios o niveles de gravedad, y extraemos muestras aleatorias de cada grupo.
- **Ejemplo**: en un estudio multicéntrico de infarto agudo de miocardio, estratificamos por sexo y edad (menores de 50, 50–70, mayores de 70) y sacamos 50 pacientes de cada estrato.
- **Ventajas**: mejora la precisión en subgrupos, asegura representación balanceada.
- **Desventajas**: requiere conocimiento de las características de la población, más complejo en diseño.
- **Usos**: estudios de prevalencia por comorbilidad, ensayos en poblaciones vulnerables.

D. **Muestreo por Conglomerados**

- **Qué es**: agrupamos unidades en conglomerados naturales (p. ej. hospitales, consultorios) y seleccionamos algunos grupos al azar, evaluando luego a todos o a una submuestra de sus integrantes.
- **Ejemplo**: elegimos aleatoriamente 10 de 50 centros de atención primaria y evaluamos la adherencia al tratamiento antihipertensivo en todos sus pacientes.
- **Ventajas**: reduce costos y logística en estudios geográficamente dispersos.
- **Desventajas**: puede aumentar el error de muestreo debido a similitudes internas al conglomerado.
- **Usos**: encuestas nacionales de salud, estudios multicéntricos de eficacia de vacunas.

6.1.2. Tipos de muestreo no probabilístico

Estos métodos no garantizan probabilidades conocidas de selección, suelen usarse cuando la representatividad estricta no es esencial o cuando la población es difícil de listar.

A. **Muestreo de Conveniencia**

- **Qué es**: seleccionamos sujetos accesibles, como pacientes que acuden al laboratorio en una mañana.
- **Ejemplo**: analizar glucemia de quien pase ese día por la clínica.
- **Ventajas**: muy barato y rápido.
- **Desventajas**: alto riesgo de sesgo, no generalizable.
- **Usos**: estudios piloto, generación de hipótesis.

B. **Muestreo por Cuotas**

- **Qué es**: fijamos cuotas por características (sexo, raza) y llenamos cada cupo con sujetos de conveniencia.
- **Ejemplo**: 50 hombres y 50 mujeres voluntarios para encuesta de salud mental.

- **Ventajas**: asegura cierto equilibrio demográfico.
- **Desventajas**: dentro de cada cuota sigue habiendo sesgo de conveniencia.
- **Usos**: encuestas de opinión rápida, estudios de mercado de fármacos.

C. Muestreo Intencional o de Juicio

- **Qué es**: elegimos casos específicos por criterio del investigador, como pacientes con complicaciones raras.
- **Ejemplo**: seleccionar 30 pacientes con shock séptico para explorar biomarcadores.
- **Ventajas**: ideal para casos de interés clínico específico.
- **Desventajas**: imposible calcular precisión estadística, no representativo.
- **Usos**: series de casos, estudios cualitativos.

D. Muestreo en Bola de Nieve

- **Qué es**: pedimos a sujetos iniciales referir a otros, formando una cadena de inclusión.
- **Ejemplo**: investigar hábitos de consumo de antibióticos en pacientes con resistencia, donde cada participante refiere a conocidos con historial similar.
- **Ventajas**: útil en poblaciones ocultas o difíciles de alcanzar.
- **Desventajas**: refuerza sesgos de red, difícil estimar tamaño de población.
- **Usos**: estudios de enfermedades poco frecuentes, investigaciones en poblaciones marginales.

6.2. Cálculo del tamaño de la muestra

Al planear un estudio clínico, elegir cuántos pacientes incluir es tan decisivo como escoger el calibre de la red antes de pescar: si los agujeros son muy grandes, nos escaparán muchas capturas, precisión insuficiente; si la red es demasiado fina, requeriremos un esfuerzo desmesurado, recursos excesivos. A continuación, describimos cómo calcular el tamaño de muestra según dos enfoques complementarios, por *precisión* y por *poder*, y cómo ajustar esa cifra cuando anticipamos pérdidas de seguimiento.

6.2.1. Tamaño de muestra por precisión

Cuando nuestro objetivo es estimar un parámetro poblacional, por ejemplo, la media de presión arterial sistólica, con un margen de error aceptable, usamos la «regla de la malla» basada en la *precisión*.

A. **Estimación de una Media**

Fórmula

$$n = \left(\frac{Z_{1-\alpha/2}\,\sigma}{E} \right)^2$$

Donde, $Z1-\alpha/2$ es el valor crítico de la distribución normal para un nivel de confianza $1-\alpha$ (por ejemplo, 1,96 para 95 %), σ la DS poblacional (o de estudios previos), E el error máximo tolerable (margen de error), p. ej. ± 5 mmHg.

Analogía

Imagina el intervalo de confianza como el ancho de un pasillo por dónde camina la verdadera media: un *pasillo* muy estrecho (E pequeño) exige una muestra grande para estar seguro de no tropezar con los muros (no exceder el error).

Ejemplo

Si σ =12 mmHg, queremos 95 % de confianza y E = 5 mmHg,

$$n = \left(\frac{1{,}96 \times 12}{5}\right)^2 \approx 22{,}1$$

redondeamos a n = 23 pacientes.

B. **Estimación de una Proporción**

Fórmula

$$n = \frac{Z^2_{1-\alpha/2}\, p\,(1-p)}{E^2}$$

Donde p es la proporción esperada (p. ej. 0,2 de eventos adversos), E el margen de error en proporción (p. ej. ± 0,05 para ± 5 %).

Analogía

Si imaginamos que cada paciente es una canica pintada de rojo (evento) o azul (no evento), el margen de error define cuán apretado atamos la bolsa antes de sacudirla, un error menor exige más canicas para que la mezcla sea representativa.

Ejemplo

Para p = 0,295 % de confianza y E = 0,05

$$n = \frac{(1{,}96)^2 \times 0{,}2 \times 0{,}8}{0{,}05^2} \approx 246$$

necesitamos n = 246 pacientes.

6.2.2. Tamaño de Muestra por Poder

Cuando comparamos dos grupos, por ejemplo, un nuevo fármaco versus placebo, queremos asegurarnos de que la *red* detecte diferencias clínicamente importantes con alta probabilidad, es decir con suficiente *poder estadístico*.

A. **Comparación de Dos Medias**

Fórmula Aproximada

$$n = 2\,\frac{\sigma^2\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)^2}{\delta^2}$$

Donde δ es la diferencia mínima relevante (p. ej. 5 mmHg), $Z_{1-\beta}$ corresponde al poder 1–β (por ejemplo, 0,84 para 80 % de poder).

Analogía

El poder es como el haz de un faro en la niebla: con menos alcance (β) podemos no ver el barco enemigo (δ), pero al ampliar el haz (n mayor) aumentamos la probabilidad de detectarlo.

Ejemplo

Suponiendo σ = 12 mmHg, α = 0,05, poder 80 %, δ = 5 mmHg,

$$n = 2\,\frac{12^2\,(1,96+0,84)^2}{5^2} \approx 2\,\frac{144\,7,84}{25} \approx 2\,\frac{1129}{25} \approx 90,3$$

B. **Comparación de Dos Proporciones**

- **Fórmula aproximada**

$$n \;=\; \frac{\left(Z_{1-\alpha/2}\sqrt{2p(1-p)} \;+\; Z_{1-\beta}\sqrt{p_1(1-p_1)+p_2(1-p_2)}\right)^2}{(p_1-p_2)^2}$$

donde p = (p1+p2) / 2

Ejemplo

Para detectar un cambio de 0,10 a 0,20 en incidencia de un efecto adverso, con 80 % de poder y α=0,05, típicamente se requieren varias centenas de pacientes por grupo.

6.2.3. Ajuste por Pérdidas de Seguimiento

Toda red pierde algunos peces mientras la izamos, así que debemos aumentar la malla para compensar *dropouts* o abandonos.

$$n_{\text{ajustado}} = \frac{n_{\text{calculado}}}{1-L}$$

donde L es la proporción anticipada de pérdidas (p. ej. 0,10 para 10 %).

Ejemplo

Si calculamos n=200 y esperamos 15 % de ausencias,

$$n_{\text{ajustado}} = \tfrac{200}{1-0,15} \approx 235$$

deberíamos reclutar 235 pacientes para terminar con los 200 necesarios.

Con nuestra malla bien tensada, es decir con el tamaño de muestra ajustado para garantizar precisión y poder, podemos pasar al siguiente eslabón del

análisis: construir los **intervalos de confianza**, que funcionan como redes de seguridad al estimar parámetros poblacionales, y **formular la H_0 y la alternativa**, el andamiaje que sustenta las pruebas estadísticas y permite discernir si un hallazgo clínico es efecto real o simple azar.

Conclusiones

1. **Puente a la Generalización:** La estadística inferencial permite pasar de la descripción de una muestra a hacer afirmaciones probabilísticas sobre la población de la que proviene.

2. **Muestreo Fundamental:** La forma en que se selecciona la muestra (muestreo) es crítica para la validez de la inferencia. Los métodos probabilísticos son la base para generalizar resultados de manera formal.

3. **Estimación con Incertidumbre:** Los parámetros poblacionales se estiman usando datos muestrales (estimadores puntuales), pero siempre se debe acompañar de un intervalo de confianza que refleje la precisión y el margen de error de la estimación.

4. **Prueba de Hipótesis Formal:** La formulación de hipótesis (nula y alternativa) proporciona un método estructurado para tomar decisiones sobre efectos o diferencias observadas, usando el p-valor como medida de evidencia contra la H_0.

5. **Errores Inherentes:** En la toma de decisiones basada en hipótesis, existen riesgos de cometer errores: Error Tipo I (rechazar H_0 cuando es verdadera) y Error Tipo II (no rechazar H_0 cuando es falsa). El diseño del estudio busca controlar y equilibrar estos errores.

6. **Importancia del Tamaño Muestral:** Calcular adecuadamente el tamaño de la muestra antes de iniciar el estudio es crucial para asegurar que tenga suficiente precisión para estimar parámetros o suficiente poder estadístico para detectar diferencias clínicamente relevantes si existen, considerando también las posibles pérdidas de participantes.

CAPÍTULO 9

INFERENCIA ESTADÍSTICA EN INVESTIGACIÓN CLÍNICA. PRUEBAS DE HIPÓTESIS, ERRORES Y ESTIMACIÓN DE LA INCERTIDUMBRE

Statistical Inference in Clinical Research: Hypothesis Testing, Errors, and Estimation of Uncertainty

Resumen

Este capítulo profundiza en la inferencia estadística aplicada a la investigación clínica, detallando el proceso de prueba de hipótesis como un método formal para decidir entre una hipótesis nula (H_0, sin efecto) y una alternativa (H_a, con efecto). Explica los pasos clave: formulación de H_0 y H_a, elección del nivel de significancia (α), selección del estadístico de prueba adecuado, cálculo del p-valor (probabilidad de observar los datos o algo más extremo si H_0 es cierta) y la toma de decisión comparando p con α. Se discute la diferencia entre pruebas de una y dos colas. Se abordan los errores inherentes al proceso: Error Tipo I (α, rechazar H_0 verdadera, falsa alarma) y Error Tipo II (β, no rechazar H_0 falsa, omisión). Se introduce el concepto de poder estadístico (1-β) como la probabilidad de detectar un efecto real si existe, los factores que lo influencian (tamaño del efecto, tamaño muestral, variabilidad, α, diseño) y estrategias para aumentarlo. Se clarifica la interpretación correcta del p-valor, enfatizando que no mide la magnitud del efecto ni la probabilidad de que H_0 sea verdadera. Finalmente, se explica la construcción e interpretación de los intervalos de confianza como rangos de valores plausibles para el parámetro poblacional, que complementan al p-valor ofreciendo una medida de la precisión de la estimación.

Palabras Clave: Prueba de hipótesis, error tipo I, error tipo II, poder estadístico, valor p, intervalo de confianza.

Abstract

This chapter delves into statistical inference applied to clinical research, detailing the hypothesis testing process as a formal method to decide between a null hypothesis (H_0, no effect) and an alternative hypothesis (H_a, effect exists). It explains the key steps: formulating H_0 and H_a, choosing the significance level (α), selecting the appropriate test statistic, calculating the p-value (probability of observing the data or more extreme results if H_0 is true), and making a decision by comparing p with α. The difference between one-tailed and two-tailed tests is discussed. It addresses the inherent errors: Type I error (α, rejecting a true H_0, false alarm) and Type II error (β, failing to reject a false H_0, omission). The concept of statistical power (1-β) is introduced as the probability of detecting a real effect if it exists, the factors influencing it (effect size, sample size, variability, α, design), and strategies to increase it. The correct interpretation of the p-value is clarified, emphasizing it does not measure effect magnitude or the probability of H_0 being true. Finally, the construction and interpretation of confidence intervals are explained as ranges of plausible values for the population parameter, complementing the p-value by providing a measure of estimation precision.

Keywords: *Hypothesis testing, type I error, type II error, statistical power, p-value, confidence interval*

7.1. Prueba de Hipótesis

La prueba de hipótesis es la brújula que guía al investigador desde la anécdota muestral hasta la certeza poblacional, pues al observar una diferencia en nuestro estudio debemos discernir si se trata de un destello azaroso o de un verdadero efecto biológico. Imagínese un tribunal clínico en el que la H_0 asume la "inocencia", por ejemplo, "el nuevo fármaco no reduce la presión arterial más que el placebo", y la H_a encarna la "acusación", es decir, "el fármaco sí produce una reducción".

1. **Formular H_0 y H_a**, definiendo claramente el efecto a evaluar. H_0 postula ausencia de cambio ($\delta = 0$), mientras que H□ puede ser bidireccional ($\delta \neq 0$, prueba de dos colas) o unidireccional ($\delta > 0$ o $\delta < 0$, prueba de una cola) según la pregunta clínica. Por ejemplo, si sospechamos que un antibiótico acelera el aclaramiento bacteriano, la cola única concentra toda la evidencia en la dirección esperada, como apuntar con un foco directo hacia la diana.

2. **Elegir el nivel de significancia (α)**, ese umbral arbitrario pero convencional que fija cuántas falsas alarmas estamos dispuestos a tolerar. Un $\alpha = 0{,}05$ equivale a aceptar que en uno de cada veinte juicios observados podríamos condenar un fármaco inocuo, un balance entre prudencia y pragmatismo que debe justificarse en cada contexto.

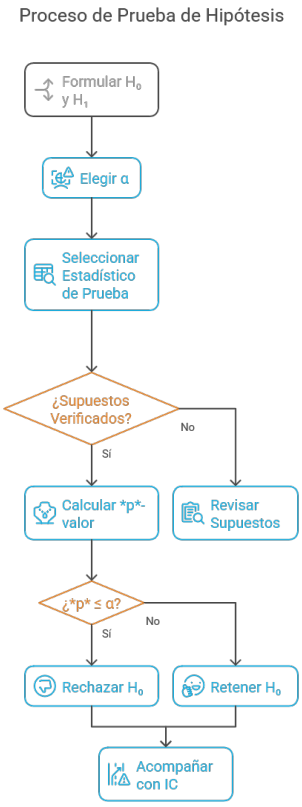
3. **Seleccionar el estadístico de prueba** adecuado, el «ajuste pericial» que resumirá nuestros datos en un solo número. Si comparamos medias de presión arterial, recurrimos a la t de Student; si contrastamos proporciones de eventos adversos, usamos la χ^2. Cada prueba trae consigo supuestos, normalidad, homogeneidad de varianzas, que conviene verificar con diagramas de caja o pruebas de Levene antes de avanzar.

4. **Calcular el *p*-valor**, esa cifra que indica la probabilidad de obtener un resultado tan extremo como el observado, o más, bajo la premisa de inocencia (H_0). Si el *p*-valor es pequeño, como una gota de tinta en un vaso de agua, la mancha sugiere que algo distinto a la pura casualidad colorea nuestros hallazgos.

5. **Tomar la decisión**: si $p \leq \alpha$, rechazamos H_0 y abrazamos la alternativa, convencidos de que el efecto es robusto; si $p > \alpha$, retenemos H_0 y reconocemos que la evidencia no basta para descartar la neutralidad, pues podría tratarse de un destello estadístico.

6. **Acompañar con un *intervalo de confianza***, ese corredor de seguridad que rodea nuestro estimador y muestra el rango donde se esconde la verdad poblacional. Un IC al 95 % para la diferencia de medias, por ejemplo [–8 mmHg, –2 mmHg], excluye el cero y refuerza un *p*-valor significativo, mientras que un IC que sí contiene el cero advierte sobre la fragilidad del efecto (ver figura 35).

Figura 35. *Flujograma del Proceso de Prueba de Hipótesis Estadística. Elaborado con napkin: https://www.napkin.ai/es/*



Nota. Este diagrama ilustra los pasos fundamentales en una prueba de hipótesis: formulación de la H_0 y Ha, elección y selección del estadístico de prueba, verificación de supuestos, cálculo del pvalor, comparación con el nivel de significancia α para decidir entre rechazar o retener H_0, y finalmente la presentación de los resultados acompañados de su intervalo de confianza (IC).

En este proceso habitan dos fantasmas: el **error tipo I**, la falsa alarma de condenar un fármaco inocuo, y el **error tipo II**, el riesgo de dejar escapar un tratamiento genuinamente eficaz. Al diseñar el estudio ponderamos ambos, ajustando el tamaño de muestra para no sacrificar ni la sensibilidad ni la especificidad de nuestras conclusiones. Con este andamiaje metodológico

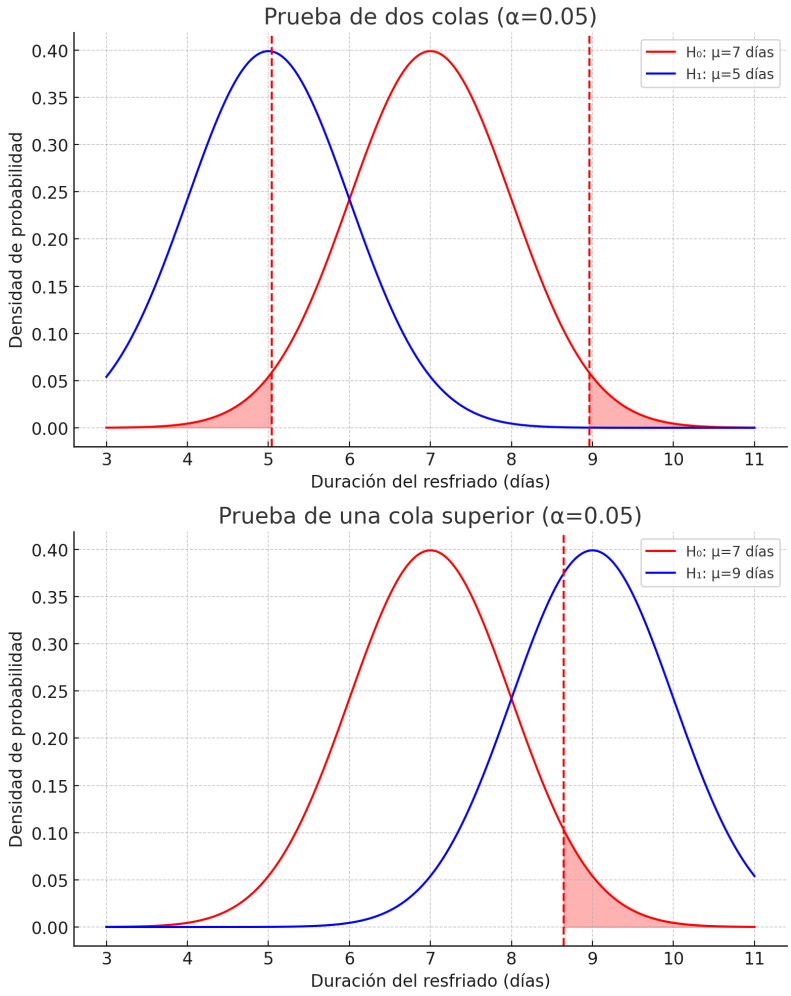
bien afianzado, estamos preparados para adentrarnos en las **distribuciones de probabilidad**, ese terreno donde la inferencia se nutre de modelos formales y donde hallaremos nuevas rutas para afinar nuestras decisiones clínicas.

Para contextualizar el concepto de una y dos colas para las pruebas de hipótesis, supongamos un ensayo clínico en el que comparamos la duración del resfriado común bajo dos condiciones, sin vitamina C, con una media esperada de 7 días, frente a un grupo tratado con vitamina C, y queremos decidir si la diferencia observada se debe al azar o a un verdadero efecto terapéutico.

Al decidir si un hallazgo clínico es fruto de la casualidad o de un verdadero efecto, la elección entre una prueba de una o de dos colas define cómo distribuimos nuestra *zona de rechazo* y, en consecuencia, dónde apuntamos el foco de evidencia. Las gráficas anteriores muestran en rojo la curva de la hipótesis nula (H_0: $\mu = 7$ días) y en azul la alternativa, con un color más saturado en las áreas críticas (ver figura 36).

- **Prueba de dos colas ($\alpha = 0{,}05$)**: dividimos el 5 % de probabilidad de rechazo en ambos extremos de la distribución nula, sombreando un 2,5 % en la cola izquierda ($x \leq \mu_0 - 1{,}96\ \sigma$) y otro 2,5 % en la cola derecha ($x \geq \mu_0 + 1{,}96\ \sigma$). Este enfoque es como disparar flechas a un blanco circular cuando ignoramos si el desplazamiento será a la derecha o a la izquierda. En el ejemplo, la alternativa propone $\mu_1 = 5$ días, y el área roja señalada en ambos extremos marca los valores tan improbables bajo H_0 que, de observarse, llevarían a rechazarla sin presuponer dirección.
- **Prueba de una cola superior ($\alpha = 0{,}05$)**: concentramos todo el 5 % de rechazo en la cola derecha, sombreando $x \geq \mu_0 + 1{,}645\ \sigma$. Así, apuntamos el haz de luz solo hacia un lado del blanco, aumentando el poder de detectar un efecto en esa dirección, pero aceptando que no podríamos detectar un cambio opuesto. Aquí la alternativa fija $\mu_1 = 9$ días, y sólo las observaciones extraordinariamente altas (rojo intenso) llevan a cuestionar H_0.

Figura 36. *Distribuciones bajo la* $\mathbf{H_0}$ *y regiones críticas en pruebas de una y dos colas* ($\alpha = 0{,}05$). Elaborado con GeminiPRO.



Nota. En el panel superior, la prueba de dos colas reparte el 5 % de probabilidad de rechazo en ambos extremos de la curva nula (H_0: $\mu = 7$ días), sombreando regiones por debajo de $\mu_0 - 1{,}96\ \sigma$ y por encima de $\mu_0 + 1{,}96\ \sigma$; la curva azul muestra la alternativa H con $\mu = 5$ días. En el panel inferior, la prueba de una cola concentra todo el 5 % en la cola derecha, sombreando $x \geq \mu_0 + 1{,}645\ \sigma$, y la curva alternativa ilustra H con $\mu = 9$ días; las líneas punteadas marcan los valores críticos de z.

Elegir entre una o dos colas implica sopesar la plausibilidad biológica del efecto contrario y el costo de falsas alarmas frente a omisiones. Con estas herramientas visuales y conceptuales, el investigador puede afinar su prueba para que el tono y la dirección de la evidencia coincidan con la pregunta clínica.

7.2. Errores tipo I y II

En el laboratorio de la estadística inferencial, los errores tipo I y tipo II son como las falsas alarmas y los silencios cómplices de un detector de incendios: si salta sin fuego (tipo I) perturba el ritmo normal, y si no avisa ante llamas genuinas (tipo II) expone al desastre. Entender ambos es crucial para afinar el equilibrio entre la vigilancia excesiva y la omisión peligrosa.

Error tipo I (α)

Es la "falsa alarma" de nuestro juicio clínico: rechazamos la $\mathbf{H_0}$ cuando en realidad no hay efecto. En términos de diagnóstico, equivale a diagnosticar enfermedad a un paciente sano, lo cual puede derivar en tratamientos innecesarios y ansiedad innecesaria.

Error tipo II (β)

Es el "silencio ante el peligro": retenemos la $\mathbf{H_0}$ cuando existe un efecto real. En medicina, sería como pasar por alto una patología presente, retrasando la intervención y empeorando el pronóstico. Su opuesto, el verdadero positivo, refleja el acierto de detectar un efecto clínico auténtico.

Para visualizar estas posibilidades, ayuda el clásico cuadro de contingencia entre la realidad poblacional y nuestra decisión:

Tabla 14. *Tabla de contingencia de las hipótesis*.

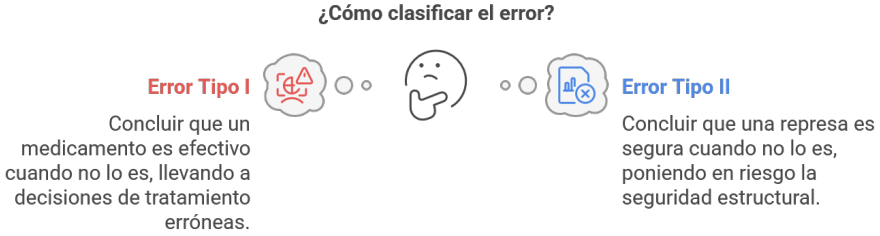
Realidad \ Decisión	**Rechazar H_0 (*efecto*)**	**Retener H_0 (*no efecto*)**
H_0 es verdadera	Error tipo I (α), falsa alarma	Acierto verdadero negativo

Realidad \ Decisión	Rechazar H_0 (*efecto*)	Retener H_0 (*no efecto*)
wH_a es verdadera	Acierto verdadero positivo	Error tipo II (β), omisión crítica

En esta tabla 14, cada celda representa un escenario posible: los "aciertos" son los casos en que la prueba coincide con la realidad, mientras que las "fallas" (α y β) marcan los dos tipos de desvío sistemático.

Controlar α y β es un acto de balance clínico: si reducimos α (menos falsas alarmas) elevamos β (más omisiones), y viceversa. Para decidir el nivel de significancia y el tamaño de muestra adecuados debemos tener clara la consecuencia de cada error en la práctica médica, pues calibrar correctamente nuestro detector estadístico marca la diferencia entre un diagnóstico seguro y un aviso apresurado o tardío. Con estas bases, el siguiente paso es aprender a diseñar estudios que maximicen el poder sin disparar alarmas infundadas (ver figura 37).

Figura 37. *¿Cómo clasificar el error? Elaborado con napkin: https://www.napkin.ai/es/*



- Imagina que un equipo de investigación desea evaluar la efectividad de un nuevo medicamento para bajar la presión arterial.
- La H_0 es que el medicamento no tiene efecto sobre la presión arterial en comparación con un placebo.

- **Error Tipo I (α):** Si el estudio concluye que el medicamento es efectivo para bajar la presión arterial cuando en realidad no lo es, el equipo ha cometido un error Tipo I.
- Por ejemplo, si se utiliza un nivel de significancia del 0.05, están dispuestos a aceptar un 5% de probabilidad de cometer este error.
- **Error Tipo II (β) y Poder:** Antes de iniciar el estudio, el equipo calcula que con un tamaño de muestra de 100 participantes, el estudio tendría un poder de 0.80 para detectar una diferencia significativa en la presión arterial si el medicamento realmente tiene un efecto.
- Esto significa que si el medicamento tiene un efecto real, hay un 80% de probabilidad de que el estudio detecte ese efecto, y por lo tanto, un 20% de probabilidad de cometer un error Tipo II.

7.3. Poder Estadístico

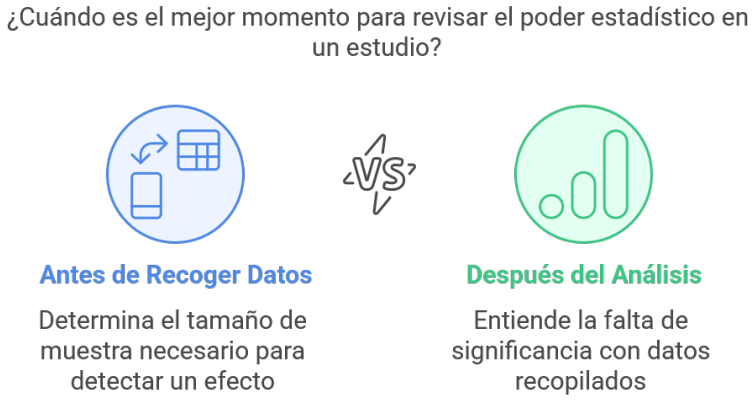
Al planear un estudio clínico, el cálculo del poder estadístico no es un mero trámite, sino la brújula que asegura que nuestra red de muestreo tenga el tamaño y la tensión adecuados para capturar el efecto que verdaderamente buscamos.

El **poder estadístico** (1 – β) es la probabilidad de evitar esa falla, de disparar la alarma cuando realmente hay fuego. Un 80 % de poder significa que en ocho de cada diez estudios similares reconoceríamos un efecto clínicamente relevante, siempre y cuando exista.

¿En qué momento evaluar el poder?

La Figura 38 muestra que la evaluación del poder estadístico puede realizarse en dos momentos complementarios del estudio: **antes de recoger datos**, cuando el investigador calcula el tamaño de muestra necesario en función del tamaño de efecto esperado, el nivel de significación deseado y la variabilidad de la medida, asegurando así que el diseño tenga la sensibilidad suficiente para detectar diferencias reales; y **después del análisis**, cuando, ante la ausencia de significación, se revisan los datos recopilados para determinar si la falta de hallazgos se debe a un verdadero efecto nulo o a una potencia insuficiente.

Figura 38. *Momentos Clave para Evaluar el Poder Estadístico en un Estudio. Elaborado con napkin: https://www.napkin.ai/es/*



Por ejemplo, imagina un profesor que va a impartir una clase magistral (la "investigación"). **Antes de la sesión**, calcula cuántos alumnos asistirán y prepara suficiente material (apuntes, ejemplos, ejercicios) para que todos participen activamente, esto equivale al cálculo de tamaño de muestra para asegurar poder. **Después de la clase**, si observa que muchos estudiantes no comprendieron el concepto central, no sabe si el problema fue la calidad de la enseñanza (un verdadero "efecto nulo") o si simplemente no hubo tiempo ni recursos suficientes (falta de "potencia"); entonces decide ajustar el ritmo, ampliar el material o repetir la sesión con más ejemplos. De igual modo, revisar el poder estadístico en ambas etapas permite planificar recursos y, a posteriori, interpretar con rigor los resultados y corregir el diseño en futuras investigaciones.

¿Cuándo Calcular el Poder?

Antes de poner en marcha el ensayo, realizamos un análisis *a priori* que fija el tamaño de muestra necesario para alcanzar, por ejemplo, un 80 % de poder. Es como afinar la lente de un microscopio antes de observar la muestra, garantizando que podamos distinguir cualquier célula rara. En cambio, el análisis *post hoc*, hacer cálculos de poder al terminar el estudio, apenas aporta valor, pues el poder dependerá del *p*-valor obtenido y no añadirá perspectiva más allá de lo ya revelado por los resultados.

Factores que afectan el poder

1. **Tamaño del efecto (δ)**. Un estudio diseñado para detectar una reducción de 10 mmHg en presión arterial tendrá mucha más probabilidad de éxito que si intentamos distinguir apenas 2 mmHg, pues un efecto mayor brilla como un faro en la niebla.

2. **Tamaño de muestra (n)**. Cada paciente adicional estrecha el intervalo de confianza, reduciendo el ruido, de modo que detectar incluso diferencias sutiles se vuelve factible, al igual que afinar el zoom de una cámara para captar un insecto diminuto.

3. **Variabilidad (σ)**. Menos dispersión en las mediciones, por ejemplo al estandarizar la posición y el aparataje para medir la presión arterial, es semejante a quitar el viento en el mar antes de lanzar un bote de rescate; sin el oleaje, navegar se hace menos incierto.

4. **Nivel de significancia (α)**. Elevar α de 0,05 a 0,10 concentra más área bajo la curva crítica, aumentando el poder, pero también incrementa el riesgo de falsas alarmas, como subir el volumen de la alarma de un detector de humo para no perder ni un crujido—tendremos más detecciones falsas.

5. **Diseño del estudio**. Métodos como medidas repetidas, ensayos pareados o estratificados actúan como estabilizadores de la nave, reduciendo la variabilidad intra-sujeto y mejorando la capacidad de detectar el efecto real (ver figura 39).

Figura 39. *Factores que afectan el poder estadístico. Elaborado con napkin: https://www.napkin.ai/es/*



Estrategias para Aumentar el Poder

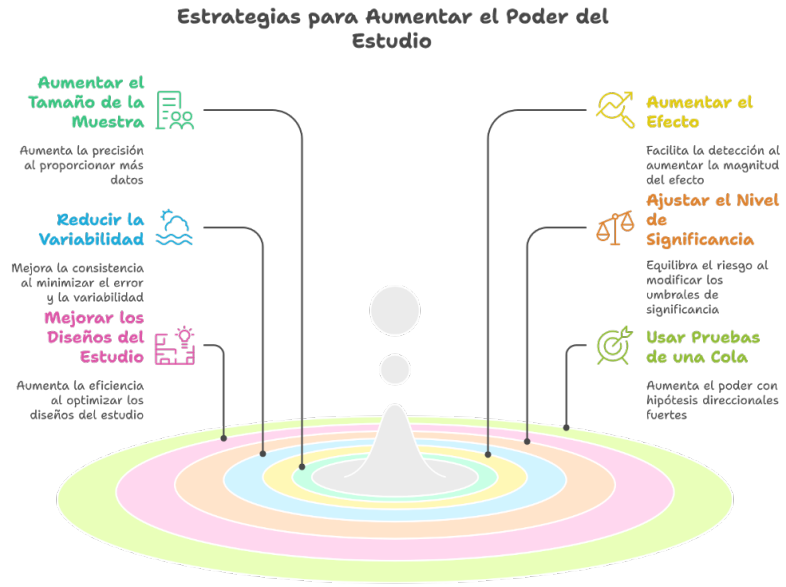
En la investigación clínica, garantizar un poder estadístico suficiente resulta tan vital como afinar la puntería antes de disparar, pues de ello depende que los efectos reales sobresalgan del ruido de fondo. Podemos identificar cinco estrategias clave para reforzar ese poder sin multiplicar indiscriminadamente recursos ni comprometer la validez del estudio (ver figura 40).

En primer lugar, aumentar el tamaño muestral, por ejemplo de 100 a 200 pacientes, equivale a tensar la red para pescar efectos pequeños, pues con más "anzuelos" surge una mayor probabilidad de capturar diferencias sutiles en la respuesta terapéutica, tal como muestra la figura que ilustra este principio. Asimismo, reducir la variabilidad estandarizando protocolos de medición, por ejemplo fijando la hora y la técnica al determinar la glucemia en sangre capilar, minimiza el ruido de fondo y permite que las señales clínicas brillen con mayor claridad.

Por otro lado, optimizar el diseño a través de ensayos cruzados o pareados, en los que cada participante actúa como su propio control, elimina gran parte de la dispersión intersujeto, de modo que las comparaciones internas, más allá de las diferencias individuales, revelan con nitidez el verdadero efecto del fármaco. En situaciones de urgencia clínica, aumentar el nivel de significación α, siempre que se acepte un mayor riesgo de error tipo I, puede dar un empujón al poder, priorizando la detección temprana de un beneficio potencial sobre la precaución estadística rígida. Finalmente, emplear pruebas dirigidas, de una cola, cuando la hipótesis tiene una dirección claramente definida, concentra la probabilidad de rechazo en el extremo relevante del *pvalue*, optimizando la sensibilidad del contraste.

Con estas palancas bien calibradas, podemos diseñar ensayos que no solo describan fenómenos, sino que los detecten con la fuerza y la precisión que demanda la práctica médica.

Figura 40. *Estrategias para aumentar el poder. Elaborado con napkin: https://www.napkin.ai/es/*



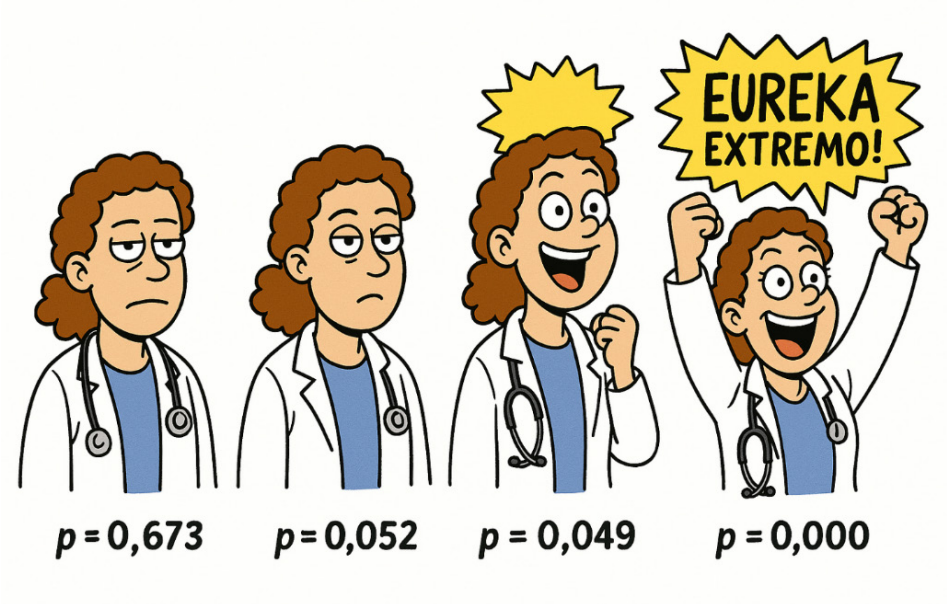
7.4. Valor P

El *p*valor es la brújula que mide cuán sorprendentes serían nuestros datos si la $\mathbf{H_0}$ fuese cierta, pues cuantifica la probabilidad de obtener, bajo la presunción de "sin efecto", un estadístico igual o más extremo que el observado. Imagine que somos buzos con un detector de metales en la playa: cuando no hay nada enterrado, el equipo debería callar; si emite un pitido muy fuerte, tan improbable como escuchar esa señal en dos de cada cien exploraciones sin metal $p = 0{,}02$, sospechamos que algo valioso yace bajo la arena. De igual modo, un p bajo indica que el hallazgo es poco compatible con la inocencia de la H_0.

En el juicio de un ensayo clínico, cada paciente es un testigo y el *p*valor mide cuán escandaloso sería su testimonio bajo la suposición de que el fármaco no funciona. Por ejemplo, si comparamos la caída de fiebre entre 50 tratados con un nuevo antipirético y 50 con placebo, y obtenemos $p = 0{,}03$, estamos diciendo que solo en tres de cada cien ensayos sin verdadero efecto veríamos una diferencia tan grande. Esa rareza nos invita a rechazar la hipótesis de "sin eficacia" y considerar que el medicamento tiene un impacto real.

Sin embargo, el pvalor no nos habla de la magnitud clínica del efecto ni de la probabilidad de que la **H_0** sea verdadera o falsa. Un $p = 0{,}049$ puede disparar el entusiasmo. como en la viñeta donde un investigador pasa de la indiferencia a exclamar "¡Eureka!", pero no nos dice si la diferencia es mayor de 2 °C de fiebre o de 0,2 °C (ver figura 41). Por ello, conviene acompañarlo siempre de un *intervalo de confianza*, que muestre el rango de valores plausibles para la verdadera diferencia, y de una medida del tamaño del efecto, para ponderar la relevancia práctica.

Figura 41. *Reacciones de la médica ante distintos umbrales de significancia estadística (valores de p).* Elaborado con GeminiPRO.



En la práctica, comparamos el pvalor con un nivel de significancia predefinido, α, habitualmente 0,05. Si $p < \alpha$, rechazamos la H_0; si $p \geq \alpha$, no disponemos de evidencia suficiente y la retenemos. Esta regla de decisión no es infalible, pues un pvalor bajo no garantiza un hallazgo clínicamente significativo, y un p alto no exculpa definitivamente la H_0: a veces la muestra es pequeña o el efecto sutil.

Además, existe confusión común al interpretar el pvalor. No es la probabilidad de que la nula sea cierta, ni el riesgo de cometer un error tipo I; simplemente mide la compatibilidad de los datos con H_0. Tampoco sustituye al

poder estadístico, que es la capacidad de nuestro diseño para detectar un efecto real. Por ello, al informar resultados, el investigador debe presentar el pvalor como una alarma estadística, pero fundamentar sus conclusiones en la magnitud del efecto, el contexto biológico y la solidez del diseño experimental.

7.5. Intervalo de Confianza

Al trazar un intervalo de confianza dibujamos no solo un punto estimado, sino todo un corredor de valores plausibles donde puede encontrarse el verdadero parámetro poblacional, de la misma manera en que un explorador marca en el mapa la franja de terreno donde cree que se halla el tesoro, en lugar de limitarse a un punto exacto. Ese corredor, habitualmente al 95 %, significa que si repitiésemos el estudio muchas veces bajo las mismas condiciones, en aproximadamente noventa y cinco ocasiones la verdadera media o proporción quedaría atrapada dentro del intervalo estimado.

Para construir un intervalo de confianza de la media, partimos de la estimación puntual $\bar{x}$, sumamos y restamos el producto de un valor crítico $Z_{1-\alpha/2}$ que para un 95 % vale 1,96 bajo la normal, por el error estándar. que refleja la dispersión dividida por la raíz del tamaño de muestra. Por ejemplo, si medimos la presión arterial sistólica en 100 pacientes y obtenemos $\bar{x} = 130$ mmHg con $\sigma = 15$ mmHg, el intervalo al 95 % será

$$130 \pm 1,96 \, \frac{15}{\sqrt{100}} = 130 \pm 2,94$$

es decir [127,06 mmHg, 132,94 mmHg], un pasillo lo suficientemente estrecho para guiar la interpretación clínica.

En el caso de una proporción, como la tasa de efectos adversos de un nuevo fármaco, empleamos la fórmula

$$\hat{p} \pm Z_{1-\alpha/2} \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$$

donde p^ es la proporción observada. Si detectamos un 20 % de pacientes con náuseas en 200 participantes, el IC al 95 % sería

$$0{,}20 \pm 1{,}96\sqrt{\frac{0{,}20\times 0{,}80}{200}} \approx 0{,}20 \pm 0{,}055$$

o sea [14,5 %, 25,5 %], un rango que permite al clínico calibrar la incertidumbre antes de conversar con el paciente.

La amplitud del intervalo depende de tres palancas: tamaño de muestra, variabilidad de los datos y nivel de confianza. Una muestra mayor o una menor dispersión afinan el pasillo, al igual que un haz de luz más intenso acota la zona oscura donde podría esconderse el objetivo. Por su parte, elevar el nivel de confianza de 95 % a 99 % ensancha el intervalo, pues exigimos más seguridad de capturar el verdadero valor, a cambio de una precisión menor.

El intervalo de confianza no solo informa sobre qué tan lejos puede estar la verdad de nuestra estimación, sino que funciona como espejo del *p*-valor: si al comparar dos tratamientos el IC de la diferencia de medias excluye el cero, coincide con un *p*-valor significativo en una prueba de dos colas. Con este instrumento en mano, el investigador transita con paso firme desde la muestra hasta la población, sabiendo que cada estimación viene acompañada de un margen de seguridad, indispensable para la toma de decisiones en salud.

Conclusiones

1. **Marco Formal de Decisión:** La prueba de hipótesis ofrece un procedimiento estructurado (H_0, H□, α, estadístico, p-valor) para evaluar si las diferencias o efectos observados en una muestra son estadísticamente significativos o probablemente debidos al azar.

2. **Tipos de Error Controlados:** El proceso reconoce y busca controlar dos tipos de error: el Error Tipo I (α), fijado por el investigador, y el Error Tipo II (β), relacionado inversamente con el poder estadístico.

3. **Poder Estadístico Esencial:** El poder (1-β) es la capacidad del estudio para detectar un efecto real y debe ser considerado *a priori* en el diseño,

calculando el tamaño muestral necesario en función del efecto esperado, α, y la variabilidad.

4. **Interpretación Cauta del p-valor:** El p-valor mide la compatibilidad de los datos con la H_0, pero no la magnitud del efecto ni la importancia clínica; valores pequeños sugieren evidencia contra H_0, pero requieren contexto.

5. **Intervalos de Confianza Informativos:** Los intervalos de confianza (IC) proporcionan un rango de valores plausibles para el parámetro poblacional (ej. media, diferencia de medias, proporción), indicando la precisión de la estimación y complementando la información del p-valor.

6. **Inferencia Completa:** Una inferencia estadística robusta en investigación clínica debe incluir no solo el resultado de la prueba de hipótesis (p-valor) sino también la estimación del efecto y su intervalo de confianza para una interpretación completa y clínicamente relevante.

CAPÍTULO 10

SELECCIÓN Y APLICACIÓN DE PRUEBAS ESTADÍSTICAS: DE LO PARAMÉTRICO A LO NO PARAMÉTRICO

Selection and Application of Statistical Tests: From Parametric to Nonparametric

Resumen

Este capítulo se centra en la selección y aplicación adecuada de pruebas estadísticas en investigación, distinguiendo entre pruebas paramétricas y no paramétricas. Las paramétricas (ej. t de Student, ANOVA) se usan para variables cuantitativas continuas que cumplen supuestos de normalidad y homogeneidad de varianzas, ofreciendo mayor potencia. Las no paramétricas (ej. U de Mann-Whitney, Kruskal-Wallis, Chi-cuadrado, Wilcoxon) se aplican a variables ordinales, nominales o cuando no se cumplen los supuestos paramétricos, siendo más robustas pero generalmente menos potentes. Se detalla la importancia de considerar el tipo de variable, la relación entre muestras (independientes vs. pareadas) y el número de grupos. Se presenta un procedimiento estructurado en cinco pasos para seleccionar la prueba correcta: 1) Identificar tipo de variable y relación muestral, 2) Verificar supuestos (normalidad, homogeneidad de varianza, frecuencias esperadas), 3) Escoger la prueba principal (usando guías como la Tabla 15), 4) Aplicar pruebas post-hoc si es necesario (ej. tras ANOVA), y 5) Interpretar el estadístico y p-valor en el contexto clínico. Se ejemplifican rutas específicas para datos cuantitativos y cualitativos.

Palabras Clave: Selección de pruebas estadísticas, pruebas paramétricas, pruebas no paramétricas, supuestos estadísticos, muestras independientes, muestras pareadas.

Abstract

This chapter focuses on the appropriate selection and application of statistical tests in research, distinguishing between parametric and non-parametric tests. Parametric tests (e.g., t-Student, ANOVA) are used for continuous quantitative variables that meet assumptions of normality and homogeneity of variances, offering greater power. Non-parametric tests (e.g., Mann-Whitney U, Kruskal-Wallis, Chi-squared, Wilcoxon) are applied to ordinal or nominal variables, or when parametric assumptions are not met for quantitative data, being more robust but generally less powerful. The importance of considering variable type, sample relationship (independent vs. paired), and the number of groups is detailed. A structured five-step procedure for selecting the correct test is presented: 1) Identify variable type and sample relationship, 2) Verify assumptions (normality, homogeneity of variance, expected frequencies), 3) Choose the main test (using guides like Table 14), 4) Apply post-hoc tests if necessary (e.g., after ANOVA), and 5) Interpret the statistic and p-value in the clinical context. Specific pathways for quantitative and qualitative data are exemplified.

Keywords: *Statistical test selection, parametric tests, non-parametric tests, statistical assumptions, independent samples, paired samples.*

8.1 Pruebas Paramétricas y no Paramétricas

En toda investigación, elegir la prueba estadística adecuada equivale a encontrar la llave precisa para abrir una cerradura, porque si falta compatibilidad entre el tipo de variable, la estructura muestral y los supuestos de la prueba, el análisis puede atascarse o arrojar resultados engañosos. Como indica la Tabla 15, distinguimos dos grandes familias. Las pruebas paramétricas, por ejemplo, el t de Student o la ANOVA, presuponen que las variables cuantitativas continuas siguen una distribución normal y mantienen varianzas homogéneas. Cuando la carretera de nuestros datos no está perfectamente nivelada, podemos intentar alisar el terreno mediante transformaciones logarítmica o raíz cuadrada, ajustando la forma de la distribución a los requisitos de la prueba. Si aun así persisten irregularidades, convocamos a las pruebas no paramétricas como Mann–Whitney, Kruskal–Wallis, Wilcoxon o Friedman, que aunque menos potentes sobre datos ideales, funcionan como un todoterreno capaz de avanzar sin problemas sobre terrenos abruptos.

Tabla 15. *Selección de pruebas estadísticas según tipo de variable, relación muestral y objetivos analíticos.*

Tipo de prueba	Tipo de variable	Relación muestral	Grupos	Ejemplos de pruebas	Medida de asociación	Predicción
Paramétricas	Cuantitativa continua	Independientes	2	t de Student (muestras independientes)	Coeficiente de Pearson	Regresión lineal
		Independientes	≥ 3	ANOVA de una vía		
		Pareadas	2	t de Student para muestras apareadas		
		Pareadas	≥ 3	ANOVA de medidas repetidas		
No paramétricas	Ordinal o discreta (cuantitativa)	Independientes	2	U de Mann–Whitney	Coeficiente de Spearman	Regresión logística
		Independientes	≥ 3	Kruskal–Wallis		
		Pareadas	2	Wilcoxon de rangos signados		
		Pareadas	≥ 3	Friedman		
	Nominal (dicotómica)	Independientes	2	Chicuadrado de independencia	Coeficiente Phi (o V de Cramer)	
				(o prueba exacta de Fisher si celdas < 5)		
		Pareadas	2	McNemar		
				Prueba Q de Cochran		
Tiempo a evento	Duración / supervivencia	Independientes	–	Curvas de Kaplan–Meier + prueba logrank	Hazard ratio	Modelo de Cox

Para comprender mejor la noción de *muestras independientes* frente a *muestras pareadas*, imaginemos dos escenarios sencillos. En un ensayo donde a un grupo de pacientes se le administra el fármaco A y a otro grupo distinto el fármaco B, las mediciones de cada persona no guardan relación directa con las del otro grupo; estas son *muestras independientes*. Por el contrario, si medimos la presión arterial de cada paciente antes y después de aplicar un mismo tratamiento, o comparamos el rendimiento del ojo derecho e izquierdo de una misma persona, cada par de observaciones está vinculado bajo condiciones similares y hablamos de *muestras pareadas*.

En el caso de variables cuantitativas continuas y muestras independientes, el t de Student para dos grupos o la ANOVA de una vía cuando son tres o más permiten medir la fuerza de asociación con el coeficiente de Pearson y proyectar valores futuros mediante regresión lineal. Para datos pareados recurrimos al t apareado o a ANOVA de medidas repetidas. Cuando la escala es ordinal o discreta, elegimos los contrastes de rangos Mann–Whitney o Kruskal–Wallis para muestras independientes y Wilcoxon o Friedman para pareadas, evaluamos correlaciones con el coeficiente de Spearman y utilizamos regresión logística para modelar resultados binarios.

Para variables nominales dicotómicas, el chicuadrado de independencia o la prueba exacta de Fisher si alguna celda tiene menos de cinco casos cubre las comparaciones en muestras independientes, mientras que McNemar o la Q de Cochran atienden los datos pareados. La fuerza de asociación se refleja en el coeficiente Phi o en V de Cramer y la regresión logística anticipa eventos dicotómicos. En estudios de tiempo a evento, las curvas de Kaplan–Meier junto con la prueba logrank permiten comparar distribuciones de supervivencia, el hazard ratio cuantifica la razón de riesgo instantáneo y el modelo de Cox incorpora covariables y censura para la predicción multivariante.

En condiciones óptimas, las pruebas paramétricas siempre serán la opción preferida gracias a su mayor potencia estadística y a la precisión de sus modelos predictivos, pero cuando el terreno de nuestros datos no admite nivelaciones razonables, las pruebas no paramétricas garantizan conclusiones confiables sin sacrificar la validez del análisis.

8.2 Procedimiento para la Selección de Pruebas Estadísticas

Partiendo de la distinción entre pruebas paramétricas y no paramétricas que exploramos en la sección anterior, conviene estructurar un protocolo claro que garantice la compatibilidad entre el tipo de variable, la relación muestral y los supuestos de cada contraste. Este procedimiento se articula en cinco etapas secuenciales. Primero, se identifica la naturaleza de la variable (cuantitativa continua u ordinal/nominal) y la relación entre las muestras independientes, pareadas o múltiples, pues en un ensayo con dos grupos distintos hablaremos de muestras independientes, mientras que si comparamos mediciones de un mismo individuo antes y después, o de dos ojos de un mismo sujeto, nos encontramos ante muestras pareadas. A continuación, se verifica que los datos cumplan los requisitos básicos: para las variables continuas se prueban la normalidad (Shapiro Wilk) y la homogeneidad de varianzas (Levene o Bartlett), y para las categóricas se comprueba que cada celda de la tabla de contingencia disponga de frecuencias esperadas suficientes. Una vez confirmados estos supuestos, el tercer paso es escoger el contraste principal según el esquema de la Tabla 15 "Pruebas estadísticas paramétricas y no paramétricas". En el bloque cuantitativo continuo con muestras independientes se recurre al t de Student para dos grupos o a la ANOVA de una vía cuando son tres o más, mientras que para datos apareados se emplea el t pareado o ANOVA de medidas repetidas. Si los supuestos no pueden restablecerse con transformaciones, como logarítmica o raíz cuadrada, se opta por pruebas no paramétricas de rangos: Mann–Whitney o Kruskal–Wallis para muestras independientes, Wilcoxon o Friedman en el caso pareado. Para variables nominales dicotómicas se eligen el chicuadrado de independencia o la prueba exacta de Fisher, y en muestras apareadas el test de McNemar o la Q de Cochran.

El cuarto paso consiste en aplicar pruebas post hoc cuando proceda, por ejemplo Tukey, Bonferroni, Scheffé o Dunnett tras un ANOVA significativo, de modo que podamos identificar con precisión qué comparaciones múltiples son responsables de las diferencias halladas sin inflar el error tipo I. Finalmente, en la quinta etapa se calculan el estadístico correspondiente y su valor p, se comparan estos resultados con el nivel de significación α establecido al inicio, habitualmente 0,05 o 0,01, y se contextualiza la magnitud y la dirección del efecto en términos clínicos o experimentales. Con este mapa

en la mano, el investigador dispone de una guía robusta, lista para adaptarse tanto a los contrastes de datos cuantitativos como a los análisis de variables cualitativas y de supervivencia.

Partiendo de esa distinción, podemos organizar el proceso de selección paso a paso:

1. **Identificar el tipo de variable y la relación entre muestras**, esto es, si los datos son cuantitativos continuos o cualitativos (ordinales o nominales), y si los grupos son independientes, pareados o múltiples.
2. **Verificar supuestos**: para cuantitativos, comprobar normalidad (*ShapiroWilk*) y varianzas homogéneas (Levene o Bartlett); para categóricos, asegurar frecuencias esperadas suficientes en cada celda de la tabla de contingencia.
3. **Escoger la prueba principal**:

- *Paramétricas independientes*: t de Student (dos grupos), ANOVA (tres o más).
- *Paramétricas pareadas*: t para muestras apareadas, ANOVA de medidas repetidas.
- *No paramétricas independientes*: U de MannWhitney, KruskalWallis.
- *No paramétricas pareadas*: Wilcoxon, Friedman.
- *Categóricas nominales*: chicuadrado o prueba exacta de Fisher; para datos apareados, McNemar o Q de Cochran.

4. **Aplicar pruebas post hoc** cuando corresponda, eligiendo el método según la complejidad de las comparaciones y el control del error global.
5. **Interpretar** el estadístico y el valor *p*, comparándolo con el nivel de significación α y contextualizando la magnitud y la dirección del efecto.

En este punto, conviene distinguir dos grandes rutas: una dirigida a datos cuantitativos continuos, donde se aprovecha la potencia de los contrastes paramétricos y sus transformaciones; y otra pensada para datos cualitativos, ordinales o nominales, que exige pruebas no paramétricas. A continuación, se presenta el procedimiento completo, primero en su versión general y luego segmentado según el tipo de dato, siempre ilustrado con ejemplos prácticos.

Procedimiento General

1. **Formulación del problema y definición de hipótesis**
   Identifica la pregunta de investigación y plantea H_0 (ausencia de efecto o asociación) frente a H□ (efecto o relación significativa).

2. **Selección del nivel de significación (α)**
   Decide el riesgo de cometer un error tipo I, comúnmente 0,05 o 0,01.

3. **Recolección y organización de datos**
   Asegúrate de que la muestra sea representativa y esté bien registrada, con variables claramente definidas.

Adaptación para Datos Cuantitativos

4a. Verificación de supuestos

- *Normalidad*: prueba ShapiroWilk.
- *Homogeneidad de varianzas*: Levene o Bartlett.
- *Transformaciones de datos*: si fallan los supuestos, aplica logarítmica, raíz cuadrada o elige pruebas no paramétricas.

5a. Elección de la prueba estadística

- *Paramétricas independientes*: t de Student (dos grupos), ANOVA (tres o más).
- *Paramétricas pareadas*: t para muestras apareadas, ANOVA de medidas repetidas.
- *No paramétricas independientes*: MannWhitney U, KruskalWallis.
- *No paramétricas pareadas*: Wilcoxon, Friedman.

6a. Realización de la prueba y cálculo del estadístico

Calcula el estadístico (t, F, etc.) y determina el valor p según su distribución.

7a. Pruebas post hoc (si aplica)

Tras un ANOVA significativo, emplea Tukey, Bonferroni, Scheffé o Dunnett para identificar grupos que difieran.

8a. Toma de decisiones e interpretación

Compara p con α, rechaza H_0 si $p \leq \alpha$, e interpreta la magnitud y dirección del efecto en el contexto clínico o experimental.

Ejemplo práctico:

- *Plantear*: H_0: las medias de glucemia antes y después de un tratamiento son iguales, H□: difieren.
- *Definir*: α = 0,05.
- *Recolectar*: medir glucemia en 20 pacientes antes y después.
- *Verificar*: ShapiroWilk y Levene; en caso de incumplimiento, aplicar logtransformación.
- *Realizar*: t de Student pareado.
- *Interpretar*: si $p \leq 0,05$, concluir cambio significativo en glucemia.

Adaptación para Datos Cualitativos (Categóricos)

4b. Verificación de condiciones

Asegura frecuencias esperadas suficientes en cada celda (al menos 5) para la prueba de chicuadrado.

5b. Elección de la prueba estadística

- *Chicuadrado*: asociación entre variables categóricas.
- *Prueba exacta de Fisher*: cuando las frecuencias son pequeñas o violan supuestos.
- *McNemar* o *Q de Cochran*: para datos apareados dicotómicos.

6b. Realización de la prueba y cálculo del estadístico

Calcula el estadístico basándote en las frecuencias observadas vs. esperadas, o usa el método exacto en Fisher.

7b. Toma de decisiones e interpretación

Compara p con α, rechaza H_0 si $p \leq \alpha$, y contextualiza la asociación en la investigación.

Ejemplo práctico:

- *Plantear*: H_0: no hay relación entre el hábito de ejercicio y la presencia de dolor lumbar, H□: sí existe.
- *Definir*: α = 0,05.
- *Recolectar*: encuesta a 120 participantes, codificar ejercicio (sí/no) y dolor (presente/ausente).
- *Verificar*: comprobar que cada celda tenga ≥ 5 casos; si no, usar Fisher.
- *Realizar*: chicuadrado.
- *Interpretar*: si $p \leq 0,05$, aceptar la asociación entre ejercicio y dolor lumbar.

Con este esquema, tanto el principiante como el investigador experimentado disponen de un mapa claro que guía la selección, aplicación e interpretación de pruebas estadísticas, listo para ampliarse hacia análisis de regresión múltiple o técnicas de supervivencia en etapas posteriores.

Conclusiones

1. **Selección Crítica:** Elegir la prueba estadística correcta es fundamental para la validez del análisis y depende del tipo de variable(s), la relación entre las muestras (independientes o pareadas) y el cumplimiento de los supuestos de la prueba.
2. **Paramétrico vs. No Paramétrico:** Las pruebas paramétricas son más potentes si se cumplen sus supuestos (normalidad, homogeneidad de varianzas) para datos cuantitativos continuos. Las no paramétricas son la alternativa robusta para datos ordinales/nominales o cuando los supuestos paramétricos no se cumplen.
3. **Supuestos Clave:** Verificar los supuestos (como normalidad con Shapiro-Wilk, homogeneidad de varianzas con Levene, o frecuencias esperadas mínimas para Chi-cuadrado) es un paso esencial antes de elegir

y aplicar una prueba. Las transformaciones de datos pueden ayudar a cumplir supuestos paramétricos.

4. **Procedimiento Estructurado:** Se debe seguir un proceso lógico para la selección: identificar variables y relación muestral, verificar supuestos, elegir la prueba adecuada (según guías como la Tabla 15), aplicar pruebas post-hoc si es necesario, y finalmente interpretar los resultados (p-valor y magnitud del efecto).

5. **Guía de Pruebas:** Existen pruebas específicas para diferentes escenarios:

- **Comparar 2 grupos independientes:** t-Student (paramétrica), U de Mann-Whitney (no paramétrica).
- **Comparar 2 grupos pareados:** t-Student pareado (paramétrica), Wilcoxon (no paramétrica).
- **Comparar ≥3 grupos independientes:** ANOVA (paramétrica), Kruskal-Wallis (no paramétrica).
- **Comparar ≥3 grupos pareados:** ANOVA medidas repetidas (paramétrica), Friedman (no paramétrica).
- **Asociación categórica:** Chi-cuadrado/Fisher (independientes), McNemar/Cochran (pareadas).
- **Correlación:** Pearson (paramétrica), Spearman (no paramétrica).
- **Supervivencia:** Kaplan-Meier, Log-rank, Cox.

6. **Interpretación Contextualizada:** El resultado de una prueba (p-valor) debe interpretarse junto con la magnitud del efecto y el contexto clínico o de investigación, no de forma aislada.

CAPÍTULO 11

ASOCIACIÓN Y PREDICCIÓN: CORRELACIONES Y REGRESIONES

Association and Prediction: Correlations and Regressions

Resumen

Este capítulo explora la asociación y predicción en investigación, diferenciando asociación estadística de causalidad y discutiendo los desafíos para establecer vínculos causa-efecto, incluyendo variables confusoras. Presenta los modelos teóricos de causalidad (necesaria/suficiente) y los criterios de Bradford Hill como guía para inferir causalidad. Se describen las medidas de correlación (Pearson para relaciones lineales y datos normales; Spearman para relaciones monotónicas, datos ordinales o no normales) para cuantificar la fuerza y dirección de la asociación entre dos variables cuantitativas. Introduce los modelos de regresión como herramientas predictivas: Regresión Lineal para predecir resultados continuos asumiendo linealidad y normalidad de errores; Regresión Logística para modelar la probabilidad de resultados binarios; y Regresión de Supervivencia (modelo de Cox) para analizar datos de tiempo hasta un evento, manejando censura y estimando riesgos proporcionales. Se enfatiza la importancia de verificar los supuestos de cada modelo y utilizar métricas adecuadas (R^2, AUC, HR, diagnósticos) para evaluar su ajuste y desempeño.

Palabras clave: Asociación, causalidad, correlación, regresión lineal, regresión logística, análisis de supervivencia.

Abstract

This chapter explores association and prediction in research, differentiating statistical association from causation and discussing the challenges in establishing cause-effect links, including confounding variables. It presents theoretical models of causality (necessary/sufficient) and the Bradford Hill criteria as guidelines for inferring causality. Correlation measures (Pearson for linear

relationships and normal data; Spearman for monotonic relationships, ordinal or non-normal data) are described to quantify the strength and direction of association between two quantitative variables. Regression models are introduced as predictive tools: Linear Regression to predict continuous outcomes assuming linearity and error normality; Logistic Regression to model the probability of binary outcomes; and Survival Regression (Cox model) to analyze time-to-event data, handling censoring and estimating proportional hazards. The importance of verifying each model's assumptions and using appropriate metrics (R^2, AUC, HR, diagnostics) to evaluate fit and performance is emphasized.

Keywords: *Association, causality, correlation, linear regression, logistic regression, survival analysis.*

Introducción

En el corazón de toda investigación científica late la búsqueda de **asociaciones**, es decir, la identificación de vínculos reproducibles entre dos o más variables para entender cómo un cambio en una influye en otra; de la misma forma que un naturalista observa cómo la llegada de la lluvia modifica el comportamiento de las aves, en medicina rastreamos, por ejemplo, la relación entre la exposición al tabaco y la incidencia de cáncer pulmonar para orientar estrategias de prevención. No obstante, la mera existencia de una asociación no implica **causalidad**: a veces la variable X y la variable Y se mueven al unísono porque un tercer factor las impulsa a ambas, el calor estival dispara tanto la venta de helados como el número de ahogamientos, o incluso puede ocurrir que la dirección del efecto se invierta, de modo que lo que parece causa sea en realidad consecuencia. Por ello, ante un coeficiente elevado, conviene preguntarse si no existe una **variable confusora** oculta que sesgue la interpretación, como cuando la edad y la fragilidad se asocian con la frecuencia de caídas en adultos mayores, pero ambas dependen a su vez de la comorbilidad crónica.

Dentro de este entramado surge la **correlación**, una medida cuantitativa que nos dice hasta qué punto dos variables, como las horas de sueño y el rendimiento en un test cognitivo, varían juntas de forma lineal o monotónica; su coeficiente, que oscila entre –1 y +1, actúa como un termómetro de la intensidad y la dirección del vínculo: un valor cercano a +1 indica que, a más

horas de descanso, tiende a aumentar la puntuación, mientras que un –1 sugeriría exactamente lo contrario.

Para ir más allá de la mera descripción de relaciones y aventurarnos en la predicción de resultados futuros, empleamos la **regresión**, un conjunto de técnicas que ajusta una "superficie" matemática capaz de estimar el valor de una variable dependiente, por ejemplo, la presión arterial, a partir de una o varias independientes, edad, índice de masa corporal o nivel de actividad física. En una **regresión lineal simple**, trazamos la línea que mejor pasa entre los puntos de un diagrama de dispersión, y con ello podemos anticipar cuántos milímetros de mercurio disminuirá la presión por cada kilo que pierda un paciente. Cuando la variable a predecir es dicotómica, como la aparición o no de una complicación quirúrgica, la **regresión logística** transforma las probabilidades en logaritmos, permitiendo calcular la probabilidad de riesgo individual. Así, mientras la correlación es la brújula que apunta hacia la existencia y fuerza de un lazo, la regresión es la carta de navegación que nos guía con precisión hacia la predicción de nuevos escenarios, siempre con la cautela de no confundir itinerario con causa primera.

9.1. Causalidad Versus Casualidad

En Epidemiología, uno de los objetivos más nobles y desafiantes es desentrañar la etiología de las enfermedades, es decir, comprender no solo si un factor de exposición está asociado con un desenlace adverso, sino si ese vínculo es verdadero o meramente fruto del azar, de un sesgo en el diseño del estudio o de una variable confundidora escondida entre bambalinas. Imaginemos que un investigador encuentra una fuerte asociación entre el consumo de café y la reducción de migrañas: antes de proclamar causalidad, debe descartar que esta relación no sea un espejismo estadístico, provocado por el estilo de vida de quienes beben café, o un "efecto rebote" debido a la edad o a la frecuencia de estrés laboral.

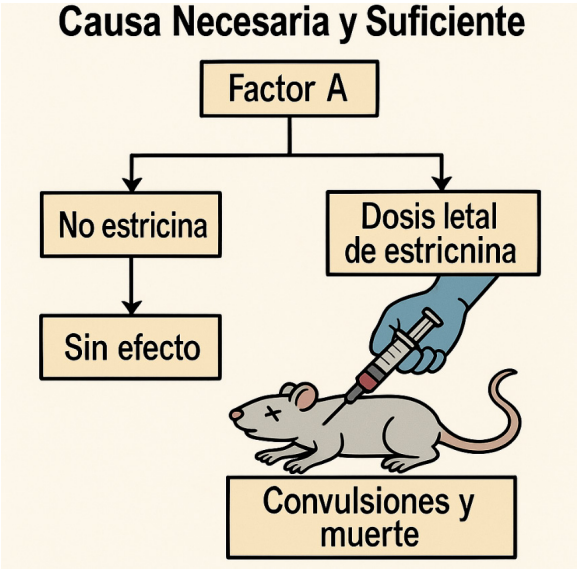
Una vez que hemos puesto a prueba el diseño y el análisis, y hemos controlado sesgos y confusores, por ejemplo, ajustando por edad, sexo o hábito de sueño, conviene preguntarnos si el factor de riesgo actúa como verdadera causa de la enfermedad. En este punto surge la distinción entre **asociación** y **causalidad**: la primera señala que dos variables varían de manera concer-

tada, mientras que la segunda implica un vínculo de causa y efecto, donde la exposición precede al desenlace y lo produce, aunque sea tras una serie de pasos intermedios.

Para describir estos matices, la teoría propone cuatro modelos de relación causal:

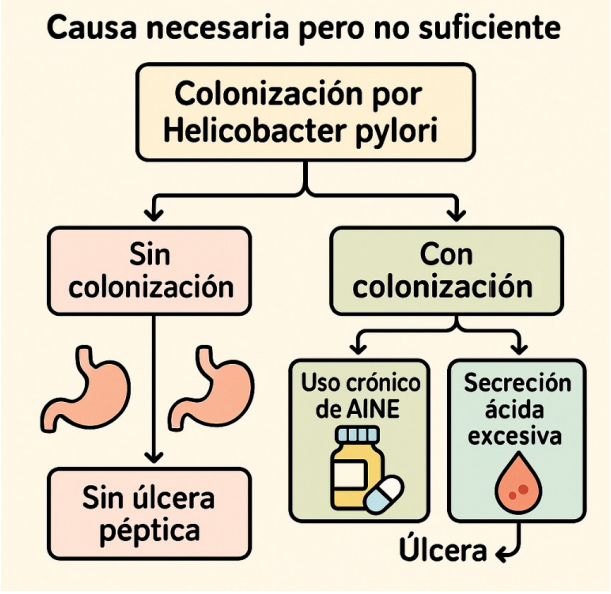
1. **Causa necesaria y suficiente.** Aquí, la enfermedad nunca se produce sin la presencia del factor, y siempre surge cuando éste está presente. Un ejemplo puramente teórico sería la inyección intracardíaca de una dosis letal de estricnina en un animal de laboratorio, que conduce invariablemente a convulsiones y muerte; sin esa dosis, el desenlace no se produce, y al administrarla, siempre aparece el efecto. En la práctica clínica humana resultan extremadamente raros estos esquemas tan rígidos, pues incluso en intoxicaciones agudas existen factores de variabilidad individual, metabolismo, estados de hidratación, que matizan el resultado (ver figura 42).

Figura 42. *Causa necesaria y suficiente*. Elaborado con GeminiPRO.



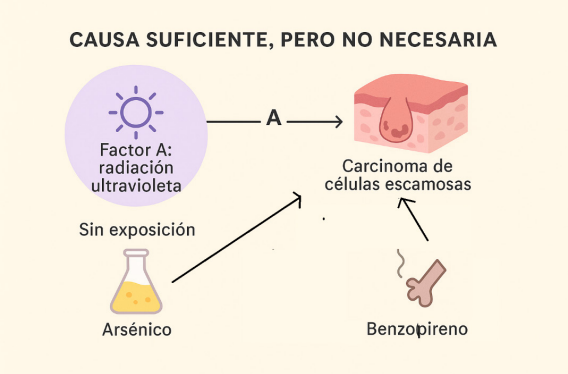
2. **Causa necesaria, pero no suficiente.** En este escenario, la presencia del factor A es imprescindible para que la enfermedad se desencadene, pero requiere de "colegas" en la cadena causal. Por ejemplo, la colonización gástrica por *Helicobacter pylori* es un requisito para la aparición de úlcera péptica en muchos pacientes, pero sin factores adicionales, uso crónico de AINE o secreción ácida excesiva, la simple presencia de la bacteria no conduce al daño mucoso grave (ver figura 43).

Figura 43. *Causa necesaria pero no suficiente*. Elaborado con GeminiPRO.



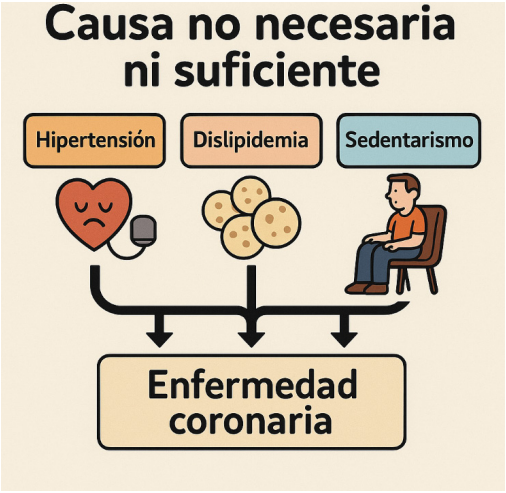
3. **Causa suficiente, pero no necesaria.** Aquí, A por sí solo puede producir la enfermedad, pero no es la única vía posible. La exposición prolongada a radiación ultravioleta de alta energía, por ejemplo, puede producir carcinoma de células escamosas en la piel sin necesidad de otro cofactor; sin embargo, compuestos químicos como el arsénico o el benzopireno también originan el mismo tipo de tumor, lo que convierte a A en una vía válida pero no exclusiva (ver figura 44).

Figura 44. *Causa suficiente, pero no necesaria*. Elaborado con GeminiPRO.



4. **Causa no necesaria ni suficiente.** Este modelo refleja la complejidad propia de las enfermedades multicausales modernas: ningún factor por sí solo explica el desenlace, y existen diversas combinaciones de A, B, C o más que pueden converger en la misma patología. Un claro ejemplo es la enfermedad coronaria, que suele desarrollarse cuando coexisten hipertensión arterial, dislipidemia y sedentarismo, pero existen individuos con solo uno o dos de estos factores que no padecen isquemia, y pacientes sin ellos que igualmente presentan la enfermedad, lo que ilustra la falta de suficiencia y necesidad de cada factor por separado (ver figura 45).

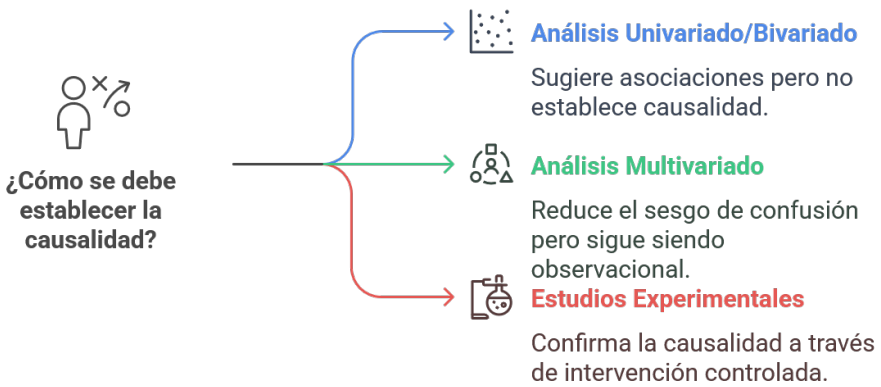
Figura 45. *Causa no necesaria ni suficiente*. Elaborado con GeminiPRO.



Este marco de relaciones sirve de brújula para diseñar intervenciones preventivas: si un factor es necesario, su erradicación abolirá la enfermedad; si resulta suficiente, bastará con minimizar la exposición; y en escenarios de multicausalidad, será preciso atacar varios eslabones simultáneamente.

La inferencia de causalidad se articula en tres niveles de rigor creciente. En primer lugar, el análisis univariado o bivariado explora de manera aislada la distribución de cada variable y las asociaciones simples entre pares, ofreciendo pistas iniciales sin descartar confundidores ni establecer un vínculo causaefecto definitivo. A continuación, el análisis multivariado ajusta simultáneamente por múltiples covariables, afinando esas asociaciones y reduciendo el sesgo de confusión, aunque sigue siendo eminentemente observacional. Finalmente, los estudios experimentales, como los ECA, manipulan directamente el factor de exposición y miden su impacto en el desenlace, cimentando la relación causal con el mayor rigor (ver figura 46).

Figura 46. *Estrategias para establecer la causalidad. Elaborado con napkin: https://www.napkin.ai/es/*



En el plano operativo, el análisis univariado se centra en caracterizar una sola variable, su media, dispersión y forma, mientras que el bivariado examina la relación entre dos medidas, contribuyendo a generar hipótesis sólidas. Al integrar tres o más factores, el análisis multivariado controla los efectos simultáneos de todos ellos, identificando las contribuciones independientes de cada predictor. Por último, el enfoque experimental, con asignación alea-

toria o controlada de la intervención, supera las limitaciones de los métodos observacionales y confirma la causalidad al observar directamente los cambios en el resultado tras modificar la exposición (ver figura 47).

Figura 47. *Métodos de análisis de datos según complejidad y objetivo. Elaborado con napkin: https://www.napkin.ai/es/*



9.2. Criterios de Bradford Hill

La tarea de distinguir causa y efecto en epidemiología es tan antigua como compleja, y los *criterios de Bradford Hill* ofrecen un faro para navegar ese mar de incertidumbres. Propuestos en 1965 por el epidemiólogo Sir Austin Bradford Hill, estos nueve principios no pretenden convertirse en un listado inflexible, sino en un conjunto de lentes que filtran la evidencia desde distintos ángulos. Así, antes de adentrarnos en cada criterio, vale entender que la fuerza de esta herramienta radica en su flexibilidad: no todos los elementos deben cumplirse a cabalidad para establecer causalidad, pero entre más criterios estén presentes, más sólida se vuelve la inferencia.

1. Fuerza de la Asociación

Un vínculo estadístico robusto, una razón de riesgos que se eleva por encima de 5 o 10, es menos probable que surja por casualidad o sesgo. Por ejemplo, la fuerte correlación entre la exposición a *Mycobacterium tuberculosis* en personal de laboratorios de microbiología y la incidencia de tuberculosis demuestra una asociación tan marcada que resulta difícil achacarla a factores de confusión. Sin embargo, una asociación pequeña no invalida la causalidad; simplemente requiere un escrutinio mayor.

2. Consistencia

Cuando distintos estudios, en lugares y momentos diversos, obtienen resultados similares, la hipótesis gana credibilidad. De este modo, la asociación entre la fumigación con insecticidas organofosforados y el riesgo de malformaciones congénitas ha sido replicada tanto en cohortes europeas como en estudios de caso–control en Sudamérica, reforzando la idea de un vínculo causal.

3. Especificidad

Aunque en salud muchas exposiciones producen múltiples efectos, la especificidad refuerza la causalidad cuando un agente se asocia principalmente con un solo desenlace. Un ejemplo clásico es la relación entre la infección por *Helicobacter pylori* y la úlcera péptica; al erradicar la bacteria, la mayoría de las úlceras remiten, como si encajaran con una llave única.

4. Temporalidad

Para que A cause B, A debe preceder en el tiempo. En estudios prospectivos de cohorte, como los que siguieron a sobrevivientes de la epidemia de SARS-CoV-2, se comprobó que los síntomas prolongados (fatiga crónica, disnea) aparecieron tras la infección aguda. Sin esta secuencia, cualquier inferencia causal se vuelve especulativa.

5. Gradiente Biológico (Respuesta Dosis–Efecto).

Cuando a mayor exposición corresponde mayor riesgo, la cadena causal se fortalece. Por ejemplo, en trabajadores expuestos a radiación ionizante, el incremento de dosis acumulada se refleja en un aumento proporcional de

casos de cáncer de tiroides[5], dibujando una curva ascendente que parece estar escrita por la propia naturaleza.

6. Plausibilidad y Coherencia

La plausibilidad se nutre de conocimientos biológicos previos: si la relación propuesta encaja en nuestro entendimiento de la fisiopatología, gana solidez. De ahí que el vínculo entre la obesidad y la resistencia a la insulina tenga respaldo en estudios de biología molecular, donde el tejido adiposo inflama y altera rutas de señalización[6]. Al mismo tiempo, la coherencia exige que no contradiga hallazgos de otras disciplinas, como la genética o la bioquímica.

7. Evidencia Experimental

Experimentos controlados, incluso en modelos animales, pueden aportar pruebas directas. Las pruebas con ratones transgénicos expuestos a altas concentraciones de partículas $PM_{2.5}$ revelan daño endotelial, lo que refuerza los datos observacionales en humanos.

8. Analogía

Cuando conocemos un agente similar que produce un efecto análogo, podemos inferir causalidad por analogía. Así, la comprobada capacidad de ciertos pesticidas para dañar el sistema nervioso central allana el camino para sospechar efectos neurotóxicos de compuestos estructuralmente parecidos.

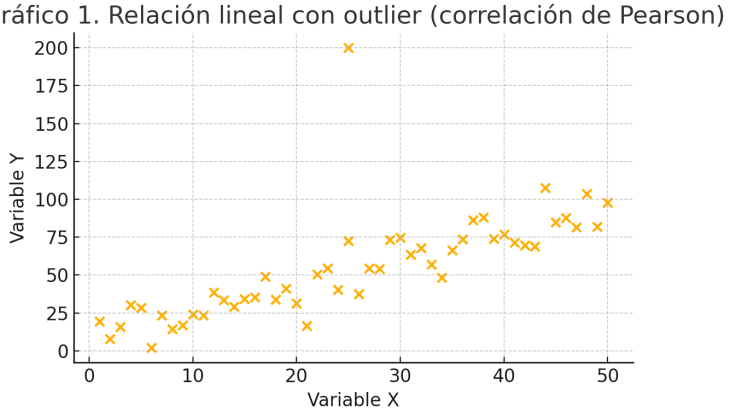
Adoptar estos criterios como guías, y no como requisitos absolutos, permite al investigador trazar rutas sólidas entre observación y conclusión. Mientras profundizamos en métodos estadísticos y técnicas de modelado, los *criterios de Bradford Hill* seguirán siendo ese tejido conectivo que vincula el mosaico de datos con la narrativa de la causalidad, invitándonos a explorar nuevas dimensiones del conocimiento epidemiológico.

9.3. Correlación de Pearson y Spearman

Cuando queremos medir hasta qué punto dos variables cuantitativas continúas viajan de la mano, recurrimos a la **correlación de Pearson**, una estadística que resume en un solo número la fuerza y la dirección de esa relación lineal. Imaginemos que registramos la edad de un grupo de pacientes y su presión arterial sistólica; antes de calcular el coeficiente, conviene compro-

bar que, al trazar ambos ejes, los puntos se distribuyen en una especie de cinta recta y no describen una curva o un cúmulo sin patrón aparente. Si ignoramos esta verificación y simplemente aplicamos Pearson, es como medir la distancia entre dos ciudades en línea recta sin reparar en que el camino real bordea montañas y valles, de modo que los *outliers*, una lectura de presión excepcionalmente alta o baja, distorsionan el resultado, ofreciendo una visión demasiado optimista o pesimista de la asociación (ver figura 48).

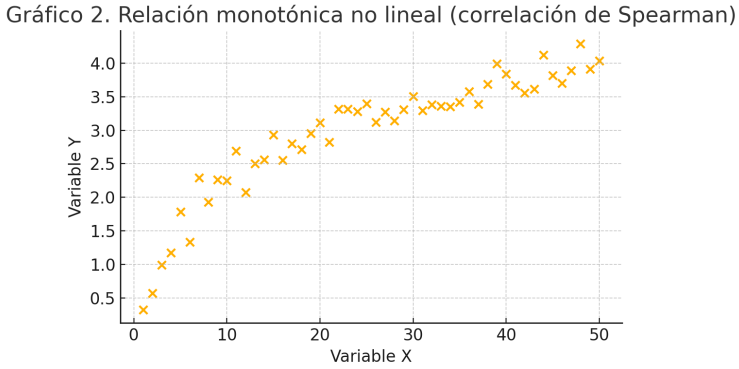
Figura 48. *Correlación de Pearson*. Elaborado con GeminiPRO.



Nota. Observamos un conjunto de datos con relación lineal positiva y un valor atípico marcado. La mayoría de los puntos sigue una cinta ascendente, a medida que la variable X crece, también lo hace la variable Y, pero ese *outlier* distorsiona la estimación de Pearson

Cuando la relación se aparta de esa rigidez lineal, o cuando las mediciones siguen distribuciones asimétricas, conviene cambiar de herramienta y apostar por la **correlación de Spearman**, que se apoya en los rangos de los datos más que en sus valores absolutos. Este método es como equipar nuestro vehículo con ruedas de barro para afrontar terrenos irregulares: funciona bien si las distancias entre puntos no son constantes, pero sí existe un orden subyacente, tal como ocurre en la escala ordinal de dolor de un paciente o en el número de episodios de insomnio registrado por noche (ver figura 49).

Figura 49. *Correlación de Spearman*. Elaborado con GeminiPRO.



Nota. En el Gráfico apreciamos una relación monotónica no lineal, por ejemplo, una evolución logarítmica, en la que los valores de Y aumentan de forma empinada al principio y luego se suavizan. Spearman detecta la dirección y solidez de ese vínculo sin verse engañada por la curvatura del camino, como un todoterreno que conserva la tracción en subidas y suaves pendientes, aunque no pueda proyectar un tramo en línea recta.

Para ilustrar la diferencia, supongamos que correlacionamos las horas de sueño con el rendimiento en un test cognitivo: Pearson detectará solo la inclinación general de la nube de puntos, mientras que Spearman capturará incluso aquellas trayectorias monotónicas en las que el avance en las puntuaciones acelera o frena de manera desigual. No obstante, al emplear Spearman hay que cuidar un aspecto esencial: al convertir valores idénticos en rangos iguales, se generan *empates*, y si no aplicamos una corrección apropiada, el coeficiente tenderá a infravalorar la fuerza real de la asociación.

En ambos casos, la interpretación de un coeficiente cercano a +1 indica una afinidad muy fuerte, a más minutos de ejercicio, menos niveles de colesterol, mientras que valores próximos a –1 señalan una relación inversa pronunciada, cuanto mayor es el consumo de tabaco, menor la función respiratoria; un punto medio alrededor de cero sugiere que la relación, si existe, es débil o no lineal. Lo verdaderamente poderoso de estas técnicas reside en combinarlas con gráficos de dispersión y con el conocimiento clínico, pues un número sin contexto puede ocultar sorpresas, pero acompañado de una narrativa sólida y un diagrama bien interpretado ilumina el camino hacia hallazgos robustos y aplicables en la práctica.

¿En qué Situaciones Conviene Emplear cada Método de Correlación?

La Tabla 16 resume de forma concisa las condiciones en que conviene optar por Pearson o por Spearman según la naturaleza de los datos y la forma de la relación, ofreciendo ejemplos prácticos que guían al investigador en la elección del método más adecuado.

Tabla 16. *Comparación de usos de la correlación de Pearson y Spearman.*

Situación	Prueba recomendada	Ejemplo práctico
Cuando ambas variables continuas cumplen normalidad, presentan relación aproximadamente lineal y carecen de *outliers* importantes	Pearson	Edad vs. presión arterial systólica
Cuando la relación es monotónica pero no lineal, los datos son ordinales o existen distribuciones asimétricas	Spearman	Horas de sueño vs. puntuación en un test cognitivo
Cuando aparecen valores extremos que distorsionan la pendiente de la nube de puntos	Spearman	Concentraciones de biomarcadores con lecturas atípicas
Cuando deseamos cuantificar directamente la proporción de variabilidad explicada por un modelo lineal	Pearson	Índice de masa corporal vs. nivel de colesterol total

9.4. Regresión Lineal, Logística y de Supervivencia

La capacidad de predecir desenlaces clínicos y cuantificar el impacto de múltiples factores sobre la salud constituye uno de los grandes logros de la estadística aplicada a la medicina. Mientras la correlación describe asociaciones, la **regresión** nos permite construir modelos con los que, partiendo de variables conocidas, estimar o clasificar resultados futuros, apoyar decisiones terapéuticas y optimizar el diseño de ensayos. En la actualidad, la disponibilidad de grandes bases de datos electrónicas y de registros de salud hace que estas herramientas sean fundamentales para la medicina de precisión, donde ya no basta describir patrones, sino anticipar efectos en poblaciones y en individuos.

Al abordar la regresión, es crucial comprender que cada técnica responde a un tipo de variable dependiente y a supuestos distintos. La **regresión lineal** funciona cuando el resultado es continuo y exhibe una relación aproximada de tipo rectilíneo con sus predictores, mientras que la **regresión logística** está diseñada para variables binarias, presencia o ausencia de un desenlace, convirtiendo probabilidades en un índice interpretable clínicamente. Por su parte, la **regresión de supervivencia**, o modelo de riesgos proporcionales, integra el tiempo hasta la aparición de un evento y la censura de datos, permitiendo estimar la probabilidad de supervivencia y comparar riesgos entre grupos.

En la práctica, el investigador recurre a estas técnicas no solo para explicar variabilidad observada sino para evaluar escenarios "¿y si...?", como proyectar, cómo varía el riesgo de complicaciones posoperatorias al cambiar la edad o la comorbilidad, o estimar el tiempo hasta recaída según distintos regímenes terapéuticos. A continuación, exploraremos cada método en detalle, describiendo su fundamento conceptual, sin recurrir a fórmulas, los tipos de datos apropiados, aplicaciones ilustrativas, ventajas y limitaciones, y las métricas clave para validar su desempeño.

9.4.1 *Regresión Lineal*

Teoría y datos

La regresión lineal ajusta una "recta" que optimiza la predicción de un resultado numérico continuo a partir de una o varias variables explicativas y asume que el efecto de cada covariable sobre el resultado es aditivo y constante. Imaginemos una vía férrea perfectamente recta, por la que un tren avanza sin desvíos ni curvas. Cada estación, nuestros valores de la variable dependiente, se sitúa a una distancia predecible de la siguiente según la velocidad constante del tren, que aquí representa el efecto de la covariable. Si la vía se mantiene lisa, el recorrido es uniforme y nuestro "tren" llega siempre al punto esperado.

Supuestos

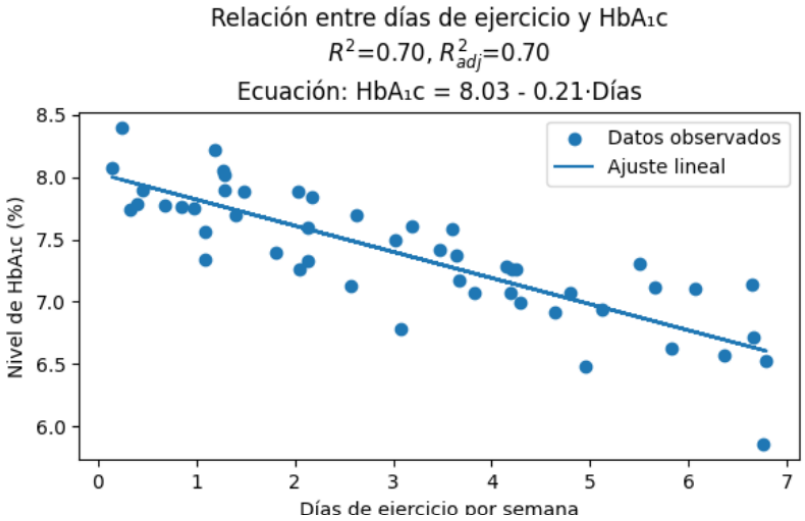
- **Linealidad**: la relación entre cada predictor y la respuesta es aproximadamente rectilínea.
- **Independencia de errores**: las desviaciones (residuos) no están correlacionadas entre sí.

- **Homoscedasticidad**: la varianza de los residuos es constante a lo largo de todos los niveles de predicción.
- **Normalidad de errores**: los residuos se distribuyen de forma aproximadamente normal.
- **Ausencia de multicolinealidad**: los predictores no están excesivamente correlacionados entre sí.

Aplicación Práctica

Un endocrinólogo usa regresión lineal para estimar el cambio en HbA_1c según dosis de insulina y días de ejercicio semanal, obteniendo una ecuación que indica la reducción esperada de décimas de HbA_1c por cada unidad de ejercicio adicional.

Figura 50. *Relación lineal entre días de ejercicio semanal y nivel de* HbA_1c. Elaborado con GeminiPRO.



Nota. El modelo de regresión lineal muestra que por cada día adicional de ejercicio por semana, el nivel de HbA_1c se reduce en promedio 0,21 puntos porcentuales, según la ecuación HbA_1c = 8,03 – 0,21·Días. El coeficiente de determinación (R^2 = 0,70) y su versión ajustada (R^2□□□ = 0,70) indican que el 70 % de la variabilidad en HbA_1c queda explicada por el número de días de ejercicio, lo que sugiere un buen ajuste del modelo y una asociación clínicamente relevante entre actividad física y control glucémico.

Ventajas y Desventajas

- *Ventajas*: interpretable, potencia estadística alta si se cumplen sus supuestos.
- *Desventajas*: sensible a *outliers* y violaciones de homoscedasticidad o normalidad, no captura no linealidades.

Métricas de Evaluación

- R^2 *y* R^2 *ajustado*
- *Error estándar de estimación*
- *Análisis gráfico de residuos*

9.4.2. *Regresión Logística*

Teoría y Datos

La regresión logística modela la probabilidad de un desenlace dicotómico, transformando odds en una escala logarítmica lineal respecto a los predictores. Pensemos en un interruptor de luz que solo puede estar en posición "encendido" o "apagado", pero cuya probabilidad de activarse depende de la intensidad de una señal eléctrica. Conforme aumenta la corriente, nuestro predictor, la posibilidad de que la bombilla se encienda crece siguiendo una curva en forma de "S", hasta acercarse a un 100 %.

Supuestos

- **Independencia de observaciones**: cada caso es independiente de los demás.
- **Linealidad en el logit**: la relación entre cada predictor continuo y el logaritmo de las probabilidades es lineal.
- **Ausencia de multicolinealidad**: predictores no excesivamente correlacionados.
- **Tamaño muestral adecuado**: suficientes eventos por covariable (regla empírica: ≥10 eventos por predictor) para evitar sobreajuste.

- **No separación completa**: no debe haber predictores que determinen 100 % un desenlace.

Aplicación Práctica

Un cirujano estima la probabilidad de complicación posoperatoria grave según edad, puntuación ASA y tiempo quirúrgico; el modelo reporta *odds ratios* para cada variable, facilitando la valoración individual de riesgo.

Ventajas y Desventajas

- *Ventajas*: admite variables mixtas, no requiere normalidad de predictores.
- *Desventajas*: interpretación menos directa que la lineal, sensible a muestras pequeñas y predictores perfectos.

Métricas de Evaluación

- *AUC (curva ROC)*
- *Test de Hosmer–Lemeshow*
- *Sensibilidad y especificidad*
- *–2 Log-Likelihood*
- R^2 *de Cox & Snell y* R^2 *de Nagelkerke (Pseudo* R^2*)*
- *Factor de Inflación de Varianza (VIF)*

Para evaluar qué tan bien se ajusta un modelo de regresión logística a los datos y qué tan efectivo es para explicar la probabilidad del resultado, se emplean diversas métricas. Además de las medidas que evalúan la capacidad del modelo para discriminar entre los dos resultados posibles (como el Área Bajo la Curva ROC, la Sensibilidad y la Especificidad), es fundamental considerar otras métricas de ajuste global y diagnóstico. A continuación, se detalla la interpretación de algunas de las más importantes mencionadas:

- **Test de Bondad de Ajuste de Hosmer-Lemeshow**: su interpretación es clave para el "ajuste general". Este test evalúa si las probabilidades predichas por el modelo se corresponden bien con las frecuencias obser-

vadas del evento a lo largo de diferentes rangos de probabilidad (agrupando las observaciones). Se busca un modelo que "se ajuste" bien a los datos. La interpretación se basa en un valor p: un valor p **alto (generalmente p > 0.05)** indica que **no hay una diferencia estadísticamente significativa** entre las frecuencias observadas y esperadas, lo que es la conclusión deseada y sugiere un **buen ajuste del modelo** a los datos. Un valor p bajo ($p \leq 0.05$) indicaría un ajuste pobre.

- **–2 Log-Likelihood (–2LL)**: Esta métrica es una medida de la variabilidad en la variable dependiente (el resultado) que no es explicada por el modelo. Deriva de la función de verosimilitud, que estima la probabilidad de observar los datos reales dado el modelo. Un valor de –2LL **más bajo** indica que el modelo se ajusta **mejor** a los datos, dejando menos variabilidad sin explicar. Se utiliza a menudo para comparar la mejora en el ajuste entre modelos anidados; una disminución significativa en el –2LL al añadir predictores sugiere que esos predictores contribuyen positivamente al ajuste del modelo.

- **Pseudo R^2 (R^2 de Cox & Snell y R^2 de Nagelkerke)**: A diferencia de la R^2 en la regresión lineal que mide la proporción de varianza explicada, en la regresión logística no hay una "varianza" directa que explicar en el mismo sentido. Estas "pseudo" R^2 intentan ofrecer una medida análoga del poder explicativo del modelo, es decir, cuánto de la "incertidumbre" en el resultado es reducida por los predictores del modelo.

  - La **R^2 de Cox & Snell** es una primera aproximación, pero su limitación es que su valor máximo casi nunca alcanza 1, incluso para modelos perfectos, lo que dificulta su interpretación como un porcentaje directo.

  - La **R^2 de Nagelkerke** es una versión ajustada de la de Cox & Snell que **sí** puede alcanzar un valor de 1. Por lo tanto, es generalmente la más utilizada para interpretar el poder explicativo: un valor **más alto** en la R^2 de Nagelkerke sugiere que los predictores del modelo explican una mayor proporción de la variabilidad en el resultado, similar a cómo interpretamos la R^2 en regresión lineal (aunque no son conceptualmente idénticas).

- **Factor de Inflación de Varianza (VIF)**: Esta es una métrica de diagnóstico utilizada para detectar multicolinealidad, que ocurre cuando dos o

más variables predictoras están altamente correlacionadas entre sí. Para cada predictor en el modelo, el VIF cuantifica cuánto se "infla" la varianza de su coeficiente estimado debido a estas correlaciones con otros predictores. Un VIF **alto** (los umbrales varían, pero a menudo se consideran problemáticos valores **por encima de 5 o 10**) para una variable indica **multicolinealidad significativa**, lo que puede hacer que la estimación e interpretación individual de ese coeficiente sea inestable. Un VIF cercano a 1 sugiere poca o ninguna colinealidad. Se calcula individualmente para cada variable predictora incluida en el modelo con el fin de verificar que se cumple el supuesto de ausencia de multicolinealidad, un requisito importante para la fiabilidad de los coeficientes del modelo.

Evaluar estas métricas en conjunto nos da una visión completa del desempeño del modelo logístico, no solo en su capacidad de clasificar, sino también en su ajuste a los datos y la estabilidad de sus estimaciones.

9.4.3. *Regresión de Supervivencia* (Cox)

Teoría y datos

El modelo de Cox analiza tiempos hasta un evento y censura de datos, estimando *hazard ratios* que cuantifican el riesgo instantáneo asociado a cada predictor. Visualicemos una carrera de larga distancia en la que algunos corredores no completan el recorrido antes del cierre del cronómetro. El modelo de Cox compara velocidades de avance instantáneas (*hazard*) entre grupos, sin exigir que todos crucen la meta, integrando la censura de quienes "se retiraron" en cualquier punto.

Supuestos

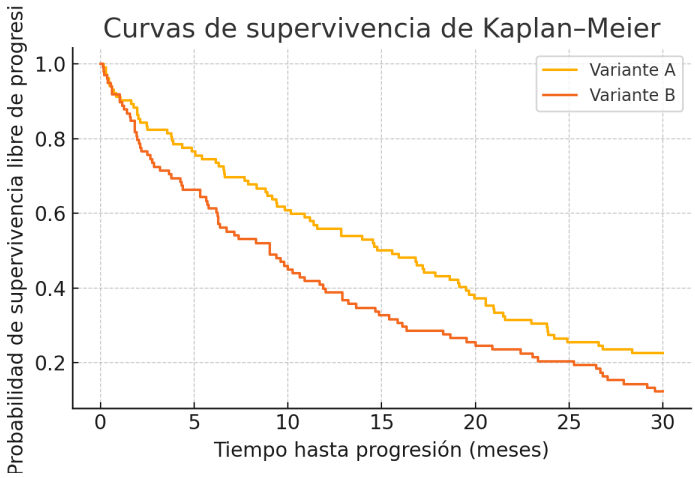
- **Proporcionalidad de riesgos**: el efecto de cada covariable sobre el riesgo es constante en el tiempo.
- **Independencia de tiempos de supervivencia**: las observaciones de tiempo hasta evento son independientes.
- **Censura no informativa**: el motivo de cese de seguimiento no se relaciona con el riesgo del evento.

- **Linealidad de covariables continuas** en el logaritmo del hazard.

Aplicación práctica

En oncología, se compara el riesgo de progresión tumoral entre perfiles genéticos ajustando por edad y estado funcional; el modelo indica cuántas veces mayor es el *hazard* para cada variante, permitiendo estratificar pacientes (ver figura 51).

Figura 51. *Curvas de supervivencia libres de progresión según perfil genético*. Elaborado con GeminiPRO.



Nota. Las curvas de Kaplan–Meier comparan la probabilidad de permanecer libre de progresión tumoral en 30 meses entre pacientes con "Variante A" y "Variante B", sin ajustar por covariables. La separación entre las curvas muestra un menor tiempo de supervivencia libre de progresión en el grupo con Variante B (línea naranja), coherente con un hazard ratio mayor a 1 obtenido en el modelo de Cox ajustado por edad y estado funcional.

Ventajas y Desventajas

- *Ventajas:* maneja censura, no exige distribución paramétrica del tiempo de supervivencia.
- *Desventajas:* vulnerable si no se cumple la proporcionalidad de riesgos; interpretación más abstracta.

Métricas de evaluación

- *Hazard ratios* con IC
- *Prueba de proporcionalidad de riesgos (residuos de Schoenfeld)*
- *Curvas de Kaplan–Meier y logrank*

Conclusiones

1. **Asociación vs. Causalidad:** Una asociación estadística entre dos variables no implica necesariamente una relación de causa-efecto; es crucial considerar posibles sesgos, factores de confusión y la dirección temporal.
2. **Inferencia Causal:** Establecer causalidad es complejo. Los modelos teóricos (causa necesaria/suficiente) y los criterios de Bradford Hill (fuerza, consistencia, especificidad, temporalidad, gradiente, plausibilidad, coherencia, evidencia experimental, analogía) ayudan a evaluar la probabilidad de un vínculo causal.
3. **Medición de Correlación:** La correlación cuantifica la asociación entre dos variables cuantitativas. Pearson mide la relación lineal (requiere normalidad), mientras que Spearman mide la relación monotónica (más robusta ante no linealidad o no normalidad).
4. **Modelos de Regresión para Predicción:** La regresión permite predecir un resultado (variable dependiente) a partir de uno o más predictores (variables independientes).
5. **Tipos Principales de Regresión:**

- **Lineal:** Para predecir variables dependientes continuas. Requiere linealidad, independencia, homoscedasticidad y normalidad de errores.
- **Logística:** Para predecir variables dependientes binarias (dicotómicas). Modela la probabilidad del evento.
- **De Supervivencia (Cox):** Para analizar datos de tiempo hasta un evento, especialmente útil cuando hay censura. Estima Hazard Ratios y asume proporcionalidad de riesgos.

6. **Verificación de Supuestos y Evaluación:** Cada modelo de regresión tiene supuestos específicos que deben verificarse. Su desempeño se evalúa mediante métricas adecuadas (ej. R^2 para lineal, AUC/Pseudo R^2 para logística, pruebas de proporcionalidad para Cox).
7. **Aplicación Cuidadosa:** Tanto la correlación como la regresión son herramientas poderosas, pero su aplicación e interpretación requieren rigor metodológico y consideración del contexto clínico o epidemiológico.

Parte III.
Epidemiología Básica

¿Por qué Empezar por Epidemiología?

Si en las secciones anteriores exploramos los cimientos teóricos y epistemológicos que sustentan la búsqueda del conocimiento en medicina y nos adentramos en el fascinante mundo del manejo y análisis de datos, sentando las bases para la formulación y evaluación de hipótesis, ahora daremos un paso crucial: aplicar todo ese conocimiento para entender la salud y la enfermedad, no en un paciente individual, sino en el vasto y complejo tapiz de una población. Aquí es donde entra en juego la epidemiología, una disciplina tan esencial para la medicina como el estetoscopio para el cardiólogo.

Quizás se pregunten, ¿por qué dedicar un capítulo completo a la epidemiología justo después de hablar de hipótesis y datos? La respuesta es sencilla y poderosa: la epidemiología es la lente a través de la cual observamos los patrones de salud y enfermedad en las poblaciones, y es la herramienta que nos permite poner a prueba nuestras hipótesis sobre qué causa la enfermedad o qué interviene para prevenirla. Recuerden la MBE que discutimos al principio; esa evidencia, en gran medida, proviene de estudios epidemiológicos bien diseñados.

Imaginen que son detectives de la salud pública. No están investigando un solo caso de infección, sino tratando de entender por qué en un barrio específico, *muchas* personas están desarrollando una extraña enfermedad. Necesitan mirar más allá del individuo para encontrar las pistas en el entorno, en los hábitos compartidos, en las exposiciones comunes. Esa es la esencia del pensamiento epidemiológico. Nos ayuda a pasar de la anécdota clínica a la comprensión de fenómenos a gran escala, lo cual es indispensable para tomar decisiones informadas, tanto en la cabecera del paciente, aplicando la evidencia, como en la planificación de políticas de salud pública.

¿Qué es la Epidemiología?

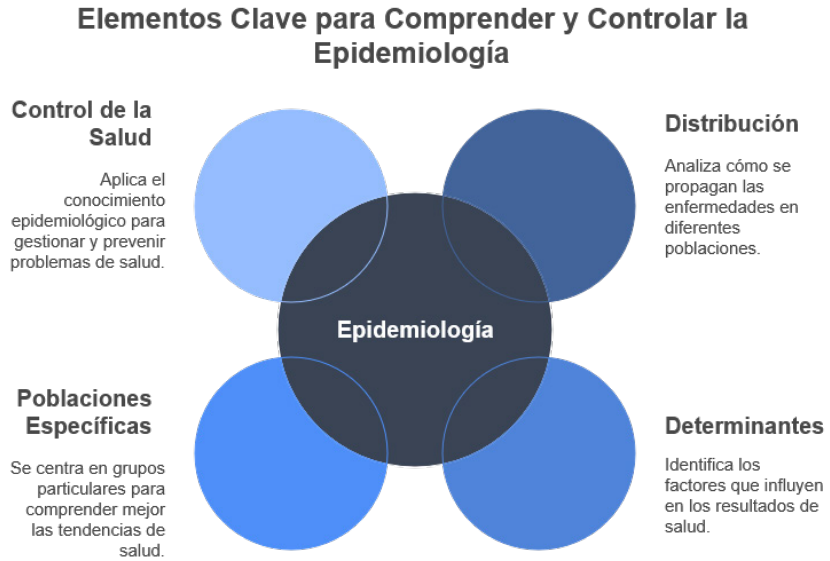
Entonces, ¿qué es exactamente esta disciplina? De forma concisa, la epidemiología es el estudio de la distribución y los determinantes de los estados o eventos relacionados con la salud en poblaciones específicas y la aplicación de este estudio para el control de los problemas de salud. Desglosemos esto un poco.

Cuando hablamos de *distribución*, nos referimos a cómo se presenta un problema de salud: ¿quiénes se enferman, edad, sexo, raza, ocupación, etc.?, ¿dónde ocurre, ubicación geográfica?, ¿cuándo sucede, época del año, tendencias a lo largo del tiempo? Es como pintar un mapa detallado de la enfermedad.

Por otro lado, los *determinantes* son los factores, causas o condiciones que influyen en la aparición de la enfermedad o estado de salud. Pueden ser biológicos, un virus, una mutación genética, conductuales, fumar, hacer ejercicio, ambientales, contaminación, acceso a agua potable, o sociales, nivel socioeconómico, acceso a servicios de salud.

Y lo crucial es que este estudio se centra en *poblaciones específicas*, no en individuos aislados. Una población puede ser tan pequeña como los estudiantes de una facultad de medicina o tan grande como los habitantes de un continente. Finalmente, el objetivo último no es solo entender, sino *aplicar* ese conocimiento para prevenir y controlar los problemas de salud, mejorando el bienestar colectivo (ver figura 52).

Figura 52. *Componentes Fundamentales de la Epidemiología*. Elaborado con GeminiPRO.



Objetivos de la Epidemiología

La epidemiología no es solo un ejercicio académico; tiene metas muy prácticas y vitales para la salud humana. Sus objetivos principales incluyen:

1. **Describir el estado de salud de las poblaciones:** Esto implica cuantificar la frecuencia de enfermedades, discapacidades, muertes, pero también de estados positivos como la salud y el bienestar. Nos dice qué tan común es un problema y quiénes son los más afectados.
2. **Identificar los factores etiológicos o de riesgo:** Busca las causas de las enfermedades y los factores que aumentan o disminuyen la probabilidad de que ocurran. Aquí es donde nuestras hipótesis sobre la relación entre variables, formuladas en la Parte II, cobran vida a nivel poblacional.
3. **Predecir la ocurrencia de la enfermedad:** Al entender los patrones y los determinantes, podemos anticipar dónde y cuándo es más probable que surjan o se propaguen los problemas de salud.
4. **Evaluar la eficacia de las intervenciones de salud:** Permite medir el impacto de programas de prevención, tratamientos, y políticas de salud en la población. ¿Funciona una nueva vacuna? ¿Es efectivo un programa para dejar de fumar? La epidemiología nos da las herramientas para responder.
5. **Proporcionar los datos necesarios para la planificación y gestión de servicios de salud:** La información epidemiológica es la base para decidir dónde asignar recursos, qué programas implementar y cómo organizar la atención médica para satisfacer las necesidades de la población.

En esencia, la epidemiología nos da la información para actuar. Como señaló una vez el renombrado epidemiólogo Geoffrey Rose, la epidemiología de poblaciones y la medicina clínica son dos caras de la misma moneda, ambas esenciales para la práctica médica completa.

Factor de exposición y evento

Dos conceptos centrales que encontraremos una y otra vez en epidemiología son el *factor de exposición* y el *evento*. Piensen en ellos como las dos variables principales que intentamos relacionar en nuestros estudios poblacionales, análogas a las variables independiente y dependiente que exploramos al hablar de datos y análisis.

El **factor de exposición** es cualquier característica, comportamiento, factor ambiental u otra variable a la que los individuos de una población están expuestos y que se cree que influye en el riesgo de desarrollar un determinado estado o evento de salud. Es la "causa" o el "predictor" potencial que estamos investigando. Ejemplos de exposiciones podrían ser: fumar cigarrillos, recibir una vacuna, vivir cerca de una fábrica contaminante, tener una dieta rica en fibra, o incluso poseer una cierta variante genética.

El **evento** (o resultado) es el estado o la enfermedad que nos interesa estudiar. Es el "efecto" o la "consecuencia" que podría estar relacionado con la exposición. Ejemplos de eventos son: desarrollar cáncer de pulmón, contraer *influenza*, sufrir un infarto de miocardio, recuperarse de una enfermedad, o morir.

La pregunta fundamental que la epidemiología intenta responder es: ¿Existe una asociación entre un factor de exposición específico y la ocurrencia de un evento de salud en una población? Y si existe, ¿qué tan fuerte es esa asociación?

Imaginemos que la salud de una población es como un barco navegando en aguas a veces turbulentas. El *evento* es si el barco llega a puerto seguro o si naufraga, enfermedad o muerte. El *factor de exposición* son las condiciones o acciones que influyen en ese viaje: el estado del casco, genética, si hay tormentas, ambiente, si la tripulación es experimentada, acceso a salud, o si hay agujeros en el bote, fumar. La epidemiología nos ayuda a identificar qué "agujeros" o "tormentas" son los más peligrosos para la navegación de la salud poblacional (ver figura 53).

Figura 53. *Relación entre Factor de Exposición y Evento de Salud.* Elaborado con GeminiPRO.



La Necesidad de Cuantificar: hacia las Medidas Epidemiológicas

Entender estos conceptos es el primer paso para adentrarnos en cómo medimos la frecuencia de los eventos, cuantificamos la relación entre exposición y evento, y diseñamos estudios para investigar estas asociaciones. Sin embargo, identificar un factor de exposición y un evento no es suficiente; necesitamos saber *cuántas* personas en la población experimentan el evento y *cuántas* están expuestas al factor de interés. Es decir, debemos cuantificar.

Piensen en la población como un gran grupo de personas, algunas de las cuales están en riesgo de sufrir un determinado evento de salud. Para entender realmente el impacto de un factor de exposición, necesitamos contar cuántos eventos ocurren en el grupo expuesto y cuántos ocurren en el grupo no expuesto, y relacionar estas cifras con el tamaño de cada grupo. Esta cuantificación nos permite comparar el riesgo entre diferentes grupos y evaluar la magnitud de un problema de salud.

Esta necesidad de ponerle números a la distribución y los determinantes de la salud nos lleva directamente al corazón del próximo capítulo: las medidas epidemiológicas. Estas medidas son las herramientas matemáticas que utilizamos para describir la frecuencia de los eventos de salud, como la prevalencia y la incidencia, y para evaluar la fuerza de la asociación entre una exposición y un evento, como el riesgo relativo o el *odds ratio*. Son el lenguaje numérico de la epidemiología, permitiéndonos pasar de la observación a la acción informada.

CAPÍTULO 12

MEDIDAS EPIDEMIOLÓGICAS

Epidemiological Measures

Resumen

Este capítulo introduce las medidas epidemiológicas como herramientas cuantitativas para entender la salud de las poblaciones, clasificándolas en medidas de ocurrencia, asociación e impacto. Se centra en las medidas de ocurrencia, que cuantifican la frecuencia de eventos de salud. Detalla la prevalencia como una "fotografía" que muestra la proporción de casos existentes (nuevos y antiguos) en una población en un momento específico, útil para medir la carga de enfermedad. Describe la incidencia como una "película" que mide la aparición de casos nuevos en una población en riesgo durante un período determinado. Explica las dos formas de incidencia: la incidencia acumulada (proporción o riesgo de desarrollar el evento en un período) y la tasa de incidencia o densidad (velocidad de aparición de nuevos casos, usando tiempo-persona en riesgo). Finalmente, utiliza el "modelo de la bañera" para ilustrar la relación dinámica entre incidencia (entrada de agua) y prevalencia (nivel de agua), donde la prevalencia depende tanto de los casos nuevos como de la duración de la enfermedad (salida de agua por curación o muerte).

Palabras Clave: Medidas epidemiológicas, prevalencia, incidencia, incidencia acumulada, tasa de incidencia, ocurrencia de enfermedad.

Abstract

This chapter introduces epidemiological measures as quantitative tools for understanding population health, classifying them into measures of occurrence, association, and impact. It focuses on measures of occurrence, which quantify the frequency of health events. Prevalence is detailed as a "snapshot" showing the proportion of existing cases (new and old) in a population at a specific point in time, useful for measuring disease burden. Incidence is described as a "movie" measuring the appearance of new cases in a population at risk over a specified period. It explains the two forms of incidence: cumulative incidence (proportion

or risk of developing the event over a period) and incidence rate or density (speed at which new cases occur, using person-time at risk). Finally, it uses the "bathtub model" to illustrate the dynamic relationship between incidence (inflow) and prevalence (water level), where prevalence depends on both new cases and disease duration (outflow through recovery or death).

Keywords: *Epidemiological measures, prevalence, incidence, cumulative incidence, incidence rate, disease occurrence.*

Introducción

Imaginen que la salud de una población es como un gran ecosistema, lleno de diversas formas de vida y complejas interacciones. Para entender este ecosistema, no basta con saber qué especies existen, necesitamos contar cuántas hay de cada una, dónde se concentran, y cómo se relacionan entre sí. Las medidas epidemiológicas son nuestras herramientas de "conteo" y "relación" en este ecosistema humano. Nos permiten pasar de la cualidad, decir "el tabaquismo causa cáncer de pulmón", a la cantidad, decir "los fumadores tienen un riesgo X veces mayor de desarrollar cáncer de pulmón que los no fumadores".

Estas medidas se agrupan generalmente en tres categorías principales, reflejando los diferentes tipos de preguntas que buscamos responder:

- **Medidas de Ocurrencia:** ¿Cuántos casos de una enfermedad existen o aparecen en una población?
- **Medidas de Asociación:** ¿Qué tan fuerte es la relación entre una exposición y un evento?
- **Medidas de Impacto:** ¿Cuánto de un evento de salud en una población se debe a una exposición particular?

Exploremos cada una de ellas.

1.1 Medidas de Ocurrencia

Las medidas de ocurrencia nos dicen con qué frecuencia se presenta un evento de salud en una población. Son fundamentales para describir el estado de salud de una comunidad y para comparar la carga de enfermedad entre diferentes poblaciones o a lo largo del tiempo. Las dos medidas de ocu-

rrencia más importantes son la prevalencia y la incidencia.

1.1.1 *Prevalencia*

Piensen en la prevalencia como una *fotografía instantánea* de la población en un momento dado. Nos dice cuántos casos de una enfermedad o condición existen *en ese preciso instante*, independientemente de cuándo comenzaron. Incluye tanto los casos nuevos como los antiguos que aún están presentes.

Formalmente, la prevalencia se define como:

$$\text{Prevalencia} = \frac{\text{Número de casos existentes de un evento en un momento dado}}{\text{Población total en ese mismo momento}}$$

Se suele expresar como una proporción, porcentaje o tasa por cada 1.000 o 10.000 personas, dependiendo de qué tan común sea el evento.

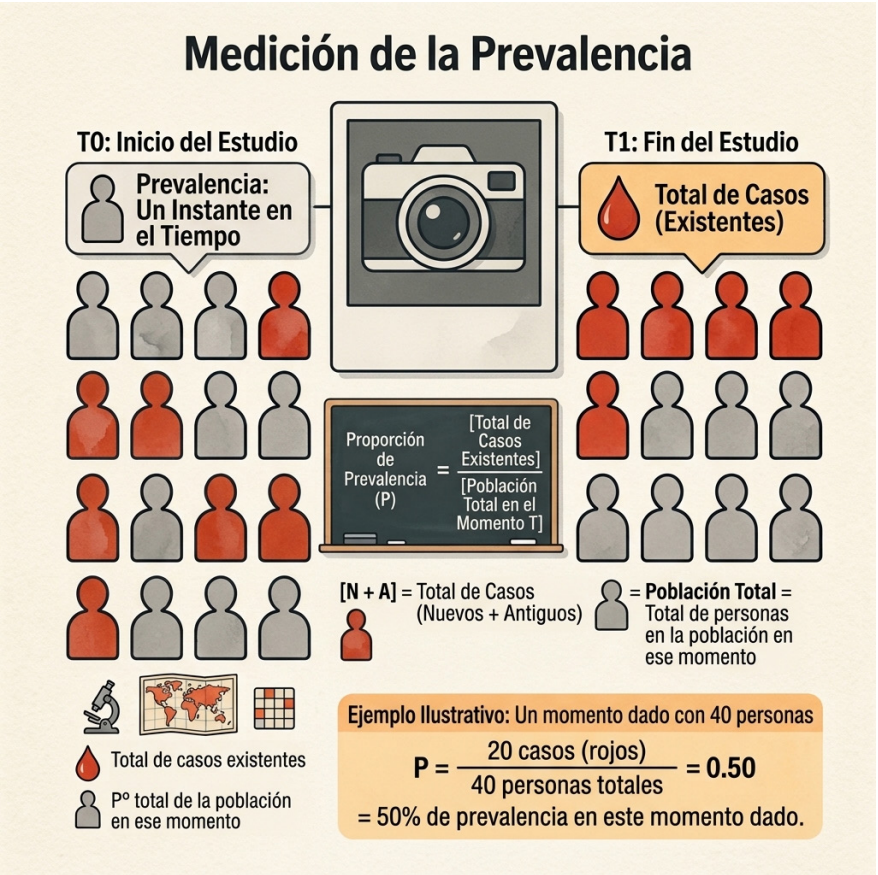
Ejemplo práctico: Supongamos que el 1 de enero de 2024, en una ciudad de 100.000 habitantes, hay 5.000 personas viviendo con diabetes. La prevalencia de diabetes en esa ciudad en esa fecha sería:

$$\text{Prevalencia} = \frac{5.000 \text{ casos de diabetes}}{100.000 \text{ habitantes}} = 0,05 \text{ o } 5$$

Esto significa que, en ese momento, el 5% de la población de esa ciudad tenía diabetes.

La prevalencia es útil para planificar servicios de salud, ya que nos da una idea de la carga actual de una enfermedad. Sin embargo, no nos dice nada sobre el riesgo de desarrollar la enfermedad o sobre la duración de la misma. Una enfermedad de alta prevalencia podría ser muy común, o podría ser una enfermedad crónica de larga duración, incluso si los casos nuevos son pocos (ver figura 54).

Figura 54. *Representación de la Prevalencia*. Elaborado con GeminiPRO.



Nota. Este diagrama ilustra la prevalencia como la proporción de individuos afectados (siluetas rojas) en relación con la población total (100 siluetas) en un momento específico, mostrando de forma clara la carga existente de casos frente a los no casos.

1.1.2 *Incidencia*

Si la prevalencia es una fotografía, la incidencia es la *película* de la población a lo largo del tiempo. Se enfoca en los *casos nuevos* de una enfermedad o condición que aparecen durante un período específico de seguimiento en una población que está en riesgo de desarrollarla.

Existen dos formas principales de medir la incidencia:

- **Incidencia** Acumulada (Proporción **de Incidencia):** Es la proporción de individuos sanos que desarrollan el evento durante un período de tiempo determinado. Asume que toda la población en riesgo fue seguida durante todo el período.

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

Se interpreta como el riesgo promedio de un individuo de desarrollar el evento durante ese período.

- **Tasa de Incidencia (Densidad de Incidencia):** Mide la velocidad a la que ocurren los casos nuevos. Es particularmente útil cuando los individuos son seguidos por diferentes lapsos de tiempo (por ejemplo, algunos se pierden durante el seguimiento o mueren por otras causas). El denominador utiliza el "tiempo persona en riesgo", que suma el tiempo que cada individuo estuvo en riesgo de desarrollar el evento.

$$\text{Tasa de incidencia} = \frac{\text{Número de casos nuevos durante el período}}{\sum \text{tiempo-persona en riesgo}}$$

La tasa se expresa por unidad de tiempo persona (por ejemplo, casos por 1.000 personas-año).

Ejemplo práctico (Incidencia Acumulada): En un estudio de seguimiento de 1.000 personas sanas durante un año para ver cuántos desarrollan *influenza*, si al final del año 100 personas han tenido *influenza*, la incidencia acumulada es:

$$\text{Incidencia Acumulada} = \frac{100 \text{ casos nuevos}}{1.000 \text{ personas en riesgo}} = 0,10 \text{ o } 10$$

Esto significa que el riesgo de desarrollar *influenza* en esa población durante ese año fue del 10%.

Ejemplo práctico (Tasa de Incidencia): Si en un estudio sobre una enfermedad rara, se siguen a 500 personas durante un total de 1.500 personas-año (sumando el tiempo que cada persona estuvo en el estudio antes de enfer-

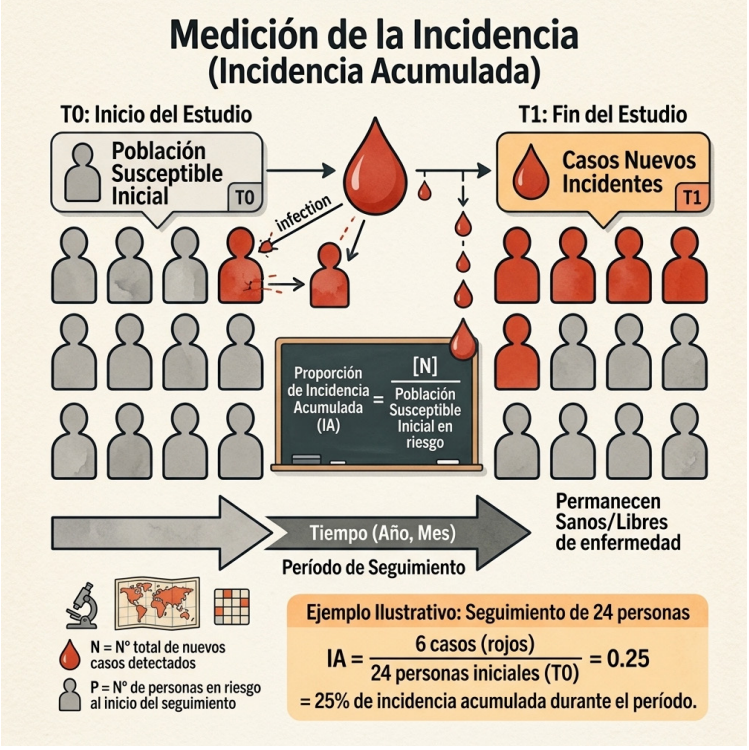
marse, perderse o terminar el seguimiento), y se observan 30 casos nuevos, la tasa de incidencia es:

$$\text{Tasa de Incidencia} = \frac{30 \text{ casos nuevos}}{1.500 \text{ personas-año}} = 0,02 \text{ casos por persona-año}$$

O 20 casos por 1.000 personas-año (0,02 * 1.000). Esto nos dice que, en promedio, ocurren 20 casos nuevos por cada 1.000 años de seguimiento en personas en riesgo.

La incidencia es fundamental para investigar las causas de las enfermedades, ya que mide el riesgo de desarrollar un evento. Un aumento en la incidencia de una enfermedad a menudo sugiere la aparición de nuevos factores de riesgo o cambios en los existentes (ver figura 55).

Figura 55. *Representación de la Incidencia*. Elaborado con GeminiPRO.

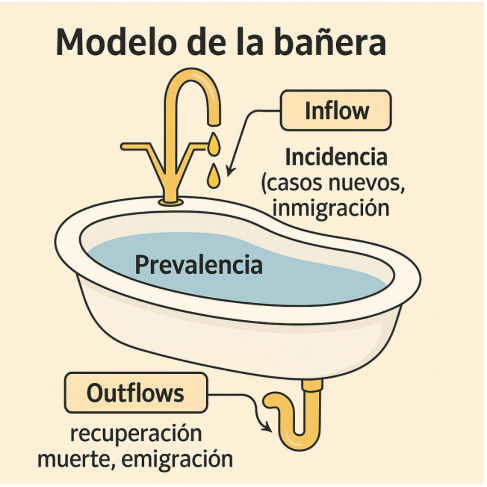


Nota. Este diagrama muestra la incidencia al resaltar los casos nuevos (siluetas rojas) que emergen de la población en riesgo (siluetas grises) a lo largo del

período indicado por la flecha, ilustrando cómo se generan eventos de salud en tiempo determinado.

La Figura 56 presenta el llamado **modelo de la bañera**, una metáfora de dinámica de sistemas que permite visualizar de forma intuitiva cómo interactúan los principales indicadores epidemiológicos. En este esquema, la **prevalencia**, el *stock* de casos existentes, equivale al volumen de agua acumulada en la bañera; su magnitud depende tanto de la **incidencia**, flujo de entrada que representa los casos nuevos e inmigración de enfermos, como de las **salidas**, recuperación, muerte o emigración de pacientes, que drenan el nivel de agua. Así, un aumento sostenido en la incidencia o una reducción de las salidas conducirá a un incremento de la prevalencia, mientras que un mayor ritmo de recuperación o mortalidad reducirá la carga de enfermedad en la población. Este modelo resulta especialmente útil para entender la evolución temporal de las enfermedades, planificar recursos de salud pública y evaluar el impacto de intervenciones orientadas a modificar las tasas de entrada o salida de casos.

Figura 56. *Modelo de la bañera*. Elaborado con GeminiPRO.



Nota. El nivel de agua dentro de la bañera representa la prevalencia (*stock* de casos existentes), la entrada de agua por el grifo (*inflow*) simboliza la incidencia (casos nuevos e inmigración) y la salida de agua por el desagüe (*outflows*) corresponde a la recuperación, muerte y emigración de casos.

Conclusiones

1. **Cuantificación en Epidemiología:** Las medidas epidemiológicas son esenciales para cuantificar la frecuencia (ocurrencia), la relación entre factores y eventos (asociación) y el efecto de exposiciones en la población (impacto).

2. **Medidas de Ocurrencia:** Describen con qué frecuencia ocurren los eventos de salud. Las principales son la prevalencia y la incidencia.

3. **Prevalencia (Fotografía):** Mide la proporción de individuos en una población que tienen una enfermedad o condición en un momento específico. Incluye casos nuevos y preexistentes. Es útil para la planificación sanitaria y la descripción de la carga de enfermedad actual.

4. **Incidencia (Película):** Mide la aparición de *nuevos* casos de una enfermedad en una población en riesgo durante un período de tiempo. Es fundamental para estudiar las causas (etiología) y el riesgo de desarrollar una enfermedad.

5. **Tipos de Incidencia:**

- **Incidencia Acumulada (Riesgo):** Proporción de personas en riesgo que desarrollan la enfermedad en un período fijo.

- **Tasa de Incidencia (Densidad):** Velocidad a la que ocurren nuevos casos, considerando el tiempo que cada persona está en riesgo (personas-tiempo). Adecuada para seguimientos dinámicos con entradas, salidas o pérdidas.

6. **Relación Prevalencia-Incidencia (Modelo Bañera):** La prevalencia está determinada por la incidencia (casos nuevos que entran) y la duración de la enfermedad (tiempo hasta la recuperación o muerte, que determina la salida). Enfermedades con alta incidencia o larga duración tienden a tener alta prevalencia.

CAPÍTULO 13

MEDIDAS DE ASOCIACIÓN

Association Measures

Resumen

Este capítulo introduce las medidas de asociación epidemiológica, herramientas para cuantificar la relación entre una exposición y un evento de salud, yendo más allá de la simple descripción de la ocurrencia de enfermedades. Se detallan dos medidas principales: el Riesgo Relativo (RR) y la Odds Ratio (OR). El RR se define como la razón entre la incidencia del evento en el grupo expuesto y la incidencia en el grupo no expuesto (IE/INE). Es la medida de elección en estudios de cohorte, ya que compara directamente el riesgo de desarrollar el evento. Un RR > 1 indica mayor riesgo en expuestos, RR < 1 indica protección y RR = 1 indica ausencia de asociación. La Odds Ratio (OR) compara las odds (probabilidad de que ocurra / probabilidad de que no ocurra) de exposición entre los casos (enfermos) y los controles (no enfermos). Se calcula como el cociente del producto cruzado (ad/bc) en una tabla 2x2 y es la medida principal en estudios de casos y controles. La OR aproxima bien al RR cuando la enfermedad es rara. Su interpretación es similar: OR > 1 sugiere asociación positiva (factor de riesgo), OR < 1 sugiere asociación negativa (factor protector) y OR = 1 sugiere ausencia de asociación.

Palabras Clave: Medidas de asociación, riesgo relativo, odds ratio, razón de momios, estudios de cohorte, estudios de casos y controles.

Abstract

This chapter introduces epidemiological measures of association, tools used to quantify the relationship between an exposure and a health event, moving beyond simply describing disease occurrence. Two main measures are detailed: Relative Risk (RR) and Odds Ratio (OR). RR is defined as the ratio of the incidence of the event in the exposed group to the incidence in the unexposed group (IE/INE). It is the measure of choice in cohort studies as it directly compares the risk

of developing the event. An RR > 1 indicates increased risk in the exposed, RR < 1 indicates protection, and RR = 1 indicates no association. The Odds Ratio (OR) compares the odds (probability of occurrence / probability of non-occurrence) of exposure between cases (diseased) and controls (non-diseased). It is calculated as the cross-product ratio (ad/bc) in a 2x2 table and is the primary measure in case-control studies. The OR approximates the RR well when the disease is rare. Its interpretation is similar: OR > 1 suggests a positive association (risk factor), OR < 1 suggests a negative association (protective factor), and OR = 1 suggests no association.

Keywords: *Measures of association, relative risk, odds ratio, cohort studies, case-control studies, risk assessment.*

Introducción

Hasta ahora hemos navegado por las aguas de la medición epidemiológica, aprendiendo a cuantificar cuánta enfermedad hay en una población (Prevalencia) y cuán rápido aparecen los nuevos casos (Incidencia). Ahora, la pregunta que naturalmente emerge es: ¿por qué ocurre esta enfermedad? ¿Qué factores o exposiciones están asociados a su aparición? Aquí es donde entran en juego las Medidas de Asociación, nuestras herramientas para desentrañar las posibles conexiones entre una exposición y un evento de salud.

Siguiendo con la metáfora de los ecosistemas, medir la prevalencia sería como hacer un censo de todas las especies presentes en un momento dado, mientras que medir la incidencia sería como seguir el nacimiento de nuevas criaturas. Las medidas de asociación, sin embargo, son los prismáticos y las redes que usamos para observar cómo la presencia de ciertos elementos en el ecosistema (digamos, un tipo particular de hongo o la ausencia de un depredador) influye en la proliferación o la disminución de una especie específica. No se trata solo de contar, sino de entender las dinámicas y relaciones subyacentes.

2.1. Riesgo Relativo

Dentro de las medidas de asociación, una de las más fundamentales y reveladoras, especialmente cuando podemos seguir a grupos de personas a lo largo del tiempo, es el Riesgo Relativo (RR). Piensen en él como un comparador de viajes. Si dos grupos de viajeros parten hacia el mismo destino (el evento de

salud), pero uno toma un camino con ciertas condiciones o "exposiciones" (quizás un sendero montañoso), y el otro toma una ruta diferente (el grupo no expuesto), el Riesgo Relativo nos dirá cuántas veces más probable es que el primer grupo llegue a su destino (experimente el evento) en comparación con el segundo.

Formalmente, el Riesgo Relativo es la razón o cociente entre la probabilidad de que el evento de interés ocurra en el grupo expuesto y la probabilidad de que ocurra en el grupo no expuesto. Dado que la probabilidad de experimentar un evento nuevo a lo largo del tiempo es precisamente lo que capturamos con la Incidencia (o la Incidencia Acumulada en periodos cerrados), el RR se calcula dividiendo la incidencia en el grupo expuesto por la incidencia en el grupo no expuesto. Como señalan algunas fuentes, es la razón de dos riesgos absolutos: el riesgo en los expuestos, que es la proporción de expuestos que enferman (a/(a+b)), dividido por el riesgo en los no expuestos, que es la proporción de no expuestos que enferman (c/(c+d)), utilizando la notación de la tabla 2x2 que ya conocemos.

La fórmula, entonces, se presenta de esta manera:

$$RR = \frac{\text{Incidencia en Expuestos}}{\text{Incidencia en No Expuestos}} = \frac{a/(a+b)}{c/(c+d)}$$

Consideremos un ejemplo clásico. Supongamos que realizamos un estudio de cohorte para investigar la asociación entre el hábito de fumar y el desarrollo de cáncer de pulmón. Seguimos a un grupo de fumadores (expuestos) y a un grupo de no fumadores (no expuestos) durante 10 años. Si al cabo de ese tiempo encontramos que la incidencia de cáncer de pulmón en los fumadores es del 10% (es decir, 10 de cada 100 fumadores desarrollaron cáncer de pulmón) y en los no fumadores es del 1% (1 de cada 100 no fumadores), el Riesgo Relativo sería:

$$RR = \frac{0.10}{0.01} = 10$$

¿Cómo interpretamos este 10? Un RR de 10 significa que los fumadores tuvieron 10 veces más riesgo de desarrollar cáncer de pulmón en ese período

de 10 años en comparación con los no fumadores. Este valor nos habla de la *magnitud* o la *fuerza* de la asociación entre fumar y el cáncer de pulmón. Si el RR hubiera sido 1, implicaría que el riesgo de cáncer de pulmón fue el mismo en ambos grupos, sugiriendo que fumar no estuvo asociado al riesgo. Si el RR fuera menor que 1, digamos 0.5, indicaría que los fumadores tuvieron solo la mitad del riesgo que los no fumadores, lo que sugeriría un efecto protector (aunque en este ejemplo es contraintuitivo y solo ilustrativo).

Como bien se menciona, el Riesgo Relativo es la medida de elección por excelencia en los **estudios de cohorte**. Estos estudios, al seguir a los individuos desde la exposición hasta el desenlace, nos permiten calcular directamente la incidencia en cada grupo. Esto establece una clara relación temporal, un pilar fundamental en la búsqueda de causalidad: la exposición debe preceder al efecto. Es como observar una semilla germinar (exposición) y luego ver si crece una planta (evento); no podemos ver la planta y luego preguntarnos si la semilla existió después.

El RR nos ofrece una medida de efecto *relativa*, indicando cuánto aumenta o disminuye el riesgo en un grupo en comparación con otro. Sin embargo, es importante recordar que un RR elevado puede tener implicaciones diferentes dependiendo del riesgo *absoluto* inicial. Un RR de 10 para una enfermedad muy rara (riesgo basal bajo) significa un aumento de riesgo menor en términos de número absoluto de casos que un RR de 2 para una enfermedad muy común (riesgo basal alto).

2.2. Razón de Momios u Odds Ratio (OR)

En los estudios de cohorte, utilizamos el **Riesgo Relativo (RR)** para comparar la incidencia de un evento entre expuestos y no expuestos a lo largo del tiempo. Sin embargo, cuando no podemos hacer un seguimiento prospectivo, por ejemplo, si la enfermedad es poco frecuente o tarda en manifestarse, se recurre al diseño de **casos y controles**. En este enfoque, partimos de quienes ya presentan el evento (casos) y los comparamos con quienes no lo tienen (controles), investigando retrospectivamente su exposición. Para estos estudios, la medida de asociación adecuada es la **Odds Ratio (OR)**, que cuantifica la razón de probabilidades (odds) de haber estado expuesto en casos frente a controles.

Piensen en las *odds* como las que se usan en las apuestas deportivas: la razón entre la probabilidad de que algo suceda y la probabilidad de que no suceda. No es lo mismo que la probabilidad (que es la razón entre los resultados favorables y el total de resultados), pero está intrínsecamente relacionada con ella.

En epidemiología, la *Odds Ratio* compara las *odds* de que los casos estuvieron expuestos con las *odds* de que los controles estuvieron expuestos. Dicho de otro modo, nos pregunta: ¿cuán más probable es encontrar la exposición en las personas enfermas (*casos*) en comparación con las personas sanas (*controles*)?

Tabla 17. *Tabla de contingencia para el cálculo de* OR.

	Enfermos (Evento)	**No Enfermos (Sin Evento)**	**Total**
Expuestos	a	b	a + b
No Expuestos	c	d	c + d
Total	a + c	b + d	N

En un estudio de casos y controles, seleccionamos a los casos (individuos con el evento, columna "Enfermos") y a los controles (individuos sin el evento, columna "No Enfermos"). Luego, investigamos retrospectivamente su historial de exposición (filas "Expuestos" y "No Expuestos"). No conocemos la población total de donde provienen los casos y controles, ni la incidencia de la enfermedad en expuestos y no expuestos en esa población. Por lo tanto, no podemos calcular directamente la Incidencia en Expuestos ni la Incidencia en No Expuestos, lo que nos impide calcular el RR.

Sin embargo, sí podemos calcular las odds de exposición en cada grupo:

- Odds de exposición en los casos: c/a (individuos expuestos entre los casos vs. no expuestos entre los casos)
- Odds de exposición en los controles: d/b (individuos expuestos entre los controles vs. no expuestos entre los controles)

La *Odds Ratio* es simplemente el cociente de estas dos *odds*:

$$OR = \frac{\text{Odds de exposición en casos}}{\text{Odds de exposición en controles}} = \frac{a/c}{b/d} = \frac{a \times d}{b \times c}$$

Observen la fórmula final: (a×d)/(b×c). ¡Es el producto cruzado de la tabla 2x2! Por eso a veces también se le llama Razón de Productos Cruzados.

¿Cómo interpretamos la *Odds Ratio*? La lógica es similar a la del RR:

- **OR = 1:** Sugiere que no hay asociación entre la exposición y el evento. Las *odds* de exposición son similares en casos y controles.
- **OR > 1:** Indica una asociación positiva. Las *odds* de exposición son mayores en los casos que en los controles. La exposición es más probable en las personas que desarrollaron el evento, lo que sugiere que es un factor asociado (posiblemente de riesgo). Un OR de, por ejemplo, 4 significa que las *odds* de exposición en los casos son 4 veces mayores que en los controles.
- **OR < 1:** Indica una asociación negativa o protectora. Las *odds* de exposición son menores en los casos que en los controles. La exposición es menos probable en las personas que desarrollaron el evento, lo que sugiere que es un factor protector. Un OR de, por ejemplo, 0.25 significa que las *odds* de exposición en los casos son una cuarta parte de las *odds* en los controles.

Es fundamental entender que la *Odds Ratio no* es lo mismo que el Riesgo Relativo. La OR compara *odds* de exposición retrospectivamente (en estudios de casos y controles), mientras que el RR compara *riesgos* o incidencias prospectivamente (en estudios de cohorte). Sin embargo, hay una conexión vital: cuando la enfermedad o el evento de interés es **raro** en la población (generalmente se considera raro si afecta a menos del 10% de la población no expuesta), la *Odds Ratio* es una muy buena **aproximación** del Riesgo Relativo. Esto convierte a la OR en una herramienta invaluable en el estudio de enfermedades poco frecuentes, donde un estudio de cohorte sería inviable por requerir un tamaño muestral gigantesco para observar suficientes casos.

La *Odds Ratio* es, por lo tanto, la medida de asociación principal en los **estudios de casos y controles**. Su capacidad para estimar la fuerza de la asociación sin necesidad de conocer las incidencias la hace indispensable en este diseño. Además, la OR es la medida natural que emerge de muchos modelos estadísticos multivariados, como la regresión logística, que nos permiten

controlar el efecto de posibles variables confusoras, añadiendo otra capa de sofisticación a nuestro análisis.

Aunque la interpretación de la OR puede ser un poco menos intuitiva que la del RR, su utilidad en el diseño de casos y controles y en el análisis multivariado la convierten en una piedra angular de la epidemiología analítica. Nos permite investigar asociaciones donde otros diseños serían prohibitivamente costosos o impracticables.

Hemos descifrado dos de las medidas de asociación más importantes, el Riesgo Relativo y la *Odds Ratio*, cada una con su ámbito de aplicación ideal, como diferentes tipos de lentes para observar nuestro ecosistema de salud. Entender cómo calcularlas e interpretarlas nos abre la puerta a cuantificar la fuerza de las relaciones entre exposiciones y eventos. Pero la epidemiología no se detiene en identificar asociaciones; también busca comprender cuánto impacto tienen estas asociaciones en la carga de enfermedad de una población y cuánto podríamos ganar si intervenimos en los factores de riesgo. Esto nos lleva directamente a las cruciales Medidas de Impacto Potencial, el siguiente paso en nuestra exploración.

Conclusiones

1. **Propósito:** Las medidas de asociación cuantifican la fuerza y dirección de la relación entre una exposición de interés y un evento de salud (enfermedad).

2. **Riesgo Relativo (RR):**

- Compara la incidencia (riesgo) del evento en expuestos vs. no expuestos (RR = IE / INE).

- Es la medida de elección para estudios de cohorte y ensayos clínicos donde se puede calcular la incidencia directamente.

- Interpretación: RR=1 (sin asociación), RR>1 (factor de riesgo), RR<1 (factor protector). Indica cuántas veces es más (o menos) probable que ocurra el evento en los expuestos comparado con los no expuestos.

3. **Odds Ratio (OR) o Razón de Momios:**

- Compara las odds de exposición en los casos (enfermos) vs. las odds de exposición en los controles (no enfermos) [OR = (a/c) / (b/d) = ad/bc].
- Es la medida de elección para estudios de casos y controles, donde no se calcula la incidencia directamente.
- Interpretación: OR=1 (sin asociación), OR>1 (mayor odds de exposición en casos, sugiere factor de riesgo), OR<1 (menor odds de exposición en casos, sugiere factor protector).

4. **Relación OR y RR:** La Odds Ratio es una buena estimación del Riesgo Relativo cuando la enfermedad estudiada es rara en la población (generalmente <10%).

5. **Contexto del Estudio:** La elección entre RR y OR depende fundamentalmente del diseño del estudio epidemiológico utilizado (cohorte para RR, casos y controles para OR).

CAPÍTULO 14

MEDIDAS DE IMPACTO

Impact Measures

Resumen

Este capítulo introduce las medidas de impacto epidemiológico, que evalúan la relevancia en salud pública de una asociación entre exposición y enfermedad, estimando cuánta enfermedad es atribuible a dicha exposición. Se describen dos medidas clave: el Riesgo Atribuible (RA) o Diferencia de Riesgos, y la Fracción Atribuible en la Población (FAP). El RA se calcula como la diferencia entre la incidencia en expuestos y la incidencia en no expuestos ($I_e - I_o$), representando el exceso absoluto de riesgo debido a la exposición entre quienes están expuestos; es útil para comunicar el beneficio individual de eliminar la exposición. La FAP, o Fracción Etiológica Poblacional, mide la proporción del total de casos en la población (expuestos y no expuestos) que se puede atribuir a la exposición. Se calcula a partir de la incidencia total y la incidencia en no expuestos [$(I_t - I_o) / It$], o usando el Riesgo Relativo (RR) y la prevalencia de la exposición en la población [$P_p(RR-1) / (P_p(RR-1) + 1)$]. La FAP es crucial para la salud pública, ya que indica el impacto potencial de eliminar la exposición a nivel comunitario y depende tanto de la fuerza de la asociación (RR) como de la frecuencia de la exposición (P_p).

Palabras clave: Medidas de impacto, riesgo atribuible, diferencia de riesgos, fracción atribuible poblacional, fracción etiológica, salud pública.

Abstract

This chapter introduces epidemiological measures of impact, which assess the public health relevance of an exposure-disease association by estimating how much disease is attributable to that exposure. Two key measures are described: Attributable Risk (AR) or Risk Difference, and Population Attributable Fraction (PAF). AR is calculated as the difference between the incidence in the exposed and the incidence in the unexposed ($I_e - I_o$), representing the absolute excess risk due to the exposure among those exposed; it is useful for communicating the in-

dividual benefit of removing the exposure. PAF, or Population Etiologic Fraction, *measures the proportion of all cases in the total population (exposed and unexposed) that can be attributed to the exposure*. It *is calculated from the total incidence and the incidence in the unexposed* [(It - Io) / It], *or using the Relative Risk* (RR) *and the prevalence of exposure in the population* [Pp(RR-1) / (Pp(RR-1) + 1)]. PAF *is crucial for public health as it indicates the potential impact of eliminating the exposure at the community level and depends on both the strength of the association* (RR) *and the frequency of the exposure* (Pp).

Keywords: *Measures of impact, attributable risk, risk difference, population attributable fraction, etiologic fraction, public health.*

Introducción

Hemos explorado cómo medir cuánta enfermedad hay (Prevalencia e Incidencia) y cómo cuantificar la fuerza con la que un factor se asocia a esa enfermedad (Riesgo Relativo y *Odds Ratio*). Pero la epidemiología, al fin y al cabo, es una ciencia aplicada orientada a mejorar la salud de las poblaciones. De poco nos serviría saber que una exposición se asocia fuertemente a una enfermedad si no podemos estimar qué ganancia en salud pública o individual podríamos obtener al modificar esa exposición. Aquí es donde entran en juego las **Medidas de Impacto**

Si las medidas de asociación son como determinar la inclinación de una pendiente (cuán empinada es la relación entre exposición y evento), las medidas de impacto nos dicen qué altura podríamos ganar o perder al subir o bajar por esa pendiente. Nos trasladan del ámbito de la asociación a la arena de la acción y el potencial preventivo. ¿Cuánto de la enfermedad que observamos es realmente "atribuible" a esa exposición que hemos identificado? ¿Cuántos casos podríamos evitar si elimin——áramos o redujéramos esa exposición?

Dentro de este conjunto de medidas de impacto, hay dos que son particularmente relevantes: el Riesgo Atribuible, centrado en los individuos expuestos, y la Fracción Atribuible en la Población, que mira el panorama completo de la comunidad.

3.1. Riesgo Atribuible (RA) o Diferencia de Riesgos

El Riesgo Atribuible es quizás la medida de impacto más directa. Nos dice cuánto del riesgo de desarrollar el evento en el grupo expuesto es *exclusiva-*

mente achacable a la exposición misma. Es la diferencia absoluta en el riesgo entre los que estuvieron expuestos y los que no. Volviendo a nuestra metáfora del viaje, si el grupo que tomó la ruta montañosa tuvo una tasa de percances del 10%, y el grupo que tomó la ruta llana tuvo una tasa de percances del 2%, la diferencia (10% - 2% = 8%) representa el exceso de percances que podemos "atribuir" al hecho de haber elegido la ruta montañosa, suponiendo que todo lo demás fuera igual.

Se calcula simplemente restando la incidencia del evento en el grupo no expuesto (I_o) de la incidencia en el grupo expuesto (I_e):

$$RA = I_e - I_o$$

Usando nuestra tabla 2x2, esto se traduce en:

$$RA = \frac{a}{a+b} - \frac{c}{c+d}$$

Retomando el ejemplo del tabaquismo y el cáncer de pulmón, donde la incidencia en fumadores era del 10% (0.10) y en no fumadores del 1% (0.01), el Riesgo Atribuible sería:

$$RA = 0.10 - 0.01 = 0.09$$

Esto significa que, de cada 100 fumadores, 9 casos de cáncer de pulmón en un período de 10 años son atribuibles al hábito de fumar. Es el riesgo adicional que enfrentan los fumadores debido a su exposición. Para un médico clínico, este valor es sumamente útil al hablar con un paciente fumador, ya que cuantifica el exceso de riesgo que podría evitar si deja de fumar. Nos da una idea del beneficio absoluto de eliminar la exposición en un individuo expuesto. Es una medida de la carga de enfermedad entre los expuestos que podría ser prevenida si la exposición no existiera.

El Riesgo Atribuible es una medida con unidades, las mismas unidades que la incidencia (por ejemplo, casos por 100 personas por año, o una proporción como en nuestro ejemplo). A diferencia del RR o la OR, que son cocientes y no tienen unidades, el RA nos da una magnitud absoluta del exceso de riesgo.

3.2. Fracción Atribuible en la Población (FAP)

Mientras que el Riesgo Atribuible nos habla del impacto de la exposición en los individuos expuestos, la Fracción Atribuible en la Población (FAP), también conocida como Fracción Etiológica Poblacional, eleva la vista para considerar el impacto de esa exposición en el *conjunto de la población*, incluyendo tanto a expuestos como a no expuestos. Esta medida responde a la pregunta: ¿Qué proporción del total de casos de la enfermedad que ocurren en toda la población se debe a esta exposición?

Piensen en la FAP como si tuviéramos un gran lago (la población total) donde nadan peces (los casos de enfermedad). Si una sustancia contaminante (la exposición) entra en el lago y causa la enfermedad en algunos peces, la FAP nos diría qué porcentaje de todos los peces enfermos en el lago se enfermaron *a causa* de ese contaminante. Aquí, no solo importa cuán tóxico sea el contaminante (la fuerza de la asociación, RR), sino también cuán extendido esté en el lago (la prevalencia de la exposición en la población). Una sustancia muy tóxica pero presente solo en una pequeña charca tendrá un impacto poblacional menor que una sustancia moderadamente tóxica pero dispersa por todo el lago.

La FAP se calcula como la proporción de la incidencia total en la población (I_t) que es atribuible a la exposición. Se puede expresar de varias maneras, pero una forma común es:

$$FAP = \frac{I_t - I_o}{I_t}$$

Donde I_t es la incidencia del evento en la población total, e I_o es la incidencia en los no expuestos. También se puede calcular utilizando el Riesgo Relativo (RR) y la prevalencia de la exposición en la población (P_p):

$$FAP = \frac{P_p(RR - 1)}{1 + P_p(RR - 1)}$$

Imaginemos que, en nuestro ejemplo de tabaquismo y cáncer de pulmón, la prevalencia de fumadores en la población total fuera del 30% (0.30), y el RR fuera 10. La Fracción Atribuible en la Población sería:

$$FAP = \frac{0.30(10-1)}{1+0.30(10-1)} = \frac{0.30\times 9}{1+0.30\times 9} = \frac{2.7}{1+2.7} = \frac{2.7}{3.7} \approx 0.73$$

Esto significa que aproximadamente el 73% de todos los casos de cáncer de pulmón que ocurren en esta población son atribuibles al tabaquismo. Este número es enormemente relevante para quienes diseñan políticas de salud pública. Nos dice qué proporción de la carga total de cáncer de pulmón podríamos hipotéticamente prevenir si lográramos eliminar completamente el hábito de fumar en la población. Es una medida del impacto potencial de una intervención a nivel comunitario.

En resumen, mientras que el Riesgo Atribuible es crucial para comprender el exceso de riesgo en los individuos expuestos y el beneficio potencial de la cesación de la exposición a nivel individual, la Fracción Atribuible en la Población es vital para estimar el impacto de una exposición en la carga de enfermedad de toda la comunidad, guiando las prioridades de salud pública y la asignación de recursos.

Hemos completado así un recorrido fundamental por las medidas epidemiológicas clave: cuantificamos la ocurrencia de enfermedades (Prevalencia e Incidencia), medimos la fuerza de la asociación entre exposiciones y eventos (Riesgo Relativo y *Odds Ratio*), y ahora estimamos el impacto potencial de esas asociaciones en la salud individual y poblacional (Riesgo Atribuible y Fracción Atribuible en la Población). Armados con estas herramientas conceptuales y de cálculo, estamos listos para dar el siguiente gran paso en la epidemiología: entender los diferentes **Diseños de Estudios Epidemiológicos**, las estructuras metodológicas que nos permiten recopilar los datos necesarios para calcular todas estas medidas y, lo más importante, generar conocimiento confiable sobre los determinantes de la salud y la enfermedad en las poblaciones.

Conclusiones

1. **Propósito:** Las medidas de impacto van más allá de la asociación y cuantifican la carga de enfermedad en los expuestos o en la población total que se debe específicamente a una exposición.

2. **Riesgo Atribuible (RA) o Diferencia de Riesgos:**

- Se calcula como: RA = Incidencia en expuestos (I_e) - Incidencia en no expuestos (I_o).
- Representa el número absoluto de casos *adicionales* por unidad de población/tiempo que ocurren en el grupo expuesto debido únicamente a la exposición.
- Es útil para estimar el beneficio individual o para los expuestos si se eliminara la exposición.

3. **Fracción Atribuible en la Población (FAP) o Fracción Etiológica Poblacional:**

- Se calcula como: FAP = (Incidencia total (I_t) - Incidencia en no expuestos (I_o)) / Incidencia total (I_t).
- Alternativamente, usando RR y Prevalencia de exposición (P_p): FAP = $[P_p(RR-1)] / [P_p(RR-1) + 1]$.
- Representa la *proporción* de todos los casos de la enfermedad en la *población total* que podría prevenirse si se eliminara la exposición.
- Depende tanto de la fuerza de la asociación (RR) como de la frecuencia de la exposición en la población (P_p).

4. **Relevancia:** El RA es más relevante para la perspectiva clínica individual (exceso de riesgo para el expuesto), mientras que la FAP es fundamental para la salud pública (impacto en la carga poblacional y potencial de prevención comunitaria).

5. **Aplicación:** Estas medidas ayudan a priorizar intervenciones, evaluar programas de prevención y comunicar el impacto real de los factores de riesgo en la salud de la población.

CAPÍTULO 15

DISEÑOS DE ESTUDIOS EPIDEMIOLÓGICOS

Epidemiological Study Designs

Resumen

Este capítulo aborda la arquitectura fundamental de la investigación epidemiológica: los diseños de estudio. Se clasifica a los estudios en dos grandes grupos: Observacionales (donde el investigador no interviene, sino que observa exposiciones y desenlaces naturales) e Intervencionales o Experimentales (donde el investigador asigna la exposición, como en los ECA). Dentro de los observacionales, se detallan los descriptivos (Reporte de Caso, Serie de Casos, Estudios Transversales descriptivos, Estudios Ecológicos) y los analíticos (Estudios Transversales analíticos, Estudios de Casos y Controles, Estudios de Cohorte). Se describe la lógica, selección de participantes, recolección de datos, análisis principal (Prevalencia, POR/PR, OR, RR), ventajas, desventajas y sesgos potenciales de cada diseño, incluyendo ejemplos detallados para reporte de caso, serie de casos, transversal, casos y controles, cohorte y ecológico. Se discute la jerarquía de la evidencia para inferir causalidad. Adicionalmente, se introducen métodos para la síntesis de evidencia (Revisión Narrativa, Análisis Bibliométrico, Revisión de Alcance, Revisión Sistemática y Metanálisis, con ejemplos y herramientas de evaluación como PRISMA y RoB). Finalmente, se abordan conceptos de tamizaje, diagnóstico, evaluación del rendimiento de pruebas (Sensibilidad, Especificidad, VPP, VPN, LR, Curva ROC/AUC, Kappa) y el Análisis de Sobrevida (censura, Kaplan-Meier, Log-Rank, Regresión de Cox/HR).

Palabras clave: Diseños epidemiológicos; estudios observacionales; estudios experimentales; estudios de cohorte; estudios de casos y controles; ensayo clínico aleatorizado.

Abstract

This chapter addresses the fundamental architecture of epidemiological research: study designs. Studies are classified into two main groups: Observational (where the researcher does not intervene but observes natural exposures and

outcomes) and Interventional or Experimental (where the researcher assigns the exposure, as in Randomized Controlled Trials - RCTs). Within observational studies, descriptive designs (Case Report, Case Series, descriptive Cross-Sectional Studies, Ecological Studies) and analytical designs (analytical Cross-Sectional Studies, Case-Control Studies, Cohort Studies) are detailed. The logic, participant selection, data collection, main analysis (Prevalence, POR/PR, OR, RR), advantages, disadvantages, and potential biases of each design are described, including detailed examples for case reports, case series, cross-sectional, case-control, cohort, and ecological studies. The hierarchy of evidence for inferring causality is discussed. Additionally, methods for evidence synthesis (Narrative Review, Bibliometric Analysis, Scoping Review, Systematic Review, and Meta-Analysis, with examples and evaluation tools like PRISMA and RoB) are introduced. Finally, concepts of screening, diagnosis, evaluation of test performance (Sensitivity, Specificity, PPV, NPV, LR, ROC Curve/AUC, Kappa), and Survival Analysis (censoring, Kaplan-Meier, Log-Rank, Cox Regression/HR) are covered.

Keywords: *Epidemiological designs; observational studies; experimental studies; cohort studies; case-control studies; randomized controlled trial.*

Introducción

Excelente. Hemos aprendido a cuantificar la presencia de la enfermedad, a medir la fuerza con la que se asocia a ciertos factores y a estimar el impacto que esas asociaciones tienen. Es como si hubiéramos aprendido a leer el mapa del territorio de la salud y la enfermedad, a identificar los picos (alta incidencia), los valles (baja prevalencia) y las rutas que conectan diferentes puntos (las asociaciones). Pero, ¿cómo se creó ese mapa? ¿Cómo recogemos la información de manera sistemática y fiable para que nuestras medidas sean precisas y nuestras conclusiones sólidas? Aquí es donde la brújula y el plan de expedición entran en juego: es la hora de adentrarnos en los Diseños de Estudios Epidemiológicos.

Piensen en el diseño de un estudio como la arquitectura fundamental de una investigación. No es solo un detalle técnico, sino la hoja de ruta que define cómo vamos a recolectar nuestros datos, qué preguntas podremos responder con ellos y con qué grado de confianza. Un diseño bien concebido es la base de la validez de nuestros hallazgos; un diseño pobre, en cambio, puede llevarnos

por caminos equivocados, arrojando resultados confusos o directamente erróneos, lo que, como mencionan los expertos, significa una lamentable pérdida de tiempo, dinero y esfuerzo, recursos preciosos en la investigación médica.

Existen diversas maneras de clasificar los estudios, pero una de las dicotomías más fundamentales y didácticas es la que los divide en dos grandes mundos: los Estudios Intervencionales (o Experimentales) y los Estudios Observacionales. Imaginen que queremos saber si un nuevo fertilizante aumenta el crecimiento de una planta. Un estudio intervencional sería como que nosotros, los investigadores, aplicáramos activamente el fertilizante a un grupo de plantas y no a otro, controlando cuidadosamente las condiciones y luego comparando el crecimiento. Estamos interviniendo, manipulando una variable para ver su efecto. Un estudio observacional, por el contrario, sería como simplemente encontrar dos campos de plantas, uno donde casualmente se ha usado ese fertilizante y otro donde no, y observar las diferencias de crecimiento, sin que nosotros hayamos decidido quién recibió qué. Simplemente observamos lo que la naturaleza o las circunstancias ya dispusieron.

Para visualizar esta clasificación y los diferentes tipos de estudios que se desprenden de ella, pueden referirse a la tabla 18, donde se presenta un diagrama de flujo que resume esta estructura y nos da un primer pantallazo de los caminos que podemos tomar al diseñar una investigación.

Tabla 18. *Clasificación de los diseños de estudio según categoría y subtipos*.

Diseño de Estudio	Subcategoría	Ejemplos
Intervencional (exposición asignada)	Asignación aleatoria individual	ECA paralelo; ECA cruzado (crossover); ECA factorial; ECA de superioridad; ECA de equivalencia; ECA de no inferioridad
	Asignación aleatoria grupal	Ensayo controlado aleatorizado por cúmulos (Cluster ECA)
	Exposición no asignada aleatoriamente	Ensayo controlado no aleatorizado

Diseño de Estudio	Subcategoría	Ejemplos
Observacional (exposición **no** asignada)	No comparativo (participantes individuales)	Estudio de caso; Series de casos
	Comparativo (participantes individuales)	Estudio transversal (cross-sectional); Estudio de casos y controles; Estudio de cohorte
	Grupal	Estudio ecológico
Investigación cualitativa	Independiente o incorporada a cuantitativo	Investigación cualitativa (realizada sola o como parte de un estudio cuantitativo)

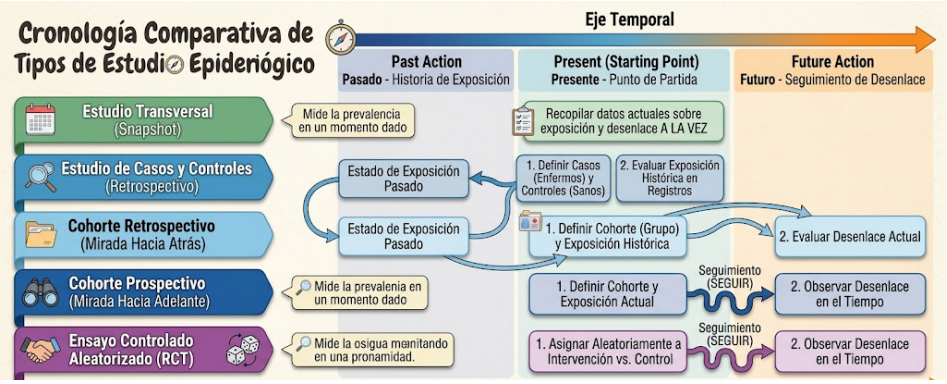
Los Estudios de Intervención, con los ECA a la cabeza, son el estándar de oro cuando queremos determinar si una intervención específica (un nuevo fármaco, una terapia, una vacuna) causa un determinado resultado. Su gran fortaleza reside en la capacidad del investigador para controlar la exposición (generalmente asignando aleatoriamente la intervención), minimizando así la influencia de otros factores y otorgando una alta validez interna a los resultados. Son ideales para poner a prueba hipótesis causales con rigor, aunque a menudo son costosos, complejos y no siempre éticamente factibles (no podemos, por ejemplo, asignar aleatoriamente a personas a exposiciones dañinas como el tabaco o la contaminación).

Por otro lado, los Estudios Observacionales son nuestra ventana a lo que ocurre en el mundo real, tal como sucede naturalmente. Aquí, el investigador no asigna la exposición, sino que observa y registra lo que ocurre con grupos de personas que ya están expuestas o no expuestas a ciertos factores, o que ya han desarrollado un determinado evento de salud. Este enfoque es invaluable para explorar la historia natural de las enfermedades, identificar posibles factores de riesgo (como el tabaquismo o ciertos hábitos dietéticos), o estudiar exposiciones que simplemente no pueden ser manipuladas experimentalmente por razones éticas o prácticas. Si bien tienen una menor capacidad para establecer causalidad directa en comparación con los ECA, debido a que el investigador tiene menos control sobre los factores de confusión, su flexibilidad y aplicabilidad a una amplia gama de preguntas los

hacen indispensables en epidemiología. Identifican patrones y sugieren asociaciones que luego pueden ser investigadas con diseños más robustos.

La forma en que se recolectan los datos en relación con el paso del tiempo es otra característica distintiva de los diferentes diseños. Algunos miran hacia el futuro (prospectivos), otros hacia el pasado (retrospectivos) y algunos toman una "foto" en un momento dado (transversales). La Figura 57 ilustra esta dimensión temporal, mostrando en una línea de tiempo cómo se sitúan los puntos de recolección de datos de exposición y desenlace para varios diseños comunes, lo que ayuda a comprender su lógica subyacente.

Figura 57. *Línea de tiempo de los diseños de estudio*. Elaborado con GeminiPRO.



La elección del diseño de estudio adecuado no es una cuestión trivial, sino una decisión crítica que depende íntimamente de la pregunta de investigación que nos planteamos, de los recursos de los que disponemos, de las consideraciones éticas y del tipo de evidencia que necesitamos generar. ¿Buscamos simplemente describir un fenómeno? ¿Explorar posibles asociaciones? ¿Establecer una relación de causalidad? Cada objetivo nos guiará hacia diseños específicos. Es como seleccionar la herramienta correcta para la tarea en cuestión, ya sea un martillo para un clavo o un bisturí para una cirugía. También es importante reconocer que los diferentes diseños contribuyen de manera distinta a la evidencia que podemos acumular sobre la causalidad, existiendo una suerte de "jerarquía de evidencia" donde unos diseños proporcionan un soporte más fuerte que otros para establecer relaciones causa-efecto.

La figura 58 ilustra la "jerarquía de evidencia" en epidemiología, presentándola como una montaña que debemos escalar para acercarnos a la cima de la **Causalidad**. En la base de la montaña, en terrenos menos elevados y quizás un poco pantanosos, encontramos los **Reportes de Caso** y las **Series de Casos**, así como los **Estudios Ecológicos** y los **Estudios Transversales**. Estos diseños son valiosos para generar hipótesis y describir patrones, son como los primeros avistamientos de un posible camino en la base de la montaña. Nos pueden sugerir una **Asociación débil**, un indicio de que algo interesante podría estar sucediendo.

Figura 58. *Nivel de evidencia de los diseños epidemiológicos*. Elaborado con GeminiPRO.



A medida que ascendemos por la ladera, ganando altura y solidez en nuestro camino, encontramos estudios que nos permiten explorar asociaciones con mayor rigor. Los **Estudios de Casos y Controles** y los **Estudios de Cohorte** se sitúan en puntos intermedios de la montaña. Estos diseños nos permiten cuantificar la fuerza de la asociación (RR u OR) y, en el caso de las cohortes prospectivas, establecer la secuencia temporal, un requisito indispensable para la causalidad. Son como tramos del sendero que nos llevan más arriba, proporcionando evidencia de una **Asociación más fuerte**.

El camino se vuelve más claro y directo cuando llegamos a la altura de los **ECA**. Estos estudios, al controlar la asignación de la exposición mediante la aleatorización, minimizan el riesgo de confusión y sesgo, ofreciendo una evidencia mucho más robusta sobre la causalidad. Son como un puente sólido que nos permite cruzar un abismo en la montaña, acercándonos significativamente a la cima.

Finalmente, en la cumbre de la montaña de la evidencia, encontramos las **Revisiones Sistemáticas** y los **Metaanálisis.** Estos no son diseños de estudio primarios, sino síntesis rigurosas de múltiples estudios (idealmente ECA bien realizados) sobre una misma pregunta. Al combinar la evidencia de diversas fuentes, nos ofrecen la visión más clara y completa posible, la panorámica desde la cima que nos permite tener la mayor confianza al hablar de **Causalidad**.

A lo largo de los próximos apartados, desglosaremos los principales tipos de estudios epidemiológicos, explorando la lógica detrás de cada uno, cómo se llevan a cabo, qué tipo de información nos proporcionan y cuáles son sus fortalezas y limitaciones. Comenzaremos nuestro viaje detallado explorando los Estudios Descriptivos, que sientan las bases al pintar un cuadro de la distribución de la enfermedad, diciéndonos quién, dónde y cuándo se ve afectado, antes de pasar a los diseños que buscan responder al porqué.

4.1. Estudios Descriptivos

Comenzamos nuestra exploración de los diseños de estudio con los cimientos de la pirámide de evidencia, los que nos permiten dar los primeros pasos en la descripción de fenómenos de salud: los **Estudios Descriptivos**. Estos estudios no buscan establecer relaciones de causa y efecto, sino simplemente pintar un cuadro detallado de cómo se distribuye una enfermedad o evento de salud en una población, respondiendo a las preguntas fundamentales de *quién*, *dónde* y *cuándo*. Son como los primeros bocetos de un artista que intenta capturar la esencia de un paisaje antes de añadir detalles o interpretaciones.

Dentro de los estudios descriptivos, encontramos la forma más elemental y a menudo el punto de partida de nuevas hipótesis: el **Reporte de Caso**.

4.1.1 *Reporte de Caso*

Imagina que estás en tu rotación clínica y te encuentras con un paciente cuya presentación es inusual, un síntoma que no encaja en los patrones comunes, una complicación inesperada de un tratamiento, o una manifestación atípica de una enfermedad conocida, especialmente en un contexto particular, como una inmunosupresión severa. Estos encuentros son como descubrir una especie rara y fascinante en nuestra selva de la salud, algo que merece ser documentado y compartido con la comunidad científica porque podría arrojar luz sobre aspectos poco conocidos de una patología o alertar a otros clínicos sobre presentaciones inusuales.

Un Reporte de Caso es precisamente eso: una descripción detallada y sistemática de un solo paciente que presenta una condición o un desenlace de interés. Su principal utilidad clínica radica en su capacidad para:

- **Documentar presentaciones raras o inusuales** de enfermedades.
- **Identificar posibles efectos adversos** inesperados de tratamientos.
- **Describir nuevas enfermedades o síndromes** en sus etapas iniciales.
- **Generar hipótesis** para futuras investigaciones.
- **Servir como herramienta de aprendizaje** para otros profesionales de la salud.

Aunque no permiten establecer causalidad ni generalizar hallazgos a poblaciones más amplias, los reportes de caso son la chispa que a menudo enciende la mecha de investigaciones más robustas. Son el primer grito de alerta ante algo novedoso o inesperado.

Para asegurar que estos reportes sean claros, completos y útiles, existen guías estandarizadas, como la lista de verificación **CARE (CAse Reporting Essentials)**, que orienta a los autores sobre la información clave que debe incluirse. Siguiendo esta estructura, un reporte de caso típico se organiza de la siguiente manera, similar al ejemplo que tenemos sobre sífilis rectal en inmunosupresión severa:

- **Resumen:** Un resumen conciso que presenta el contexto del caso, la presentación clínica principal, los hallazgos clave, el diagnóstico, el trata-

miento y la principal conclusión o "mensaje para llevar a casa". Es como la sinopsis que te engancha a leer el resto de la historia clínica.

- **Introducción:** Establece el contexto de la enfermedad o condición, su relevancia, lo que se sabe al respecto y por qué este caso particular es único o importante. Aquí se justifica la necesidad de reportar el caso. En el ejemplo de sífilis rectal, se destaca la rareza de la presentación rectal, especialmente en pacientes inmunocomprometidos, y la dificultad diagnóstica.

- **Presentación del Caso:** Esta es la narrativa central, el corazón del reporte. Se describe detalladamente la historia del paciente, incluyendo datos demográficos, antecedentes médicos relevantes, la cronología de los síntomas, los hallazgos del examen físico, los resultados de las pruebas diagnósticas (laboratorio, imagenología, histopatología) y el manejo recibido. Es crucial ser preciso y objetivo.

Aquí es donde la incorporación de imágenes se vuelve fundamental para ilustrar los hallazgos clave. Al presentar imágenes radiológicas (como tomografías o resonancias) o histológicas (biopsias), es vital seguir ciertas pautas para que sean informativas:

- **Imágenes Radiológicas:** Deben mostrar las vistas relevantes (axial, sagital, coronal), indicar el tipo de estudio (ej. CT contrastada), y si es posible, incluir flechas o marcadores que señalen las áreas de interés (ej. engrosamiento de la pared rectal, "signo del blanco"). La leyenda de la figura debe describir claramente lo que se observa en cada panel (A, B, C, etc.).

- **Imágenes Histológicas:** Es esencial incluir la magnificación utilizada (ej. ×40, ×200, ×400) para que el lector comprenda el nivel de detalle que se está observando. Las flechas pueden señalar características celulares o estructurales específicas (ej. infiltrado plasmocitario, capilares). La leyenda debe describir los hallazgos microscópicos relevantes en cada panel y, si aplica, el tipo de tinción utilizada (ej. Hematoxilina y Eosina, inmunohistoquímica). Idealmente, se incluye una barra de escala en la imagen para dar una referencia visual del tamaño real de las estructuras observadas (ej. 100 μm, 50 μm).

- **Discusión:** En esta sección, se interpreta el caso en el contexto de la literatura existente. Se compara la presentación del paciente con casos previamente reportados, se discute la relevancia de los hallazgos clave, se explican las posibles razones detrás de la presentación atípica (como la inmunosupresión en nuestro ejemplo) y se reflexiona sobre las implicaciones clínicas y fisiopatológicas. También se pueden mencionar las dificultades diagnósticas o terapéuticas encontradas.
- **Conclusión:** Se resume brevemente el mensaje principal del reporte y se destacan las lecciones aprendidas o las recomendaciones para la práctica clínica.

Ventajas y Limitaciones

Las principales **ventajas** de los reportes de caso son su simplicidad, bajo costo y rapidez para documentar observaciones novedosas. Son accesibles y pueden ser publicados relativamente rápido, alertando a la comunidad médica sobre nuevos fenómenos.

Sin embargo, sus **limitaciones** son significativas: al basarse en un solo individuo, no permiten generalizar los hallazgos a una población. Son susceptibles a sesgos de selección (el "caso" reportado puede no ser representativo) y no pueden establecer relaciones causales definitivas. Un hallazgo en un reporte de caso es una "señal", no una "prueba" de causalidad.

A pesar de sus limitaciones, los reportes de caso son una parte valiosa del panorama de la investigación médica, sirviendo como un "sistema de vigilancia" informal para detectar lo inusual y estimular futuras investigaciones. A menudo, cuando varios reportes de caso describen un fenómeno similar, se justifica la realización de estudios más estructurados, como las **Series de Casos**, que agrupan y analizan las características de múltiples pacientes con una condición similar, lo que nos permite dar un paso más allá en la descripción de patrones y la generación de hipótesis.

4.1.2. *Series de Casos*

Una Serie de Casos es una colección de reportes de caso que describen a varios pacientes que comparten una condición, una exposición, o un desenlace similar. No se trata solo de juntar reportes individuales, sino de analizar y pre-

sentar las características comunes y las variaciones entre ellos. Esto nos permite ir un paso más allá en la descripción: no solo decimos "esto ocurrió en un paciente", sino "esto ocurrió en *varios* pacientes con estas características".

La utilidad clínica de las series de casos es considerable, ya que nos permiten:

- **Describir el espectro clínico** de una enfermedad, incluyendo sus diferentes manifestaciones, severidad y evolución.
- **Identificar posibles factores de riesgo** o exposiciones comunes entre los casos.
- **Documentar la experiencia con un tratamiento** o intervención particular en un grupo de pacientes.
- **Generar hipótesis más sólidas** que las de un solo reporte de caso, aunque todavía sin la capacidad de establecer causalidad.
- **Alertar a la comunidad médica** sobre la aparición de nuevos problemas de salud o presentaciones inusuales que podrían estar pasando desapercibidas.

Piensen en una serie de casos como si varios exploradores, cada uno en una parte diferente de la selva, encontraran el mismo tipo de planta rara. Al juntar sus observaciones, pueden empezar a notar que la planta siempre crece en un tipo particular de suelo o cerca de un cuerpo de agua, información valiosa que un solo hallazgo no habría revelado.

Para ilustrar la estructura y el contenido de una serie de casos, tomemos como ejemplo el estudio sobre Actinomicosis (https://www.mdpi.com/1648-9144/61/2/256) que imita malignidades en múltiples sitios anatómicos. Este trabajo, siguiendo la guía CARE adaptada para series, nos muestra cómo se presenta la información de varios pacientes para destacar un tema común: la capacidad de *Actinomyces* spp. de simular procesos cancerosos, un desafío diagnóstico significativo.

La estructura de esta serie de casos, como la de muchas otras, sigue un formato similar al del reporte de caso, pero adaptado para describir a múltiples individuos:

- **Resumen:** Resume el contexto de la enfermedad (*Actinomyces* spp. y su mimetismo con malignidades), los objetivos de la serie (presentar desafíos diagnósticos y manifestaciones diversas), una breve descripción de los materiales y métodos (revisión de casos confirmados), los resultados clave (la dificultad diagnóstica y la confirmación histopatológica), y las conclusiones principales (necesidad de alta sospecha y examen histopatológico).

- **Introducción:** Proporciona un trasfondo sobre *Actinomyces* spp., la bacteria causante, su naturaleza crónica y su capacidad de imitar otras condiciones, particularmente malignidades. Se discute la epidemiología, los factores de riesgo y la variabilidad de la presentación clínica, destacando la dificultad del diagnóstico oportuno y la importancia de la histopatología. Se justifica por qué esta serie de casos es relevante para arrojar luz sobre este desafío.

- **Materiales y Métodos:** Describe cómo se seleccionaron los pacientes para la serie (casos diagnosticados con *Actinomyces* spp. en un período específico), la fuente de los datos (revisión de historias clínicas), las variables que se recopilaron (edad, sexo, antecedentes, localización de la infección, síntomas, duración, procedimientos diagnósticos, tratamiento y resultados). Se menciona la aprobación ética y el consentimiento informado, así como el uso de la lista de verificación CARE para su elaboración. Se detallan los diferentes exámenes utilizados en cada caso, como cultivos, imágenes (resonancia magnética, PET-CT, CT abdominal, entero-RM) y, crucialmente, la histopatología.

- **Resultados:** Presenta los hallazgos de la serie. A menudo, se inicia con un resumen general de las características de los pacientes. Luego, se describe cada caso individualmente, como si fueran pequeños reportes de caso dentro de la serie. Para cada paciente, se detalla su historia clínica, los hallazgos del examen físico, los resultados de las pruebas de imagenología y laboratorio, y especialmente, los hallazgos histopatológicos que llevaron al diagnóstico de *Actinomyces* spp. Se pueden incluir imágenes relevantes (radiológicas o histológicas) para ilustrar los puntos clave, con las descripciones detalladas y las escalas o magnificaciones apropiadas, como vimos con las imágenes histológicas de la sífilis rectal. En el

ejemplo de Actinomicosis, se describen tres casos con presentaciones en la amígdala, la base de la lengua y el antro gástrico, resaltando cómo en todos ellos se sospechó inicialmente malignidad y la histopatología fue clave para el diagnóstico correcto.

- **Discusión:** Se analizan los hallazgos de la serie en conjunto y se comparan con la literatura existente sobre *Actinomyces* spp. Se discuten las similitudes y diferencias entre los casos presentados, se profundiza en el desafío diagnóstico que representa el mimetismo con malignidades y otras infecciones, y se reflexiona sobre las implicaciones clínicas de estos hallazgos. Se pueden explorar los posibles mecanismos detrás de las presentaciones atípicas y la importancia de un enfoque multidisciplinario.
- **Conclusiones:** Se resumen los puntos más importantes de la serie de casos, enfatizando las lecciones aprendidas y las recomendaciones para la práctica clínica. En el ejemplo de Actinomicosis, se concluye que la enfermedad sigue siendo un desafío diagnóstico, subraya la necesidad de una alta sospecha y un examen histopatológico exhaustivo, y sugiere futuras investigaciones para mejorar el diagnóstico y tratamiento.

Ventajas y Limitaciones:

Las **ventajas** de las series de casos, en comparación con un reporte de caso único, incluyen una mayor capacidad para describir el rango de manifestaciones clínicas, identificar patrones y generar hipótesis más robustas. Son relativamente fáciles y económicas de realizar, y pueden ser particularmente útiles para enfermedades raras o presentaciones inusuales.

Sin embargo, al igual que los reportes de caso, las series de casos tienen **limitaciones** importantes. Siguen siendo estudios descriptivos, lo que significa que no pueden establecer causalidad. La selección de casos puede estar sesgada, y la falta de un grupo control impide comparar la frecuencia de la exposición o el desenlace con un grupo de referencia. Por lo tanto, aunque sugieren asociaciones, no pueden confirmarlas.

A pesar de estas limitaciones, las series de casos son un escalón importante en la jerarquía de la evidencia. Nos permiten consolidar observaciones iniciales y, si los patrones observados son lo suficientemente convincentes, pueden impulsar la realización de estudios analíticos, como los estudios de casos y con-

troles o de cohorte, que sí están diseñados para investigar relaciones causales. Así, la documentación cuidadosa de estas colecciones de casos es fundamental para el avance del conocimiento médico, abriendo la puerta a investigaciones más profundas sobre los misterios de la salud y la enfermedad.

4.1.3. *Ejemplo de Reporte de Caso*

Antes de sumergirnos en la redacción, el primer y crucial paso es identificar y seleccionar el caso adecuado para reportar. No todos los casos, por interesantes que parezcan en el momento, justifican un reporte formal. Para guiar esta selección, podemos recurrir a criterios que nos ayuden a evaluar la relevancia y la novedad de nuestra observación.

Selección del Caso: Criterios FINER y la Información Diagnóstica

Piensa en el proceso de selección como la búsqueda de esa joya verdaderamente rara y significativa en la vasta mina de la práctica clínica. ¿Qué hace que un caso sea digno de ser compartido con la comunidad científica? Aquí es donde los criterios **FINER** (que ya hemos explorado en la formulación de preguntas de investigación) vuelven a ser útiles, adaptados a la evaluación de un caso individual:

- **Factible:** ¿Tenemos acceso completo a la historia clínica del paciente, incluyendo resultados de laboratorio, imágenes, reportes de patología y cualquier otra información relevante? ¿Podemos obtener el consentimiento informado del paciente o sus representantes legales para publicar los datos, asegurando la anonimidad? Si la respuesta es no a cualquiera de estas preguntas, el caso, por interesante que sea, no es factible de reportar formalmente.
- **Interesante:** ¿El caso presenta una característica que lo hace notable? Puede ser una presentación clínica extremadamente rara de una enfermedad común, una complicación inesperada de un procedimiento o tratamiento, una asociación inusual entre dos condiciones, una respuesta excepcional a una terapia, o la primera descripción de un fenómeno en un contexto particular (geográfico, demográfico). Debe ser algo que, al ser conocido por otros clínicos, pueda cambiar su forma de pensar o actuar.
- **Novedoso:** ¿El caso aporta información que no ha sido ampliamente documentada en la literatura médica? ¿Es la primera vez que se describe

esta presentación, complicación o asociación? Incluso si hay reportes previos, ¿este caso añade un detalle crucial, una nueva perspectiva diagnóstica (como el uso de inmunohistoquímica específica en el ejemplo de sífilis rectal) o terapéutica?

- **Ético:** ¿Podemos reportar este caso respetando la privacidad y confidencialidad del paciente? ¿Estamos seguros de que la publicación no causará daño o estigma al paciente? La obtención del consentimiento informado es un pilar fundamental de la ética en la publicación de casos clínicos.
- **Relevante:** ¿Tiene este caso implicaciones prácticas para otros clínicos, investigadores o pacientes? ¿Ayuda a mejorar el diagnóstico, el manejo o la comprensión de una enfermedad? Un caso relevante ofrece lecciones que van más allá de la experiencia individual del paciente.

La **información diagnóstica** que se tenga del paciente es crucial en este proceso de selección. Un caso es particularmente fuerte para un reporte si cuenta con documentación diagnóstica robusta, como:

- Confirmación histopatológica (biopsias, citologías).
- Hallazgos de imagenología distintivos (CT, RM, ecografía, radiografías).
- Resultados de laboratorio o microbiológicos concluyentes.
- Respuesta documentada a un tratamiento específico.

Un caso con un diagnóstico bien establecido y documentado permite a los autores presentar hallazgos concretos y fiables, aumentando la credibilidad y utilidad del reporte.

Planteamiento de la Pregunta y el Objetivo de Investigación

Una vez que hemos seleccionado un caso que cumple con los criterios FINER y cuenta con información diagnóstica sólida, el siguiente paso es articular claramente por qué este caso merece ser reportado. Esto se traduce en plantear una pregunta de investigación implícita o explícita y definir el objetivo del reporte.

- **Pregunta de Investigación (implícita):** Aunque los reportes de caso no suelen presentar una pregunta de investigación formal como en estudios analíticos, hay una pregunta subyacente que el reporte busca responder

o explorar. Por ejemplo, en el caso de sífilis rectal: "¿Cómo se presenta la sífilis rectal en pacientes con inmunosupresión severa por VIH, y qué herramientas diagnósticas avanzadas son útiles en este contexto?"

- **Objetivo del Reporte:** El objetivo es el propósito explícito de describir el caso. Generalmente se relaciona con destacar la presentación inusual, el desafío diagnóstico, el manejo particular, o la lección aprendida. Se formula de manera concisa.

Ejemplos de objetivos:

- "Presentar un caso de [condición] con una presentación atípica para aumentar la conciencia sobre su variabilidad clínica."
- "Describir el abordaje diagnóstico de [condición] en un paciente con [contexto particular] y resaltar la utilidad de [técnica diagnóstica específica]."
- "Reportar una complicación inusual de [tratamiento/procedimiento] para alertar a los clínicos sobre este riesgo potencial."

El proceso de selección basado en FINER y la información diagnóstica, seguido por la formulación de la pregunta y el objetivo, nos asegura que el reporte de caso no sea solo una anécdota clínica, sino una contribución valiosa al conocimiento médico.

Estructura General del Reporte de Caso:

Un reporte de caso típico se estructura en las siguientes secciones principales:

1. **Título**
2. **Autores y Afiliaciones**
3. **Resumen**
4. **Palabras clave**
5. **Introducción**
6. **Presentación del caso**
7. **Discusión**

8. **Conclusión**
9. **Agradecimientos (Opcional)**
10. **Financiación (Si aplica)**
11. **Declaración del Comité de Ética**
12. **Declaración de Consentimiento Informado**
13. **Declaración de Disponibilidad de Datos (Si aplica)**
14. **Abreviaciones (Si aplica)**
15. **Referencias**
16. **Figuras y Tablas (Insertadas en el texto o al final)**

Ahora, veamos el contenido clave para cada una de estas secciones:

1. Título:

- Debe ser conciso, claro y reflejar el contenido principal del caso.
- Idealmente, debe mencionar la condición o el hallazgo inusual y el contexto clínico relevante.
- Puede incluir palabras clave que ayuden a la búsqueda.
- *Ejemplo de estructura*: "[Condición Inusual] en [Contexto Clínico Específico]: Reporte de Caso y [Hallazgo/Aspecto Destacado]".

2. Autores y Afiliaciones:

- Lista de los autores que contribuyeron significativamente al reporte.
- Afiliación institucional de cada autor (Departamento, Institución, Ciudad, País).
- Información de contacto para el autor de correspondencia.

3. Resumen:

- Es un resumen estructurado que sigue el flujo del reporte. Suele tener un límite de palabras (ej. 150-250 palabras).

- **Contexto y objetivo:** Breve contexto sobre la condición, por qué el caso es interesante o importante de reportar (relacionado con el objetivo).
- **Presentación del caso:** Descripción concisa de los datos demográficos del paciente, los síntomas principales, los hallazgos clave del examen físico, diagnóstico, intervenciones y desenlace.
- **Discusión:** Destacar la principal lección aprendida o la relevancia del caso.
- **Conclusión:** Resumir la implicación clínica o la recomendación principal derivada del caso.

4. Palabras clave:

- Lista de 3 a 8 palabras o frases clave que describan los aspectos más importantes del caso.
- Ayudan a indexar el reporte y facilitan su búsqueda en bases de datos.
- Utiliza términos estandarizados si es posible (ej. MeSH terms).

5. Introducción:

- Generalmente consta de 2 a 4 párrafos.
- **Párrafo 1:** Introducir la enfermedad o condición general, su importancia, prevalencia típica o características conocidas.
- **Párrafo 2 (y siguientes):** Describir el aspecto específico y raro del caso que se va a reportar (ej. presentación atípica, complicación inusual, respuesta a tratamiento, asociación con otra condición). Explicar por qué esta presentación es inusual o desafiante basándose en la literatura existente. Aquí se relaciona con la **novedad** y el **interés** del caso según los criterios FINER.
- **Último Párrafo:** Indicar el propósito del reporte de caso (ej. "Presentamos este caso para ilustrar…", "Este reporte tiene como objetivo destacar…"). Esto refleja el **objetivo** planteado.

6. Presentación del caso:

- Esta sección narra la historia clínica del paciente de manera cronológica y detallada.

- **Datos Demográficos:** Edad, sexo, origen étnico (si es relevante y no compromete la anonimidad).
- **Historia y Presentación:** Descripción detallada de los síntomas, su inicio, evolución, factores que los modifican. Antecedentes médicos, quirúrgicos, familiares, sociales, hábitos (ej. tabaquismo, alcohol, ocupación) relevantes.
- **Línea de Tiempo Clave de Eventos (Opcional pero recomendado por CARE):** Una lista o figura que resume las fechas importantes de síntomas, diagnósticos, intervenciones y desenlaces. (Similar a la Figura 1 del ejemplo de sífilis rectal).
- **Evaluación Inicial:** Hallazgos del examen físico al momento de la presentación. Resultados de laboratorio iniciales relevantes.
- **Hallazgos de Imagenología, Endoscopia u Otros Procedimientos:** Describir los resultados de estudios clave (ej. CT, RM, radiografías, ecografías, endoscopias). **Para las figuras de imagenología, incluir:** tipo de estudio, vista (axial, sagital, coronal), si es con contraste, y usar flechas o marcadores para señalar los hallazgos relevantes en la figura. La leyenda debe describir lo que se ve.
- **Trabajo Microbiológico y/o Histopatológico:** Describir los resultados de cultivos, serologías, pruebas moleculares, y especialmente, los hallazgos de biopsias o citologías. **Para las figuras de histopatología, incluir:** tipo de tejido/órgano, tinción utilizada (ej. H&E, inmunohistoquímica), **magnificación** (ej. ×40, ×100, ×400) y una **barra de escala** (ej. 100 µm, 50 µm). Usar flechas para señalar células o estructuras específicas relevantes (ej. granulomas, infiltrado celular, microorganismos, células malignas). La leyenda debe describir detalladamente los hallazgos microscópicos a cada magnificación. La solidez de esta información se relaciona directamente con la **factibilidad** y **relevancia** del caso.
- **Tratamiento y Profilaxis:** Describir las intervenciones terapéuticas realizadas (fármacos, dosis, vía, duración), procedimientos quirúrgicos, u otras medidas de manejo. Mencionar profilaxis si fue relevante.
- **Investigaciones y Manejo Adicionales:** Cualquier otro estudio o manejo que se realizó y fue importante para el caso.

- **Desenlace:** Describir la evolución del paciente tras el tratamiento, la respuesta clínica, los resultados de seguimiento (ej. mejoría, estabilidad, progresión, efectos adversos).

7. Discusión:

- Interpretar el caso en el contexto de la literatura científica.
- **Párrafo 1:** Resumir los puntos clave del caso y el diagnóstico.
- **Párrafo 2 (y siguientes):** Comparar la presentación de tu paciente con casos similares reportados en la literatura. ¿En qué se parece? ¿En qué difiere? ¿Por qué crees que la presentación fue inusual? Discutir los desafíos diagnósticos o terapéuticos específicos que se presentaron. Explicar la fisiopatología o los mecanismos subyacentes relevantes para el caso. Esto se relaciona con el **interés** y la **novedad** del caso.
- **Limitaciones:** Reconocer las limitaciones inherentes a un reporte de caso (no generalizable, no establece causalidad, posible sesgo de selección).
- **Recomendaciones Futuras:** Sugerir qué tipo de investigación adicional sería necesaria para comprender mejor la condición o responder preguntas pendientes (ej. estudios de series de casos más amplias, estudios analíticos).
- **Lecciones Aprendidas:** Destacar los mensajes clave para la práctica clínica que otros médicos pueden aprender de este caso. Esto se relaciona con la **relevancia** del caso.

8. Conclusión:

- Breve resumen de la importancia del caso y las principales conclusiones o recomendaciones derivadas de él. Debe ser conciso y claro. Evita repetir la discusión.

9. Agradecimientos (Opcional):

- Agradecer a personas o departamentos que colaboraron en el manejo del paciente o en la elaboración del reporte (ej. personal de enfermería, otros especialistas, patólogos, radiólogos).

10. Financiación (Si aplica):

- Declarar si el reporte de caso recibió alguna financiación específica.

11. Declaración del Comité de Ética:

- Confirmar que el estudio cumple con los principios éticos y fue aprobado por un comité de ética (ver ejemplo arriba).

12. Declaración de Consentimiento Informado:

- Confirmar que se obtuvo el consentimiento informado del paciente para la publicación (ver ejemplo arriba).

13. Declaración de Disponibilidad de Datos (Si aplica):

- Indicar si los datos del caso están disponibles y cómo acceder a ellos (si es relevante y permitido).

14. Abreviaciones (Si aplica):

- Lista de todas las abreviaciones utilizadas en el manuscrito y su significado.

15. Referencias:

- Lista de las fuentes citadas en el texto, siguiendo un formato de citación estándar (ej. Vancouver, AMA, etc.).

16. Figuras y Tablas:

- Las figuras (imágenes radiológicas, histológicas, diagramas de línea de tiempo) y tablas (resúmenes de datos, comparaciones con otros casos) deben ser claras, tener títulos descriptivos y leyendas detalladas que permitan su comprensión sin necesidad de leer el texto completo.

Siguiendo esta estructura, podrás organizar la información de tu caso de manera lógica y completa, aumentando las posibilidades de que tu reporte sea aceptado y útil para la comunidad médica.

4.1.4 Estudios *Transversales* (*Cross-sectional Studies*)

Imagina que quieres saber cuántas personas en tu ciudad tienen bicicleta y cuántas de ellas usan casco en un día cualquiera. No necesitas seguir a la

gente a lo largo del tiempo; simplemente puedes ir a varios puntos de la ciudad en un día específico, observar a los ciclistas y registrar si llevan casco o no. Un estudio transversal funciona de manera similar: recolecta datos de una población objetivo en un *único punto en el tiempo*. Es como tomar una fotografía de la salud de una comunidad en un momento determinado, capturando quién tiene una enfermedad o característica de salud particular y quién tiene una exposición de interés, todo al mismo tiempo.

Esta metodología es particularmente útil para:

- **Medir la verdadera carga de una enfermedad** o la prevalencia de características relacionadas con la salud en una población definida. Como señala el material, nos ayuda a descubrir el "iceberg clínico", esa gran proporción de enfermedad que existe en la comunidad pero que no llega a los servicios de salud.
- **Evaluar las necesidades de salud** de una población, lo cual es fundamental para la planificación y asignación de recursos sanitarios.
- **Realizar encuestas** sobre opiniones, actitudes o comportamientos relacionados con la salud, como hábitos de fumar o consumo de alcohol.
- **Explorar la asociación entre una exposición y una enfermedad crónica**, especialmente cuando no se conoce con precisión el momento de inicio de la enfermedad.

Teoría y Tipos:

Los estudios transversales son, por naturaleza, **observacionales**. El investigador no interviene ni manipula la exposición o el desenlace, simplemente observa y registra lo que encuentra en la población en el momento del estudio. Se clasifican principalmente en dos tipos, según su objetivo:

- **Estudios Transversales Descriptivos:** Su objetivo principal es medir la prevalencia y describir la distribución de una enfermedad o característica de salud en una población definida. Responden a las preguntas de *quién*, *dónde* y *cuándo* ocurre el fenómeno, en ese momento específico.
- **Estudios Transversales Analíticos:** Van un paso más allá al investigar la interrelación entre variables de interés, es decir, exploran la asocia-

ción entre una exposición y un desenlace. Por ejemplo, podrían examinar si la prevalencia de una enfermedad es mayor en un grupo expuesto a un factor que en un grupo no expuesto. Sin embargo, es crucial recordar que, debido a que la exposición y el desenlace se miden simultáneamente, estos estudios tienen limitaciones importantes para establecer relaciones causales definitivas. No pueden determinar si la exposición precedió al desenlace o viceversa.

Los estudios transversales son inherentemente **transversales** en cuanto al tiempo de recolección de datos del individuo. No son prospectivos (siguiendo a los individuos hacia el futuro) ni retrospectivos (mirando hacia atrás en el tiempo desde un desenlace), aunque pueden utilizar datos recogidos previamente si estos representan un punto específico en el tiempo.

Población y Muestreo

El primer paso en el diseño de un estudio transversal es definir claramente la **población objetivo**, es decir, el grupo de personas al cual se pretenden generalizar los resultados. A partir de esta población objetivo, se selecciona una **población de estudio** (los individuos invitados a participar) y, finalmente, el **tamaño de la muestra** que realmente participa.

La selección de una **muestra representativa** de la población objetivo es fundamental para la validez externa del estudio, es decir, para poder generalizar los hallazgos. El **muestreo aleatorio** es una técnica común para intentar lograr esta representatividad. Además, el **tamaño de la muestra** debe ser lo suficientemente grande para asegurar que la estimación de la prevalencia (o de la asociación) tenga una precisión adecuada, lo que se relaciona con el concepto de potencia estadística que vimos en capítulos anteriores.

Un desafío común es la **baja tasa de respuesta** o participación, lo que puede introducir un **sesgo de selección** (específicamente, sesgo de participación o no respuesta) si las características de quienes participan difieren sistemáticamente de quienes no lo hacen, y si estas características están relacionadas con la exposición o el desenlace. Esto puede llevar a subestimar o sobreestimar la prevalencia real en la población objetivo.

Recolección y Análisis de Datos:

La recolección de datos en estudios transversales puede emplear diversos **instrumentos** y **técnicas**:

- **Cuestionarios:** Pueden ser autoadministrados o aplicados por entrevistadores. Son relativamente económicos para muestras grandes, pero pueden tener tasas de respuesta más bajas y están sujetos a sesgo de información (ej. sesgo de memoria, deseabilidad social).
- **Entrevistas:** Permiten obtener información más detallada y aclarar dudas, pero son más costosas y requieren personal capacitado. Pueden estar sujetas a sesgo del entrevistador.
- **Exámenes Médicos o Pruebas de Laboratorio:** Proporcionan datos objetivos sobre el estado de salud o la presencia de exposiciones biológicas. Son más fiables para ciertas variables, pero también pueden ser costosos y logísticamente complejos.
- **Revisión de Registros Existentes:** Utilizar datos ya recopilados (ej. registros hospitalarios, bases de datos administrativas) puede ser eficiente, pero la calidad y la disponibilidad de la información pueden ser limitadas y no siempre se ajustan perfectamente a la pregunta de investigación.

Es importante notar que, aunque se habla de un "único punto en el tiempo", la recolección de datos no siempre ocurre de manera simultánea para todos los participantes. Puede extenderse por un período corto (días o semanas), siempre y cuando la condición o exposición de interés no cambie drásticamente en ese lapso y la recolección represente efectivamente la situación en un momento dado.

El **análisis de datos** en un estudio transversal se centra principalmente en el cálculo de la **Prevalencia**. Como vimos, la prevalencia de una enfermedad en la muestra de estudio se calcula dividiendo el número de casos existentes en el momento del estudio por el total de personas en la muestra en ese mismo momento.

$$Prevalencia = \frac{\text{Número de casos del evento en la muestra en un momento dado}}{\text{Total de personas en la muestra en el mismo momento}}$$

En los estudios transversales analíticos, también se calculan **medidas de asociación** para explorar la relación entre la exposición y el desenlace. Las más comunes son:

- ***Prevalence Odds Ratio* (POR):** Es análogo al *Odds Ratio* de los estudios de casos y controles. Se calcula como el cociente de las *odds* de tener la enfermedad entre los expuestos y las *odds* de tener la enfermedad entre los no expuestos en el momento del estudio. Es la medida preferida para enfermedades crónicas o exposiciones de larga duración.

$$POR = \frac{\text{Odds de enfermedad en expuestos}}{\text{Odds de enfermedad en no expuestos}} = \frac{a/b}{c/d} = \frac{a \times d}{b \times c}$$

Su interpretación es similar a la del OR: un POR > 1 sugiere una mayor *odds* de tener la enfermedad en los expuestos.

- ***Prevalence Ratio* (PR):** Es análogo al Riesgo Relativo de los estudios de cohorte. Se calcula como el cociente de la prevalencia de la enfermedad entre los expuestos y la prevalencia de la enfermedad entre los no expuestos en el momento del estudio. Es la medida preferida para enfermedades agudas o de corta duración.

$$PR = \frac{\text{Prevalencia en expuestos}}{\text{Prevalencia en no expuestos}} = \frac{a/(a+b)}{c/(c+d)}$$

- Su interpretación es similar a la del RR: un PR > 1 sugiere una mayor probabilidad de tener la enfermedad en los expuestos en ese momento.

Cuando la prevalencia de la enfermedad es baja (generalmente < 15%), el POR es una buena aproximación del PR y, por lo tanto, del Riesgo Relativo, lo cual es una propiedad útil a recordar.

Sesgos y Variables Confusoras:

Como todos los estudios observacionales, los estudios transversales son susceptibles a **sesgos** y a la influencia de **variables confusoras**.

- **Variables Confusoras:** Una variable confusora es aquella que está asociada tanto con la exposición como con el desenlace, pero no es un paso intermedio en la cadena causal. Su presencia puede distorsionar la verdadera asociación entre la exposición y el desenlace. En un estudio transversal, si identificamos y medimos posibles variables confusoras

(ej. edad, sexo, nivel socioeconómico), podemos intentar controlarlas durante el análisis estadístico para obtener una estimación más precisa de la asociación.

- **Sesgos:** Los sesgos son errores sistemáticos en el diseño, conducción o análisis de un estudio que llevan a una estimación incorrecta de la asociación. En los estudios transversales, algunos sesgos son particularmente relevantes:

  - **Sesgo de Selección:** Ocurre si la muestra no es representativa de la población objetivo debido a la forma en que se seleccionan o participan los individuos. El **sesgo de no respuesta** (si quienes no participan difieren de quienes sí lo hacen) y el **sesgo de acceso a la atención médica** (si la probabilidad de ser identificado con una enfermedad depende de si se busca atención médica) son ejemplos. El **sesgo de incidencia-prevalencia** (también conocido como sesgo de supervivencia) es crucial en estudios transversales de enfermedades crónicas: al tomar una "foto" en un momento dado, solo capturamos a los casos prevalentes, es decir, a aquellos que han sobrevivido hasta ese momento. Si la exposición afecta la duración de la enfermedad (supervivencia), la asociación observada en los casos prevalentes puede no reflejar la asociación en los casos incidentes. La **migración selectiva** (si las personas enfermas o expuestas se mudan del área de estudio) también puede introducir sesgo de selección.

  - **Sesgo de Medición:** Ocurre si la información sobre la exposición o el desenlace es inexacta. La **clasificación errónea aleatoria** (errores de medición que ocurren al azar) generalmente diluye la asociación real, acercándola a 1. La **clasificación errónea no aleatoria** (errores que dependen del estado de exposición o desenlace) puede sesgar la asociación en cualquier dirección. Ejemplos incluyen el **sesgo de recuerdo** (si los casos recuerdan la exposición de manera diferente a los no casos) o el **sesgo del entrevistador** (si el entrevistador influye en las respuestas).

Consideraciones Éticas:

Como en cualquier investigación que involucre seres humanos, los estudios transversales deben cumplir con principios éticos rigurosos. Es fundamental obtener la aprobación de un **Comité de Ética de Investigación** antes de

iniciar la recolección de datos. Además, se debe obtener el **consentimiento informado** de todos los participantes, explicándoles claramente el propósito del estudio, los procedimientos, los riesgos y beneficios potenciales, la confidencialidad de sus datos y su derecho a retirarse en cualquier momento.

Usos, Aplicaciones, Ventajas y Desventajas:

En resumen, los estudios transversales son herramientas valiosas para describir la salud de una población, estimar la prevalencia de enfermedades y exposiciones, y generar hipótesis. Sus **ventajas** incluyen ser relativamente rápidos y económicos de realizar (ya que no requieren un seguimiento a largo plazo), permitir la medición simultánea de múltiples exposiciones y desenlaces, y ser útiles para planificar servicios de salud.

Sin embargo, sus **desventajas** son importantes: no pueden establecer la secuencia temporal entre exposición y desenlace, lo que limita su capacidad para inferir causalidad. No son eficientes para estudiar enfermedades raras o de muy corta duración. Son susceptibles a varios sesgos, especialmente el sesgo de selección y el sesgo de incidencia-prevalencia en enfermedades crónicas. Tampoco permiten calcular tasas de incidencia.

4.1.5. *Ejemplo de estudio transversal analítico*

Continuamos nuestro viaje por los diseños de estudio, ascendiendo un poco más en la jerarquía de la evidencia. Dejamos atrás la mera descripción para adentrarnos en la exploración de relaciones entre variables. Es hora de hablar de los **Estudios Transversales Analíticos**, esos que toman una "foto" de la población en un momento dado, pero no solo para describir, sino para empezar a indagar si ciertas exposiciones se asocian con ciertos desenlaces.

Imagina que te encuentras en medio de un bullicioso campus universitario y te preguntas: ¿existe alguna relación entre el nivel de estrés que experimentan los estudiantes de medicina y su rendimiento académico? No necesitas seguirlos durante años; puedes, en un momento específico del semestre, preguntarles sobre su nivel de estrés actual y obtener su promedio de calificaciones. Al analizar estos datos conjuntamente, puedes explorar si los estudiantes con mayor estrés tienden a tener un rendimiento académico diferente al de aquellos con menor estrés. Este es el corazón de un estudio transversal analítico.

Para ilustrar cómo se gesta y estructura un estudio de este tipo, sigamos un caso ejemplo, desde la pregunta inicial hasta la interpretación de los resultados, teniendo en cuenta el rigor metodológico y ético necesario.

Caso Ejemplo: Estrés Percibido y Rendimiento Académico en Estudiantes de Medicina

1. Formulación del Problema y Definición de Hipótesis

Todo estudio comienza con una pregunta, una curiosidad que queremos satisfacer de manera sistemática. En nuestro caso, la pregunta podría ser:

- **Pregunta de Investigación:** ¿Existe una asociación entre el nivel de estrés percibido y el rendimiento académico en estudiantes de medicina de [Nombre de la Universidad] durante el semestre académico [Número del Semestre] de [Año]?

Para evaluar si esta pregunta es adecuada para un estudio transversal y si vale la pena investigarla, aplicamos los criterios FINER:

- **Factible:** ¿Podemos acceder a los estudiantes de medicina? ¿Podemos medir su nivel de estrés percibido (ej. con un cuestionario validado) y obtener datos sobre su rendimiento académico (ej. promedio de calificaciones)? ¿Tenemos los recursos y el tiempo para recolectar y analizar los datos en un momento dado? (Sí, parece factible en un entorno universitario).
- **Interesante:** ¿Es interesante para la comunidad académica y para los propios estudiantes entender esta relación? (Sí, el estrés es un problema común en estudiantes de medicina y su impacto en el rendimiento es relevante).
- **Novedoso:** Aunque hay estudios sobre estrés y rendimiento, ¿este estudio aporta algo nuevo? Quizás se enfoca en una población específica, utiliza un instrumento particular, o explora variables confusoras no consideradas antes. (Podría ser novedoso en el contexto específico de esta universidad o al explorar ciertos factores).
- **Ético:** ¿Podemos recolectar esta información de manera confidencial y voluntaria, sin coerción y respetando la privacidad de los estudiantes?

(Sí, con un diseño ético adecuado, consentimiento informado y anonimización de datos).

- **Relevante:** ¿Los hallazgos podrían informar intervenciones para reducir el estrés o mejorar el rendimiento académico? (Sí, los resultados podrían guiar programas de apoyo estudiantil).

Si nuestra pregunta cumple con los criterios FINER, procedemos a plantear nuestras hipótesis:

- **H_0:** No existe una asociación estadísticamente significativa entre el nivel de estrés percibido y el rendimiento académico en estudiantes de medicina de [Nombre de la Universidad] durante el semestre académico [Número del Semestre] de [Año].
- **Ha:** Existe una asociación estadísticamente significativa entre el nivel de estrés percibido y el rendimiento académico en estudiantes de medicina de [Nombre de la Universidad] durante el semestre académico [Número del Semestre] de [Año].

2. Selección del Nivel de Significancia (α)

Antes de recolectar datos, definimos nuestro umbral para decidir si un resultado es estadísticamente significativo. Comúnmente, se establece un nivel de significancia (α) de **0.05**. Esto significa que estamos dispuestos a aceptar un 5% de probabilidad de rechazar la H_0 cuando en realidad es verdadera (cometer un error Tipo I).

3. Recolección y Organización de Datos

- **Población Objetivo:** Todos los estudiantes de medicina de [Nombre de la Universidad].
- **Población de Estudio:** Estudiantes de medicina matriculados en el semestre [Número del Semestre] de [Año].
- **Muestreo:** Idealmente, se utilizaría un **muestreo aleatorio simple** o **estratificado** (ej. por año de estudio) para obtener una muestra representativa de la población de estudio. Si se utiliza un **muestreo por conveniencia** (ej. encuestar a los estudiantes disponibles en ciertas clases),

se debe reconocer esta limitación en la discusión. Se debe calcular un tamaño de muestra adecuado para tener suficiente potencia estadística para detectar una asociación si existe.

- **Recolección de Datos:** Se recolectan datos en un punto específico del semestre. Los **instrumentos** pueden incluir:
  - Un **cuestionario validado** para medir el estrés percibido (ej. Escala de Estrés Percibido - PSS).
  - Solicitud del **promedio de calificaciones** del semestre anterior o acumulado como medida de rendimiento académico.
  - Preguntas sobre posibles **variables confusoras** (ej. horas de estudio semanales, apoyo social, situación laboral, antecedentes de problemas de salud mental).
- **Método:** La recolección podría ser mediante **encuestas en línea** (más eficientes para grandes números, pero pueden tener menor tasa de respuesta) o **encuestas en papel** (mayor tasa de respuesta en algunos contextos, pero más costosas y laboriosas).
- **Organización:** Los datos recolectados se ingresan en una base de datos (ej. Excel, SPSS, R) para su limpieza y análisis.

Estructura del Artículo (Adaptada para Estudio Transversal Analítico):

Siguiendo una estructura similar a la de un reporte de caso, pero con secciones adaptadas para un estudio analítico:

Título:

- Debe reflejar las variables principales y el diseño del estudio.
- *Ejemplo:* "Asociación entre Estrés Percibido y Rendimiento Académico: Un Estudio Transversal en Estudiantes de Medicina de [Nombre de la Universidad]".

Autores y Afiliaciones: (Similar al reporte de caso)

Resumen:

- **Contexto y Objetivo:** Contexto del problema (estrés y rendimiento en estudiantes de medicina) y el objetivo del estudio (explorar la asociación).

- **Metodología:** Breve descripción del diseño (transversal), población, muestreo, instrumentos y análisis estadístico principal.
- **Resultados:** Presentar los hallazgos descriptivos clave y el resultado del análisis de asociación principal (ej. POR, PR, diferencia de medias, p-valor).
- **Conclusión:** Resumir la principal conclusión y su implicación.

Palabras clave: (Similar al reporte de caso, incluyendo las variables clave y el diseño)

1. Introducción: (Similar al reporte de caso, pero enfocándose en la prevalencia del problema y la justificación para explorar la asociación)

- **Párrafo 1:** Introducir el estrés en el contexto académico y su relevancia.
- **Párrafo 2:** Hablar sobre el rendimiento académico y los factores que pueden influirlo.
- **Párrafo 3:** Discutir la posible relación entre estrés y rendimiento, mencionando estudios previos y la brecha de conocimiento que este estudio busca abordar.
- **Último Párrafo:** Indicar el objetivo del estudio (ej. "El objetivo de este estudio transversal fue examinar la asociación entre el nivel de estrés percibido y el rendimiento académico en…").

2. Métodos:

- **Diseño del Estudio:** Describir claramente que se trata de un estudio transversal analítico.
- **Población y Muestreo:** Definir la población objetivo y de estudio. Explicar el método de muestreo y cómo se seleccionó la muestra. Mencionar el tamaño de la muestra final y la tasa de respuesta.
- **Recolección de Datos:** Describir los instrumentos utilizados (ej. nombre del cuestionario de estrés, cómo se midió el rendimiento académico). Explicar el procedimiento de recolección y el período en que se realizó.
- **Variables:** Definir claramente las variables de exposición (estrés percibido, quizás categorizado en niveles), el desenlace (rendimiento acadé-

mico, medido como promedio de calificaciones) y las posibles variables confusoras medidas.

- **Consideraciones Éticas:** Declarar la aprobación por el Comité de Ética y el proceso de Consentimiento Informado (similar a la plantilla de reporte de caso).
- **Análisis Estadístico:** Describir el plan de análisis, siguiendo una lógica de lo general a lo específico:

  - **Estadística Descriptiva:** Describir las características de la muestra utilizando frecuencias y porcentajes para variables categóricas (ej. sexo, año de estudio) y medidas de tendencia central (media, mediana) y dispersión (DS, IQR) para variables continuas (ej. edad, promedio de calificaciones, puntuación de estrés). Esto se suele presentar en una **Tabla 1: Características de la Muestra de Estudio**.
  - **Análisis Bivariado (Exploración de la Asociación):** Aquí se aplica el proceso que mencionaste:

    - **4A/4B. Verificación de Supuestos/Condiciones:** Antes de aplicar pruebas paramétricas, verificar la normalidad de los datos continuos (ej. Shapiro-Wilk) y la homogeneidad de varianzas (ej. Levene). Para tablas de contingencia, verificar las frecuencias esperadas (condición para Chi-cuadrado). Si no se cumplen los supuestos, considerar transformaciones o pruebas no paramétricas.
    - **5A/5B. Elección de la Prueba Estadística Adecuada:** Seleccionar la prueba según el tipo de variables y si se cumplen los supuestos. Si el estrés se categoriza (ej. bajo vs. alto) y el rendimiento es continuo, se podría usar una **Prueba t de Student** (si cumple supuestos) o **Mann-Whitney U** (si no). Si el estrés se categoriza en >2 grupos (ej. bajo, moderado, alto) y el rendimiento es continuo, se usaría **ANOVA** (si cumple supuestos) o **Kruskal-Wallis** (si no). Si ambas variables (estrés y rendimiento) se tratan como continuas, se podría usar un **coeficiente de correlación de Pearson** (si cumplen supuestos de normalidad y linealidad) o **Spearman** (si no). Si ambas variables se categorizan, se usaría la **Prueba de Chi-Cuadrado** o la **Prueba Exacta de Fisher** para tablas pequeñas.

- **6A/6B. Realización de la Prueba y Cálculo del Estadístico:** Ejecutar la prueba estadística seleccionada para obtener el estadístico de prueba y el **valor p**.
- **7A. Pruebas Post Hoc (si aplica):** Si se usa ANOVA y el resultado es significativo, realizar pruebas *post hoc* (ej. Tukey) para identificar entre qué grupos específicos existen diferencias significativas en el rendimiento académico.
- **Análisis Multivariado (Control de Confusores):** Describir cómo se abordarán las variables confusoras. Esto puede ser mediante **análisis estratificado** (analizar la asociación entre estrés y rendimiento por separado dentro de subgrupos de la variable confusora, ej. por horas de estudio) o utilizando **modelos de regresión** (ej. regresión lineal si el rendimiento es continuo) que permiten ajustar por el efecto de múltiples variables simultáneamente.

3. Resultados:

- Presentar los resultados de manera clara y objetiva, apoyándose en tablas y figuras.
- **Descripción de la Muestra:** Presentar la **Tabla 1**, resumiendo las características demográficas y otras variables de la muestra (frecuencias/porcentajes, medias/desviaciones estándar).
- **Resultados del Análisis Bivariado:** Presentar los resultados de la exploración inicial de la asociación. Esto podría ser una **Tabla 2** que muestre el rendimiento académico (media ± DE) por categoría de estrés, junto con el resultado de la prueba estadística utilizada (ej. valor t, grados de libertad, **valor p**).
- **Resultados del Análisis Multivariado:** Presentar los resultados de los modelos de regresión o análisis estratificados, mostrando la asociación entre estrés y rendimiento ajustada por los confusores. Esto podría ser una **Tabla 3** que muestre los coeficientes de regresión (o *Odds Ratios* si se modela un desenlace binario), sus intervalos de confianza del 95% y los valores p.

- Complementar las tablas con texto que describa los hallazgos principales.

4. Discusión:

- Interpretar los resultados en el contexto de la pregunta de investigación y la literatura existente.
- **Párrafo 1:** Resumir los hallazgos principales del estudio (descriptivos y analíticos).
- **Párrafo 2 (y siguientes):** Discutir la asociación encontrada entre estrés y rendimiento. Comparar los resultados con estudios previos. ¿Apoyan o contradicen hallazgos anteriores? Explorar posibles explicaciones para la asociación observada.
- **Fortalezas del Estudio:** Mencionar los aspectos positivos del estudio (ej. uso de cuestionario validado, tamaño de muestra razonable, control de confusores).
- **Limitaciones:** Es crucial reconocer las limitaciones inherentes a un estudio transversal:
  - **No permite establecer causalidad:** La asociación observada no implica que el estrés cause el bajo rendimiento (o viceversa), ya que se midieron al mismo tiempo. Podría haber causalidad inversa o un tercer factor (confusor residual no medido) que explique la asociación.
  - **Sesgos:** Discutir los posibles sesgos que pudieron afectar los resultados (ej. sesgo de no respuesta, sesgo de memoria en los cuestionarios, sesgo de supervivencia si el estrés severo lleva a abandonar la carrera).
  - **Medición en un punto en el tiempo:** La "foto" no captura la variabilidad del estrés o el rendimiento a lo largo del semestre.
- **Recomendaciones Futuras:** Sugerir cómo futuras investigaciones podrían superar las limitaciones de este estudio (ej. realizar estudios longitudinales para evaluar la secuencia temporal, utilizar medidas de estrés más objetivas, explorar otros posibles confusores).
- **Implicaciones:** Discutir las posibles implicaciones de los hallazgos para la práctica (ej. necesidad de programas de manejo del estrés para estudiantes de medicina) o para futuras investigaciones.

5. Conclusión:

- Resumir la principal conclusión del estudio, reiterando la asociación encontrada (si la hubo) y su relevancia, pero **sin afirmar causalidad**.

Secciones Finales: (Agradecimientos, Financiación, Ética, Consentimiento, Datos, Abreviaciones, Referencias - similar a la plantilla de reporte de caso)

Este enfoque estructurado, desde la formulación cuidadosa de la pregunta y la hipótesis hasta el análisis estadístico paso a paso y la discusión reflexiva de los hallazgos y sus limitaciones, es fundamental para realizar y reportar un estudio transversal analítico con rigor científico. Nos permite, aunque no lleguemos a la cima de la causalidad, dar un paso firme en la comprensión de las asociaciones entre factores de salud en un momento dado de la vida de una población.

4.2. Estudios analíticos observacionales

Estos estudios, a diferencia de los descriptivos, buscan responder al *porqué*: ¿Por qué algunas personas desarrollan una enfermedad y otras no? ¿Qué factores o exposiciones están asociados a la aparición de un evento de salud? Son como detectives que, ante un suceso (la enfermedad), retroceden en el tiempo para buscar pistas (las exposiciones) que puedan explicar lo ocurrido.

4.2.1. *Estudios de Casos y Controles*

Imagina que una extraña enfermedad aparece en una comunidad. Como epidemiólogo-detective, no puedes esperar años a ver quién se enferma en el futuro. Necesitas actuar rápido. ¿Qué haces? Identificas a las personas que ya tienen la enfermedad (los **casos**) y, crucialmente, seleccionas un grupo comparable de personas que *no* tienen la enfermedad (los **controles**). Luego, miras hacia atrás en el tiempo para determinar si la exposición al factor de interés fue más común en los casos que en los controles. Este es el diseño fundamental de un estudio de casos y controles.

Es como si investigaras un brote de intoxicación alimentaria. Identificas a los enfermos (casos) y a un grupo de amigos o familiares que no enfermaron (controles). Luego, les preguntas a ambos grupos qué comieron en una reunión reciente. Si descubres que una comida específica fue consumida con mucha más frecuencia por los enfermos que por los sanos, tienes una fuerte pista sobre la fuente del brote.

Selección de Casos y Controles:

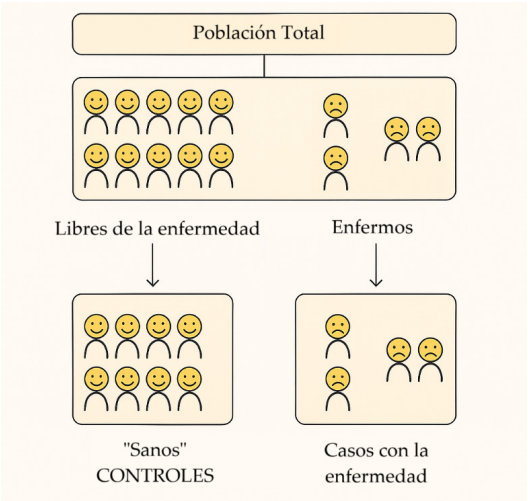
Este es un paso crítico y a menudo el más desafiante en un estudio de casos y controles.

- **Casos:** Deben ser individuos con la enfermedad o el desenlace de interés claramente definido. Es importante establecer criterios diagnósticos precisos para asegurar que todos los casos sean homogéneos. Los casos pueden ser incidentes (recién diagnosticados) o prevalentes (existentes en un momento dado), aunque los casos incidentes son generalmente preferibles para evitar el sesgo de incidencia-prevalencia (o sesgo de supervivencia), donde solo se incluyen los casos que han sobrevivido hasta el momento del estudio, y la exposición podría estar relacionada con la supervivencia.
- **Controles:** Deben ser individuos que *no* tienen la enfermedad o el desenlace de interés, pero que provienen de la misma **población fuente** que los casos. Es decir, si los controles hubieran desarrollado la enfermedad, habrían sido identificados y seleccionados como casos. La selección adecuada de controles es fundamental para evitar el sesgo de selección. Los controles pueden seleccionarse de diversas fuentes:
    - **Población general:** Idealmente, seleccionados aleatoriamente de la población de la que surgieron los casos.
    - **Hospitales:** Pacientes hospitalizados por otras causas. Es conveniente seleccionar controles con una variedad de diagnósticos no relacionados con la exposición o el desenlace de interés para minimizar el sesgo de admisión hospitalaria (sesgo de Berkson), que puede ocurrir si la probabilidad de ser hospitalizado está relacionada tanto con la exposición como con la enfermedad en estudio.
    - **Vecinos o amigos:** Personas cercanas a los casos. Esto puede ayudar a controlar variables socioeconómicas o de estilo de vida, pero también puede introducir sesgo si la exposición tiende a agruparse en redes sociales.

A menudo se utiliza el **apareamiento (*matching*)** para seleccionar controles que sean similares a los casos en ciertas características (ej. edad, sexo, raza)

que podrían ser variables confusoras. Esto ayuda a asegurar la comparabilidad entre los grupos, aunque un apareamiento excesivo (*overmatching*) con variables no confusoras puede diluir la asociación real (ver figura 59).

Figura 59. *Esquema de Selección de Casos y Controles*. Elaborado con Gemini-PRO.



Nota. Diagrama que ilustra la selección de los grupos de casos y controles a partir de la población total según su estado de enfermedad.

Identificación de Exposición:

Una vez seleccionados los casos y los controles, se recopila información sobre su historial de **exposición** al factor de interés *antes* de que los casos desarrollaran la enfermedad. Esta información puede obtenerse a través de:

- **Entrevistas o cuestionarios:** Preguntando directamente a los participantes sobre sus exposiciones pasadas. Esto es susceptible a **sesgo de recuerdo** (*recall bias*), donde los casos, al estar enfermos, pueden recordar o reportar sus exposiciones de manera diferente a los controles.
- **Registros existentes:** Revisando historias clínicas, registros laborales, bases de datos farmacéuticas, etc. Esto puede reducir el sesgo de recuerdo, pero la información puede ser incompleta o inexacta.
- **Muestras biológicas almacenadas:** Analizando muestras de sangre, tejido, etc., recolectadas previamente. Esto proporciona información ob-

jetiva sobre la exposición, pero la disponibilidad de las muestras puede ser limitada.

Es crucial que la recolección de información sobre la exposición se realice de manera similar para casos y controles para minimizar el **sesgo de medición** (*measurement bias*), incluyendo el **sesgo del entrevistador** (*interviewer bias*) o el **sesgo de sospecha de exposición** (*exposure suspicion bias*).

Análisis (OR):

En los estudios de casos y controles, no podemos calcular directamente la incidencia de la enfermedad en expuestos y no expuestos porque la población de estudio se selecciona en función del desenlace (enfermos y no enfermos), no de la exposición. Por lo tanto, no podemos calcular el Riesgo Relativo (RR). La medida de asociación utilizada es la ***Odds Ratio* (OR)**, que, como vimos anteriormente, es la razón entre las *odds* de exposición en los casos y las *odds* de exposición en los controles.

Utilizando la tabla 19:

Tabla 19. *Tabla de contingencia para casos y controles.*

	Enfermos (Casos)	No Enfermos (Controles)
Expuestos	a	b
No Expuestos	c	d

La *Odds Ratio* se calcula como:

$$OR = \frac{a/c}{b/d} = \frac{a \times d}{b \times c}$$

La interpretación de la OR es similar a la del RR:

- **OR = 1:** No hay asociación entre la exposición y la enfermedad.
- **OR > 1:** La exposición está asociada a un mayor riesgo de enfermedad (factor de riesgo).
- **OR < 1:** La exposición está asociada a un menor riesgo de enfermedad (factor protector).

Una propiedad fundamental de la OR en estudios de casos y controles es que, cuando la enfermedad o el desenlace de interés es **raro** en la población fuente de la que provienen los casos y controles, la **OR es una buena aproximación del Riesgo Relativo (RR)**. Esto es increíblemente útil porque nos permite estimar el RR (una medida más intuitiva del riesgo) incluso en diseños donde no podemos calcular directamente las incidencias. Como se muestra en el ejemplo de pericarditis constrictiva, donde la complicación es rara, el riesgo (0.018) y las *odds* (0.183) son valores diferentes, pero si tuviéramos un grupo de no expuestos, la razón de riesgos y la razón de *odds* serían similares.

Ventajas y Desventajas:

Los estudios de casos y controles tienen **ventajas** significativas:

- **Eficientes para enfermedades raras:** Permiten investigar exposiciones asociadas a enfermedades poco frecuentes sin necesidad de seguir a miles de personas durante años.
- **Eficientes para enfermedades con largos períodos de latencia:** Útiles cuando el tiempo entre la exposición y el desenlace es muy largo.
- **Relativamente rápidos y económicos:** Comparados con los estudios de cohorte prospectivos, son más rápidos y menos costosos de realizar.
- **Permiten investigar múltiples exposiciones:** Se puede preguntar sobre varios factores de exposición en relación con el desenlace de interés.

Sin embargo, también presentan **desventajas** importantes:

- **No permiten calcular la incidencia:** No se puede determinar la tasa de nuevos casos en la población.
- **Susceptibles a sesgos:** Son particularmente vulnerables al sesgo de selección (especialmente en la elección de controles) y al sesgo de recuerdo en la identificación de la exposición.
- **Dificultad para establecer la secuencia temporal:** Como la exposición se mide después del desenlace, puede ser difícil asegurar que la exposición precedió a la enfermedad, lo que limita la capacidad para inferir causalidad.

- **No eficientes para exposiciones raras:** Si la exposición es muy poco común, será difícil encontrar suficientes casos y controles expuestos.

Sesgos Específicos en Estudios de Casos y Controles:

Además de los sesgos de selección y medición generales, algunos son particularmente relevantes aquí:

- **Sesgo de Admisión Hospitalaria (Sesgo de Berkson):** Ocurre en estudios basados en hospitales si la probabilidad de ser hospitalizado está influenciada tanto por la exposición como por la enfermedad en estudio, distorsionando la asociación observada entre pacientes hospitalizados.
- **Sesgo de Incidencia-Prevalencia (Sesgo de Supervivencia o Sesgo de Neyman):** Ocurre si los casos seleccionados no son representativos de todos los casos que surgen en la población, a menudo porque solo se incluyen los casos prevalentes (sobrevivientes), y la exposición está relacionada con la supervivencia.
- **Sesgo de Detección:** Ocurre si la probabilidad de que se diagnostique la enfermedad está relacionada con la exposición. Por ejemplo, si una exposición causa síntomas leves que llevan a un diagnóstico más temprano de la enfermedad (sesgo de desenmascaramiento), o si el conocimiento de la exposición influye en la intensidad de la búsqueda diagnóstica (sesgo de sospecha diagnóstica).
- **Sesgo de Recuerdo (*Recall Bias*):** Los casos pueden recordar o reportar sus exposiciones pasadas de manera diferente a los controles, a menudo recordando mejor las exposiciones potenciales debido a su enfermedad.

Consideraciones Éticas:

Al igual que en otros estudios con participación humana, se requiere la aprobación de un **Comité de Ética de Investigación** y la obtención del **Consentimiento Informado** de los participantes (casos y controles), explicando el propósito del estudio, los procedimientos, la confidencialidad y su derecho a no participar.

A pesar de sus limitaciones, especialmente en la inferencia causal directa, los estudios de casos y controles son herramientas esenciales en la caja de he-

rramientas epidemiológica, particularmente valiosas para investigar enfermedades raras y generar hipótesis. Sin embargo, la búsqueda de evidencia más sólida nos lleva a diseños que pueden establecer la secuencia temporal de manera más clara, como los **Estudios de Cohorte**, nuestro próximo destino en este recorrido por los diseños de investigación.

4.2.2. *Ejemplo de Casos y Controles*

Caso Ejemplo: Uso Previo del Antibiótico X y Riesgo de Infección por *Clostridium difficile* (ICD) en Pacientes Hospitalizados

1. Formulación del Problema y Definición de Hipótesis

Nuestra curiosidad inicial nos lleva a la pregunta:

- **Pregunta de Investigación:** ¿Se asocia el uso previo del antibiótico X con un mayor riesgo de desarrollar infección por *Clostridium difficile* en pacientes hospitalizados en [Nombre del Hospital]?

Evaluemos esta pregunta con los criterios FINER:

- **Factible:** ¿Podemos identificar y acceder a pacientes con ICD en nuestro hospital (casos)? ¿Podemos identificar pacientes hospitalizados sin ICD que sean comparables (controles)? ¿Podemos obtener información fiable sobre el uso de antibióticos previos en ambos grupos (ej. de registros médicos)? ¿Podemos obtener consentimiento o una exención ética si aplica? (Sí, parece factible en un entorno hospitalario).
- **Interesante:** ¿Es interesante para los clínicos y epidemiólogos entender la asociación entre un antibiótico específico y la ICD? (Sí, tiene implicaciones directas para la práctica clínica y el control de infecciones).
- **Novedoso:** ¿Se ha estudiado esta asociación para el antibiótico X en particular o en este contexto hospitalario específico? (Podría ser novedoso si el antibiótico X es relativamente nuevo o si las tasas de ICD varían entre hospitales).
- **Ético:** ¿Podemos obtener información sobre el uso de antibióticos y el estado de ICD de manera confidencial y con el debido proceso ético (consentimiento o aprobación de comité)? (Sí, con la aprobación del comité de ética y consentimiento informado o exención justificada).

- **Relevante:** ¿Los hallazgos podrían informar las guías de uso de antibióticos en el hospital para reducir el riesgo de ICD? (Sí, tiene alta relevancia clínica y de salud pública).

Si la pregunta es FINER, planteamos nuestras hipótesis:

- **H_0:** No existe una asociación estadísticamente significativa entre el uso previo del antibiótico X y el riesgo de desarrollar infección por *Clostridium difficile* en pacientes hospitalizados en [Nombre del Hospital].
- **H□:** Existe una asociación estadísticamente significativa entre el uso previo del antibiótico X y el riesgo de desarrollar infección por *Clostridium difficile* en pacientes hospitalizados en [Nombre del Hospital].

2. Selección del Nivel de Significancia (α)

Definimos nuestro umbral de significancia, comúnmente α=0.05.

Estructura del Artículo (Adaptada para Estudio de Casos y Controles):

Siguiendo una estructura similar a la de los ejemplos anteriores:

Título:

- Debe reflejar la exposición, el desenlace y el diseño del estudio.
- *Ejemplo*: "Uso Previo del Antibiótico X como Factor de Riesgo para Infección por *Clostridium difficile*: Un Estudio de Casos y Controles".

Autores y Afiliaciones: (Similar a ejemplos anteriores)

Resumen:

- **Contexto y Objetivo:** Contexto del problema (ICD, factores de riesgo, interés en antibiótico X) y el objetivo del estudio (investigar la asociación).
- **Métodos:** Breve descripción del diseño (casos y controles), población fuente, criterios de selección de casos y controles, cómo se midió la exposición (uso de antibiótico X) y el análisis estadístico principal (cálculo de OR).
- **Resultados:** Presentar el número de casos y controles, la frecuencia de exposición en cada grupo, el cálculo de la OR (con IC 95% y p-valor) y los resultados del análisis ajustado si se realizó.

- **Conclusión:** Resumir la principal conclusión sobre la asociación y su implicación.

Palabras clave: (Similar a ejemplos anteriores, incluyendo la exposición, el desenlace y el diseño)

1. Introducción:

- **Párrafo 1:** Introducir la Infección por *Clostridium difficile* (ICD), su importancia clínica y epidemiológica, y su asociación conocida con el uso de antibióticos en general.
- **Párrafo 2:** Discutir que no todos los antibióticos confieren el mismo riesgo de ICD y que hay interés en identificar antibióticos de alto riesgo. Mencionar el antibiótico X y la justificación para investigarlo (ej. uso frecuente en el hospital, sospecha clínica).
- **Párrafo 3:** Explicar brevemente el diseño de casos y controles como una metodología adecuada para investigar factores de riesgo de enfermedades raras como la ICD.
- **Último Párrafo:** Indicar el objetivo del estudio (ej. "El objetivo de este estudio de casos y controles fue evaluar la asociación entre el uso previo del antibiótico X y el desarrollo de ICD en pacientes hospitalizados...").

2. Métodos:

- **Diseño del Estudio:** Describir claramente que se trata de un estudio de casos y controles.
- **Población Fuente:** Definir la población de la cual surgieron los casos y se seleccionaron los controles (ej. todos los pacientes hospitalizados en [Nombre del Hospital] durante un período específico).
- **Selección de Casos:** Definir los criterios de inclusión y exclusión para los casos (ej. diagnóstico confirmado de ICD durante la hospitalización, edad > 18 años). Describir cómo se identificaron los casos (ej. revisión de registros de laboratorio, base de datos de control de infecciones).
- **Selección de Controles:** Definir los criterios de inclusión y exclusión para los controles (ej. pacientes hospitalizados en el mismo período que

los casos, sin diagnóstico de ICD). Describir cómo se seleccionaron los controles y la estrategia de apareamiento si se utilizó (ej. apareados por edad, sexo y servicio de hospitalización para asegurar comparabilidad). Mencionar la razón de controles por caso (ej. 1:1, 2:1, 4:1).

- **Identificación y Medición de la Exposición:** Describir cómo se obtuvo la información sobre el uso previo del antibiótico X (ej. revisión de la historia clínica electrónica, registros de farmacia). Definir qué se considera "uso previo" (ej. haber recibido el antibiótico X en los 30 días previos al diagnóstico de ICD para los casos, o en un período de tiempo comparable para los controles). Es crucial estandarizar la ventana de exposición.
- **Variables:** Definir claramente la variable de desenlace (ICD: Sí/No) y la variable de exposición (Uso de antibiótico X: Sí/No). Identificar y describir las posibles **variables confusoras** que se midieron y que podrían estar asociadas tanto con el uso del antibiótico X como con el riesgo de ICD (ej. edad, comorbilidades [índice de Charlson], duración de la estancia hospitalaria, uso de otros antibióticos, uso de inhibidores de la bomba de protones).
- **Consideraciones Éticas:** Declarar la aprobación por el Comité de Ética y el proceso de Consentimiento Informado (o justificar la exención si se usan datos retrospectivos anonimizados).
- **Análisis Estadístico:** Describir el plan de análisis:
  - **Estadística Descriptiva:** Describir las características de los casos y controles (frecuencias/porcentajes, medias/desviaciones estándar) para evaluar la comparabilidad entre los grupos. Presentar una **Tabla 1: Características de Casos y Controles**.
  - **Análisis Bivariado (Exploración de la Asociación):** Presentar los datos de exposición en una **Tabla 2x2** (Tabla 2). Calcular la ***Odds Ratio* (OR)** cruda como medida de asociación entre el uso del antibiótico X y la ICD. Calcular el **Intervalo de Confianza del 95% (IC 95%)** para la OR y el **valor p**.
  - **Análisis Multivariado (Control de Confusores):** Describir cómo se ajustó por las variables confusoras identificadas. Esto se realiza típica-

mente utilizando **regresión logística no condicional** (si no hay apareamiento) o **regresión logística condicional** (si hay apareamiento). Presentar la **OR ajustada**, su IC 95% y valor p. Esto se suele presentar en una **Tabla 3: Odds Ratios (Cruda y Ajustada) para la Asociación entre el Uso del Antibiótico X y la ICD**.

3. Resultados:

- Presentar los hallazgos de manera clara y objetiva.
- Describir el número de casos y controles incluidos en el análisis.
- Presentar la **Tabla 1** con las características de los casos y controles, comentando si los grupos fueron comparables en las variables importantes (especialmente si hubo apareamiento).
- Presentar la **Tabla 2x2** (Tabla 2) mostrando la distribución de la exposición (uso de antibiótico X) en casos y controles.
- Presentar los resultados del análisis bivariado: el valor de la **OR cruda**, su IC 95% y el valor p.
- Presentar los resultados del análisis multivariado: la **OR ajustada**, su IC 95% y el valor p, indicando por qué variables se ajustó.
- Complementar las tablas con texto que describa los hallazgos principales (ej. "El uso del antibiótico X fue significativamente más frecuente en los casos de ICD (X%) que en los controles (Y%)", "La OR cruda fue de Z (IC 95% A-B, p=C)", "Después de ajustar por edad y comorbilidades, la OR ajustada fue de W (IC 95% D-E, p=F)").

4. Discusión:

- Interpretar los resultados en el contexto de la pregunta de investigación y la literatura.
- **Párrafo 1:** Resumir los hallazgos principales, enfocándose en la OR ajustada y su significancia estadística.
- **Párrafo 2 (y siguientes):** Discutir la interpretación de la OR (ej. "Los pacientes que usaron el antibiótico X tuvieron aproximadamente W ve-

ces la *odds* de desarrollar ICD en comparación con quienes no lo usaron, después de considerar [variables ajustadas]"). Comparar estos hallazgos con estudios previos sobre el mismo antibiótico o clases de antibióticos. Discutir posibles mecanismos biológicos que expliquen la asociación (si se conoce).

- **Fortalezas del Estudio:** Mencionar los aspectos positivos (ej. diseño adecuado para enfermedad rara, control de confusores importantes, fuente de datos fiable para exposición).
- **Limitaciones:** Discutir las limitaciones inherentes al diseño de casos y controles en el contexto de este estudio:
  - **No permite calcular incidencia:** No se puede determinar el riesgo absoluto de ICD asociado al antibiótico X.
  - **Susceptibilidad a sesgos:** Discutir los sesgos potenciales relevantes para este estudio (ej. **sesgo de recuerdo** si la información sobre antibióticos se obtuvo por entrevista, aunque menos probable si se usan registros; **sesgo de selección** en la elección de controles; **sesgo de incidencia-prevalencia** si se incluyeron casos prevalentes y el uso del antibiótico X afecta la supervivencia a la ICD; **sesgo de detección** si el uso del antibiótico X llevó a una mayor búsqueda de ICD).
  - **Dificultad para establecer la secuencia temporal precisa:** Aunque se busca la exposición *previa*, determinar el momento exacto de la exposición y el inicio de la enfermedad puede ser un desafío.
  - **Variables confusoras residuales:** Reconocer que puede haber variables confusoras no medidas o no controladas adecuadamente.
- **Recomendaciones Futuras:** Sugerir futuras investigaciones para abordar las limitaciones (ej. estudios de cohorte para evaluar incidencia y secuencia temporal, estudios con diferentes poblaciones, investigación de mecanismos).
- **Implicaciones:** Discutir las implicaciones clínicas o de salud pública de los hallazgos (ej. considerar el riesgo de ICD al prescribir el antibiótico X, necesidad de estrategias de manejo de antibióticos).

5. Conclusión:

- Resumir la principal conclusión del estudio sobre la asociación entre el antibiótico X y la ICD, reiterando la OR ajustada y su significancia, pero **sin afirmar causalidad directa** (ya que es un estudio observacional).

Secciones Finales: (Agradecimientos, Financiación, Ética, Consentimiento, Datos, Abreviaciones, Referencias - similar a ejemplos anteriores)

4.2.3. *Cohortes*

Si los estudios de casos y controles son como detectives que miran hacia atrás para encontrar pistas sobre un evento que ya ocurrió, los estudios de cohorte son como profetas que miran hacia adelante. Comienzan seleccionando un grupo de personas (la **cohorte**) que están libres de la enfermedad o el desenlace de interés al inicio del estudio. Luego, clasifican a estas personas según su nivel de **exposición** a un factor particular (ej. fumadores vs. no fumadores, expuestos a un químico vs. no expuestos). Finalmente, siguen a ambos grupos a lo largo del tiempo para ver quién desarrolla la enfermedad y a qué velocidad.

Piensen en una cohorte como un grupo de barcos que zarpan juntos en un largo viaje. Algunos barcos llevan una carga particular (la exposición), mientras que otros no. Los estudios de cohorte siguen a estos barcos a través de mares a veces tormentosos (el tiempo de seguimiento) para ver cuáles llegan a un puerto específico (el desenlace, la enfermedad) y si la carga que llevaban influyó en su probabilidad de llegar (o naufragar).

Un estudio de cohorte es un tipo de estudio **observacional** que sigue a un grupo de individuos a lo largo del tiempo para investigar si la exposición a un factor determinado influye en la **incidencia** de una enfermedad o desenlace en el futuro. Sus propiedades de diseño clave son:

- La **cohorte** (el grupo de estudio) debe ser representativa de la población a la que queremos generalizar nuestros resultados (ver figura 60).

- Al inicio del estudio, todos los participantes deben estar **libres de la enfermedad** o el desenlace que se va a estudiar. Esto es fundamental para poder medir la incidencia (casos nuevos).

- Se mide la **exposición** de interés al inicio del estudio (línea de base). Las exposiciones pueden ser de diversa naturaleza: comportamientos (fumar, dieta), factores ambientales (contaminación), características biológicas (presión arterial, colesterol), tratamientos, etc.
- Los sujetos son **seguidos a lo largo del tiempo** para documentar la aparición de nuevos casos de la enfermedad o desenlace en los grupos expuestos y no expuestos. El tiempo de seguimiento debe ser suficiente para que se desarrollen suficientes eventos de interés.
- Existe la posibilidad de **pérdida de seguimiento** (*loss-to-follow-up*), donde los participantes dejan de ser contactables. Minimizar esta pérdida es crucial para la validez del estudio.

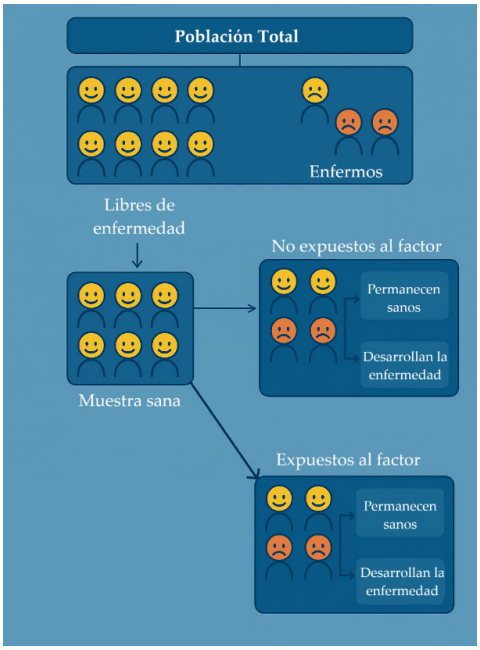


Figura 60. *Esquema de Selección de Cohorte*. Elaborado con GeminiPRO.

Cohortes descriptivas vs. analíticas:

- **Cohortes Descriptivas:** Simplemente describen la incidencia o la historia natural de una enfermedad en una cohorte a lo largo del tiempo, sin comparar grupos de exposición. Por ejemplo, seguir a un grupo de

pacientes recién diagnosticados con una enfermedad para describir su tasa de supervivencia a 5 años.

- **Cohortes Analíticas:** Son las más comunes en epidemiología etiológica. Comparan la incidencia de la enfermedad entre un grupo expuesto y un grupo no expuesto (o con diferentes niveles de exposición) dentro de la misma cohorte para investigar la asociación entre la exposición y el desenlace. Nuestro enfoque principal aquí son las cohortes analíticas.

Prospectivas y Retrospectivas:

La distinción entre cohortes prospectivas y retrospectivas se basa en el momento en que se inicia el estudio en relación con la ocurrencia de la exposición y el desenlace:

Estudios de Cohorte Prospectivos: La exposición se mide en el **presente** (al inicio del estudio) en individuos libres de la enfermedad, y se les sigue hacia el **futuro** para observar quién desarrolla el desenlace. Son como plantar semillas (medir exposición) y esperar a ver qué plantas crecen (observar desenlace). La secuencia temporal (exposición antes del desenlace) es clara, lo que fortalece la inferencia causal.

Estudios de Cohorte Retrospectivos (o Históricos): Tanto la exposición como el desenlace ya han ocurrido cuando se inicia el estudio. El investigador utiliza **registros preexistentes** (ej. historias clínicas, registros laborales, bases de datos) para identificar una cohorte en el pasado, determinar su estado de exposición en ese momento y luego rastrear su desenlace hasta el presente. Son como encontrar registros de barcos que zarparon hace tiempo, ver qué carga llevaban y determinar si llegaron a puerto. Son más rápidos y económicos que los prospectivos, pero dependen de la calidad y disponibilidad de los datos históricos. La claridad de la secuencia temporal puede ser menos robusta que en los estudios prospectivos si los registros no son lo suficientemente detallados.

Cálculo de Incidencia y RR:

La gran ventaja de los estudios de cohorte es que, al seguir a individuos libres de la enfermedad, podemos calcular directamente la **incidencia** (la tasa de nuevos casos) en los grupos expuestos y no expuestos.

- Incidencia en el Grupo Expuesto (I_e):

$$I_e = \frac{\text{Número de nuevos casos en el grupo expuesto}}{\text{Total de personas en riesgo en el grupo expuesto al inicio}}$$

- Incidencia en el Grupo No Expuesto (I_o):

$$I_o = \frac{\text{Número de nuevos casos en el grupo no expuesto}}{\text{Total de personas en riesgo en el grupo no expuesto al inicio}}$$

La medida de asociación principal en los estudios de cohorte analíticos es el **Riesgo Relativo (RR)**, que, como vimos en el Capítulo 2 de esta parte III, es el cociente de estas dos incidencias:

$$RR = \frac{I_e}{I_o} = \frac{\text{Incidencia en Expuestos}}{\text{Incidencia en No Expuestos}}$$

El RR nos dice cuántas veces más (o menos) probable es que los individuos expuestos desarrollen la enfermedad en comparación con los no expuestos. Un RR de 2 significa que los expuestos tienen el doble de riesgo; un RR de 0.5 significa que tienen la mitad del riesgo.

Además del RR (una medida de efecto *relativa*), los estudios de cohorte también permiten calcular la **Diferencia de Riesgos (DR)** o **Riesgo Atribuible (RA)**, una medida de efecto *absoluta*:

$$DR = RA = I_e - I_o$$

La DR nos indica cuántos casos adicionales de la enfermedad por cada "X" personas expuestas son atribuibles a la exposición, en comparación con los no expuestos. Es útil para estimar el impacto potencial de una intervención a nivel individual.

Para evaluar la precisión de nuestras estimaciones de RR y DR, calculamos sus **Intervalos de Confianza del 95% (IC 95%)** y el **valor p** asociado, para determinar si la asociación observada es estadísticamente significativa y probablemente no se deba solo al azar.

Variables confusoras, causalidad, y sesgo:

Como estudios observacionales, las cohortes son susceptibles a **variables confusoras** y **sesgos**.

- **Variables Confusoras:** Factores asociados tanto a la exposición como al desenlace. En un estudio de cohorte, es crucial identificar y medir posibles confusores al inicio del estudio para poder controlarlos en el análisis (ej. mediante estratificación o modelos de regresión), obteniendo una estimación más precisa del RR ajustado. El ejemplo clásico es la asociación aparente entre la terapia de reemplazo hormonal (TRH) y la reducción del riesgo de enfermedad coronaria observada en algunas cohortes, que luego se vio confundida por el hecho de que las mujeres que usaban TRH tendían a tener estilos de vida más saludables ("efecto del usuario saludable").

- **Causalidad:** Los estudios de cohorte, especialmente los prospectivos, proporcionan la evidencia observacional más sólida para inferir causalidad, principalmente porque la **secuencia temporal** es clara: la exposición se mide *antes* de la aparición del desenlace. Sin embargo, no pueden probar causalidad con la misma certeza que un ensayo clínico aleatorizado, ya que el investigador no controla la asignación de la exposición y pueden existir confusores no medidos. Los criterios de Bradford Hill (discutidos anteriormente) son útiles para evaluar la probabilidad de causalidad a partir de la evidencia de cohortes.

- **Sesgos:**

  - **Sesgo de Selección:** Ocurre si la forma en que se selecciona la cohorte o la participación en el estudio introduce diferencias sistemáticas entre los grupos de exposición o entre la muestra y la población objetivo. El **sesgo del trabajador saludable** es un ejemplo específico de cohortes ocupacionales, donde los trabajadores tienden a ser más sanos que la población general, lo que puede subestimar el riesgo asociado a exposiciones laborales si se compara a los trabajadores expuestos con la población general no expuesta. El **sesgo de no respuesta** también puede afectar la representatividad.

- **Sesgo de Implementación del Estudio:**

  - **Sesgo de Pérdida de Seguimiento (*Loss-to-Follow-up Bias*):** Es un riesgo importante en cohortes prospectivas. Ocurre si la proporción de individuos perdidos durante el seguimiento difiere entre los grupos de expo-

sición, y si la pérdida está relacionada con el desenlace. Si los individuos expuestos que desarrollan el desenlace tienen más probabilidades de perderse, se subestimará la incidencia en los expuestos y el RR.

- **Sesgo de Detección o Seguimiento (*Detection Bias* o *Follow-up Bias*):** Ocurre si la probabilidad de que se diagnostique el desenlace difiere entre los grupos de exposición. Por ejemplo, si los individuos expuestos son seguidos de manera más intensiva o se les realizan más pruebas diagnósticas, es más probable que se detecten casos en este grupo, sobreestimando la incidencia y el RR. Mantener "ciego" al personal de seguimiento sobre el estado de exposición de los participantes puede ayudar a reducir este sesgo.

- **Sesgo de Medición (*Measurement Bias*):** Inexactitudes en la medición de la exposición o el desenlace. La **clasificación errónea aleatoria** (errores al azar) tiende a diluir la asociación hacia el nulo. La **clasificación errónea no aleatoria** (errores que dependen del estado de exposición o desenlace) puede sesgar la asociación en cualquier dirección. El **sesgo de recuerdo** puede ser un problema en cohortes retrospectivas si la información sobre la exposición se obtiene retrospectivamente por entrevista. El **sesgo del entrevistador** o el **sesgo de sospecha de exposición** también pueden ocurrir si la persona que mide la exposición o el desenlace conoce el estado de la otra variable.

Ventajas y Desventajas:

Las **ventajas** de los estudios de cohorte:

- Permiten calcular directamente la **incidencia** y el **Riesgo Relativo (RR)**.
- Son el mejor diseño observacional para establecer la **secuencia temporal** (exposición antes del desenlace), fortaleciendo la inferencia causal (especialmente los prospectivos).
- Permiten investigar **múltiples desenlaces** asociados a una única exposición.
- Son eficientes para estudiar **exposiciones raras**.
- Los estudios de cohorte prospectivos minimizan el sesgo de recuerdo sobre la exposición.

Las **desventajas** de los estudios de cohorte:

- **Costosos y requieren mucho tiempo** (especialmente los prospectivos), sobre todo para enfermedades raras o con largos períodos de latencia.
- **No eficientes para enfermedades raras:** Aunque permiten calcular incidencia, se necesitaría un tamaño muestral enorme para observar suficientes casos de una enfermedad muy rara. Para estas, los casos y controles son más eficientes.
- **Susceptibles a sesgos:** Particularmente a la pérdida de seguimiento, sesgo de selección y sesgo de medición.
- Los estudios de cohorte retrospectivos dependen de la disponibilidad y calidad de los datos históricos y pueden tener más dificultades para controlar confusores no registrados.

Consideraciones Éticas:

Como en otros estudios observacionales, se requiere la aprobación de un **Comité de Ética de Investigación**. En los estudios prospectivos, es indispensable obtener el **Consentimiento Informado** de los participantes. En los estudios retrospectivos basados en registros, puede ser posible obtener una exención ética si los datos están anonimizados y no implican riesgos para los participantes.

Los estudios de cohorte, con su capacidad para seguir a los individuos en el tiempo, son herramientas poderosas para investigar la etiología de las enfermedades, cuantificar el riesgo y proporcionar evidencia valiosa sobre las asociaciones entre exposiciones y desenlaces. Su fortaleza en la claridad temporal los sitúa por encima de los estudios transversales y de casos y controles en la jerarquía de la evidencia para la inferencia causal, aunque siempre debemos ser conscientes de sus limitaciones y sesgos potenciales.

4.2.4. *Ejemplo de Estudio de Cohortes*

Piensa que, en tu universidad, con tantos estudiantes pasando horas frente a pantallas de computadoras y dispositivos móviles, te preguntas si este tiempo de pantalla prolongado está relacionado con el desarrollo del síndrome de ojo seco. Un estudio transversal solo te diría cuántos estudiantes tienen

ojo seco y pasan mucho tiempo frente a pantallas *en este momento*, pero no sabrías qué fue primero. Un estudio de casos y controles sería difícil porque el ojo seco, aunque común, puede ser fluctuante y el "inicio" no siempre es claro, además de que la exposición (tiempo de pantalla) es casi universal en cierta medida, dificultando la selección de controles no expuestos. Aquí es donde un estudio de cohorte brilla: identificas a estudiantes sin ojo seco, mides cuánto tiempo pasan frente a pantallas y los sigues para ver quién desarrolla el síndrome.

Aquí te presento un ejemplo estructurado de cómo abordarías la investigación y el reporte de un estudio de cohorte prospectivo sobre esta pregunta:

Caso Ejemplo: Tiempo de Pantalla Prolongado y la Incidencia de Síndrome de Ojo Seco en Estudiantes Universitarios: Un Estudio de Cohorte Prospectivo

1. Formulación del Problema y Definición de Hipótesis

Nuestra observación del uso generalizado de pantallas nos lleva a la pregunta:

- **Pregunta de Investigación:** ¿Se asocia el tiempo de pantalla prolongado (≥ 4 horas/día) con una mayor incidencia de síndrome de ojo seco en estudiantes universitarios de [Nombre de la Universidad] durante un período de seguimiento de 1 año?

Evaluemos esta pregunta con los criterios FINER:

- **Factible:** ¿Podemos reclutar una cohorte de estudiantes universitarios sin ojo seco? ¿Podemos medir de manera fiable su tiempo de pantalla al inicio y seguirlos durante 1 año para identificar nuevos casos de ojo seco (ej. mediante cuestionarios de seguimiento)? ¿Contamos con los recursos para el seguimiento? (Sí, parece factible en un entorno universitario con una cohorte accesible).
- **Interesante:** ¿Es interesante para los estudiantes, profesores y profesionales de la salud visual entender el impacto del tiempo de pantalla en la salud ocular? (Sí, es un tema de creciente preocupación dada la digitalización de la vida estudiantil).
- **Novedoso:** ¿Se ha estudiado esta asociación específicamente en estudiantes universitarios con este diseño y con una definición clara de

tiempo de pantalla y ojo seco? (Podría ser novedoso dependiendo de la literatura existente y el contexto específico).

- **Ético:** ¿Podemos obtener el consentimiento informado de los estudiantes, asegurar la confidencialidad de sus datos y ofrecerles información sobre salud visual al finalizar el estudio? (Sí, con un protocolo ético adecuado).
- **Relevante:** ¿Los hallazgos podrían informar recomendaciones sobre hábitos de uso de pantallas para estudiantes universitarios y guiar intervenciones preventivas? (Sí, tiene alta relevancia para la salud pública y la salud ocupacional/visual).

Si la pregunta es FINER, planteamos nuestras hipótesis:

- **H_0:** No existe una asociación estadísticamente significativa entre el tiempo de pantalla prolongado (≥ 4 horas/día) y la incidencia de síndrome de ojo seco en estudiantes universitarios de [Nombre de la Universidad] durante un período de seguimiento de 1 año.
- **H_a:** Existe una asociación estadísticamente significativa entre el tiempo de pantalla prolongado (≥ 4 horas/día) y una mayor incidencia de síndrome de ojo seco en estudiantes universitarios de [Nombre de la Universidad] durante un período de seguimiento de 1 año.

2. Selección del Nivel de Significancia (α)

Definimos nuestro umbral de significancia, comúnmente α=0.05.

Estructura del Artículo (Adaptada para Estudio de Cohorte):

Siguiendo una estructura similar a la de los ejemplos anteriores:

Título:

- Debe reflejar la exposición, el desenlace, la población y el diseño del estudio.
- *Ejemplo*: "Tiempo de Pantalla Prolongado y la Incidencia de Síndrome de Ojo Seco: Un Estudio de Cohorte Prospectivo en Estudiantes Universitarios".

Autores y Afiliaciones: (Similar a ejemplos anteriores)

Resumen:

- **Contexto y Objetivo:** Contexto del problema (uso de pantallas, ojo seco, brecha de conocimiento) y el objetivo del estudio (evaluar la asociación prospectivamente).
- **Methods:** Breve descripción del diseño (cohorte prospectiva), población, criterios de inclusión/exclusión (libres de ojo seco al inicio), medición de exposición (tiempo de pantalla) y desenlace (incidencia de ojo seco), tiempo de seguimiento, y análisis estadístico principal (cálculo de incidencia y RR).
- **Resultados:** Presentar el tamaño de la cohorte, la proporción de expuestos/no expuestos, la incidencia de ojo seco en cada grupo, el cálculo del RR (con IC 95% y p-valor) y los resultados del análisis ajustado.
- **Conclusión:** Resumir la principal conclusión sobre la asociación y su implicación.

Palabras clave: (Similar a ejemplos anteriores, incluyendo la exposición, el desenlace, la población y el diseño)

1. Introducción:

- **Párrafo 1:** Introducir el síndrome de ojo seco, su prevalencia, síntomas y factores de riesgo conocidos (ej. edad, sexo, uso de lentes de contacto).
- **Párrafo 2:** Discutir el aumento del uso de dispositivos electrónicos y el tiempo de pantalla en la población general, y específicamente en estudiantes universitarios. Mencionar la hipótesis de que el tiempo de pantalla prolongado podría ser un factor de riesgo para el ojo seco.
- **Párrafo 3:** Explicar por qué un estudio de cohorte prospectivo es un diseño adecuado para investigar esta asociación, destacando su capacidad para establecer la secuencia temporal y calcular la incidencia.
- **Último Párrafo:** Indicar el objetivo del estudio (ej. "El objetivo de este estudio de cohorte prospectivo fue investigar la asociación entre el tiempo de pantalla prolongado y la incidencia de síndrome de ojo seco en una cohorte de estudiantes universitarios libres de la condición al inicio del estudio").

2. Métodos:

- **Diseño del Estudio:** Describir claramente que se trata de un estudio de cohorte prospectivo.
- **Población de Estudio y Reclutamiento:** Definir la población de la cual se reclutó la cohorte (ej. estudiantes de pregrado y posgrado de [Nombre de la Universidad]). Describir los criterios de inclusión y exclusión, asegurando que todos los participantes estén libres de síndrome de ojo seco al inicio del estudio. Describir el proceso de reclutamiento. Mencionar el tamaño de la cohorte inicial.
- **Medición de la Exposición:** Describir cómo se midió el tiempo de pantalla al inicio del estudio (ej. cuestionario autoadministrado preguntando horas promedio por día frente a diferentes dispositivos). Definir la categoría de exposición de interés (ej. "tiempo de pantalla prolongado" definido como ≥ 4 horas/día, comparado con "< 4 horas/día"). Mencionar la fiabilidad o validación del instrumento de medición si aplica.
- **Medición del Desenlace:** Describir cómo se definió y midió el síndrome de ojo seco durante el seguimiento (ej. diagnóstico clínico por un oftalmólogo, o mediante un cuestionario validado de síntomas de ojo seco con un punto de corte definido). Explicar la frecuencia y el método de seguimiento (ej. cuestionarios de seguimiento cada 6 meses, examen oftalmológico al final del seguimiento).
- **Variables:** Definir claramente la variable de exposición (Tiempo de Pantalla: Prolongado/No Prolongado) y la variable de desenlace (Desarrollo de Síndrome de Ojo Seco: Sí/No durante el seguimiento). Identificar y describir las posibles **variables confusoras** que se midieron al inicio del estudio (ej. edad, sexo, uso de lentes de contacto, tabaquismo, condiciones médicas preexistentes, uso de ciertos medicamentos, tiempo de lectura no digital).
- **Seguimiento:** Describir la duración del seguimiento y las estrategias utilizadas para mantener el contacto con los participantes y minimizar la **pérdida de seguimiento** (ej. recordatorios por correo electrónico, llamadas telefónicas, incentivos). Reportar la proporción de participantes que completaron el seguimiento.

- **Consideraciones Éticas:** Declarar la aprobación por el Comité de Ética y el proceso de **Consentimiento Informado** detallado para un estudio prospectivo.
- **Análisis Estadístico:** Describir el plan de análisis:
  - **Estadística Descriptiva:** Describir las características de la cohorte total y por grupo de exposición (tiempo de pantalla prolongado vs. no prolongado) al inicio del estudio (Tabla 1).
  - **Cálculo de Incidencia:** Calcular la **incidencia acumulada** de síndrome de ojo seco durante el período de seguimiento en el grupo expuesto (I_e) y en el grupo no expuesto (I_o). Si el seguimiento fue variable para algunos participantes, se podría calcular la **tasa de incidencia** (número de casos nuevos por persona-tiempo en riesgo).
  - **Análisis Bivariado (Exploración de la Asociación):** Presentar los datos en una tabla 2x2 (Tabla 2) mostrando el número de participantes en cada grupo de exposición que desarrollaron o no el desenlace. Calcular el **Riesgo Relativo (RR)** crudo (I_e/I_o). Calcular el **Intervalo de Confianza del 95% (IC 95%)** para el RR y el **valor p**. Opcionalmente, calcular la **Diferencia de Riesgos (DR)** cruda ($I_e - I_o$).
  - **Análisis Multivariado (Control de Confusores):** Describir cómo se ajustó por las variables confusoras. Esto se realiza típicamente utilizando **regresión de Poisson** o **regresión logística** (si el desenlace es relativamente común) para calcular el **RR ajustado**, o **regresión de Cox** para calcular el **Hazard Ratio (HR)** si se analiza el tiempo hasta el evento. Presentar el RR o HR ajustado, su IC 95% y valor p. Esto se suele presentar en una **Tabla 3: Riesgo Relativo (Crudo y Ajustado) para la Asociación entre Tiempo de Pantalla Prolongado e Incidencia de Síndrome de Ojo Seco**.

3. Resultados:

- Presentar los hallazgos de manera clara y objetiva.
- Describir el tamaño de la cohorte inicial y final, y la proporción de pérdida de seguimiento.

- Presentar la **Tabla 1** con las características de la cohorte al inicio, por grupo de exposición.
- Presentar la **Tabla 2x2** (Tabla 2) mostrando la incidencia del desenlace en cada grupo de exposición.
- Presentar los resultados del análisis bivariado: la **incidencia** en cada grupo, el valor del **RR crudo**, su IC 95% y el valor p. Opcionalmente, la DR cruda.
- Presentar los resultados del análisis multivariado: el **RR o HR ajustado**, su IC 95% y el valor p, indicando por qué variables se ajustó.
- Complementar las tablas con texto que describa los hallazgos principales (ej. "La incidencia de síndrome de ojo seco fue mayor en el grupo con tiempo de pantalla prolongado (X%) que en el grupo con menor tiempo de pantalla (Y%)", "El RR crudo fue de Z (IC 95% A-B, p=C)", "Después de ajustar por [variables ajustadas], el RR ajustado fue de W (IC 95% D-E, p=F)").

4. Discusión:

- Interpretar los resultados en el contexto de la pregunta de investigación y la literatura.
- **Párrafo 1:** Resumir los hallazgos principales, enfocándose en el RR o HR ajustado y su significancia estadística.
- **Párrafo 2 (y siguientes):** Discutir la interpretación del RR o HR (ej. "Los estudiantes con tiempo de pantalla prolongado tuvieron W veces el riesgo de desarrollar síndrome de ojo seco en 1 año en comparación con aquellos con menor tiempo de pantalla, después de considerar [variables ajustadas]"). Comparar estos hallazgos con estudios previos. Discutir posibles mecanismos biológicos que expliquen la asociación (ej. menor frecuencia de parpadeo, evaporación de la lágrima).
- **Fortalezas del Estudio:** Mencionar los aspectos positivos (ej. diseño prospectivo que permite establecer secuencia temporal y calcular incidencia, reclutamiento de una cohorte libre del desenlace, medición de exposición al inicio, control de confusores importantes).

- **Limitaciones:** Discutir las limitaciones del estudio:
  - **Pérdida de Seguimiento:** Evaluar si la proporción de pérdida fue significativa y si hubo diferencias entre grupos de exposición, lo que podría introducir sesgo.
  - **Medición de la Exposición:** ¿Fue precisa la medición del tiempo de pantalla (autoinforme)? ¿Varió significativamente durante el seguimiento?
  - **Medición del Desenlace:** ¿Fue precisa la definición y medición del ojo seco? ¿Se detectaron todos los casos nuevos?
  - **Variables Confusoras Residuales:** Reconocer que puede haber confusores no medidos o no controlados adecuadamente.
  - **Generalizabilidad:** ¿La cohorte de estudiantes universitarios es representativa de otros grupos de edad o poblaciones?
- **Recomendaciones Futuras:** Sugerir futuras investigaciones para abordar las limitaciones (ej. estudios con mediciones más objetivas de tiempo de pantalla y ojo seco, estudios con seguimientos más largos, investigación de intervenciones preventivas).
- **Implicaciones:** Discutir las implicaciones clínicas o de salud pública de los hallazgos (ej. necesidad de educación sobre higiene visual, recomendaciones sobre pausas en el uso de pantallas).

5. Conclusión:

- Resumir la principal conclusión del estudio sobre la asociación entre el tiempo de pantalla prolongado y la incidencia de síndrome de ojo seco, reiterando el RR o HR ajustado y su significancia. Aunque es un estudio observacional, la claridad temporal de un diseño prospectivo permite una inferencia causal más fuerte que en estudios transversales o de casos y controles.

Secciones Finales: (Agradecimientos, Financiación, Ética, Consentimiento, Datos, Abreviaciones, Referencias - similar a ejemplos anteriores)

Este ejemplo ilustra cómo un estudio de cohorte prospectivo nos permite seguir el curso natural de los eventos, desde la exposición hasta el desenlace, proporcionando evidencia robusta para cuantificar el riesgo y fortalecer la

inferencia causal en la investigación epidemiológica. Es un diseño poderoso, aunque exigente en términos de tiempo y recursos, fundamental para comprender la etiología de las enfermedades en las poblaciones.

4.2.5. *Estudios Ecológicos*

Hemos recorrido el camino de la epidemiología observacional desde la descripción individual y grupal (reportes y series de casos), pasando por la instantánea poblacional (estudios transversales) y la mirada hacia atrás desde el desenlace (casos y controles), hasta la observación del futuro desde la exposición (estudios de cohorte). Cada diseño nos ha ofrecido una perspectiva única, con sus propias fortalezas y limitaciones. Ahora, nos adentramos en un tipo de estudio que mira el panorama completo, analizando datos a nivel de grupos o poblaciones enteras, no de individuos: los **Estudios Ecológicos**.

Imagina que quieres explorar si el consumo per cápita de chocolate en diferentes países se relaciona con el número de premios Nobel ganados por esos países. No vas a preguntar a cada persona cuánto chocolate come y cuántos premios Nobel tiene. En cambio, buscas datos agregados: el consumo promedio de chocolate por país y el número total de premios Nobel por país. Luego, comparas estas cifras entre diferentes países. Este es el enfoque de un estudio ecológico: la unidad de análisis no es el individuo, sino el **grupo** (un país, una región, una ciudad, una escuela).

Un estudio ecológico es un tipo de estudio **observacional** en el que la unidad de análisis es una **población** o un **grupo de personas**, en lugar de individuos. Se utilizan datos agregados de estas poblaciones para examinar la relación entre el nivel promedio de exposición a un factor y la frecuencia de un desenlace de salud. Es como si viéramos un bosque desde un helicóptero: podemos observar patrones generales en diferentes áreas (grupos), pero no podemos distinguir las características o experiencias de cada árbol individual.

Tipos:

Los estudios ecológicos pueden clasificarse según la dimensión que exploran:

- **Estudios Ecológicos Multigrupo (o Geográficos):** Comparan la frecuencia de un desenlace de salud entre diferentes grupos o poblaciones en un **único punto o período en el tiempo**. El ejemplo del consumo de chocolate y los premios Nobel es un estudio ecológico geográfico, comparando diferentes países en un momento dado.

- **Estudios Ecológicos de Tendencias Temporales (o Series de Tiempo):** Examinan los cambios en la frecuencia de un desenlace de salud en una **misma población** a lo largo del **tiempo**, y cómo estos cambios se correlacionan con cambios en el nivel de exposición promedio de esa población a lo largo del tiempo. Por ejemplo, analizar si la tasa de tabaquismo en un país ha disminuido a lo largo de las décadas y si esa disminución se correlaciona con una disminución en las tasas de cáncer de pulmón en el mismo país.

Recolección y Análisis de Datos:

La característica distintiva de los estudios ecológicos es que utilizan **datos agregados** o grupales. Estos datos suelen provenir de fuentes secundarias, como:

- Registros de estadísticas vitales (tasas de mortalidad, natalidad).
- Censos poblacionales.
- Registros de enfermedades (registros de cáncer, sistemas de vigilancia epidemiológica).
- Encuestas de salud a nivel poblacional.
- Datos ambientales (niveles promedio de contaminación del aire en una ciudad).
- Datos de consumo o ventas (ventas per cápita de cigarrillos, consumo de alimentos).

El **análisis de datos** implica comparar las medidas resumen de exposición y desenlace entre los diferentes grupos o a lo largo del tiempo. Esto a menudo se realiza calculando coeficientes de correlación o utilizando modelos de regresión donde la unidad de análisis es el grupo. Por ejemplo, se podría calcular la correlación entre el porcentaje de fumadores en diferentes provincias y la tasa de mortalidad por cáncer de pulmón en esas mismas provincias.

Usos y Aplicaciones:

Los estudios ecológicos son particularmente útiles para:

- **Generar hipótesis:** Son una forma rápida y económica de explorar posibles asociaciones entre exposiciones y desenlaces a nivel poblacional,

sugiriendo áreas para una investigación más detallada con diseños individuales.

- **Examinar tendencias a gran escala:** Permiten observar patrones geográficos o temporales en la distribución de enfermedades y exposiciones que podrían no ser evidentes a nivel individual.
- **Evaluar el impacto de políticas o intervenciones a nivel poblacional:** Por ejemplo, evaluar si una política de salud pública implementada a nivel nacional se asocia con cambios en las tasas de una enfermedad.
- **Investigar exposiciones para las cuales los datos a nivel individual son difíciles o imposibles de obtener:** Como la calidad del aire en una ciudad o las políticas de salud de un país.

Ventajas y Desventajas:

Las principales **ventajas** de los estudios ecológicos son:

- **Rapidez y bajo costo:** Utilizan datos ya existentes, lo que los hace relativamente rápidos y económicos de realizar.
- **Permiten investigar exposiciones a nivel grupal:** Son adecuados para factores que varían a nivel de grupo (ej. políticas, medio ambiente).
- **Útiles para generar hipótesis:** Pueden identificar patrones que justifican estudios más rigurosos.

Sin embargo, sus **desventajas** son significativas y limitan seriamente su capacidad para establecer causalidad:

- **La Falacia Ecológica:** Esta es la principal limitación y el gran peligro de los estudios ecológicos. Ocurre cuando se infiere que una asociación observada a nivel de grupo necesariamente se aplica a nivel individual. Por ejemplo, si encontramos que las áreas con mayor venta de helados tienen mayores tasas de ahogamiento, sería una falacia ecológica concluir que comer helado causa ahogamiento. La verdadera explicación es que ambos eventos están asociados a un tercer factor: el clima cálido. No podemos saber si los individuos que compran helado son los mismos que se ahogan. Los datos agregados ocultan la variabilidad individual.

- **No permiten establecer causalidad a nivel individual:** Debido a la falacia ecológica, los estudios ecológicos no pueden determinar si la exposición y el desenlace están asociados en los mismos individuos.
- **Susceptibles a sesgos y confusores no controlados:** Al usar datos agregados, es difícil controlar el efecto de variables confusoras a nivel individual.
- **Dependencia de la calidad de los datos agregados:** La precisión de los resultados depende de la fiabilidad y comparabilidad de las estadísticas utilizadas.

Variables confusoras y sesgo:

- **Variables Confusoras:** El control de confusores es un desafío importante en los estudios ecológicos. Una asociación observada a nivel de grupo puede ser espuria debido a diferencias en la distribución de variables confusoras entre los grupos, que no pueden ser controladas adecuadamente con datos agregados.
- **Sesgos:** Además de la falacia ecológica (que es un tipo de sesgo de inferencia), los estudios ecológicos pueden estar sujetos a sesgos relacionados con la calidad y comparabilidad de los datos agregados (ej. diferencias en los sistemas de registro de enfermedades entre regiones o a lo largo del tiempo).

Consideraciones Éticas:

Aunque los estudios ecológicos utilizan datos agregados y generalmente no involucran contacto directo con individuos, aún existen consideraciones éticas relacionadas con el uso de datos poblacionales, la privacidad y la posible interpretación errónea de los resultados que podría llevar a estigmatización de grupos. La transparencia sobre las fuentes de datos y las limitaciones del diseño es fundamental.

En conclusión, los estudios ecológicos son una herramienta útil para explorar patrones de salud a gran escala, generar hipótesis y evaluar el impacto de intervenciones a nivel poblacional, aprovechando la disponibilidad de datos agregados. Sin embargo, su incapacidad para vincular la exposición

y el desenlace a nivel individual y el riesgo inherente de la falacia ecológica limitan severamente su capacidad para establecer relaciones causales. Son un punto de partida en la investigación etiológica, que debe ser seguido por diseños de estudio a nivel individual para confirmar o refutar las asociaciones sugeridas.

4.2.6. Ejemplo de Estudio Ecológico

En un Colombia, observas que las tasas de mortalidad por enfermedades cardiovasculares varían considerablemente entre diferentes regiones. Sospechas que esta variación podría estar relacionada con diferencias en la calidad del aire, específicamente los niveles de partículas finas ($PM_{2.5}$), ya que sabes que la contaminación del aire puede afectar la salud cardiovascular. No tienes datos individuales sobre la exposición a $PM_{2.5}$ de cada persona ni sobre su historial de salud cardiovascular. Sin embargo, sí tienes datos públicos sobre los niveles promedio anuales de PM_2□$_5$ en las principales ciudades o regiones del país y las tasas de mortalidad por enfermedades cardiovasculares reportadas para esas mismas áreas. Puedes utilizar estos datos agregados para explorar si existe una asociación a nivel de grupo.

Aquí te presento un ejemplo estructurado de cómo abordarías la investigación y el reporte de un estudio ecológico sobre esta pregunta:

Caso Ejemplo: Niveles Promedio de Partículas Finas ($PM_{2.5}$) y Tasas de Mortalidad por Enfermedades Cardiovasculares: Un Estudio Ecológico Multigrupo

1. Formulación del Problema y Definición de Hipótesis

Nuestra observación de la variación en las tasas de mortalidad cardiovascular y la sospecha sobre la contaminación del aire nos lleva a la pregunta:

- **Pregunta de Investigación:** ¿Existe una asociación a nivel de región entre los niveles promedio anuales de partículas finas ($PM_{2.5}$) y las tasas de mortalidad por enfermedades cardiovasculares en las principales regiones de [Nombre del País] en el año [Año Específico]?

Evaluemos esta pregunta con los criterios FINER, adaptados a la unidad de análisis grupal:

- **Factible:** ¿Podemos obtener datos agregados fiables sobre los niveles promedio de $PM_{2.5}$ y las tasas de mortalidad cardiovascular para las regiones de interés y el año específico? (Sí, si existen registros públicos o bases de datos accesibles).
- **Interesante:** ¿Es interesante para las autoridades de salud pública y los investigadores explorar esta asociación a nivel regional? (Sí, puede guiar políticas ambientales y de salud).
- **Novedoso:** ¿Se ha estudiado esta asociación en este país o con este nivel de agregación de datos? (Podría ser novedoso dependiendo del contexto).
- **Ético:** ¿El uso de datos agregados públicos plantea riesgos significativos para la privacidad de los individuos? (Generalmente no, ya que los datos no son identificables a nivel personal).
- **Relevante:** ¿Los hallazgos podrían tener implicaciones para la salud pública, como la necesidad de implementar medidas para reducir la contaminación del aire en ciertas regiones? (Sí, tiene alta relevancia para la salud pública).

Si la pregunta es FINER, planteamos nuestras hipótesis a nivel de grupo:

- **H_0:** No existe una asociación estadísticamente significativa a nivel de región entre los niveles promedio anuales de partículas finas ($PM_{2.5}$) y las tasas de mortalidad por enfermedades cardiovasculares en las principales regiones de [Nombre del País] en el año [Año Específico].
- **H□:** Existe una asociación estadísticamente significativa a nivel de región entre los niveles promedio anuales de partículas finas ($PM_{2.5}$) y las tasas de mortalidad por enfermedades cardiovasculares en las principales regiones de [Nombre del País] en el año [Año Específico].

2. Selección del Nivel de Significancia (α)

Definimos nuestro umbral de significancia, comúnmente α=0.05.

Estructura del Artículo (Adaptada para Estudio Ecológico):

Siguiendo una estructura similar a la de los ejemplos anteriores:

Título:

- Debe reflejar la exposición, el desenlace, la unidad de análisis y el diseño del estudio.
- *Ejemplo*: "Asociación Ecológica entre Niveles de $PM_{2.5}$ y Mortalidad Cardiovascular en Regiones de [Nombre del País]".

Autores y Afiliaciones: (Similar a ejemplos anteriores)

Resumen:

- **Contexto y Objetivo:** Contexto del problema (contaminación del aire, salud cardiovascular, variación regional) y el objetivo del estudio (explorar la asociación a nivel de región).
- **Metodología:** Breve descripción del diseño (ecológico multigrupo), unidad de análisis (regiones), fuentes de datos (registros públicos), variables ($PM_{2.5}$ promedio, tasa de mortalidad) y análisis estadístico (correlación, regresión).
- **Resultados:** Presentar el rango de niveles de $PM_{2.5}$ y tasas de mortalidad por región, y el resultado del análisis de asociación (ej. coeficiente de correlación, resultado de regresión, p-valor).
- **Conclusión:** Resumir la principal conclusión sobre la asociación observada a nivel de grupo y su posible implicación.

Palabras clave: (Similar a ejemplos anteriores, incluyendo la exposición, el desenlace, la unidad de análisis y el diseño).

1. Introducción:

- **Párrafo 1:** Introducir las enfermedades cardiovasculares como un problema de salud pública importante y mencionar algunos de sus factores de riesgo conocidos a nivel individual.
- **Párrafo 2:** Discutir la contaminación del aire, específicamente las partículas finas ($PM_{2.5}$), como un posible factor de riesgo ambiental para la salud cardiovascular, basándose en evidencia previa (posiblemente de estudios individuales).

- **Párrafo 3:** Mencionar que las tasas de enfermedades cardiovasculares y los niveles de contaminación del aire varían geográficamente. Explicar que un estudio ecológico es un diseño adecuado para explorar si esta variación a nivel de grupo está relacionada.
- **Último Párrafo:** Indicar el objetivo del estudio (ej. "El objetivo de este estudio ecológico fue examinar la asociación entre los niveles promedio anuales de $PM_{2.5}$ y las tasas de mortalidad por enfermedades cardiovasculares en las principales regiones de [Nombre del País]").

2. Metodología:

- **Diseño del Estudio:** Describir claramente que se trata de un estudio ecológico multigrupo.
- **Unidad de Análisis:** Definir la unidad de análisis (ej. las 20 principales regiones administrativas o ciudades con datos disponibles en [Nombre del País]).
- **Fuentes de Datos:** Especificar las fuentes de los datos agregados (ej. Ministerio de Salud para tasas de mortalidad, Ministerio de Medio Ambiente o estaciones de monitoreo para niveles de $PM_{2.5}$). Mencionar el año específico de los datos.
- **Variables:** Definir claramente la variable de exposición a nivel de grupo (ej. Nivel promedio anual de $PM_{2.5}$ en $\mu g/m^3$) y la variable de desenlace a nivel de grupo (ej. Tasa de mortalidad por enfermedades cardiovasculares por 100.000 habitantes). Identificar y describir posibles **variables confusoras a nivel de grupo** que podrían estar asociadas tanto con los niveles de $PM_{2.5}$ como con la mortalidad cardiovascular (ej. porcentaje de población urbana, nivel socioeconómico promedio de la región, acceso a servicios de salud, prevalencia de tabaquismo a nivel regional si hay datos disponibles).
- **Consideraciones Éticas:** Mencionar que se utilizaron datos agregados disponibles públicamente y que no se requirió consentimiento individual o aprobación de comité de ética para el uso de datos anonimizados a nivel de grupo (si este es el caso y está justificado éticamente).

- **Análisis Estadístico:** Describir el plan de análisis:

  - **Estadística Descriptiva:** Describir las variables de exposición y desenlace para cada unidad de análisis (región). Presentar el rango, media y DS de los niveles de $PM_{2.5}$ y las tasas de mortalidad cardiovascular en las regiones estudiadas. Esto se puede presentar en una **Tabla 1: Características de las Regiones Estudiadas**.

  - **Análisis de Asociación:** Explorar la relación entre el nivel promedio de $PM_{2.5}$ y la tasa de mortalidad cardiovascular utilizando:

    - **Coeficiente de Correlación:** Calcular el coeficiente de correlación de Pearson o Spearman para evaluar la fuerza y dirección de la asociación lineal entre las dos variables a nivel de región. Presentar el coeficiente y su valor p. Esto se puede visualizar con un **Gráfico de Dispersión** (Figura 1) donde cada punto representa una región, con el nivel de $PM_{2.5}$ en un eje y la tasa de mortalidad en el otro.

    - **Modelos de Regresión Ecológica:** Utilizar un modelo de regresión lineal simple o múltiple (si se incluyen confusores a nivel de grupo) donde la tasa de mortalidad es la variable dependiente y el nivel de $PM_{2.5}$ es la variable independiente. Presentar los coeficientes de regresión, sus intervalos de confianza y valores p.

3. Resultados:

- Presentar los hallazgos de manera clara y objetiva, apoyándose en tablas y figuras.
- Presentar la **Tabla 1** con las características descriptivas de las regiones.
- Presentar el **Gráfico de Dispersión** (Figura 1) mostrando la relación entre $PM_{2.5}$ y tasas de mortalidad por región.
- Presentar el resultado del análisis de correlación (coeficiente y p-valor).
- Presentar los resultados del modelo de regresión, incluyendo los coeficientes y valores p, tanto crudos como ajustados por confusores a nivel de grupo si se realizó.

- Complementar las tablas y figuras con texto que describa los hallazgos principales (ej. "Se observó una correlación positiva significativa entre los niveles promedio de $PM_{2.5}$ y las tasas de mortalidad cardiovascular a nivel regional (r = X, p = Y)", "Por cada aumento de Z μg/m^3 en los niveles promedio de $PM_{2.5}$, la tasa de mortalidad cardiovascular aumentó en W por 100.000 habitantes, después de ajustar por [confusores a nivel de grupo]").

Figura 1. Gráfico de Dispersión: Relación entre Niveles Promedio de $PM_{2.5}$ y Tasas de Mortalidad por Enfermedades Cardiovasculares por Región.

(Cada punto representa una región. El eje X muestra el nivel promedio anual de $PM_{2.5}$ y el eje Y muestra la tasa de mortalidad por enfermedades cardiovasculares por 100.000 habitantes).

4. Discusión:

- Interpretar los resultados en el contexto de la pregunta de investigación y la literatura.
- **Párrafo 1:** Resumir el hallazgo principal: la asociación observada a nivel de grupo entre $PM_{2.5}$ y mortalidad cardiovascular.
- **Párrafo 2 (y siguientes):** Discutir la interpretación de esta asociación. **Es fundamental ser cauteloso y no hacer inferencias a nivel individual.** Explicar que este hallazgo sugiere que las regiones con mayor contaminación del aire tienden a tener mayores tasas de mortalidad cardiovascular, lo cual es consistente con la hipótesis de que la contaminación del aire es un factor de riesgo. Mencionar estudios a nivel individual que apoyan esta hipótesis biológicamente plausible.
- **Fortalezas del Estudio:** Mencionar los aspectos positivos (ej. uso de datos a nivel de grupo para explorar tendencias poblacionales, bajo costo, rapidez, utilidad para generar hipótesis).
- **Limitaciones:** Discutir las limitaciones cruciales de un estudio ecológico:
  - **La Falacia Ecológica:** Explicar claramente que la asociación observada a nivel de grupo no garantiza que los individuos más expuestos dentro de esas regiones sean los que mueren por enfermedades cardiovasculares.

- **Control Limitado de Confusores Individuales:** No se puede controlar por factores de riesgo a nivel individual (ej. tabaquismo individual, dieta, genética) que podrían explicar la asociación observada a nivel de grupo. Solo se pueden controlar confusores para los cuales se tienen datos agregados a nivel de grupo, y estos pueden no ser suficientes.
- **Calidad y Comparabilidad de los Datos Agregados:** Discutir posibles problemas con la precisión o estandarización de los datos de $PM_{2.5}$ o mortalidad entre regiones o fuentes.
- **Medición Imprecisa de la Exposición:** El nivel promedio de $PM_{2.5}$ en una región puede no reflejar la exposición real de los individuos dentro de esa región.

- **Recomendaciones Futuras:** Sugerir futuras investigaciones para abordar las limitaciones (ej. realizar estudios de cohorte o casos y controles a nivel individual para confirmar la asociación, utilizar mediciones de exposición a nivel individual, controlar por confusores individuales).
- **Implicaciones:** Discutir las posibles implicaciones para la salud pública, como la necesidad de investigar más a fondo en las regiones con alta contaminación y mortalidad, y la posible justificación para políticas de reducción de la contaminación del aire.

5. Conclusión:

- Resumir la principal conclusión del estudio, indicando que se encontró una asociación a nivel de región entre los niveles de $PM_{2.5}$ y las tasas de mortalidad cardiovascular. Reiterar que este hallazgo **genera una hipótesis** que debe ser investigada con estudios a nivel individual para confirmar si la asociación existe en las personas.

4.3. Estudios experimentales

Hemos ascendido en nuestra comprensión de los diseños de estudio, pasando de la observación de lo que ocurre naturalmente a la capacidad de tomar una "instantánea" o mirar hacia atrás en el tiempo. Ahora, llegamos a la cima de la evidencia individual: los **Estudios Experimentales**, donde el investigador toma un papel activo, interviniendo para probar una hipótesis. Dentro de este grupo, el diseño más riguroso para determinar una relación de causa y efecto es el **Ensayo Clínico Aleatorizado (ECA)**.

Si los estudios observacionales son como detectives que investigan un crimen después de que ha ocurrido, los ECA son como científicos que diseñan un experimento en un laboratorio controlado para probar si una sustancia causa una reacción específica. La clave está en el **control** y la **manipulación** de la exposición (la intervención).

4.3.1. *Estudios Experimentales: Ensayos Clínicos Aleatorizados*

Los estudios intervencionista, como los ECA, se eligen cuando queremos probar si una **intervención** específica (un tratamiento, una práctica, una estrategia de salud) **causa** un cambio medible en un **desenlace**. La intervención puede ser de naturaleza preventiva (buscando reducir el riesgo de desarrollar una enfermedad en individuos en riesgo) o terapéutica (buscando mejorar el desenlace en individuos que ya tienen una enfermedad).

La característica definitoria de un ECA, y lo que lo sitúa en la cima de la jerarquía de la evidencia para la causalidad (solo superado por las revisiones sistemáticas y metaanálisis de ECA), es el proceso de **aleatorización**.

Las Fases de la Investigación Clínica

El desarrollo de una nueva intervención, especialmente un fármaco, es un proceso largo y riguroso que avanza a través de distintas etapas, conocidas como fases clínicas. Estas fases están diseñadas para evaluar progresivamente la seguridad, la dosificación, la eficacia y los efectos secundarios de la intervención antes de que pueda ser aprobada para su uso generalizado.

- **Fase Preclínica:** Esta etapa ocurre *antes* de que la intervención se pruebe en humanos. Involucra investigación de laboratorio y estudios en animales. El objetivo es entender cómo funciona la intervención, evaluar su seguridad inicial y determinar si es prometedora para pasar a estudios en humanos. Si los resultados preclínicos son favorables, se solicita la aprobación regulatoria (ej. a la FDA en Estados Unidos) para iniciar los ensayos clínicos en personas.
- **Fase I:** Son los primeros estudios en humanos. Generalmente involucran a un pequeño grupo de voluntarios sanos (o a veces pacientes con la enfermedad avanzada si la intervención es muy tóxica). El objetivo principal es evaluar la **seguridad** y la **tolerabilidad** de la intervención,

determinar la dosis segura y cómo el cuerpo absorbe, distribuye, metaboliza y excreta la sustancia (farmacocinética y farmacodinámica). La eficacia no es el objetivo primario en esta fase.

- **Fase II:** Si la Fase I demuestra que la intervención es segura en dosis iniciales, se pasa a la Fase II. Estos estudios involucran a un grupo más grande de personas que *tienen* la enfermedad o condición que se busca tratar o prevenir. El objetivo principal es evaluar la **eficacia preliminar** de la intervención (¿funciona para lo que está diseñada?) y seguir evaluando la seguridad y los efectos secundarios en una población de pacientes. A menudo se comparan diferentes dosis para encontrar la más efectiva con menos efectos secundarios.
- **Fase III:** Si la Fase II muestra evidencia de eficacia y seguridad aceptable, se avanza a la Fase III. Estos son ensayos a gran escala que involucran a cientos o miles de participantes. Son típicamente **ECA** que comparan la nueva intervención con un placebo o con el tratamiento estándar actual. El objetivo es confirmar la **eficacia** de la intervención, monitorear los efectos secundarios en una población más grande y diversa, y recopilar información para permitir que la intervención se use de manera segura. Si los resultados de la Fase III son positivos, se puede solicitar la aprobación regulatoria para comercializar la intervención.
- **Fase IV:** También conocidos como estudios de post-comercialización. Estos estudios se realizan *después* de que la intervención ha sido aprobada y está disponible para el público. El objetivo es monitorear la seguridad a largo plazo, recopilar información sobre los beneficios y riesgos en poblaciones más amplias y diversas, comparar la intervención con otras disponibles, y explorar nuevos usos potenciales.

La mayoría de los ECA que se discuten en el contexto de la evaluación de la eficacia de tratamientos corresponden a la **Fase III**.

Ensayo Controlado Aleatorizado Paralelo

El tipo más común de ECA es el **ECA Paralelo**. En este diseño, los participantes son asignados aleatoriamente a uno de dos o más "brazos" de tratamiento (generalmente un grupo de intervención y un grupo de control) y son seguidos prospectivamente en paralelo para comparar los desenlaces entre los grupos.

Los pasos esenciales para diseñar un ECA paralelo incluyen:

1. **Formular la hipótesis:** Definir claramente qué se espera que la intervención logre en comparación con el control (ej. reducir la mortalidad en un X%).
2. **Definir métodos de reclutamiento:** Establecer quién es elegible para participar.
3. **Definir la intervención:** Describir detalladamente el tratamiento o la práctica que se va a probar.
4. **Definir el grupo de comparación (Control):** Establecer con qué se comparará la intervención.
5. **Determinar el tamaño de la muestra:** Calcular cuántos participantes se necesitan.
6. **Especificar las medidas de desenlace:** Definir cómo se medirá el efecto de la intervención.
7. **Obtener aprobación ética:** Asegurar que el estudio sea moralmente aceptable.
8. **Obtener consentimiento informado:** Asegurar que los participantes entiendan y acepten participar.
9. **Generar y ocultar la secuencia de asignación:** Garantizar la aleatorización.
10. **Indicar si se utiliza cegamiento (*blinding*):** Especificar quién no conoce la asignación del tratamiento.
11. **Realizar análisis por intención de tratar (*intention-to-treat*):** Analizar los datos según la asignación inicial, no según el tratamiento recibido.

Definición de Población, Eventos y Asignación:

- **Población:** Se define una **población de referencia (o objetivo)** a la que se pretenden generalizar los resultados. A partir de esta, se selecciona una **muestra de estudio** que cumple con los **criterios de inclusión y exclusión**.
    - **Criterios de Inclusión:** Características que deben tener los participantes para ser elegibles (ej. edad, diagnóstico específico, gravedad de la enfermedad).

- **Criterios de Exclusión:** Características que impiden la participación (ej. comorbilidades que podrían interferir con la intervención o el desenlace, contraindicaciones para la intervención, incapacidad para dar consentimiento). Establecer criterios estrictos puede aumentar la validez interna (asegurar que el efecto se deba a la intervención), pero limita la **validez externa** (generalizabilidad) a pacientes que no cumplen esos criterios (ej. embarazadas, niños, ancianos, pacientes con múltiples comorbilidades, a menudo subrepresentados en los ECA).

- **Eventos (Desenlaces):** Son las medidas que se utilizan para evaluar el efecto de la intervención. Deben estar **precisamente definidos**, ser **medibles**, **repetibles**, **fiables** y **relevantes** (tanto para clínicos como para pacientes). Se define un **desenlace primario** (el de mayor importancia y en el que se basa el cálculo del tamaño muestral) y puede haber **desenlaces secundarios**. Los desenlaces pueden ser clínicos (mortalidad, complicaciones), de laboratorio, reportados por el paciente (síntomas, calidad de vida) o económicos.

- **Asignación (Aleatorización):** Es el proceso por el cual cada participante tiene la misma probabilidad de ser asignado a cualquiera de los grupos de tratamiento. Su propósito es crear grupos comparables al inicio del estudio, distribuyendo equitativamente las características de los participantes (conocidas y desconocidas) entre los brazos del estudio. Esto minimiza el **sesgo de selección** y el efecto de **variables confusoras**.

- **Aleatorización Simple:** Equivalente a lanzar una moneda por cada participante. Puede resultar en desbalances en el tamaño de los grupos, especialmente en estudios pequeños.

- **Aleatorización por Bloques:** Asegura que el número de participantes en cada brazo sea similar en cualquier punto del reclutamiento. Se crean bloques de un tamaño predefinido con un número igual de asignaciones a cada brazo dentro de cada bloque.

- **Aleatorización Estratificada:** Se utiliza para asegurar que variables pronósticas importantes (confusoras potenciales) estén distribuidas equitativamente entre los grupos. Los participantes se dividen en estratos según la variable (ej. edad > 65 años vs. ≤ 65 años) y luego se aleatoriza dentro de cada estrato, a menudo usando aleatorización por bloques dentro de cada estrato.

- **Aleatorización Factorial:** Permite evaluar dos o más intervenciones simultáneamente en un solo ECA. Los participantes son aleatorizados a combinaciones de las intervenciones (ej. Intervención A + Intervención B, Intervención A + Placebo B, Placebo A + Intervención B, Placebo A + Placebo B).
- **Diseño Cruzado (*Cross-over*):** En este diseño, cada participante recibe secuencialmente la intervención y el control en diferentes períodos de tiempo, actuando cada participante como su propio control. Se requiere un período de "lavado" (*washout*) entre los tratamientos para evitar efectos residuales. Es eficiente para estudios pequeños y condiciones crónicas estables, pero no es adecuado para desenlaces que son curados por la primera intervención o para condiciones que cambian rápidamente.

Placebo/Control y Seguimiento:

- **Placebo/Control:** El grupo de comparación (control) es fundamental. Puede recibir:
- **Placebo:** Una sustancia o procedimiento inerte que se ve, siente y administra de manera idéntica a la intervención activa. Se utiliza para diferenciar los efectos específicos de la intervención de los efectos inespecíficos (ej. efecto placebo, atención recibida). Éticamente, solo es aceptable usar placebo si no existe un tratamiento estándar probado para la condición, o si el uso de placebo es necesario metodológicamente y no expone a los participantes a riesgos serios o irreversibles (según la Declaración de Helsinki).
- **Tratamiento Estándar (Control Activo):** El mejor tratamiento disponible actualmente. Es la comparación ética cuando existe un tratamiento probado.
- **No Tratamiento:** Cuando no hay intervención activa ni placebo.
- **Seguimiento:** Los participantes en todos los brazos del estudio son seguidos prospectivamente durante un período predefinido para documentar la aparición de los desenlaces de interés. El seguimiento debe ser estandarizado y riguroso en todos los grupos para minimizar el **sesgo de detección** o **seguimiento**.

Análisis e Interpretación:

El análisis primario en un ECA compara la frecuencia de los desenlaces entre el grupo de intervención y el grupo de control. Dependiendo del tipo de desenlace, se pueden calcular:

- **Riesgo Relativo (RR)** o **Hazard Ratio (HR)** para desenlaces binarios (ej. mortalidad, desarrollo de enfermedad).
- **Diferencia de Riesgos (DR)** para desenlaces binarios.
- **Diferencia de Medias** para desenlaces continuos (ej. cambio en la presión arterial).

Estos estadísticos se acompañan de sus **Intervalos de Confianza del 95% (IC 95%)** y **valores p** para evaluar la magnitud, dirección y significancia estadística del efecto de la intervención.

El análisis principal en un ECA debe ser por **intención de tratar (ITT)**. Esto significa que los participantes se analizan en el grupo al que fueron asignados aleatoriamente, independientemente de si realmente recibieron o completaron la intervención. El análisis ITT preserva los beneficios de la aleatorización y proporciona una estimación más realista del efecto de la intervención en la práctica clínica.

La interpretación se centra en si la diferencia observada en los desenlaces entre los grupos es estadísticamente significativa ($p < \alpha$) y clínicamente relevante. Si el RR o HR es significativamente diferente de 1 (o la diferencia de medias es significativamente diferente de 0), y el IC 95% no incluye el valor nulo (1 para RR/HR, 0 para diferencia de medias), se rechaza la H_0 y se concluye que la intervención tuvo un efecto.

Ventajas y Desventajas:

Las principales **ventajas** de los ECA:

- **Mayor capacidad para establecer causalidad:** La aleatorización es la mejor herramienta para controlar variables confusoras (conocidas y desconocidas), minimizando el sesgo de selección y fortaleciendo la inferencia causal.

- **Minimizan sesgos:** El cegamiento (si es posible) reduce el sesgo de detección y performance.
- Permiten calcular la **incidencia** y el **Riesgo Relativo** (o medidas de efecto similares).
- Proporcionan la evidencia más sólida para evaluar la **eficacia y seguridad** de nuevas intervenciones.

Las **desventajas** de los ECA:

- **Costosos y requieren mucho tiempo:** Especialmente para desenlaces raros o con largos períodos de latencia.
- **Limitaciones éticas:** No siempre es ético aleatorizar a las personas a una exposición potencialmente dañina o negarles un tratamiento estándar probado.
- **Limitaciones de generalizabilidad (validez externa):** Los criterios de inclusión/exclusión estrictos y el contexto controlado del estudio pueden limitar la aplicabilidad de los resultados a la práctica clínica habitual o a poblaciones de pacientes más diversas.
- **Pueden no reflejar la efectividad en el "mundo real":** La adherencia al tratamiento y el manejo en un ECA pueden ser diferentes a los de la práctica clínica habitual.
- **Pérdida de seguimiento:** Puede introducir sesgo si es diferencial entre grupos.
- **Contaminación:** Los participantes del grupo control pueden acceder a la intervención, diluyendo el efecto observado.

Consideraciones Éticas:

Los ECA plantean importantes consideraciones éticas:

- **Comité de Ética de Investigación:** La aprobación por un comité de ética es obligatoria.
- **Consentimiento Informado:** Es fundamental obtener un consentimiento informado válido de todos los participantes, asegurando que

comprendan el estudio, los riesgos, beneficios y alternativas, y que su participación sea voluntaria.

- ***Equipoise* Clínico (*Clinical Equipoise*):** Existe *equipoise* clínico cuando hay una genuina incertidumbre en la comunidad médica sobre cuál de los tratamientos que se comparan es superior. No es ético aleatorizar si ya se sabe que un tratamiento es claramente mejor o peor que otro.

En resumen, los ECA son el estándar de oro para evaluar la eficacia y seguridad de las intervenciones debido a su capacidad para controlar el sesgo y la confusión a través de la aleatorización. Aunque presentan desafíos éticos, logísticos y de generalizabilidad, la evidencia que proporcionan es fundamental para la toma de decisiones en salud y se sitúa en la cima de la jerarquía de la evidencia individual.

4.3.2. Ejemplo de Ensayo Clínico Aleatorizado

Imagina que se ha desarrollado un nuevo fármaco experimental, llamémoslo "Cardio-Plus", que se cree que podría reducir la presión arterial de manera más efectiva y con menos efectos secundarios que el tratamiento estándar actual para la hipertensión. Para confirmar esta hipótesis con el mayor rigor científico posible, la herramienta ideal es un ECA.

Aquí te presento un ejemplo estructurado de cómo se diseñaría y reportaría un ECA para evaluar la eficacia y seguridad de Cardio-Plus:

Caso Ejemplo: Eficacia y Seguridad del Fármaco Experimental Cardio-Plus para la Reducción de la Presión Arterial en Pacientes con Hipertensión Arterial Esencial: Un Ensayo Clínico Aleatorizado, Doble Ciego, Controlado con Placebo

1. Formulación del Problema y Definición de Hipótesis

Nuestra pregunta central es si Cardio-Plus es mejor que un placebo para reducir la presión arterial:

- **Pregunta de Investigación:** En pacientes adultos con hipertensión arterial esencial, ¿es el tratamiento con el fármaco experimental Cardio-Plus más efectivo que un placebo para reducir la presión arterial sistólica después de 12 semanas de tratamiento?

Aplicamos los criterios FINER para evaluar la pregunta y la factibilidad del ECA:

- **Factible:** ¿Podemos reclutar pacientes con hipertensión esencial que cumplan los criterios? ¿Podemos administrar el fármaco experimental y el placebo de manera segura y cegada? ¿Podemos medir la presión arterial de forma estandarizada? ¿Tenemos los recursos (personal, instalaciones, financiación) para llevar a cabo un seguimiento de 12 semanas? ¿Podemos obtener la aprobación regulatoria y ética? (Sí, si la infraestructura y los recursos están disponibles).
- **Interesante:** ¿Es interesante para la comunidad médica y los pacientes tener un nuevo tratamiento potencialmente más efectivo para la hipertensión? (Sí, la hipertensión es una condición prevalente y un nuevo tratamiento sería de gran interés).
- **Novedoso:** ¿Es Cardio-Plus un fármaco nuevo o con un mecanismo de acción novedoso? ¿Se ha probado ya su eficacia y seguridad en un ECA riguroso? (Sí, asumimos que es un fármaco en desarrollo que requiere esta evaluación).
- **Ético:** ¿Es ético comparar Cardio-Plus con placebo si existe un tratamiento estándar probado para la hipertensión? (Esto es una consideración crítica. Podría ser ético si el ECA incluye un tercer brazo con el tratamiento estándar, o si se reclutan pacientes que no responden o no toleran el tratamiento estándar, o si el placebo se usa como adyuvante al tratamiento estándar en ambos brazos. Para simplificar el ejemplo, asumiremos que es ético en este contexto particular, quizás porque se enfoca en pacientes con hipertensión leve que no requieren tratamiento inmediato según guías, o se compara como terapia adyuvante).
- **Relevante (Relevante):** ¿Los resultados podrían cambiar la práctica clínica y mejorar el manejo de la hipertensión? (Sí, un fármaco más efectivo y seguro tendría un impacto significativo).

Si la pregunta es FINER y las consideraciones éticas se abordan adecuadamente, planteamos nuestras hipótesis:

- **H_0:** No existe una diferencia estadísticamente significativa en la reducción de la presión arterial sistólica después de 12 semanas entre los pacientes tratados con Cardio-Plus y los pacientes que reciben placebo, en pacientes adultos con hipertensión arterial esencial.

- **H_a:** El tratamiento con el fármaco experimental Cardio-Plus resulta en una mayor reducción estadísticamente significativa de la presión arterial sistólica después de 12 semanas en comparación con el placebo, en pacientes adultos con hipertensión arterial esencial.

2. Selección del Nivel de Significancia (α)

Definimos nuestro umbral de significancia, comúnmente α=0.05.

Estructura del Artículo (Reporte de un ECA):

El reporte de un ECA sigue guías estandarizadas como la declaración CONSORT (*Consolidated Standards of Reporting Trials*) para asegurar transparencia y completitud. La estructura típica es:

Título:

- Debe incluir la intervención, el control, el desenlace principal, la población y el diseño del estudio.
- *Ejemplo*: "Eficacia y Seguridad de Cardio-Plus vs. Placebo en Pacientes con Hipertensión Esencial: Un Ensayo Clínico Aleatorizado, Doble Ciego, Controlado con Placebo".

Autores y Afiliaciones: (Similar a ejemplos anteriores)

Resumen:

- **Contexto:** Contexto del problema (hipertensión, necesidad de tratamientos efectivos).
- **Objetivo:** El objetivo principal del ECA.
- **Metodología:** Diseño (ECA paralelo, doble ciego, controlado con placebo), población, criterios de inclusión/exclusión, intervención y control, aleatorización, cegamiento, desenlace primario, tamaño muestral, duración del seguimiento, análisis estadístico principal (ITT).
- **Resultados:** Número de participantes aleatorizados y analizados, características basales de los grupos, resultado del desenlace primario en cada grupo (ej. reducción promedio de PA con DE), diferencia entre grupos (con IC 95% y p-valor), resultados de desenlaces secundarios clave y eventos adversos.

- **Conclusión:** Resumir la principal conclusión sobre la eficacia y seguridad de la intervención.
- **Registro del ensayo:** Número de registro del ensayo en una base de datos pública (ej. ClinicalTrials.gov).

Palabras clave: (Similar a ejemplos anteriores)

1. Introducción:

- **Párrafo 1:** Introducir la hipertensión arterial como un problema de salud global, sus consecuencias y la necesidad de tratamientos efectivos.
- **Párrafo 2:** Describir el fármaco experimental Cardio-Plus, su mecanismo de acción propuesto y la evidencia preclínica o de Fases I/II que justifica su evaluación en un ECA de Fase III.
- **Párrafo 3:** Explicar por qué un ECA es el diseño apropiado para evaluar la eficacia y seguridad de Cardio-Plus, destacando los beneficios de la aleatorización y el cegamiento.
- **Último Párrafo:** Indicar el objetivo principal y los objetivos secundarios del ECA.

2. Metodología:

- **Diseño del Estudio:** Describir claramente que se trata de un ECA paralelo, la asignación (ej. 1:1 a Cardio-Plus o placebo), el cegamiento (doble ciego: participantes e investigadores/personal del estudio), y el control (placebo). Mencionar el período de reclutamiento y la duración del tratamiento/seguimiento.
- **Población de Estudio y Reclutamiento:** Describir los centros participantes (si es multicéntrico). Definir los **criterios de inclusión** (ej. adultos entre 18-75 años, diagnóstico de hipertensión esencial, PA sistólica entre 140-160 mmHg después de un período de "lavado" de antihipertensivos previos) y **criterios de exclusión** (ej. hipertensión secundaria, comorbilidades cardiovasculares graves, enfermedad renal o hepática significativa, embarazo, uso de ciertos medicamentos). Describir el proceso de reclutamiento y selección.

- **Intervención y Control:** Describir detalladamente la intervención (Cardio-Plus: dosis, frecuencia, vía de administración) y el control (Placebo: apariencia idéntica, frecuencia, vía). Especificar si se permite el uso de otras medicaciones concomitantes (ej. tratamiento estándar de base que se mantiene igual en ambos grupos).

- **Aleatorización:** Describir el método de aleatorización utilizado (ej. aleatorización simple, por bloques, estratificada por edad o gravedad de la hipertensión). Explicar cómo se generó la secuencia de asignación aleatoria y, crucialmente, cómo se mantuvo la **ocultación de la asignación** (*allocation concealment*) para evitar que los investigadores que reclutan sepan a qué grupo será asignado el próximo participante (ej. sistema de aleatorización centralizado por web o telefónico, sobres opacos sellados secuenciales).

- **Cegamiento (*Blinding*):** Describir quiénes estuvieron cegados a la asignación del tratamiento (doble ciego: participantes, personal del estudio que administra el tratamiento, personal que evalúa los desenlaces). Explicar cómo se logró el cegamiento (ej. uso de placebo idéntico, etiquetado ciego de los envases). Mencionar si hubo alguna dificultad para mantener el cegamiento.

- **Desenlaces:** Definir el **desenlace primario** (ej. cambio promedio en la presión arterial sistólica desde el inicio hasta la semana 12). Definir los **desenlaces secundarios** (ej. cambio en la presión arterial diastólica, proporción de pacientes que alcanzan la meta de PA, incidencia de eventos cardiovasculares mayores [infarto, stroke], incidencia de eventos adversos, cambios en marcadores bioquímicos). Describir cómo y cuándo se midieron los desenlaces.

- **Tamaño de la Muestra:** Explicar cómo se calculó el tamaño de la muestra para tener suficiente **potencia estadística** (ej. 80% o 90%) para detectar una diferencia clínicamente significativa en el desenlace primario con un nivel de significancia α predefinido. Mencionar los supuestos utilizados en el cálculo (ej. DS esperada del cambio en PA, diferencia esperada entre grupos).

- **Seguimiento:** Describir el calendario de visitas de seguimiento, los procedimientos realizados en cada visita (ej. medición de PA, evaluación de síntomas, recolección de eventos adversos, adherencia al tratamiento). Documentar la **pérdida de seguimiento** y las razones.
- **Consideraciones Éticas:** Declarar que el protocolo del estudio fue aprobado por el **Comité de Ética de Investigación** de cada centro participante y por las autoridades regulatorias correspondientes. Confirmar que se obtuvo **Consentimiento Informado** escrito de todos los participantes antes de cualquier procedimiento del estudio. Mencionar que el estudio se llevó a cabo en adherencia a la Declaración de Helsinki y las Buenas Prácticas Clínicas (BPC).
- **Análisis Estadístico:** Describir el plan de análisis, incluyendo:
  - **Análisis Primario:** Especificar que el análisis principal será por **intención de tratar (ITT)**. Describir la prueba estadística para comparar el desenlace primario entre grupos (ej. prueba t de Student para diferencia de medias, análisis de covarianza [ANCOVA] ajustando por el valor basal del desenlace).
  - **Análisis de Desenlaces Secundarios:** Describir las pruebas estadísticas para los desenlaces secundarios.
  - **Análisis de Seguridad:** Describir cómo se compararán los eventos adversos entre grupos (ej. frecuencias, riesgos relativos).
  - **Manejo de Datos Faltantes:** Describir cómo se abordarán los datos faltantes (ej. imputación).
  - **Análisis de Subgrupos (Si planeado):** Si se planean análisis en subgrupos preespecificados, describirlos.

3. Resultados:

- Presentar los hallazgos de manera clara y objetiva, siguiendo el flujo del diagrama CONSORT.
- **Flujo de Participantes:** Presentar un **Diagrama de Flujo CONSORT** (Figura 1) que muestre el número de participantes evaluados para elegibilidad, excluidos (con razones), aleatorizados a cada grupo, que re-

cibieron la intervención, que completaron el seguimiento y que fueron incluidos en el análisis.

- **Características Basales:** Presentar una **Tabla 1: Características Basales de los Participantes por Grupo de Tratamiento**. Comparar las características demográficas, clínicas y otras variables clave al inicio del estudio entre los grupos para demostrar que la aleatorización fue exitosa en crear grupos comparables.
- **Desenlace Primario:** Presentar los resultados del desenlace primario en cada grupo (ej. media ± DE de la reducción de PA sistólica). Presentar la **diferencia entre grupos**, su **IC 95%** y el **valor p** del análisis primario (ITT).
- **Desenlaces Secundarios:** Presentar los resultados de los desenlaces secundarios clave en cada grupo, con medidas de efecto (RR, DR, diferencia de medias) e IC 95% y valores p.
- **Seguridad y Eventos Adversos:** Presentar una tabla resumiendo los eventos adversos más comunes y graves en cada grupo (Tabla 2). Comparar la incidencia de eventos adversos totales y específicos entre grupos.
- **Adherencia al Tratamiento:** Reportar la adherencia al tratamiento en cada grupo.

Figura 1. Diagrama de Flujo del Ensayo (Adaptado de CONSORT). (Muestra el número de participantes en cada etapa del estudio).

Tabla 1. Características Basales de los Participantes por Grupo de Tratamiento. (Compara la distribución de variables clave entre los grupos Cardio-Plus y Placebo al inicio).

Tabla 2. Eventos Adversos por Grupo de Tratamiento. (Lista la frecuencia de eventos adversos en cada grupo).

4. Discusión:

- Interpretar los resultados en el contexto de la pregunta de investigación y la literatura.
- **Párrafo 1:** Resumir los hallazgos principales del estudio, enfocándose en el resultado del desenlace primario y los hallazgos de seguridad.

- **Párrafo 2 (y siguientes):** Discutir la interpretación de los resultados. Si el resultado fue significativo, ¿qué significa la magnitud del efecto (ej. cuántos mmHg se redujo la PA adicionalmente con Cardio-Plus)? Comparar los hallazgos con estudios previos (Fases I/II, otros tratamientos). Discutir la plausibilidad biológica del efecto.
- **Fortalezas del Estudio:** Destacar los aspectos que aumentan la validez interna (ej. aleatorización rigurosa, ocultación de la asignación, doble cegamiento, análisis ITT) y la validez externa (si la muestra es representativa).
- **Limitaciones:** Discutir las limitaciones del estudio (ej. duración del seguimiento, tamaño de la muestra [si no se alcanzó el tamaño calculado o si el efecto fue menor de lo esperado], generalizabilidad a otras poblaciones [ej. pacientes con comorbilidades excluidas], eventos adversos a largo plazo no evaluados, adherencia al tratamiento en el "mundo real").
- **Recomendaciones Futuras:** Sugerir futuras investigaciones (ej. estudios de extensión a largo plazo, estudios en poblaciones diferentes, estudios comparativos directos con otros tratamientos estándar, investigación de mecanismos).
- **Implicaciones:** Discutir las implicaciones clínicas de los hallazgos para el manejo de la hipertensión.

5. Conclusión:

- Resumir la principal conclusión del estudio sobre la eficacia y seguridad de Cardio-Plus, indicando si apoya su uso en pacientes con hipertensión esencial.

Secciones Finales: (Agradecimientos, Financiación, Ética, Consentimiento, Datos, Abreviaciones, Referencias - similar a ejemplos anteriores)

Este ejemplo ilustra cómo un ECA, mediante la aleatorización y el control riguroso, busca establecer con la mayor confianza posible si una intervención causa un efecto. Es un diseño complejo y exigente, pero la solidez de la evidencia que genera lo convierte en la base de la MBE para la evaluación de intervenciones terapéuticas y preventivas.

4.4. Síntesis de evidencia

Hemos recorrido la jerarquía de la evidencia desde los estudios descriptivos hasta los ECA, que nos ofrecen la mejor evidencia individual sobre causa y efecto. Sin embargo, incluso un ECA riguroso es solo una pieza del rompecabezas. Frecuentemente, existen múltiples estudios que investigan la misma pregunta, arrojando resultados que pueden variar o incluso ser contradictorios. Imaginemos que queremos saber si un nuevo medicamento realmente reduce el riesgo de infarto en pacientes con alto riesgo y encontramos diez ECA relevantes con conclusiones diversas. La simple lectura de resúmenes no nos proporcionará una imagen clara ni fiable. ¿Cómo integramos, entonces, toda esta información dispersa para llegar a una conclusión robusta? Necesitamos métodos estructurados para recopilar, evaluar críticamente y sintetizar la evidencia existente. Existen diferentes enfoques para esta síntesis, desde resúmenes más tradicionales hasta metodologías altamente rigurosas. Exploraremos estas herramientas comenzando por las **revisiones narrativas**, seguidas por los **análisis bibliométricos** que mapean cuantitativamente la producción científica, las **revisiones de alcance (*scoping reviews*)** que delimitan la extensión de la literatura, y culminando en las **revisiones sistemáticas** y los **metaanálisis**, diseñados para responder preguntas específicas con el máximo rigor.

4.4.1. *Revisión Narrativa* (*Narrative Review*)

Una revisión narrativa es un resumen tradicional de la literatura existente sobre un tema particular. A diferencia de las revisiones sistemáticas y de alcance, las revisiones narrativas no siguen una metodología explícita o un protocolo predefinido para la búsqueda, selección y evaluación de los estudios. Se basan en el conocimiento y la experiencia del autor para recopilar y sintetizar la literatura que consideran relevante.

¿Por qué utilizamos Revisiones Narrativas?

Las revisiones narrativas son útiles para:

- Proporcionar una descripción general del estado actual del conocimiento sobre un tema.
- Introducir un tema a lectores no familiarizados con él.
- Explorar la historia y evolución de un concepto o campo de investigación.

- Identificar la base del conocimiento existente y sugerir nuevas áreas de estudio.
- Contextualizar hallazgos de investigación individuales.

Principios y Conducta de una Revisión Narrativa:

Aunque no existe una metodología estandarizada para las revisiones narrativas, generalmente implican los siguientes pasos:

1. **Definir el tema y alcance:** El autor selecciona un tema de interés y determina la amplitud de la cobertura.
2. **Búsqueda de literatura:** El autor busca literatura relevante basándose en su conocimiento del campo, bases de datos conocidas o referencias clave. La exhaustividad de la búsqueda puede variar considerablemente.
3. **Selección de estudios:** El autor selecciona los estudios que considera más pertinentes o influyentes para incluirlos en la revisión. Los criterios de selección a menudo no se explicitan.
4. **Síntesis de la evidencia:** El autor resume, discute y sintetiza los hallazgos de los estudios seleccionados de manera narrativa, a menudo organizando la información por temas o cronología.

Ventajas y Desventajas:

Las revisiones narrativas tienen algunas ventajas:

- Son relativamente más rápidas y sencillas de realizar que las revisiones sistemáticas o de alcance.
- Permiten una mayor flexibilidad en la selección y presentación de la literatura.
- Pueden ofrecer una perspectiva experta y contextualización profunda de un tema.

Sin embargo, presentan desventajas significativas:

- Son inherentemente subjetivas, ya que dependen del criterio del autor.
- No son reproducibles, ya que la búsqueda y selección no siguen un protocolo explícito.

- Pueden estar sujetas a sesgos del autor, como el sesgo de selección o de publicación.
- No evalúan formalmente la calidad de los estudios incluidos, lo que limita la solidez de sus conclusiones.
- Se consideran una forma de evidencia más débil para la toma de decisiones informada.

4.4.2. *Análisis Bibliométrico*

Un análisis bibliométrico es un método cuantitativo utilizado para examinar y evaluar la literatura académica dentro de un campo de estudio específico. Implica el uso de herramientas estadísticas y matemáticas para analizar patrones de publicación, citación, colaboración y tendencias temáticas. Este tipo de análisis proporciona una visión objetiva del desarrollo histórico y el estado actual de un área de investigación, identificando a los autores, instituciones, países y publicaciones más influyentes, así como los temas emergentes.

¿Por qué realizamos Análisis Bibliométricos?

En un entorno académico donde el volumen de publicaciones crece constantemente, comprender la estructura y la dinámica de un campo de investigación puede ser un desafío. Los análisis bibliométricos son esenciales porque:

- **Miden la productividad científica:** Cuantifican el número de publicaciones por autores, instituciones, países o revistas a lo largo del tiempo.
- **Identifican la influencia:** Utilizan métricas de citación para determinar el impacto de publicaciones, autores o revistas individuales.
- **Visualizan redes de colaboración:** Muestran las conexiones entre autores, instituciones o países que colaboran en la investigación.
- **Detectan tendencias temáticas:** Analizan la co-ocurrencia de palabras clave para identificar los temas de investigación más activos y emergentes.
- **Evalúan el rendimiento:** Permiten comparar la productividad e impacto de diferentes entidades (autores, instituciones, países).
- **Informan la toma de decisiones:** Ayudan a investigadores, financiadores y responsables políticos a comprender el panorama de la investigación y dirigir recursos de manera efectiva.

Principios y Conducta de un Análisis Bibliométrico:

La realización de un análisis bibliométrico riguroso, como el descrito en el artículo adjunto sobre Interfaces Cerebro-Computadora (BCI) basadas en Electroencefalografía (EEG) en rehabilitación (https://www.mdpi.com/1424-8220/24/22/7125), sigue un proceso metódico:

1. **Definir el objetivo y alcance:** Establecer claramente la pregunta de investigación y el período de tiempo a analizar (por ejemplo, 2013-2023 en el artículo de referencia).
2. **Seleccionar las bases de datos:** Identificar las bases de datos académicas más relevantes y completas para el campo de estudio (como Web of Science, PubMed, Science Direct y Scopus en el artículo de referencia).
3. **Desarrollar la estrategia de búsqueda:** Crear una combinación de términos de búsqueda y operadores booleanos que capture de manera efectiva la literatura relevante, asegurando que la estrategia sea reproducible.
4. **Recopilar los datos:** Ejecutar la estrategia de búsqueda en las bases de datos seleccionadas y exportar los metadatos de las publicaciones (título, autores, afiliaciones, resumen, palabras clave, citas, etc.).
5. **Limpiar y organizar los datos:** Eliminar duplicados y estandarizar la información (nombres de autores, afiliaciones, etc.) para preparar los datos para el análisis. En el artículo de referencia, esto incluyó la eliminación de duplicados y un proceso de selección basado en criterios de inclusión y exclusión.
6. **Realizar el análisis:** Aplicar técnicas bibliométricas utilizando software especializado (como Microsoft Excel y R con paquetes como ggplot2 y Bibliometrix en el artículo de referencia). Esto incluye análisis de tendencias temporales, distribución geográfica, redes de coautoría, análisis de coocurrencia de palabras clave, etc.
7. **Visualizar e interpretar los resultados:** Presentar los hallazgos de manera clara utilizando gráficos, mapas y tablas (como se muestra en las Figuras 1-10 del artículo de referencia) e interpretar su significado en el contexto del campo de estudio.

8. **Reportar los hallazgos:** Documentar la metodología y los resultados del análisis de manera transparente.

Metodologías en Análisis Bibliométricos:

Las metodologías específicas empleadas en un análisis bibliométrico dependen de los objetivos del estudio. Algunas técnicas comunes incluyen:

- **Análisis de productividad:** Conteo simple de publicaciones por diferentes categorías.
- **Análisis de citación:** Examen de las citas recibidas para evaluar el impacto. Métricas como el número total de citas, el índice h, o el análisis de co-citación y acoplamiento bibliográfico.
- **Análisis de coautoría:** Mapeo de redes de colaboración entre autores, instituciones o países.
- **Análisis de coocurrencia de palabras clave:** Identificación de temas de investigación mediante el análisis de palabras clave que aparecen juntas en las publicaciones. Esto puede incluir la creación de mapas temáticos.
- **Análisis de rendimiento:** Comparación de métricas de productividad e impacto entre diferentes entidades.

El artículo de referencia utiliza varias de estas metodologías, incluyendo el análisis de tendencias anuales de publicación y citación (Figura 2), la distribución geográfica y redes de colaboración (Figura 4), la productividad de autores e instituciones (Tabla 2, Figuras 6 y 7), y el análisis de palabras clave y temático (Figuras 9 y 10).

Ventajas y Desventajas:

Las ventajas de los análisis bibliométricos incluyen:

- Proporcionan una visión cuantitativa y objetiva del campo.
- Permiten identificar tendencias, actores clave y áreas de investigación activas o emergentes.
- Son útiles para comparar el rendimiento de investigación.

Las desventajas son:

- La calidad del análisis depende de la calidad y cobertura de las bases de datos utilizadas.
- Las métricas de citación pueden variar entre disciplinas y no siempre reflejan la calidad intrínseca de la investigación.
- No capturan la investigación no publicada o la literatura gris de manera exhaustiva.
- La interpretación de los resultados debe hacerse con cautela, considerando el contexto del campo.

Consideraciones Éticas:

Aunque los análisis bibliométricos trabajan con datos públicos de publicaciones, es importante ser transparente sobre la metodología utilizada y evitar interpretaciones sesgadas de los resultados. Al identificar a los actores clave, se debe reconocer que las métricas bibliométricas son solo un aspecto del impacto y la contribución a un campo.

En conclusión, el análisis bibliométrico es una herramienta poderosa para comprender la estructura y evolución de un campo de investigación. Al aplicar métodos cuantitativos a la literatura académica, proporciona información valiosa sobre la productividad, influencia, colaboración y tendencias temáticas, complementando las revisiones narrativas y de alcance. El estudio de referencia sobre BCI en rehabilitación es un excelente ejemplo de cómo se puede utilizar esta metodología para mapear un campo de rápido crecimiento.

4.4.3. *Revisión de Alcance (Scoping review)*

Una revisión de alcance es un tipo de síntesis de conocimiento que tiene como objetivo identificar y mapear rápidamente la evidencia clave sobre un tema o pregunta de investigación amplia. A diferencia de las revisiones sistemáticas, que buscan responder preguntas específicas y delimitadas, las revisiones de alcance se utilizan para explorar la extensión, el rango y la naturaleza de la investigación existente, determinar si una revisión sistemática es factible, o resumir y difundir hallazgos. Se caracterizan por su flexibilidad y amplitud, lo que las hace adecuadas para temas complejos o heterogéneos.

¿Por qué necesitamos Revisiones de Alcance?

En campos de investigación en rápida evolución o con una vasta y diversa literatura, puede ser un desafío comprender el panorama de la evidencia. Las revisiones de alcance son esenciales porque:

- **Mapean la evidencia:** Proporcionan una visión general de lo que se ha investigado sobre un tema, identificando los tipos de estudios existentes, las poblaciones estudiadas y los conceptos explorados.
- **Identifican vacíos de conocimiento:** Ayudan a determinar áreas donde la investigación es escasa o inexistente, orientando futuras agendas de investigación.
- **Clarifican conceptos:** Pueden ayudar a definir y refinar conceptos clave en campos emergentes.
- **Informan revisiones futuras:** Se utilizan a menudo como un paso preliminar para identificar la necesidad y la viabilidad de una revisión sistemática más enfocada.
- **Sintetizan hallazgos preliminares:** Pueden resumir hallazgos de estudios diversos que no serían adecuados para un metaanálisis.

Principios y Conducta de una Revisión de Alcance:

Realizar una revisión de alcance sigue un proceso estructurado, aunque con mayor flexibilidad que una revisión sistemática. Generalmente, incluye los siguientes pasos:

1. **Identificar la pregunta de investigación:** Definir claramente el objetivo y la pregunta de la revisión. A diferencia de las revisiones sistemáticas, la pregunta de una revisión de alcance es típicamente más amplia. A menudo, se utiliza el formato PCC (Población, Concepto, Contexto) para estructurar esta pregunta.
2. **Identificar estudios relevantes:** Desarrollar una estrategia de búsqueda integral para identificar estudios de diversas fuentes, incluyendo bases de datos electrónicas, literatura gris y otras fuentes relevantes. La estrategia de búsqueda tiende a ser más amplia para capturar la mayor cantidad de evidencia posible.

3. **Seleccionar estudios:** Filtrar los estudios identificados según los criterios de inclusión y exclusión predefinidos. Este proceso a menudo implica varias rondas de cribado por título, resumen y texto completo.
4. **Extraer datos:** Desarrollar un formulario de extracción de datos para recopilar la información relevante de los estudios incluidos. Los datos extraídos suelen centrarse en las características del estudio, metodologías, poblaciones y conceptos clave.
5. **Colacionar, resumir e informar los resultados:** Analizar y sintetizar los datos extraídos para proporcionar una descripción general de la evidencia. Los resultados se presentan a menudo de forma tabular o narrativa, destacando los hallazgos clave y los vacíos de conocimiento. El marco PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for *Scoping reviews*) proporciona una guía para reportar revisiones de alcance de manera transparente.

Metodologías en Revisiones de Alcance:

La metodología de una revisión de alcance debe ser rigurosa y transparente, aunque adaptada a la naturaleza exploratoria de este tipo de síntesis.

- **Algoritmos de Búsqueda:** Se emplean estrategias de búsqueda amplias utilizando múltiples bases de datos y términos de búsqueda flexibles para capturar una variedad de estudios relevantes.
- **Selección de Estudios:** El proceso de selección debe ser realizado por al menos dos revisores para minimizar el sesgo, aunque la naturaleza amplia de la revisión puede requerir un enfoque más iterativo.
- **Extracción de Datos:** La extracción de datos se enfoca en mapear las características de los estudios y los conceptos clave, en lugar de evaluar críticamente la calidad metodológica de cada estudio individual en detalle.
- **Síntesis de la Evidencia:** La síntesis es predominantemente cualitativa o narrativa, describiendo el alcance y la naturaleza de la investigación existente. No se realiza metaanálisis.

Ventajas y Desventajas:

Las revisiones de alcance ofrecen varias ventajas:

- Proporcionan una descripción general rápida de la literatura existente sobre un tema.
- Son flexibles y pueden abordar preguntas de investigación amplias y complejas.
- Ayudan a identificar vacíos de conocimiento y orientar futuras investigaciones.
- Pueden informar la necesidad y el alcance de futuras revisiones sistemáticas.

Sin embargo, también presentan desventajas:

- No evalúan críticamente la calidad metodológica de los estudios incluidos en profundidad.
- No proporcionan una síntesis cuantitativa de los hallazgos.
- La amplitud de la búsqueda puede resultar en un gran volumen de literatura para gestionar.

Consideraciones Éticas:

Aunque las revisiones de alcance no suelen involucrar directamente a participantes humanos, es importante mantener la transparencia en el proceso de revisión y reportar los hallazgos de manera imparcial.

En resumen, las revisiones de alcance son una herramienta valiosa para mapear y resumir la literatura existente sobre un tema de investigación amplio. Son particularmente útiles en campos emergentes o cuando la heterogeneidad de la evidencia impide una revisión sistemática tradicional. Al identificar los vacíos de conocimiento y proporcionar una visión general del panorama de la investigación, las revisiones de alcance desempeñan un papel importante en la planificación de futuras investigaciones y síntesis de evidencia.

Tabla 20. *Comparativo de Tipos de Revisiones de Literatura*

Característica	Revisión Narrativa	Análisis Bibliométrico	Revisión de Alcance (*Scoping review*)	Revisión Sistemática (Systematic Review)
Objetivo Principal	Resumir y describir la literatura existente; Proveer perspectiva experta.	Examinar y evaluar cuantitativamente la literatura; Identificar patrones de publicación, citación, colaboración y tendencias temáticas.	Mapear el alcance y la naturaleza de la investigación; Identificar vacíos.	Responder una pregunta específica y delimitada; Sintetizar la mejor evidencia disponible.
Pregunta de Investigación	Amplia.	Puede variar (Ej: ¿Quiénes son los autores más productivos? ¿Cuáles son los temas emergentes?).	Amplia y exploratoria (a menudo PCC).	Específica y focalizada (a menudo PICO).
Alcance	Amplio, definido por el autor.	Definido por la estrategia de búsqueda y las bases de datos.	Amplio para cubrir todo el conocimiento relevante.	Estrecho y bien definido.
Estrategia de Búsqueda	No explícita ni exhaustiva; Basada en el conocimiento del autor.	Rigurosa y reproducible para recopilar metadatos.	Rigurosa y amplia para identificar la mayor cantidad de evidencia posible.	Rigurosa, exhaustiva y reproducible; Uso de múltiples bases de datos.
Selección de Estudios	Subjetiva, sin criterios explícitos o reproducibles.	Basada en criterios de inclusión/exclusión para la recopilación de datos; El análisis se aplica a los metadatos de los estudios incluidos.	Iterativa, basada en criterios de inclusión/exclusión explícitos.	Rigurosa, basada en criterios de inclusión/exclusión explícitos; Realizada por duplicado.

Característica	Revisión Narrativa	Análisis Bibliométrico	Revisión de Alcance (*Scoping review*)	Revisión Sistemática (Systematic Review)
Evaluación Crítica de Estudios	Generalmente no formal o limitada.	No evalúa la calidad intrínseca de los estudios; Analiza métricas de impacto (ej. citaciones).	No es un objetivo principal; Se enfoca en caracterizar los estudios.	Rigurosa evaluación de la calidad metodológica y el riesgo de sesgo.
Síntesis	Narrativa y cualitativa; Basada en la interpretación del autor.	Cuantitativa (análisis de datos bibliométricos); Visualizaciones (gráficos, mapas).	Narrativa y descriptiva; Mapeo de la evidencia.	Cuantitativa (Metaanálisis) o cualitativa rigurosa.
Reproducibilidad	Baja.	Alta (si la estrategia de búsqueda y los métodos de análisis son explícitos).	Moderada a alta (dependiendo de la explicitación de pasos).	Alta.
Protocolo Predefinido	No requerido.	Recomendado para la transparencia de la metodología.	Recomendado (Ej: JBI, PRISMA-ScR).	Esencial (Ej: PROSPERO).
Sesgo	Alto riesgo de sesgo del autor.	Puede estar influenciado por la selección de bases de datos o la interpretación de métricas; Menor riesgo de sesgo de selección de estudios individuales que en la narrativa.	Menor riesgo que la narrativa, pero posible sesgo de mapeo.	Minimizado mediante métodos rigurosos.

4.4.4. *Revisiones Sistemáticas*

Una revisión sistemática es un tipo de estudio secundario que utiliza métodos rigurosos y explícitos para identificar, seleccionar, evaluar críticamente y sintetizar la evidencia de estudios primarios relevantes sobre una pregunta de investigación específica. A diferencia de las revisiones narrativas tradicionales (que pueden ser subjetivas y no reproducibles), una revisión sistemática sigue un **protocolo predefinido** para minimizar el sesgo y proporcionar un resumen imparcial de la mejor evidencia disponible.

¿Por qué necesitamos Revisiones Sistemáticas?

En la era de la información, la cantidad de investigación publicada es abrumadora. Para los profesionales de la salud, es imposible mantenerse al día con todos los estudios relevantes en su campo. Además, necesitamos confiar en **evidencia de alta calidad** para tomar decisiones clínicas informadas. Las revisiones sistemáticas son esenciales porque:

- **Sintetizan la evidencia:** Agregan los hallazgos de múltiples estudios, proporcionando una visión general más completa y fiable que un solo estudio.
- **Reducen el sesgo:** Al seguir un protocolo explícito y riguroso, minimizan el riesgo de seleccionar selectivamente estudios que apoyen una conclusión particular.
- **Resuelven controversias:** Pueden ayudar a aclarar resultados contradictorios entre estudios individuales.
- **Identifican brechas en el conocimiento:** Al revisar la literatura existente, pueden señalar áreas donde se necesita más investigación.
- **Informan la toma de decisiones:** Son una fuente clave de evidencia para guías de práctica clínica, políticas de salud y decisiones individuales de tratamiento.

Principios y Conducta de una Revisión Sistemática:

Realizar una revisión sistemática es un proceso metódico que generalmente sigue estos pasos:

1. **Formular la Pregunta de Investigación:** Definir claramente la pregunta que se va a responder, a menudo utilizando el formato **PICOS** (Participantes, Intervención, Comparación, Desenlace, Diseño del Estudio). Esto asegura que la búsqueda de estudios sea enfocada.

2. **Identificar Estudios:** Realizar una búsqueda exhaustiva y sistemática de la literatura en múltiples bases de datos electrónicas, registros de ensayos clínicos, listas de referencias de estudios relevantes y, si es posible, literatura gris (tesis, informes no publicados) para encontrar *todos* los estudios potencialmente relevantes.

3. **Seleccionar Estudios:** Revisar los títulos y resúmenes de los estudios identificados para determinar si cumplen con los criterios de inclusión predefinidos (basados en la pregunta PICOS). Luego, obtener los textos completos de los estudios potencialmente elegibles y revisarlos nuevamente para la inclusión final. Este proceso de selección debe ser realizado por al menos dos revisores de forma independiente para reducir el sesgo.

4. **Evaluar Críticamente los Estudios:** Evaluar la calidad metodológica y el riesgo de sesgo de cada estudio incluido utilizando herramientas estandarizadas. Estudios con alto riesgo de sesgo pueden ser excluidos o analizados por separado.

5. **Extraer Datos:** Extraer sistemáticamente los datos relevantes de cada estudio incluido utilizando formularios de extracción de datos estandarizados. Los datos incluyen características de los participantes, la intervención, el control, el diseño del estudio, el tiempo de seguimiento y los resultados del desenlace (medidas de efecto, tamaño muestral, etc.).

6. **Sintetizar la Evidencia:** Combinar y resumir los hallazgos de los estudios incluidos. Si los estudios son suficientemente homogéneos en cuanto a su diseño, población, intervención y desenlace, se puede realizar un **Metaanálisis** (síntesis cuantitativa). Si los estudios son demasiado diversos, se realiza una síntesis cualitativa o narrativa.

7. **Interpretar los Hallazgos:** Interpretar los resultados de la síntesis (metaanálisis o síntesis cualitativa) en el contexto de la pregunta de investi-

gación, considerando la calidad de la evidencia, la heterogeneidad entre estudios y los posibles sesgos.

8. **Reportar la Revisión:** Documentar todo el proceso y los hallazgos de manera transparente y detallada, siguiendo guías de reporte como la declaración **PRISMA** (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*).

Metodologías en Revisiones Sistemáticas:

La metodología de una revisión sistemática debe estar preespecificada en un **protocolo** antes de iniciar la búsqueda de estudios. Este protocolo detalla cada paso del proceso para asegurar la transparencia y reproducibilidad.

Algoritmos de Búsqueda:

La identificación de estudios relevantes es un paso crítico que requiere una estrategia de búsqueda exhaustiva y sensible. Los algoritmos de búsqueda se construyen combinando términos relacionados con cada componente de la pregunta PICOS (Participantes, Intervención, Comparación, Desenlace) y el diseño del estudio (si se restringe a un tipo específico, como ECA).

- **Bases de Datos:** Se utilizan múltiples bases de datos electrónicas, como:
  - **PubMed/MEDLINE:** Amplia cobertura de literatura biomédica.
  - **Embase:** Fuerte en literatura sobre fármacos y farmacología.
  - **Cochrane Central Register of Controlled Trials (CENTRAL):** Especializada en ensayos clínicos.
  - **Web of Science:** Cubre diversas disciplinas científicas.
  - **Scopus:** Base de datos multidisciplinaria.
  - Otras bases de datos especializadas según el tema (ej. PsycINFO para psicología, CINAHL para enfermería).
- **Términos de Búsqueda:** Se utilizan tanto **términos controlados** (vocabulario estandarizado, como MeSH en PubMed o Emtree en Embase) como **términos de texto libre** (palabras clave que aparecen en títulos o resúmenes).

- **Operadores Booleanos:** Se combinan los términos de búsqueda utilizando operadores lógicos:

  - **AND:** Para combinar conceptos diferentes (ej. "hipertensión" AND "tratamiento").

  - **OR:** Para combinar sinónimos o términos relacionados dentro de un mismo concepto (ej. "infarto de miocardio" OR "ataque cardíaco" OR "infarto").

  - **NOT:** Para excluir conceptos (usar con precaución, ya que puede eliminar estudios relevantes).

- **Estrategias Adicionales:** Se complementa la búsqueda en bases de datos con:

  - Revisión de las listas de referencias de los estudios incluidos y de otras revisiones relevantes.

  - Búsqueda en registros de ensayos clínicos (ej. ClinicalTrials.gov, WHO ICTRP) para identificar estudios en curso o no publicados.

  - Contacto con expertos en el campo.

  - Búsqueda de literatura gris.

El algoritmo de búsqueda debe ser documentado detalladamente en el protocolo y el reporte de la revisión para asegurar su transparencia.

Herramientas de Evaluación del Riesgo de Sesgo:

Evaluar el riesgo de sesgo en los estudios primarios es fundamental porque la validez de una revisión sistemática depende de la calidad de los estudios que incluye. Diferentes herramientas se utilizan según el diseño del estudio primario:

- **Para Estudios de Intervención (principalmente ECA):** Estas herramientas se centran en aspectos del diseño y la ejecución que son cruciales para minimizar el sesgo en los estudios aleatorizados.

  - ***Cochrane Risk of Bias* Tool (RoB 1.0):** Una herramienta ampliamente utilizada que evalúa el riesgo de sesgo en siete dominios específicos (ge-

neración de secuencia aleatoria, ocultación de la asignación, cegamiento de participantes y personal, cegamiento de la evaluación de resultados, datos de resultados incompletos, reporte selectivo y otros sesgos). Para cada dominio, se califica el riesgo como bajo, alto o incierto.

- ***Cochrane Risk of Bias Tool for Randomized Trials* (RoB 2.0):** Una versión actualizada de la herramienta Cochrane, con un enfoque basado en dominios y señalización para guiar el juicio del revisor sobre el riesgo de sesgo global.
- **Escala de Jadad:** Una escala numérica más antigua que evalúa la calidad metodológica basándose en tres ítems principales: aleatorización, cegamiento y manejo de abandonos/pérdidas de seguimiento. La puntuación varía de 0 a 5, donde puntuaciones más altas indican mayor calidad. Aunque es más simple, la herramienta Cochrane RoB es generalmente preferida por su evaluación más detallada de los diferentes dominios de sesgo.

- **Para Estudios Observacionales (ej. Cohortes, Casos y Controles, Transversales):** Evaluar el riesgo de sesgo en estudios observacionales es más complejo debido a la falta de aleatorización y la mayor susceptibilidad a confusores y otros sesgos.
- **ROBINS-I (*Risk Of Bias In Non-randomised Studies - of Interventions*):** Una herramienta diseñada específicamente para evaluar el riesgo de sesgo en estudios no aleatorizados que evalúan el efecto de una intervención. Considera el riesgo de sesgo en diferentes dominios a lo largo del tiempo, comparando el grupo de intervención con el grupo de control en diferentes puntos.
- **Otras Herramientas:** Existen otras listas de verificación o escalas para evaluar la calidad de estudios observacionales, aunque ROBINS-I es una de las más completas para estudios de intervención no aleatorizados. Para estudios de asociación etiológica (cohortes, casos y controles), a menudo se adaptan principios de evaluación de sesgo relevantes para esos diseños.

- **Visualización del Riesgo de Sesgo:**
- **Gráficos de Resumen del Riesgo de Sesgo (*Risk of Bias Summary/Graph*):** Se utilizan para visualizar el riesgo de sesgo de los estudios incluidos, a menudo mostrando el porcentaje de estudios con bajo, alto o riesgo incierto para cada dominio o un resumen visual para cada estudio.

- **ROBVIS (*Risk of Bias Visualization*):** Una herramienta para crear visualizaciones gráficas del riesgo de sesgo evaluado con herramientas como ROBINS-I.

Síntesis de la Evidencia:

La síntesis es el proceso de combinar los resultados de los estudios individuales.

- **Síntesis Cualitativa (o Narrativa):** Se utiliza cuando los estudios son demasiado diferentes para combinarlos estadísticamente. Implica resumir y describir los hallazgos de los estudios en palabras, identificando patrones y diferencias.
- **Síntesis Cuantitativa (Metaanálisis):** Se utiliza cuando los estudios son suficientemente similares (homogéneos). Implica combinar estadísticamente los resultados de los estudios individuales para obtener una medida resumen del efecto (ej. una *Odds Ratio* o *Riesgo Relativo* ponderado) con mayor precisión que la de cualquier estudio individual. (El Metaanálisis se detallará en la siguiente sección).

Ventajas y Desventajas:

Las **ventajas** de las revisiones sistemáticas:

- **Proporcionan la mejor evidencia disponible:** Al sintetizar múltiples estudios, ofrecen una estimación más precisa y fiable del efecto de una intervención o de una asociación.
- **Reducen el sesgo:** El proceso sistemático minimiza la selección sesgada de estudios.
- **Aumentan la potencia estadística:** Al combinar datos de varios estudios, pueden detectar efectos que un solo estudio no podría.
- **Son más rápidas y económicas que iniciar un nuevo estudio grande:** Utilizan datos ya existentes.
- **Identifican brechas en la investigación:** Ayudan a planificar estudios futuros.

Las **desventajas** de las revisiones sistemáticas:

- **Requieren tiempo y esfuerzos considerables:** Realizar una revisión sistemática rigurosa es un proceso complejo.
- **La calidad depende de la calidad de los estudios incluidos:** Una revisión sistemática de estudios de baja calidad seguirá siendo de baja calidad ("basura entra, basura sale").
- **Susceptibles a sesgos:** Aunque buscan reducir el sesgo, pueden verse afectados por sesgos en los estudios primarios o sesgos relacionados con la publicación.
- **Pueden ser difíciles de interpretar si hay mucha heterogeneidad** entre los estudios.

Sesgos en Revisiones Sistemáticas:

Además de los sesgos presentes en los estudios primarios, las revisiones sistemáticas pueden verse afectadas por sesgos relacionados con la **diseminación de la evidencia**:

- **Sesgo de Publicación (*Publication Bias*):** Ocurre si la probabilidad de que un estudio sea publicado depende de la dirección o significancia estadística de sus resultados. Los estudios con resultados positivos y significativos tienen más probabilidades de ser publicados que aquellos con resultados nulos o negativos. Esto puede llevar a que una revisión sistemática que solo incluya estudios publicados sobreestime el efecto de una intervención. Se puede explorar gráficamente con **gráficos de embudo (*funnel plots*)**.
- **Sesgo de Lenguaje (*Language Bias*):** Ocurre si solo se incluyen estudios publicados en ciertos idiomas (ej. solo en inglés), excluyendo estudios relevantes publicados en otros idiomas, que podrían tener resultados diferentes.
- **Sesgo de Retraso en la Publicación (*Time Lag Bias*):** Los estudios con resultados positivos tienden a publicarse más rápidamente que los estudios con resultados nulos o negativos.

- **Sesgo de Publicación Múltiple (*Multiple Publication Bias*):** Los mismos resultados de un estudio pueden publicarse en múltiples lugares, lo que podría llevar a que se les dé un peso excesivo en una revisión.
- **Sesgo de Reporte de Desenlaces (*Outcome Reporting Bias*):** Dentro de un estudio, solo se reportan selectivamente los desenlaces que mostraron resultados significativos, omitiendo otros desenlaces medidos.

Consideraciones éticas:

Aunque las revisiones sistemáticas no involucran directamente a participantes humanos, existen consideraciones éticas relacionadas con la transparencia, la integridad y la interpretación responsable de la evidencia. Es importante declarar cualquier conflicto de interés de los revisores y asegurarse de que la revisión se realice de manera imparcial.

En resumen, las revisiones sistemáticas son herramientas fundamentales en la MBE. Al sintetizar rigurosamente la evidencia de estudios primarios, nos proporcionan la mejor información disponible para guiar la práctica clínica y la toma de decisiones en salud, aunque siempre debemos ser conscientes de sus limitaciones y los posibles sesgos que pueden afectarlas. Son el paso previo esencial para realizar un Metaanálisis cuando la síntesis cuantitativa es apropiada.

4.4.5. Ejemplo de revisión sistemática de ensayos clínicos aleatorizados

Hemos recorrido la jerarquía de la evidencia, llegando a la cima con los ECA. Sin embargo, para obtener la imagen más completa y fiable sobre una pregunta de investigación, a menudo necesitamos sintetizar la evidencia de múltiples estudios primarios. Este es el rol de la **Revisión Sistemática**, el diseño que se sitúa en la cúspide de la jerarquía de la evidencia, especialmente cuando se combina con un metaanálisis.

Reportar una revisión sistemática de manera transparente y completa es esencial para que otros investigadores y clínicos puedan evaluar su fiabilidad y aplicabilidad. La declaración **PRISMA** es la guía estándar para este reporte, asegurando que se incluyan todos los elementos clave.

Aquí te presento la estructura típica de un reporte de revisión sistemática, basándonos en los principios de PRISMA y lo que hemos discutido sobre este diseño:

Estructura del Reporte de Revisión Sistemática:

Un reporte de revisión sistemática generalmente incluye las siguientes secciones principales:

1. **Título**
2. **Autores y Afiliaciones**
3. **Resumen**
4. **Palabras clave**
5. **Introducción**
6. **Metodología**
7. **Resultados**
8. **Discusión**
9. **Conclusión**
10. **Secciones Finales** (Agradecimientos, Financiación, Conflictos de Interés, Contribuciones de los Autores, Declaración de Disponibilidad de Datos, etc.)
11. **Referencias**
12. **Tablas y Figuras** (Diagrama de Flujo PRISMA, Tablas de Características de los Estudios, Evaluación del Riesgo de Sesgo, Gráficos de Bosque/Funnel Plots si aplica)

Ahora, veamos el contenido clave para cada una de estas secciones:

1. Título:

- Debe identificar claramente el reporte como una revisión sistemática (y metaanálisis si aplica).
- Debe reflejar la pregunta de investigación, incluyendo idealmente los componentes **PICOS** (Participantes, Intervención, Comparación, Desenlace) y el tipo de estudios primarios incluidos (ej. ECA).

- *Ejemplo*: "Efectividad de Probióticos, Prebióticos y Simbióticos en el Manejo de la Resistencia a la Insulina y el Desbalance Hormonal en Mujeres con Síndrome de Ovario Poliquístico (SOP): Una Revisión Sistemática de ECA".

2. Autores y Afiliaciones:

- Lista de los autores que contribuyeron significativamente a la revisión sistemática.
- Afiliación institucional de cada autor.
- Información de contacto para el autor de correspondencia.

3. Resumen:

- Un resumen estructurado que sigue el flujo de la revisión, generalmente con secciones como:
  - **Contexto/Objetivos:** Contexto del problema de salud y el objetivo de la revisión.
  - **Metodología:** Diseño de la revisión (sistemática, metaanálisis), fuentes de datos, criterios de elegibilidad de los estudios, proceso de selección, evaluación del riesgo de sesgo, y métodos de síntesis de datos.
  - **Resultados:** Número de estudios incluidos, características clave de los estudios y participantes, resultados principales de la síntesis (cualitativa o cuantitativa, incluyendo medidas de efecto e IC 95% si se realizó metaanálisis).
  - **Conclusiones:** Resumen de las principales conclusiones de la revisión y sus implicaciones.
  - ***Systematic Review Registration*:** Número de registro del protocolo de la revisión (ej. en PROSPERO).

4. Palabras clave:

- Lista de términos clave que describen el contenido principal de la revisión, incluyendo el tipo de estudio (revisión sistemática, metaanálisis) y los componentes PICOS.

5. Introducción:

- Generalmente consta de 2 a 4 párrafos.
- **Párrafo 1:** Introducir el problema de salud o la condición de interés, su importancia y la incertidumbre o controversia existente que justifica la revisión.
- **Párrafo 2:** Describir la intervención o exposición de interés y la justificación para investigarla.
- **Párrafo 3:** Explicar por qué una revisión sistemática es el diseño apropiado para abordar la pregunta de investigación, destacando las limitaciones de los estudios individuales y la necesidad de sintetizar la evidencia.
- **Último Párrafo:** Indicar claramente el objetivo principal y los objetivos secundarios de la revisión, a menudo formulados como una pregunta PICOS explícita.

6. Metodología:

- Esta es una sección crucial que describe detalladamente el proceso de la revisión para asegurar su transparencia y reproducibilidad. Debe incluir:
  - **Protocolo y registro:** Declarar si existe un protocolo de revisión (y dónde está disponible) y si la revisión fue registrada prospectivamente en una base de datos (ej. PROSPERO), proporcionando el número de registro.
  - **Criterios de elegibilidad:** Describir detalladamente los criterios de inclusión y exclusión para los estudios primarios (basados en PICOS y diseño del estudio), tal como se definió en el protocolo.
  - **Fuentes de información:** Enumerar todas las bases de datos electrónicas consultadas (ej. PubMed, Embase, Cochrane CENTRAL, Web of Science, Scopus), las fechas de la última búsqueda, y otras fuentes (ej. registros de ensayos clínicos, literatura gris, listas de referencias, contacto con expertos).
  - **Estrategia de búsqueda:** Presentar la estrategia de búsqueda completa utilizada para al menos una base de datos, incluyendo los términos de búsqueda (controlados y de texto libre) y los operadores booleanos. Explicar cómo se adaptó la estrategia para otras bases de datos.

- **Selección de estudios:** Describir el proceso de selección de estudios (ej. revisión inicial de títulos/resúmenes, revisión a texto completo), cómo se realizó (ej. por dos revisores independientes) y cómo se resolvieron las discrepancias.
- **Extracción de datos:** Describir el proceso de extracción de datos, el formulario estandarizado utilizado, y cómo se realizó (ej. por dos revisores independientes).
- **Riesgo de sesgo:** Describir la herramienta(s) utilizada(s) para evaluar el riesgo de sesgo de los estudios primarios y cómo se realizó la evaluación (ej. por dos revisores independientes). **Es crucial mencionar cómo varían las herramientas según el diseño del estudio primario:**

- **Para Estudios de Intervención (ECA):** Herramientas como la ***Cochrane Risk of Bias Tool* (RoB 1.0 o RoB 2.0)** o la escala de **Jadad** (aunque menos detallada). Evalúan dominios como aleatorización, ocultación de la asignación, cegamiento, datos incompletos, reporte selectivo, etc.
- **Para Estudios Observacionales (Cohortes, Casos y Controles, Transversales):** Herramientas como **ROBINS-I (*Risk Of Bias In Non-randomised Studies - of Interventions*)** para estudios de intervención no aleatorizados, o listas de verificación adaptadas para estudios de asociación etiológica que evalúan aspectos como selección de participantes, medición de exposición y desenlace, control de confusores, etc.

- **Síntesis de los datos:** Describir cómo se sintetizaron los datos. Indicar si se realizó un **metaanálisis** (síntesis cuantitativa) o una **síntesis cualitativa/narrativa**. Si se realizó metaanálisis, describir los métodos estadísticos utilizados (ej. modelo de efectos fijos o aleatorios, cómo se evaluó la heterogeneidad [Prueba Q, I^2], cómo se manejó la heterogeneidad si estaba presente [análisis de subgrupos, metaregresión]). Si se realizó síntesis cualitativa, explicar cómo se agruparon y resumieron los hallazgos.
- **Riesgo de sesgo:** Describir cómo se evaluó el sesgo de publicación u otros sesgos de reporte (ej. inspección visual de gráficos de embudo, pruebas estadísticas de asimetría si aplica).

- **Análisis de sensibilidad:** Describir cualquier análisis de sensibilidad planificado para evaluar la robustez de los resultados (ej. excluyendo estudios de baja calidad, utilizando diferentes modelos estadísticos).

7. Resultados:

- Presentar los hallazgos de la revisión de manera clara y objetiva, siguiendo el flujo de los estudios a través de la revisión.
- **Selección de estudios:** Presentar el **Diagrama de Flujo PRISMA** (Figura 1) que documenta el número de registros identificados, examinados, elegibles, incluidos y excluidos (con razones).
- **Características de los estudios:** Presentar una tabla resumen (Tabla 1) con las características clave de cada estudio incluido (ej. primer autor, año, país, diseño, población, intervención, comparador, duración, desenlaces).
- **Riesgo de sesgo de los estudios incluidos:** Presentar un resumen de la evaluación del riesgo de sesgo, a menudo utilizando figuras (Figura 2: gráfico de resumen del riesgo de sesgo por dominio y por estudio; Figura 3: gráfico de riesgo de sesgo por dominio a través de todos los estudios). Describir las principales preocupaciones sobre el riesgo de sesgo.
- **Resultados de la síntesis:** Presentar los resultados de la síntesis de datos. Si se realizó metaanálisis, presentar los **gráficos de bosque (Forest Plots)** (Figura 4) para cada desenlace principal, mostrando los resultados de cada estudio individual y la medida resumen global con su IC 95% y valor p. Reportar las estadísticas de heterogeneidad (Prueba Q, I^2). Si se realizó síntesis cualitativa, resumir los hallazgos de los estudios de manera narrativa.
- **Riesgo de sesgo:** Presentar los resultados de la evaluación del sesgo de publicación (ej. gráfico de embudo - Figura 5 - e interpretación de su simetría/asimetría).
- **Análisis de sensibilidad:** Reportar los resultados de cualquier análisis de sensibilidad realizado.

8. Discusión:

- Interpretar los hallazgos de la revisión sistemática en el contexto de la pregunta de investigación y la literatura existente.
  - **Resumen de los hallazgos:** Resumir los hallazgos principales de la revisión, enfocándose en los resultados de la síntesis para los desenlaces primarios y secundarios clave, y los hallazgos de seguridad.
  - **Fortalezas y limitaciones de la revisión:** Discutir las fortalezas de la revisión (ej. búsqueda exhaustiva, métodos rigurosos, evaluación del riesgo de sesgo) y sus limitaciones (ej. calidad de los estudios primarios, heterogeneidad no explicada, posible sesgo de publicación, limitaciones de la síntesis cualitativa si aplica).
  - **Fortalezas y limitaciones de la evidencia:** Discutir la calidad general de la evidencia para cada desenlace principal, considerando el riesgo de sesgo de los estudios primarios, la heterogeneidad, la inconsistencia, la imprecisión y el sesgo de publicación (siguiendo el enfoque GRADE si se utilizó).
  - **Interpretación de los resultados:** Interpretar los resultados en el contexto de la pregunta de investigación. Si se realizó metaanálisis, discutir la magnitud y precisión del efecto estimado. Si hubo heterogeneidad, discutir sus posibles fuentes (clínicas, metodológicas) y si se pudieron explicar (ej. con análisis de subgrupos).
  - **Comparación con otra evidencia:** Comparar los hallazgos de esta revisión con otras revisiones o metaanálisis previos sobre el mismo tema, discutiendo similitudes y diferencias.
  - **Implicaciones en la práctica:** Discutir las posibles implicaciones de los hallazgos para la práctica clínica o las políticas de salud.
  - **Implicaciones en la investigación:** Identificar las brechas en el conocimiento y sugerir direcciones para futuras investigaciones primarias.

9. Conclusión:

- Resumir la conclusión principal de la revisión sistemática de manera concisa, respondiendo a la pregunta de investigación principal basada en la evidencia sintetizada.

10. Secciones Finales:

- Incluir secciones estándar como:
  - **Agradecimientos:** Agradecer a personas o instituciones que colaboraron.
  - **Financiación:** Declarar las fuentes de financiación para la revisión.
  - **Conflictos de Interés:** Declarar cualquier conflicto de interés de los autores.
  - **Contribuciones de los autores:** Describir la contribución específica de cada autor.
  - **Declaración de disponibilidad de datos:** Indicar si los datos de la revisión están disponibles y cómo acceder a ellos.

11. Referencias:

- Lista de todas las fuentes citadas en el texto, incluyendo los estudios primarios incluidos en la revisión sistemática y otras referencias relevantes (ej. guías de reporte, herramientas de riesgo de sesgo).

12. Tablas y Figuras:

- Incluir las tablas y figuras mencionadas en las secciones de Métodos y Resultados, con títulos y leyendas claras.

Siguiendo esta estructura y las guías PRISMA, se puede reportar una revisión sistemática de manera completa y transparente, permitiendo que la comunidad científica evalúe la solidez de sus hallazgos y confíe en las conclusiones derivadas de la síntesis de la evidencia.

4.4.6. *Metanálisis*

Hemos visto cómo las revisiones sistemáticas nos permiten recopilar y evaluar críticamente toda la evidencia disponible sobre una pregunta específica. Ahora, si los estudios incluidos en una revisión sistemática son lo suficientemente similares, podemos dar un paso más y combinarlos estadísticamente para obtener una estimación global del efecto. Este proceso se conoce como **Metaanálisis**.

Un metaanálisis es una técnica estadística que combina los resultados de dos o más estudios primarios que abordan la misma pregunta de investigación

de manera similar. Es la **síntesis cuantitativa** de una revisión sistemática. Su objetivo es obtener una medida resumen del efecto (ej. *Odds Ratio*, *Riesgo Relativo*, diferencia de medias) que sea más precisa y potente que la obtenida por cualquiera de los estudios individuales.

¿Por qué hacer un Metaanálisis?

Los metaanálisis se realizan por varias razones importantes:

- **Aumentar la potencia estadística:** Al combinar datos de múltiples estudios, el tamaño muestral total aumenta considerablemente, lo que incrementa la capacidad para detectar un efecto si realmente existe (reducir el riesgo de un error Tipo II).
- **Mejorar la precisión de la estimación del efecto:** La medida resumen del efecto obtenida de un metaanálisis suele tener un intervalo de confianza más estrecho que el de los estudios individuales, proporcionando una estimación más precisa del verdadero efecto poblacional.
- **Resolver controversias:** Pueden ayudar a aclarar situaciones en las que estudios individuales han arrojado resultados inconsistentes.
- **Proporcionar una estimación global del efecto:** Ofrecen una medida resumen del efecto de una intervención o exposición en una población más amplia y diversa que la de cualquier estudio individual.
- **Generar nuevas hipótesis:** La exploración de la heterogeneidad (ver más abajo) puede sugerir que el efecto varía en diferentes subgrupos de pacientes o bajo diferentes condiciones, generando nuevas preguntas de investigación.

¿Cuándo realizar un Metaanálisis?

La decisión de realizar o no un metaanálisis dentro de una revisión sistemática es crucial y depende principalmente de la **comparabilidad** o **homogeneidad** de los estudios primarios incluidos, así como de la disponibilidad de datos. No siempre es apropiado o incluso posible combinar estadísticamente los resultados.

Se debe considerar realizar un metaanálisis cuando:

1. **Los estudios son clínicamente homogéneos:** Tienen poblaciones de participantes, intervenciones, comparaciones y desenlaces lo suficientemente similares como para que tenga sentido combinar sus resultados. Por ejemplo, si varios ECA evalúan el mismo fármaco (misma dosis, duración) en la misma población de pacientes con la misma condición y miden el mismo desenlace principal.

2. **Los estudios son metodológicamente similares:** Tienen diseños de estudio comparables (ej. todos son ECA paralelos) y un riesgo de sesgo similar.

3. **Existe suficiente número de estudios:** Generalmente se requieren al menos dos estudios, pero se recomienda tener al menos cinco o diez estudios para que un metaanálisis sea robusto y para poder evaluar el sesgo de publicación.

4. **Los datos necesarios están disponibles:** Los estudios primarios deben reportar los datos de desenlace de una manera que permita su inclusión en el metaanálisis (ej. medias y desviaciones estándar para desenlaces continuos, número de eventos y participantes para desenlaces binarios).

5. **La heterogeneidad estadística es manejable:** Si bien cierta heterogeneidad es esperable, una heterogeneidad estadística muy alta (ej. I^2 consistentemente por encima del 75-90%) que no puede ser explicada por análisis de subgrupos o metaregresión puede indicar que los estudios son demasiado diferentes para combinarse estadísticamente.

No se debe realizar un metaanálisis cuando:

- **Los estudios son demasiado heterogéneos:** Las diferencias clínicas o metodológicas entre los estudios son tan grandes que combinar sus resultados sería como "mezclar peras con manzanas", llevando a una medida resumen que no representa adecuadamente el efecto en ninguna población o intervención específica.
- **Hay un número insuficiente de estudios:** Si solo hay uno o dos estudios, un metaanálisis no aportará valor estadístico adicional a la revisión sistemática.
- **El riesgo de sesgo es consistentemente alto y variable:** Si los estudios incluidos tienen un riesgo de sesgo significativo y diferente, la combinación de sus resultados podría producir una estimación sesgada.

- **Los datos necesarios no están disponibles** en los estudios primarios.

Ejemplo para entender cuándo NO se hace un Metaanálisis:

Consideremos el artículo proporcionado: "*Effectiveness of Probiotics, Prebiotics, and Synbiotics in Managing Insulin Resistance and Hormonal Imbalance in Women with Polycystic Ovary Syndrome (PCOS): A Systematic Review of Randomized Clinical Trials*" (https://www.mdpi.com/2072-6643/16/22/3916).

En la sección de Métodos (2.7. *Data Synthesis*) y en la Discusión (4.5. *Limitations of the Review*), los autores explican por qué **decidieron no realizar un metaanálisis cuantitativo** para la mayoría de los desenlaces a pesar de haber identificado 11 ECA. Las razones clave mencionadas son:

- **Heterogeneidad en las Intervenciones:** Los estudios utilizaron diferentes tipos de intervenciones (probióticos, prebióticos, simbióticos), diferentes cepas bacterianas, diferentes dosis y diferentes formas de administración (cápsulas, jugos).
- **Heterogeneidad en los Desenlaces Medidos:** Aunque abordaban resistencia a la insulina y hormonas, los marcadores específicos medidos y cómo se reportaron variaron entre los estudios.
- **Heterogeneidad Poblacional Limitada:** Aunque todos eran mujeres con SOP, las características específicas de las poblaciones (ej. fenotipos de SOP, gravedad de la resistencia a la insulina) pudieron variar.

Debido a esta **heterogeneidad clínica y metodológica**, los autores concluyeron que combinar estadísticamente los resultados sería inapropiado y podría llevar a conclusiones engañosas. En su lugar, optaron por una **síntesis cualitativa** (narrativa), resumiendo los hallazgos de cada estudio individual en tablas y texto para identificar patrones generales y discutir las diferencias entre los estudios.

Este ejemplo ilustra perfectamente que una revisión sistemática es el proceso riguroso de identificar, seleccionar y evaluar estudios, y el metaanálisis es solo una *posible* herramienta para sintetizar los resultados si los estudios son lo suficientemente comparables. La decisión de realizar un metaanálisis debe ser justificada y explícita en el reporte de la revisión sistemática.

Combinando Estimaciones en un Metaanálisis:

El metaanálisis no es simplemente un promedio simple de los resultados de los estudios. Utiliza un **promedio ponderado**, donde a cada estudio se le asigna un "peso" basado en su precisión. Los estudios más grandes y con menor variabilidad (mayor precisión) reciben un mayor peso en el cálculo de la medida resumen global. El peso suele ser inversamente proporcional a la varianza (el cuadrado del error estándar) de la medida de efecto del estudio.

Es crucial que los estudios que se combinan en un metaanálisis sean **comparables**. Aunque no tienen que ser idénticos, deben ser lo suficientemente similares en cuanto a la población, la intervención, la comparación y el desenlace para que tenga sentido combinarlos estadísticamente. Aquí es donde la evaluación crítica de los estudios primarios en la revisión sistemática es fundamental.

Heterogeneidad:

La **heterogeneidad** se refiere a la variabilidad o inconsistencia en los resultados de los estudios incluidos en un metaanálisis. Puede ser de diferentes tipos:

- **Heterogeneidad Clínica:** Diferencias en las características de los participantes, la intervención (dosis, duración), el comparador, el desenlace (definición, medición) o el contexto del estudio (ubicación, entorno clínico). En el ejemplo (https://www.mdpi.com/1648-9144/61/3/489) sobre Fibrosis Quística (FQ), esto podría incluir diferencias en la edad de los pacientes, la gravedad de la FQ, las cepas específicas de probióticos utilizadas, las dosis o la duración del tratamiento.
- **Heterogeneidad Metodológica:** Diferencias en el diseño del estudio, el riesgo de sesgo, la calidad metodológica o los métodos de análisis de los estudios primarios. En el artículo sobre FQ, se menciona la inclusión de ensayos cruzados (*crossover*) junto con ensayos paralelos como una fuente potencial de heterogeneidad metodológica.
- **Heterogeneidad Estadística:** Variabilidad en las medidas de efecto observadas entre los estudios que es mayor de lo que se esperaría solo por el azar.

Pruebas y Estimación de la Heterogeneidad Estadística:

Se utilizan métodos estadísticos para evaluar si existe heterogeneidad estadística significativa:

- **Prueba Q de Cochran:** Una prueba estadística (basada en la distribución Chi-cuadrado) que evalúa la H_0 de que no hay heterogeneidad entre los estudios (es decir, que todos los estudios están estimando un único efecto verdadero). Un valor p bajo (tradicionalmente < 0.10 en el contexto de heterogeneidad, no < 0.05) sugiere evidencia contra la H_0, indicando heterogeneidad estadística. Sin embargo, esta prueba tiene baja potencia con pocos estudios. En el artículo sobre FQ, se menciona el valor de chi-cuadrado y el valor p asociado para la heterogeneidad en los metaanálisis realizados.
- **Estadístico I^2:** Este estadístico describe la *magnitud* de la heterogeneidad. Representa el porcentaje de la variación total en las medidas de efecto que se debe a la heterogeneidad real entre los estudios, más allá de lo esperado por el azar. Varía de 0% a 100%. Un I^2 alto (ej. > 50% o 75%) sugiere una heterogeneidad sustancial. En el metaanálisis del artículo sobre la función pulmonar (FEV1) en FQ, se reporta un I^2 = 95%, lo que indica una heterogeneidad muy alta.

Cómo Lidiar con la Heterogeneidad:

Si se detecta heterogeneidad estadística significativa (especialmente si es alta según el I^2), hay varias opciones:

1. **No realizar un Metaanálisis:** Si la heterogeneidad es muy alta y no se pueden identificar sus fuentes, combinar los resultados puede ser inapropiado y llevar a conclusiones engañosas. En este caso, una síntesis cualitativa puede ser más adecuada. El artículo sobre probióticos y SOP (el otro documento mencionado) decidió no realizar metaanálisis para la mayoría de los desenlaces debido a la heterogeneidad clínica y metodológica.
2. **Realizar un Metaanálisis de Efectos Aleatorios (*Random-Effects Meta-analysis*):** Este modelo asume que los estudios no solo difieren debido al azar (variabilidad dentro del estudio), sino también porque el verdadero efecto puede variar entre los estudios (variabilidad entre estudios).

Incorpora esta variabilidad adicional en el cálculo de la medida resumen y su intervalo de confianza, lo que generalmente resulta en intervalos de confianza más amplios en comparación con el modelo de efectos fijos si hay heterogeneidad. Es más conservador y a menudo preferido cuando hay evidencia de heterogeneidad. El artículo sobre FQ menciona el uso de modelos de efectos aleatorios cuando el I^2 excede el 50% o en estudios con múltiples cepas probióticas.

3. **Realizar Análisis de Subgrupos (*Subgroup Analysis*):** Si se sospecha que la heterogeneidad se debe a diferencias clínicas o metodológicas específicas, se pueden realizar metaanálisis separados para subgrupos de estudios (ej. estudios con diferente duración de tratamiento, diferente tipo de producto [probiótico vs. simbiótico], o diferente número de cepas). Esto puede ayudar a identificar si el efecto varía en diferentes subgrupos. El artículo sobre FQ realiza análisis de subgrupos por duración de la intervención, tipo de producto y número de cepas.

4. **Realizar Meta-regresión:** Una técnica estadística que permite explorar si las diferencias en ciertas características de los estudios (a nivel de estudio, no individual) explican parte de la heterogeneidad observada.

Metaanálisis de Efectos Fijos (*Fixed-Effects Meta-analysis*):

Este modelo es más simple y se utiliza cuando se asume que no hay heterogeneidad estadística significativa (o es muy baja). Asume que todos los estudios están estimando un único efecto verdadero subyacente y que cualquier variación entre los resultados se debe únicamente al azar (error de muestreo dentro del estudio). Da más peso a los estudios más grandes. El artículo sobre FQ menciona el uso de modelos de efectos fijos cuando el I^2 es inferior al 50%.

Análisis de Sensibilidad (*Sensitivity Analysis*):

Es una técnica importante para evaluar la robustez de los resultados del metaanálisis. Implica repetir el metaanálisis bajo diferentes supuestos o excluyendo ciertos estudios para ver si la medida resumen global cambia significativamente. Por ejemplo, se puede realizar un análisis de sensibilidad excluyendo estudios de baja calidad, estudios con alto riesgo de sesgo de publicación, o estudios con características atípicas (ej. ensayos cruzados

si la pregunta principal se basa en ensayos paralelos). El artículo sobre FQ realiza análisis de sensibilidad excluyendo ensayos cruzados para evaluar si el diseño metodológico influye en los resultados.

Presentación de Metaanálisis: El Gráfico de Bosque (*Forest Plot*)

Los resultados de un metaanálisis se presentan típicamente en un **Gráfico de Bosque (*Forest Plot*)**. Este gráfico resume visualmente los resultados de cada estudio individual y la medida resumen global del metaanálisis.

- Cada **línea horizontal** representa un estudio individual.
- El **cuadrado o círculo** en cada línea representa la medida de efecto estimada para ese estudio (ej. OR, RR, MD). El **tamaño del cuadrado/círculo** es proporcional al peso que ese estudio recibió en el metaanálisis.
- La **línea horizontal** que atraviesa el cuadrado/círculo representa el **Intervalo de Confianza del 95% (IC 95%)** para la medida de efecto de ese estudio.
- La **línea vertical sólida** (generalmente en 1 para OR/RR o 0 para diferencia de medias) representa el **valor nulo** (sin efecto).
- El **diamante** en la parte inferior del gráfico representa la **medida resumen global** del metaanálisis. El **centro del diamante** es la estimación puntual del efecto combinado, y la **anchura del diamante** representa el **IC 95%** para esa estimación global.
- Si el IC 95% del diamante (el diamante completo) **no cruza la línea del valor nulo**, el resultado del metaanálisis es estadísticamente significativo.

El artículo sobre FQ incluye varios gráficos de bosque (Figuras 3, 4, 5, 6, 7, 8, 9) que ilustran estos componentes para los metaanálisis realizados sobre FEV1 y exacerbaciones pulmonares.

Evaluación e Interpretación de Metaanálisis:

Al interpretar los resultados de un metaanálisis, es importante considerar:

- La **medida resumen del efecto** (magnitud y dirección).
- El **Intervalo de Confianza del 95%** (precisión de la estimación).
- El **valor p** (significancia estadística).

- La **heterogeneidad** (magnitud y posibles fuentes).
- El **riesgo de sesgo** de los estudios primarios incluidos.
- La **aplicabilidad** de los resultados a la práctica clínica.

Si el metaanálisis muestra un efecto significativo, se debe evaluar si este efecto es también clínicamente relevante. Si hay heterogeneidad significativa, se debe intentar explicarla y considerar si la combinación de estudios fue apropiada.

Sesgo en Metaanálisis:

Además de los sesgos en los estudios primarios, los metaanálisis son susceptibles a sesgos relacionados con la **diseminación de la evidencia**:

- **Sesgo de Publicación (*Publication Bias*):** Ocurre si la probabilidad de que un estudio sea publicado depende de sus resultados. Los estudios con resultados positivos y significativos tienen más probabilidades de ser publicados, lo que puede llevar a una sobreestimación del efecto en el metaanálisis. Se puede explorar visualmente con **gráficos de embudo (*funnel plots*)**. Un gráfico de embudo simétrico sugiere ausencia de sesgo de publicación, mientras que uno asimétrico (falta de estudios pequeños con resultados no significativos) sugiere su presencia. También se pueden utilizar pruebas estadísticas de asimetría, como la **prueba de Egger**. En el artículo sobre FQ, se menciona que no se evaluó el sesgo de publicación con gráfico de embudo o prueba de Egger porque el metaanálisis incluía menos de 10 estudios, un número insuficiente para la aplicación fiable de estos métodos.

Ventajas y Desventajas del Metaanálisis:

Las **ventajas** del metaanálisis (además de las de la revisión sistemática):

- **Aumenta la potencia y precisión:** Permite detectar efectos pequeños con mayor confianza y obtener estimaciones más precisas.
- **Proporciona una estimación más robusta y generalizable** del efecto si los estudios son comparables.
- **Puede resolver inconsistencias** entre estudios individuales.

Las **desventajas** del metaanálisis:

- **Susceptible a la heterogeneidad:** Si los estudios son demasiado diferentes ("mezclar peras con manzanas"), el resultado del metaanálisis puede ser engañoso.
- **Susceptible a los sesgos de los estudios primarios:** Un metaanálisis de estudios de baja calidad o con alto riesgo de sesgo producirá resultados sesgados.
- **Susceptible a sesgos de publicación y otros sesgos de diseminación.**
- **Requiere experiencia estadística** para realizarse e interpretarse correctamente.

Consideraciones Éticas:

Las consideraciones éticas son similares a las de las revisiones sistemáticas, enfocándose en la transparencia, la integridad y la interpretación responsable de la evidencia agregada.

En resumen, el metaanálisis es una herramienta estadística poderosa que, cuando se aplica adecuadamente dentro del marco de una revisión sistemática rigurosa, nos permite sintetizar cuantitativamente la evidencia de múltiples estudios para obtener estimaciones más precisas y fiables del efecto de una intervención o asociación. Es una pieza clave en la MBE, que nos ayuda a ir más allá de los hallazgos de estudios individuales para obtener una comprensión más completa y robusta de la evidencia científica.

4.4.7. *Ejemplo de metanálisis*

La Decisión Crucial: ¿Hacer o No Hacer un Metaanálisis?

La clave para decidir si se debe realizar un metaanálisis es evaluar la **comparabilidad** o **homogeneidad** de los estudios primarios incluidos en la revisión sistemática. Si los estudios son demasiado diferentes, combinar sus resultados estadísticamente puede ser engañoso y producir una estimación sin sentido ("mezclar peras con manzanas").

Se debe considerar un metaanálisis si los estudios comparten suficiente similitud en:

1. **Homogeneidad Clínica:**

- **Población:** ¿Son los participantes similares en cuanto a edad, sexo, gravedad de la enfermedad, comorbilidades, etc.?
- **Intervención:** ¿Es la intervención (o exposición) la misma o muy similar en cuanto a tipo, dosis, duración, vía de administración?
- **Comparación:** ¿Es el grupo de comparación similar (ej. placebo, tratamiento estándar)?
- **Desenlace:** ¿Se mide el mismo desenlace de interés (ej. mortalidad, presión arterial, FEV1) utilizando métodos y unidades de medida comparables?

2. **Homogeneidad Metodológica:**

- **Diseño del Estudio:** ¿Son todos los estudios del mismo tipo de diseño (ej. todos ECA paralelos, todos estudios de cohorte)?
- **Riesgo de Sesgo:** ¿Tienen los estudios un riesgo de sesgo similar? (Aunque el metaanálisis puede explorar el impacto del riesgo de sesgo, una alta variabilidad en la calidad puede ser una razón para no combinar).

3. **Disponibilidad de Datos:** ¿Los estudios reportan los datos de desenlace en un formato que permita la combinación estadística (ej. medias y desviaciones estándar para desenlaces continuos, número de eventos y tamaño total del grupo para desenlaces binarios)?

4. **Número Suficiente de Estudios:** Aunque técnicamente se pueden combinar dos estudios, se necesita un número mayor (idealmente ≥5 o 10) para obtener estimaciones robustas y poder evaluar el sesgo de publicación.

Ejemplo: Aplicando la Decisión a la Revisión Sistemática sobre Probióticos y Fibrosis Quística

Consideremos la revisión sistemática sobre la efectividad de probióticos, prebióticos y simbióticos en pacientes con Fibrosis Quística (FQ), basada en el artículo "Effectiveness of Probiotics, Prebiotics, and Symbiotic Supplementation in Cystic Fibrosis Patients: A Systematic Review and Meta-Analysis of Clinical Trials" (https://www.mdpi.com/1648-9144/61/3/489).

Después de realizar la búsqueda y seleccionar los estudios primarios (13 ECA en este caso), los autores deben decidir para cada desenlace si es apropiado realizar un metaanálisis.

Desenlaces donde SÍ se consideró Metaanálisis:

- **Función Pulmonar (principalmente FEV1):** Varios estudios (8 de 13) midieron este desenlace. Aunque hubo cierta variabilidad en cómo se reportó (ej. valor absoluto vs. porcentaje predicho), era un desenlace cuantitativo común.
- **Exacerbaciones Pulmonares:** Varios estudios (5 de 13) reportaron el número de pacientes con exacerbaciones o la frecuencia de exacerbaciones. Este es un desenlace binario o de recuento que puede combinarse.

Para estos desenlaces, los autores del artículo encontraron que había un número suficiente de estudios que reportaban datos de manera que permitía la combinación estadística.

Desenlaces donde NO se realizó Metaanálisis (o solo síntesis cualitativa):

- **Hospitalizaciones:** Aunque 6 estudios evaluaron hospitalizaciones, la forma de reportar los datos pudo variar (ej. número de hospitalizaciones, número de pacientes hospitalizados, duración de la estancia).
- **Biomarcadores Inflamatorios (Calprotectina, Interleukinas, etc.):** Varios estudios evaluaron estos marcadores (7 de 13), pero la gran diversidad de marcadores específicos medidos, las unidades, los métodos de medición y la forma de reportar los resultados hacen que la combinación estadística sea muy difícil o inapropiada.
- **Síntomas Gastrointestinales:** Pocos estudios evaluaron este desenlace de manera comparable.
- **Calidad de Vida:** Solo dos estudios evaluaron este desenlace, un número insuficiente para un metaanálisis robusto.
- **Eventos Adversos:** Reportados de manera variable y a menudo cualitativa.

En estos casos, a pesar de que varios estudios podían haber investigado el mismo concepto (ej. inflamación), la **heterogeneidad clínica y metodoló-**

gica en la forma en que se definió y midió el desenlace, o el **número insuficiente de estudios** con datos combinables, llevaron a la decisión de no realizar un metaanálisis para estos desenlaces, sino una **síntesis cualitativa** (describir los hallazgos de cada estudio en texto y tablas, como se ve en la sección 3.5 del artículo).

Realizando el Metaanálisis:

Una vez que se decide que un metaanálisis es apropiado para un desenlace específico, se siguen estos pasos:

1. **Extracción de Datos:** Se extraen los datos numéricos relevantes de cada estudio para el desenlace de interés. Para FEV1 (continuo), se necesitan la media, la DS y el tamaño de la muestra para cada grupo (intervención y control). Para Exacerbaciones (binario), se necesita el número de pacientes con el evento y el número total de pacientes en cada grupo.
2. **Cálculo de la Medida de Efecto Individual:** Para cada estudio, se calcula una medida del efecto (ej. Diferencia de Medias [MD] para FEV1, Riesgo Relativo [RR] para Exacerbaciones) y su error estándar.
3. **Evaluación de la Heterogeneidad Estadística:** Se utilizan la Prueba Q y el estadístico I^2 para cuantificar la heterogeneidad.

- En el artículo, para FEV1, se reporta $I^2 = 95\%$ ($p < 0.00001$), indicando una heterogeneidad muy alta.
- Para Exacerbaciones, se reporta $I^2 = 57\%$ ($p = 0.05$), indicando heterogeneidad sustancial.

4. **Elección del Modelo Estadístico:**

- Si la heterogeneidad es baja (ej. $I^2 < 50\%$), se usa un modelo de **Efectos Fijos**. Este modelo asume que hay un único efecto verdadero y las diferencias entre estudios se deben solo al azar.
- Si la heterogeneidad es moderada a alta (ej. $I^2 \geq 50\%$), se usa un modelo de **Efectos Aleatorios**. Este modelo asume que el efecto verdadero puede variar entre los estudios (debido a diferencias clínicas o metodológicas reales) y que las diferencias observadas se deben tanto a esta variabilidad real como al azar.

- En el artículo, dado el alto I^2 para ambos desenlaces, se utilizó un modelo de efectos aleatorios (aunque no se especifica explícitamente en el Resumen o la sección de Resultados, es la práctica estándar con alta heterogeneidad).

5. **Cálculo de la Medida Resumen Global:** Se combinan las medidas de efecto individuales de cada estudio utilizando un promedio ponderado (dando más peso a los estudios más precisos) según el modelo elegido (efectos fijos o aleatorios). Esto produce la medida resumen global del metaanálisis (ej. el MD o RR combinado) y su IC 95% y valor p.

- En el artículo, para FEV1, el MD combinado fue 4.7 (IC 95%: -5.4 a 14.8; p = 0.37).

- Para Exacerbaciones, el RR combinado fue 0.81 (IC 95%: 0.48–1.37; p = 0.43).

6. **Presentación de los Resultados: Gráfico de Bosque (Forest Plot):** Los resultados se visualizan en un gráfico de bosque, que muestra el efecto y el IC 95% para cada estudio individual y la medida resumen global (el diamante). Los gráficos de bosque en el artículo (Figuras 3, 4, 5, 6, 7, 8, 9) son ejemplos de esta presentación.

7. **Exploración de la Heterogeneidad (Si es Alta):** Si la heterogeneidad es alta, se intenta identificar sus fuentes mediante:

- **Análisis de Subgrupos:** Realizar metaanálisis separados para grupos de estudios con características similares (ej. duración del tratamiento < 6 meses vs. ≥ 6 meses; probióticos vs. simbióticos; cepa única vs. múltiples cepas). El artículo realiza estos análisis de subgrupos (Figuras 4, 5, 6, 8, 9).

- **Metaregresión:** Analizar si ciertas características de los estudios (variables a nivel de estudio) explican la variabilidad en los resultados.

8. **Evaluación del Sesgo de Publicación:** Se utilizan gráficos de embudo (funnel plots) y pruebas estadísticas (ej. Prueba de Egger) para evaluar si hay evidencia de sesgo de publicación. En el artículo, esto no se realizó debido al bajo número de estudios (< 10).

9. **Análisis de Sensibilidad:** Repetir el metaanálisis excluyendo ciertos estudios (ej. de baja calidad, ensayos cruzados) para ver si los resultados

cambian significativamente. El artículo realiza análisis de sensibilidad excluyendo ensayos cruzados.

Interpretación del Metaanálisis (Basado en el artículo):

La interpretación se basa en la medida resumen global, su significancia estadística y la heterogeneidad:

- Para FEV1, el MD combinado (4.7) sugiere una pequeña mejora en la función pulmonar con probióticos/simbióticos, pero el amplio IC 95% (-5.4 a 14.8) cruza el valor nulo (0), y el valor p (0.37) no es significativo. La heterogeneidad es muy alta (I^2 = 95%). Los análisis de subgrupos no explicaron esta heterogeneidad ni mostraron efectos significativos en ningún subgrupo.
- Para Exacerbaciones, el RR combinado (0.81) sugiere una ligera reducción en el riesgo de exacerbaciones, pero el IC 95% (0.48–1.37) cruza el valor nulo (1), y el valor p (0.43) no es significativo. La heterogeneidad es sustancial (I^2 = 57%). Los análisis de subgrupos tampoco mostraron efectos significativos.

La conclusión del metaanálisis en el artículo es que, a pesar de la combinación estadística, no hay evidencia robusta de que los probióticos o simbióticos mejoren significativamente la función pulmonar o reduzcan las exacerbaciones en pacientes con FQ. La alta heterogeneidad sugiere que, si hay un efecto, podría variar mucho dependiendo de factores no explorados o de las diferencias entre los estudios.

En resumen, el metaanálisis es una poderosa herramienta de síntesis cuantitativa, pero su aplicación debe ser cuidadosa y justificada por la comparabilidad de los estudios primarios. La decisión de realizarlo es un paso crítico en el proceso de la revisión sistemática.

4.4.8. *Tamizaje, Diagnóstico y Evaluación del Rendimiento de Pruebas*

El camino de un individuo a través del espectro de una enfermedad a menudo comienza con la posibilidad de su presencia. La detección puede ocurrir en diferentes momentos y con diferentes propósitos. Comprender la distinción entre las pruebas utilizadas para la detección temprana en poblaciones

(tamizaje) y aquellas utilizadas para confirmar o descartar una enfermedad en individuos con síntomas (diagnóstico) es fundamental. Igualmente, es importante saber cómo evaluar cuán buenas son estas pruebas para cumplir su propósito.

4.4.8.1.*Tamizaje vs. Diagnóstico: Propósitos Diferentes*

Aunque ambos involucran la aplicación de pruebas para identificar la presencia de una condición de salud, el tamizaje y el diagnóstico tienen propósitos y poblaciones objetivo distintos:

- **Screening (Tamizaje):**

  - **Propósito:** Identificar individuos **asintomáticos** en una población aparentemente sana que tienen una alta probabilidad de tener una enfermedad. El objetivo es la detección temprana para permitir una intervención oportuna y potencialmente mejorar el pronóstico o reducir la transmisión.
  - **Población Objetivo:** Grupos de personas que se consideran en riesgo de una enfermedad particular, pero que no presentan síntomas.
  - **Características de la Prueba:** Las pruebas de tamizaje suelen ser rápidas, relativamente económicas, fáciles de administrar y **altamente sensibles**. Priorizan la identificación de la mayoría de los casos verdaderos, incluso si esto significa tener un número considerable de falsos positivos. Un resultado positivo en una prueba de tamizaje no confirma el diagnóstico; indica la necesidad de pruebas adicionales (pruebas confirmatorias).
  - **Ejemplo:** Mamografía para detectar cáncer de mama en mujeres de cierta edad, prueba de Papanicolaou para detectar lesiones precancerosas en el cuello uterino, medición de glucosa en ayunas para detectar diabetes en individuos con factores de riesgo.

- **Diagnóstico:**

  - **Propósito:** Confirmar o descartar la presencia de una enfermedad específica en individuos que ya presentan **síntomas** o signos sugestivos de la enfermedad, o que han tenido un resultado positivo en una prueba de tamizaje.

- **Población Objetivo:** Individuos con sospecha clínica de tener una enfermedad.
- **Características de la Prueba:** Las pruebas diagnósticas pueden ser más complejas, costosas o invasivas que las de tamizaje. A menudo priorizan la **especificidad** para minimizar los falsos positivos y evitar diagnósticos erróneos y tratamientos innecesarios. El "estándar de oro" o prueba de referencia es la prueba más precisa y definitiva para confirmar o descartar la enfermedad.
- **Ejemplo:** Biopsia para confirmar cáncer después de una mamografía anormal, cultivo bacteriano para diagnosticar una infección, angiografía coronaria para confirmar enfermedad arterial coronaria.

4.4.8.2. *Evaluación del Rendimiento de las Pruebas Diagnósticas y de tamizaje*

Para saber cuán buena es una prueba para clasificar correctamente a los individuos como enfermos o sanos, comparamos sus resultados con el verdadero estado de la enfermedad, determinado por el estándar de oro. Utilizamos una tabla de contingencia de 2x2 para organizar los resultados (ver tabla 21):

Tabla 21. *Tabla contingencia 2x2 para prueba de oro y resultado del nuevo test o prueba.*

		Estado de la enfermedad (Según estándar de oro)		
		Enfermo	**Sano**	**Total**
Resultado de la prueba	**Positivo**	Verdadero Positivo (VP)	Falso Positivo (FP)	VP + FP
	Negativo	Falso Negativo (FN)	Verdadero Negativo (VN)	FN + VN
	Total	VP + FN	FP + VN	**N = VP+FP+-FN+VN**

Nota. **VP**: Verdadero Positivo; **FP**: Falso Positivo; **FN**: Falso Negativo; **VN**: Verdadero Negativo

A partir de esta tabla, calculamos las métricas clave:

- **Sensibilidad:** La proporción de verdaderos positivos entre todos los que realmente tienen la enfermedad. Mide la capacidad de la prueba para detectar a los enfermos.

$$\text{Sensibilidad} = \frac{VP}{VP + FN}$$

Especificidad: La proporción de verdaderos negativos entre todos los que realmente no tienen la enfermedad. Mide la capacidad de la prueba para descartar la enfermedad en los sanos.

$$\text{Especificidad} = \frac{VN}{VN + FP}$$

• Existe una **compensación** entre sensibilidad y especificidad: aumentar una generalmente disminuye la otra al cambiar el umbral de la prueba (punto de corte para un resultado positivo).

• Valor **Predictivo Positivo (VPP):** La probabilidad de que una persona con un resultado positivo en la prueba realmente tenga la enfermedad. Es la proporción de verdaderos positivos entre todos los que dieron positivo en la prueba.

$$\text{VPP} = \frac{VP}{VP + FP}$$

Valor Predictivo Negativo (VPN): La probabilidad de que una persona con un resultado negativo en la prueba realmente no tenga la enfermedad. Es la proporción de verdaderos negativos entre todos los que dieron negativo en la prueba.

$$\text{VPN} = \frac{VN}{VN + FN}$$

• Importante**:** A diferencia de la sensibilidad y la especificidad, el VPP y el VPN **dependen de la prevalencia de la enfermedad** en la población donde se aplica la prueba. Una mayor prevalencia generalmente aumenta el VPP y disminuye el VPN, asumiendo que la sensibilidad y especificidad de la prueba son constantes.

• Razón **de Verosimilitud Positiva (LR+):** Indica cuánto más probable es un resultado positivo en la prueba en una persona enferma que en una persona sana. Se calcula a partir de la sensibilidad y la especificidad.

$$LR+ = \frac{\text{Sensibilidad}}{1 - \text{Especificidad}} = \frac{VP/(VP + FN)}{FP/(FP + VN)}$$

Un LR+ > 1 aumenta la probabilidad post-prueba de la enfermedad; cuanto mayor sea el LR+ (ej. > 10), mayor será el aumento.

Razón de Verosimilitud Negativa (LR-): Indica cuánto menos probable es un resultado negativo en la prueba en una persona enferma que en una persona sana.

$$LR- = \frac{1 - \text{Sensibilidad}}{\text{Especificidad}} = \frac{FN/(VP + FN)}{VN/(FP + VN)}$$

Un LR- < 1 disminuye la probabilidad posprueba de la enfermedad; cuanto menor sea el LR- (ej. < 0.1), mayor será la disminución. Los LR son útiles porque, a diferencia de los valores predictivos, **no dependen de la prevalencia**.

4.4.8.3. Del *Pretest* al Post-test: *Probabilidad y Odds*

En la práctica clínica, a menudo queremos saber la probabilidad de que un paciente tenga la enfermedad *después* de realizar una prueba, basándonos en la probabilidad que teníamos *antes* de la prueba.

- **Probabilidad Pre-prueba:** La probabilidad de que un paciente tenga la enfermedad antes de realizar la prueba. Se basa en la información clínica (síntomas, signos, factores de riesgo) y la prevalencia de la enfermedad en poblaciones similares.
- **Odds Pre-prueba:** Una forma alternativa de expresar la probabilidad pre-prueba.

$$Odds_{pre} = \frac{Probabilidad_{pre}}{1 - Probabilidad_{pre}}$$

Odds Post-prueba: Los odds de que el paciente tenga la enfermedad después de conocer el resultado de la prueba. Se calculan multiplicando los odds pre-prueba por la razón de verosimilitud apropiada (LR+ si el resultado es positivo, LR- si el resultado es negativo).

$$Odds_{post} = Odds_{pre} \times LR$$

Probabilidad Post-prueba: La probabilidad de que el paciente tenga la enfermedad después de conocer el resultado de la prueba. Se convierte a partir de los odds post-prueba:

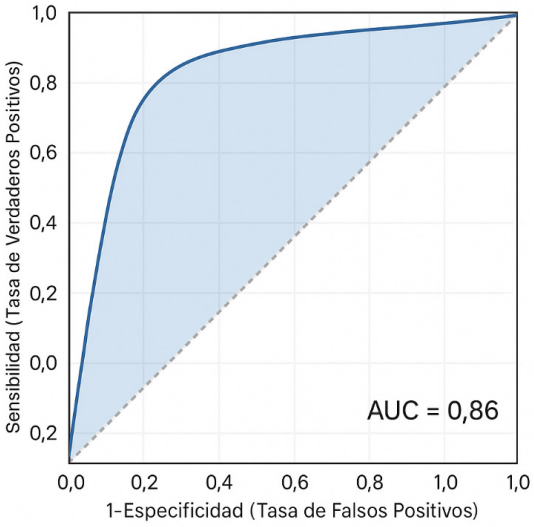
$$Probabilidad_{post} = \frac{Odds_{post}}{Odds_{post}+1}$$

Nomograma de Fagan: Una herramienta gráfica que permite convertir rápidamente la probabilidad pre-prueba en probabilidad post-prueba utilizando el LR de la prueba, sin necesidad de calcular los odds intermedios.

Curva ROC y Área Bajo la Curva:

Para pruebas con resultados continuos (ej. un nivel de biomarcador), podemos evaluar el rendimiento de la prueba en todos los posibles puntos de corte. La **Curva ROC** es un gráfico que representa la sensibilidad (Tasa de Verdaderos Positivos) en el eje Y frente a 1-Especificidad (Tasa de Falsos Positivos) en el eje X para todos los umbrales posibles. Una curva que se acerca a la esquina superior izquierda indica un mejor rendimiento. El **Área Bajo la Curva (AUC)** resume el rendimiento general de la prueba en una sola métrica, variando de 0.5 (rendimiento aleatorio) a 1.0 (rendimiento perfecto). Un AUC más cercano a 1.0 indica que la prueba discrimina mejor entre individuos enfermos y sanos (ver figura 61).

Figura 61. *Curva ROC con AUC*. Elaborado con GeminiPRO.



Nota. El eje Y representa la sensibilidad (tasa de verdaderos positivos) y el eje X representa 1-especificidad (tasa de falsos positivos). La curva ilustra el rendimiento del test en todos los puntos de corte posibles. El AUC es 0,86, lo que indica una buena capacidad discriminativa del test para diferenciar entre individuos enfermos y sanos.

4.4.8.4. Fiabilidad (Confiabilidad) y Repetibilidad de las Pruebas

Además de ser válida (medir lo que se supone que debe medir), una prueba diagnóstica debe ser **fiable** o **confiable**. La fiabilidad se refiere a la **consistencia** de los resultados de la prueba cuando se repite bajo las mismas condiciones. Una prueba fiable produce resultados similares en mediciones repetidas del mismo individuo o muestra.

- **Repetibilidad (*Test-retest reliability*):** Se refiere a la consistencia de los resultados cuando la misma prueba se administra al mismo individuo en diferentes ocasiones, asumiendo que el estado del individuo no ha cambiado. Evalúa si la prueba produce resultados consistentes a lo largo del tiempo.
- **Confiabilidad Inter-observador (*Inter-rater reliability*):** Se refiere a la consistencia de los resultados cuando la misma prueba es administrada o interpretada por diferentes personas (observadores, evaluadores, técnicos). Es crucial para pruebas que implican cierto grado de juicio subjetivo en su administración o interpretación (ej. lectura de imágenes radiológicas, evaluación de síntomas).
- **Confiabilidad Intra-observador (*Intra-rater reliability*):** Se refiere a la consistencia de los resultados cuando la misma persona administra o interpreta la misma prueba en diferentes ocasiones.

Evaluando la Confiabilidad: El Coeficiente Kappa de Cohen

Para evaluar la confiabilidad inter-observador o intra-observador para variables categóricas (ej. resultado de la prueba: positivo/negativo; diagnóstico: presente/ausente), se utiliza comúnmente el **Coeficiente Kappa de Cohen (κ)**.

Kappa mide el grado de acuerdo entre dos evaluadores (o entre dos mediciones del mismo evaluador) más allá de lo que se esperaría por puro azar.

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

Donde:

- P_o es la proporción de acuerdo observado (el número de casos en los que ambos evaluadores están de acuerdo, dividido por el total de casos).
- P_e es la proporción de acuerdo esperado por azar.

El valor de Kappa varía típicamente entre 0 y 1:

- κ=0: El acuerdo observado es igual al acuerdo esperado por azar (no hay acuerdo real).
- κ=1: Acuerdo perfecto (más allá del azar).
- Valores negativos de Kappa indican un acuerdo peor que el esperado por azar.

Interpretación común de los valores de Kappa:

- < 0.00: Pobre acuerdo
- 0.00 - 0.20: Ligero acuerdo
- 0.21 - 0.40: Acuerdo justo
- 0.41 - 0.60: Acuerdo moderado
- 0.61 - 0.80: Acuerdo sustancial
- 0.81 - 1.00: Acuerdo casi perfecto

Una prueba fiable debe tener un valor de Kappa aceptable para asegurar que los resultados no varían significativamente dependiendo de quién la realice o la interprete.

4.4.8.5. *Validez Interna y Externa en Estudios de Pruebas Diagnósticas/tamizaje*

- **Validez Interna:** Se refiere a cuán válidos son los resultados del estudio *dentro* del estudio mismo. En el contexto de la evaluación de pruebas,

la validez interna se ve amenazada por sesgos que afectan la estimación precisa de la sensibilidad, especificidad y otras métricas de rendimiento. Un estudio con alta validez interna proporciona estimaciones no sesgadas del rendimiento de la prueba en la población estudiada.

- **Validez Externa (Generalizabilidad):** Se refiere a cuán aplicables son los resultados del estudio a otras poblaciones, entornos o momentos. Incluso si un estudio tiene alta validez interna, sus resultados pueden no ser generalizables si la población estudiada (ej. pacientes de un centro de referencia altamente especializado) difiere significativamente de la población en la que se aplicará la prueba en la práctica clínica habitual. El sesgo de espectro es una amenaza importante para la validez externa.

4.4.8.6. *Sesgos Específicos en Estudios de Pruebas Diagnósticas*

Los estudios que evalúan el rendimiento de las pruebas son susceptibles a sesgos que pueden inflar o desinflar artificialmente las estimaciones de sensibilidad, especificidad y valores predictivos, afectando su validez interna:

- **Sesgo de Espectro (*Spectrum Bias*):** Ocurre cuando la población de pacientes enfermos y sanos incluida en el estudio no representa el rango completo de la enfermedad o la salud en la población donde se usará la prueba. Por ejemplo, incluir solo casos muy graves y controles muy sanos puede sobreestimar la sensibilidad y especificidad.
- **Sesgo de Verificación (*Verification Bias*):** Ocurre cuando no todos los participantes se someten al estándar de oro, o la decisión de aplicar el estándar de oro depende del resultado de la prueba en estudio. Esto puede llevar a que los VP, FN, FP y VN no se cuenten correctamente.
- **Sesgo de Pérdida de Seguimiento (*Loss-to-Follow-up Bias*):** Relevante en estudios que siguen a los participantes a lo largo del tiempo, si la pérdida de seguimiento difiere entre grupos o resultados de prueba.
- **Sesgo de Reporte (*Reporting Bias*):** Ocurre si el conocimiento del resultado de la prueba en estudio influye en la evaluación del estándar de oro, o viceversa. El cegamiento de quienes administran e interpretan las pruebas es crucial para minimizar este sesgo.

4.4.8.7. *Ventajas y Desventajas de los Programas de Tamizaje*

Los programas de tamizaje, aunque prometedores para la detección temprana, también tienen inconvenientes:

- **Ventajas:** Potencial para reducir la morbilidad y mortalidad mediante la detección temprana y el tratamiento oportuno, identificar factores de riesgo, proporcionar tranquilidad a los individuos con resultados negativos.
- **Desventajas:**
  - **Falsos Positivos:** Generan ansiedad, estrés y llevan a pruebas de confirmación innecesarias (que pueden ser costosas, invasivas o tener riesgos).
  - **Falsos Negativos:** Proporcionan una falsa sensación de seguridad, lo que puede retrasar el diagnóstico y el tratamiento.
  - **Sobrediagnóstico (*Overdiagnosis*):** Identificación de condiciones que cumplen los criterios diagnósticos pero que nunca habrían causado síntomas o daño clínico en la vida del individuo. Esto lleva a tratamiento innecesario y sus riesgos.
  - Costos del programa y las pruebas de seguimiento.
  - Posibles efectos adversos de la prueba de tamizaje misma.

4.4.8.8. *Sesgos Específicos en Estudios de Efectividad de tamizaje*

Evaluar si un programa de tamizaje realmente mejora los resultados de salud a nivel poblacional (ej. reduce la mortalidad por la enfermedad) es desafiante debido a sesgos que pueden hacer que el tamizaje parezca más beneficioso de lo que realmente es:

- **Sesgo de Selección (*Selection Bias*):** Los individuos que eligen participar en programas de tamizaje pueden ser diferentes de los que no participan (ej. más saludables, mayor acceso a atención médica), lo que puede influir en los resultados independientemente del tamizaje.
- **Sesgo de Tiempo de Adelanto (*Lead-Time Bias*):** Los pacientes diagnosticados por tamizaje parecen sobrevivir más tiempo simplemente porque su diagnóstico se hizo antes, no porque el tamizaje haya prolon-

gado su vida real. Es el tiempo adicional entre el diagnóstico por tamizaje y el diagnóstico por síntomas.

- **Sesgo de Tiempo de Duración (*Length Time Bias*):** Las pruebas de tamizaje tienen más probabilidades de detectar casos de la enfermedad de progresión lenta (que tienen un pronóstico naturalmente mejor) que casos de progresión rápida. Esto puede inflar la supervivencia aparente en el grupo de tamizaje.

Conclusión

El tamizaje y el diagnóstico son componentes esenciales de la atención médica, pero la interpretación de sus resultados requiere una comprensión sólida de las métricas de rendimiento (validez) y fiabilidad de las pruebas, así como los sesgos que pueden afectarlas. La elección de la prueba adecuada, la interpretación de sus resultados en el contexto clínico del paciente (probabilidad pre-prueba) y la consideración de los posibles sesgos son fundamentales para una toma de decisiones informada y para evaluar críticamente la literatura sobre pruebas diagnósticas y programas de tamizaje.

4.4.9. *Análisis de Sobrevida y Tiempo al Evento*

Una vez que se ha diagnosticado una enfermedad o se ha aplicado una intervención, a menudo estamos interesados en el tiempo que tarda en ocurrir un evento específico. Este evento puede ser la muerte (de ahí el término "sobrevida"), pero también puede ser la recurrencia de una enfermedad, la aparición de una complicación, la remisión de los síntomas, o cualquier otro resultado que ocurra en un punto definido en el tiempo. El análisis de sobrevida nos proporciona herramientas estadísticas para estudiar estos tiempos hasta el evento.

4.4.9.1. *Introducción al Análisis de Sobrevida y Datos de Tiempo al Evento*

El análisis de sobrevida es un conjunto de métodos estadísticos diseñados para analizar datos donde la variable de interés es el **tiempo hasta que ocurre un evento específico**. Estos datos se conocen como datos de **tiempo al evento**.

A diferencia de los análisis que solo consideran si un evento ocurrió o no (variables binarias) o mediciones en un solo punto en el tiempo, el análisis de sobrevida incorpora explícitamente la **dimensión temporal**. Considera cuánto tiempo vivió un individuo *hasta* el evento o *hasta* el final del seguimiento.

4.4.9.2. ¿Por qué el Análisis de Sobrevida es Diferente? Manejando la Censura

La característica distintiva y principal razón por la que necesitamos métodos especiales para el análisis de sobrevida es la presencia de **censura**. La censura ocurre cuando no conocemos el tiempo exacto hasta el evento para todos los individuos en el estudio. Esto puede suceder por varias razones:

- **Censura Administrativa:** El estudio finaliza antes de que el evento haya ocurrido en todos los participantes.
- **Pérdida de Seguimiento:** Un participante abandona el estudio antes de que ocurra el evento o antes de que finalice el estudio.
- **Ocurrencia de un Evento Competitivo:** Un participante experimenta otro evento que impide que el evento de interés ocurra (ej. un paciente en un estudio sobre tiempo hasta la recurrencia del cáncer muere por una causa no relacionada antes de que el cáncer recurra).

En todos estos casos, sabemos que el individuo estuvo "en riesgo" durante un cierto período de tiempo y que el evento *no* ocurrió durante ese tiempo. Ignorar a estos individuos o tratarlos como si no hubieran experimentado el evento subestimaría la verdadera tasa de eventos. El análisis de sobrevida permite incorporar la información de estos individuos censurados de manera adecuada.

4.4.9.3. Conceptos Clave en el Análisis de Sobrevida

- **Tiempo Cero (Origen):** El punto de inicio desde el cual se mide el tiempo. Debe estar claramente definido (ej. fecha del diagnóstico, fecha de inicio del tratamiento, fecha de reclutamiento en el estudio).
- **Evento:** El resultado de interés que ocurre en un punto específico en el tiempo (ej. muerte, recurrencia de la enfermedad, infarto de miocardio).

- **Tiempo al Evento:** La duración desde el Tiempo Cero hasta que ocurre el Evento.
- **Censura:** Cuando el tiempo al evento no se conoce exactamente. La mayoría de las veces se trata de **censura por la derecha**, donde sabemos que el evento no ocurrió hasta un cierto tiempo, pero no sabemos si ocurrió después.

4.4.9.4. *Métodos Descriptivos: Tablas y Curvas de Sobrevida*

Estos métodos se utilizan para describir la distribución del tiempo al evento en una población o en diferentes grupos.

- **Tablas de Sobrevida (Tablas de Vida):** Dividen el período de seguimiento en intervalos de tiempo discretos. Para cada intervalo, calculan:
  - El número de individuos que estaban vivos y en riesgo al comienzo del intervalo.
  - El número de eventos que ocurrieron durante el intervalo.
  - El número de individuos censurados durante el intervalo.
  - La probabilidad de sobrevivir durante ese intervalo.
  - La probabilidad de sobrevida acumulada desde el inicio del estudio hasta el final del intervalo.

Son útiles para resumir datos de sobrevida, especialmente en estudios con un número moderado de eventos.

- **Curvas de Sobrevida (Curvas de Kaplan-Meier):** Son la forma más común de visualizar la distribución del tiempo al evento. Estiman la probabilidad de sobrevida a lo largo del tiempo, teniendo en cuenta la censura.
  - El eje X representa el tiempo desde el inicio del estudio.
  - El eje Y representa la probabilidad de sobrevida acumulada (la proporción de individuos que aún no han experimentado el evento).
  - La curva comienza en 1.0 (100% de probabilidad de sobrevida al tiempo cero) y desciende gradualmente a medida que ocurren eventos.

- Las "escaleras" o descensos en la curva ocurren cada vez que se observa un evento.
- Las marcas en la curva (*ticks*) indican los puntos en el tiempo en que los individuos fueron censurados.

Las curvas de Kaplan-Meier son muy útiles para comparar visualmente la sobrevida entre diferentes grupos (ej. grupo de tratamiento vs. grupo control).

4.4.9.5. *Comparando Curvas de Sobrevida: La Prueba de Log-Rank*

Cuando se compara la sobrevida entre dos o más grupos (ej. en un ECA o un estudio de cohortes), la **Prueba de Log-Rank** es una prueba estadística no paramétrica comúnmente utilizada.

- **Propósito:** Evaluar si existen diferencias estadísticamente significativas en las curvas de sobrevida entre los grupos que se comparan. La H_0 es que no hay diferencia en la sobrevida entre los grupos.
- **Funcionamiento:** Compara el número observado de eventos en cada grupo con el número esperado de eventos si la H_0 fuera cierta, en cada momento en que ocurre un evento. Suma estas diferencias a lo largo del tiempo.
- **Resultado:** Proporciona un estadístico de prueba y un valor p. Un valor p bajo (ej. < 0.05) sugiere que hay evidencia para rechazar la H_0 y concluir que existen diferencias significativas en la sobrevida entre los grupos.

4.4.9.6. *Cuantificando el Efecto de Covariables: Regresión de Cox*

Si queremos evaluar el efecto de una o más variables (covariables) en el tiempo hasta el evento, controlando por otras variables, utilizamos modelos de regresión. El modelo más común en el análisis de sobrevida es la **Regresión de Riesgos Proporcionales de Cox (*Cox Proportional Hazards Regression*)**.

- **Propósito:** Estimar la relación entre una o más covariables y el riesgo instantáneo de experimentar el evento en un momento dado (el "riesgo" o "hazard"). Permite evaluar el efecto de una intervención (ej. tratamiento) o de factores pronósticos (ej. edad, etapa de la enfermedad) en la sobrevida, ajustando por el efecto de otras variables.
- **Funcionamiento:** Modela el riesgo instantáneo del evento en función de las covariables. La característica clave del modelo de Cox es que asu-

me que el efecto de las covariables es **proporcional** a lo largo del tiempo (es decir, el riesgo relativo asociado a una covariable es constante en el tiempo - Asunción de Riesgos Proporcionales).

- **Resultado:** El resultado principal del modelo de Cox es el ***Hazard* Ratio (HR)** para cada covariable.

4.4.9.7. *Interpretando el Hazard* Ratio (HR)

El **HR** es la medida de efecto clave obtenida de un modelo de regresión de Cox.

- **Definición:** Es la razón de los riesgos instantáneos de experimentar el evento entre dos grupos o para un cambio de una unidad en una covariable continua.
- **Interpretación:**
  - **HR = 1:** No hay diferencia en el riesgo instantáneo del evento entre los grupos o para el cambio en la covariable. Las curvas de sobrevida serían idénticas.
  - **HR > 1:** El riesgo instantáneo del evento es mayor en el grupo de interés o a medida que aumenta la covariable. Esto implica una **menor sobrevida**. Por ejemplo, un HR de 2.0 para un tratamiento significa que los pacientes que reciben ese tratamiento tienen el doble de riesgo instantáneo de experimentar el evento (ej. morir) en cualquier momento dado en comparación con el grupo control.
  - **HR < 1:** El riesgo instantáneo del evento es menor en el grupo de interés o a medida que aumenta la covariable. Esto implica una **mayor sobrevida**. Por ejemplo, un HR de 0.50 para un tratamiento significa que los pacientes que reciben ese tratamiento tienen la mitad del riesgo instantáneo de experimentar el evento en cualquier momento dado en comparación con el grupo control.
- **Intervalo de Confianza (IC 95%):** Al igual que otras medidas de efecto, el HR se acompaña de un IC 95%. Si el IC 95% para el HR incluye el valor 1, el resultado no es estadísticamente significativo (no podemos descartar que el verdadero HR sea 1).

- **Valor p:** La significancia estadística del HR se evalúa con un valor p. Un valor p bajo (ej. < 0.05) sugiere que el HR observado es estadísticamente significativo.

4.4.9.8. *Asunciones del Modelo de Regresión de Cox*

La validez de los resultados de un modelo de Cox depende de que se cumplan ciertas asunciones, la más importante es la **Asunción de Riesgos Proporcionales**.

- **Asunción de Riesgos Proporcionales:** Asume que el HR entre dos grupos o asociada a una covariable es constante a lo largo del tiempo. En otras palabras, si un tratamiento reduce el riesgo a la mitad al inicio del estudio, se asume que también reduce el riesgo a la mitad en momentos posteriores. Esta asunción debe ser evaluada (ej. gráficamente o con pruebas estadísticas) antes de confiar en los resultados del modelo de Cox. Si no se cumple, se pueden usar modelos más complejos (ej. modelos de Cox estratificados o con riesgos dependientes del tiempo).

Conclusión

El análisis de sobrevida y tiempo al evento es esencial para estudiar resultados que ocurren a lo largo del tiempo. Los métodos descriptivos como las tablas de sobrevida y las curvas de Kaplan-Meier nos permiten visualizar la distribución del tiempo al evento y comparar grupos. La prueba de Log-Rank nos ayuda a determinar si las diferencias observadas en las curvas de sobrevida son estadísticamente significativas. Finalmente, el modelo de regresión de Cox y el HR nos permiten cuantificar el efecto de las covariables en el riesgo del evento, ajustando por otros factores. Comprender estos métodos es fundamental para interpretar la investigación que evalúa el tiempo hasta desenlaces importantes en salud.

Puntos Clave del Capítulo

1. **Clasificación Fundamental:** Los diseños de estudio se dividen principalmente en Observacionales (el investigador observa) y Experimentales/Intervencionales (el investigador asigna la exposición). La elección depende de la pregunta, recursos y ética.

2. **Estudios Observacionales Descriptivos:** Describen la distribución de la salud (quién, dónde, cuándo). Incluyen Reportes/Series de Casos (novedades, hipótesis), Estudios Transversales (prevalencia "instantánea") y Estudios Ecológicos (datos grupales, riesgo de falacia ecológica).
3. **Estudios Observacionales Analíticos:** Exploran asociaciones entre exposición y desenlace. Incluyen Estudios Transversales (asociación en un punto temporal), Estudios de Casos y Controles (retrospectivos, comparan exposición pasada en enfermos vs. sanos, usan OR, eficientes para enfermedades raras) y Estudios de Cohorte (prospectivos o retrospectivos, siguen grupos según exposición para medir incidencia y RR, establecen temporalidad).
4. **Estudios Experimentales (ECA):** Son el estándar de oro para inferir causalidad. La aleatorización minimiza sesgos y confusión. Comparan una intervención con un control (placebo o estándar). Siguen fases (Preclínica, I, II, III, IV). El análisis principal es por Intención de Tratar (ITT).
5. **Jerarquía de Evidencia:** Existe una jerarquía en la capacidad de los diseños para establecer causalidad: Metaanálisis/Revisiones Sistemáticas > ECA > Cohorte > Casos y Controles > Transversal > Ecológico > Serie/Reporte de Caso.
6. **Síntesis de Evidencia:** Las Revisiones (Narrativa, Alcance, Sistemática) y el Metanálisis integran resultados de múltiples estudios. Los Análisis Bibliométricos estudian patrones de publicación. PRISMA y herramientas RoB guían la calidad de las revisiones.
7. **Screening y Diagnóstico:** El tamizaje busca detectar enfermedades en asintomáticos (alta sensibilidad); el diagnóstico confirma la enfermedad en sintomáticos (alta especificidad). Su rendimiento se evalúa con Sensibilidad, Especificidad, VPP, VPN, LR y Curvas ROC (AUC). La fiabilidad se mide con Kappa. Requieren evaluación cuidadosa de sesgos (espectro, verificación, adelanto, duración).
8. **Análisis de Sobrevida:** Analiza el tiempo hasta un evento, manejando datos censurados. Kaplan-Meier describe la sobrevida, Log-Rank compara curvas entre grupos, y la Regresión de Cox estima HR ajustando por covariables, asumiendo riesgos proporcionales.

CAPÍTULO 16

INTERPRETACIÓN CRÍTICA DE LA LITERATURA CIENTÍFICA

Critical Interpretation of Scientific Literature

Resumen

Este capítulo enfatiza la necesidad de realizar una interpretación crítica de la literatura científica para discernir la fiabilidad y aplicabilidad de la evidencia. Introduce los conceptos clave de sesgo (error sistemático), confusión (distorsión por un tercer factor), validez interna (corrección de los resultados dentro del estudio) y validez externa (generalizabilidad). Propone un proceso general de evaluación crítica en pasos, aplicable a cualquier estudio: comprender la pregunta, identificar el diseño, evaluar población, mediciones, manejo de confusores, análisis estadístico y otros sesgos, e interpretar resultados en contexto. Luego, detalla cómo aplicar este análisis crítico a diferentes tipos de estudios: de desenlace (outcome studies, evaluando relevancia y medición del resultado, usando métricas como DR, RRR, NNT), reportes/series de casos (evaluando descripción, novedad, sesgos), estudios descriptivos, ecológicos (cuidado con la falacia ecológica), transversales analíticos (limitación temporal), casos y controles (sesgo de selección/recuerdo), cohortes (pérdida de seguimiento), estudios de sobrevivencia (manejo de censura, riesgos proporcionales) y ECA (aleatorización, ocultación, cegamiento, ITT). Finalmente, aborda la evaluación crítica de síntesis de evidencia (revisiones narrativas, sistemáticas, metanálisis), contenido de divulgación (YouTubers/expertos) y el uso emergente de IA para asistir en este análisis.

Palabras clave: Lectura crítica, evaluación crítica, sesgo, confusión, validez interna, validez externa.

Abstract

This chapter emphasizes the need for critical appraisal of scientific literature to discern the reliability and applicability of evidence. It introduces key concepts like bias (systematic error), confounding (distortion by a third factor), internal

validity (correctness of results within the study), and external validity (generalizability). It proposes a general step-by-step critical appraisal process applicable to any study: understanding the research question, identifying the design, evaluating the population, measurements, handling of confounders, statistical analysis, and other biases, and interpreting results in context. It then details how to apply this critical analysis to different study types: outcome studies (assessing outcome relevance and measurement, using metrics like RD, RRR, NNT), case reports/series (evaluating description, novelty, bias), descriptive studies, ecological studies (beware of ecological fallacy), analytical cross-sectional studies (temporal limitation), case-control studies (selection/recall bias), cohort studies (loss to follow-up), survival studies (handling censoring, proportional hazards), and RCTs (randomization, concealment, blinding, ITT). Finally, it addresses the critical appraisal of evidence synthesis (narrative/systematic reviews, meta-analysis), health disclosure content (YouTubers/experts), and the emerging use of AI to assist in this analysis.

Keywords: *Critical appraisal, bias, confounding, internal validity, external validity, evidence assessment.*

5.1. Principios Generales de la Evaluación Crítica (Sesgo, Confusión, Validez Interna y Externa)

Hemos recorrido el fascinante camino de la investigación, desde la chispa inicial de una pregunta hasta la construcción de estudios rigurosos y la síntesis de la evidencia. Ahora, con el mapa de los diseños de estudio en nuestras manos y las herramientas estadísticas a nuestra disposición, nos enfrentamos a una tarea igualmente crucial, la de discernir la fiabilidad y la relevancia de la información que encontramos en la vasta y creciente marea de la literatura científica. No toda publicación, por muy bien intencionada que sea, ofrece evidencia igualmente sólida o aplicable a nuestros pacientes. Como detectives de la verdad médica, debemos aprender a mirar más allá del titular o el resumen, a evaluar con ojo crítico la metodología, a identificar las trampas ocultas y a sopesar el verdadero valor de un estudio.

La medicina, en su esencia más noble, se nutre de la evidencia, pero esta evidencia no es un monolito uniforme; más bien, se presenta como un mosaico de piezas con diferentes formas, colores y texturas, cada una aportando una parte de la imagen completa, pero con distintos grados de nitidez y fiabili-

dad. Comprender los principios que nos permiten evaluar la calidad de esas piezas es fundamental para construir un edificio de conocimiento robusto y, sobre todo, para tomar las mejores decisiones en el cuidado de la salud. No basta con leer, es indispensable interpretar, cuestionar y contextualizar.

En el corazón de esta evaluación crítica residen varios conceptos fundamentales, pilares sobre los que se asienta la fiabilidad de cualquier estudio. Ignorarlos es como intentar construir sobre arena movediza; el resultado, por muy impresionante que parezca, carecerá de solidez. Estos principios nos alertan sobre posibles distorsiones en los hallazgos, nos ayudan a diferenciar las asociaciones genuinas de las coincidencias fortuitas y nos permiten juzgar cuán aplicables son los resultados a nuestra propia realidad clínica. Son, en esencia, nuestro escudo contra la información engañosa y nuestra brújula en la búsqueda de la verdad.

Aunque ya hemos tocado estos temas, es indispensable reiterar aquí la importancia cardinal del **sesgo** y la **confusión**, pues son las amenazas más insidiosas a la validez de cualquier estudio y su identificación es el primer paso en la interpretación crítica. El sesgo, recordemos, es un error sistemático que desvía los resultados de la verdad, como un error de calibración constante en un instrumento de medición. Se introduce en cualquier fase del estudio, desde la selección de los participantes hasta la forma en que se recogen o analizan los datos, llevando a una estimación consistentemente incorrecta del efecto o la asociación que se investiga. Por su parte, la confusión ocurre cuando la asociación observada entre una exposición y un desenlace está distorsionada por la influencia de un tercer factor, una variable confusora, que está asociada tanto con la exposición como con el desenlace, pero no es un paso intermedio en la relación causal. Este confusor "mezcla" sus efectos, haciendo que parezca que la exposición causa el desenlace, o que la asociación es más fuerte o más débil de lo que realmente es.

Estos dos conceptos, sesgo y confusión, son amenazas directas a la **validez interna** de un estudio. La validez interna se refiere a cuán correctos son los resultados del estudio para la población específica y las condiciones bajo las cuales se realizó. Un estudio con alta validez interna ha minimizado el sesgo y controlado adecuadamente la confusión, de modo que podemos confiar en que la asociación o el efecto observado es real dentro de ese estudio

particular. Es como asegurar que el experimento de laboratorio se realizó correctamente, sin errores sistemáticos en la medición o en la asignación de tratamientos.

Sin embargo, que un estudio tenga alta validez interna no garantiza que sus hallazgos sean aplicables a todos los pacientes o en todos los entornos. Aquí entra en juego la **validez externa** o generalizabilidad. La validez externa se refiere a cuán aplicables son los resultados de un estudio a poblaciones, entornos y condiciones diferentes a las del estudio original. Si un ensayo clínico sobre un nuevo fármaco para la diabetes se realizó exclusivamente en hombres jóvenes con diabetes tipo 2 de reciente diagnóstico y sin comorbilidades, ¿podemos generalizar esos resultados a mujeres mayores con diabetes de larga evolución y múltiples complicaciones? Probablemente no sin cautela. La validez externa nos invita a considerar si la población del estudio es representativa de la población a la que queremos aplicar los resultados, si la intervención se administró de una manera que sea reproducible en la práctica clínica habitual, y si el contexto del estudio refleja el entorno del "mundo real".

Evaluar críticamente un estudio implica, por tanto, un ejercicio dual: primero, juzgar su validez interna, buscando posibles fuentes de sesgo y confusión que puedan invalidar los resultados dentro del estudio mismo; y segundo, si la validez interna es aceptable, juzgar su validez externa, considerando cuán aplicables son esos resultados a nuestra propia práctica o a la población que nos interesa. Ambos aspectos son indispensables para traducir la evidencia científica en decisiones clínicas informadas y responsables.

Para facilitar este proceso de evaluación crítica, podemos seguir un enfoque estructurado. No se trata de memorizar una lista interminable de sesgos, sino de desarrollar una mentalidad inquisitiva que cuestione cada paso del estudio, desde cómo se seleccionaron los participantes hasta cómo se analizaron e interpretaron los datos.

Paso a Paso: El Proceso General de Evaluación Crítica

Evaluar críticamente un artículo científico puede parecer una tarea abrumadora al principio, pero se vuelve manejable si se aborda de manera sistemática. Aquí te presentamos un paso a paso general que puedes aplicar a cualquier tipo de estudio:

1. **Comprende la Pregunta de Investigación:** ¿Qué querían realmente investigar los autores? ¿Cuál es el objetivo principal del estudio? Asegúrate de entender la pregunta antes de evaluar cómo intentaron responderla.
2. **Identifica el Diseño del Estudio:** ¿Qué tipo de estudio es, ensayo clínico aleatorizado, cohorte, casos y controles, transversal, etc.? Reconocer el diseño te da una idea inicial de las fortalezas y limitaciones inherentes del estudio y qué tipo de preguntas puede responder de manera más fiable.
3. **Evalúa la Población de Estudio:** ¿Quiénes fueron incluidos en el estudio? ¿Son representativos de la población a la que quieres aplicar los resultados, validez externa? ¿Cómo fueron reclutados los participantes? ¿Podría haber sesgo de selección?
4. **Evalúa la Medición de la Exposición y el Desenlace:** ¿Cómo se definieron y midieron la exposición, o intervención, y el desenlace? ¿Fueron las mediciones precisas y fiables? ¿Podría haber sesgo de medición, ej. sesgo de recuerdo, clasificación errónea?
5. **Identifica Posibles Confusores:** ¿Qué otros factores podrían estar asociados tanto con la exposición como con el desenlace? ¿Midieron los autores estos posibles confusores?
6. **Evalúa Cómo se Manejaron los Confusores:** ¿Utilizaron los autores técnicas en el diseño, ej. aleatorización, apareamiento, o en el análisis, ej. ajuste en modelos de regresión, estratificación, para controlar el efecto de los confusores? ¿Fue adecuado el control?
7. **Evalúa el Análisis Estadístico:** ¿Fueron apropiados los métodos estadísticos utilizados para el tipo de datos y el diseño del estudio? ¿Se interpretaron correctamente los resultados estadísticos, ej. valor p, intervalos de confianza?
8. **Busca Otros Tipos de Sesgo:** Considera otros sesgos relevantes para el diseño específico, ej. pérdida de seguimiento en cohortes y ECA, sesgo de verificación en estudios diagnósticos, sesgo de publicación en revisiones sistemáticas.
9. **Interpreta los Resultados en Contexto:** ¿Cuál es la magnitud del efecto o la asociación observada? ¿Es clínicamente relevante? ¿Son los resultados consistentes con otros estudios?

10. **Considera la Validez Externa:** Si el estudio tiene buena validez interna, ¿a quiénes y en qué entornos se pueden aplicar estos resultados? ¿Hay características de la población del estudio o del contexto que limiten la generalizabilidad?

11. **Saca Conclusiones:** Basado en tu evaluación crítica, ¿cuán fiable es la evidencia proporcionada por este estudio? ¿Qué implicaciones tiene para la práctica clínica o futuras investigaciones?

Este proceso, que iremos detallando al evaluar críticamente los diferentes tipos de estudios, es una habilidad que se perfecciona con la práctica. Nos permite ir más allá de la simple aceptación de los resultados publicados para convertirnos en consumidores informados y juiciosos de la literatura científica, una habilidad indispensable para navegar con éxito en el complejo y cambiante panorama de la MBE.

5.2. Análisis Crítico de Estudios de Desenlace (Outcome Studies)

En la práctica clínica, nuestro objetivo último es mejorar los resultados que realmente importan para los pacientes: aliviar síntomas, prevenir complicaciones, prolongar la vida, mejorar la calidad de vida. Los estudios de desenlace, u *outcome studies*, son precisamente aquellos que se centran en evaluar estos resultados finales, a menudo en entornos más cercanos a la práctica clínica habitual. A diferencia de los estudios que evalúan solo marcadores subrogados, como un nivel de laboratorio o una medida fisiológica que se espera que prediga un desenlace importante, los estudios de desenlace miden directamente lo que le sucede al paciente.

La evaluación crítica de estos estudios requiere aplicar los principios generales de sesgo, confusión y validez, interna y externa, que ya hemos discutido, pero con un enfoque particular en cómo estos conceptos se aplican a la definición, medición e interpretación de los resultados finales. No basta con que un estudio muestre una diferencia estadísticamente significativa en un desenlace; debemos juzgar si ese desenlace es realmente relevante para el paciente y si la forma en que se midió es fiable y libre de sesgos.

Piensen en un estudio de desenlace como si evaluáramos el éxito de un programa de entrenamiento para corredores. No nos interesa solo si los corredores mejoraron su consumo de oxígeno en una prueba de laboratorio, un marcador subrogado, sino si realmente pudieron correr una maratón más

rápido o con menos fatiga, un desenlace importante. Al evaluar el estudio, nos preguntaríamos: ¿Cómo definieron "correr más rápido"? ¿Fue una mejora clínicamente significativa? ¿Midieron la fatiga de manera objetiva o solo preguntaron a los corredores, susceptible a sesgo? ¿El programa funcionó para todo tipo de corredores o solo para un grupo particular, validez externa?

Paso a Paso: Evaluación Crítica de Estudios de Desenlace

Al abordar un estudio de desenlace, considera los siguientes puntos clave, además de los pasos generales de evaluación crítica:

1. **Claridad y Relevancia del Desenlace:**

- ¿El desenlace principal está claramente definido? ¿Es un evento específico, una medición continua, un resultado reportado por el paciente?
- ¿Es el desenlace realmente importante para los pacientes y para la toma de decisiones clínicas o de salud pública? ¿Es un desenlace final, duro, como mortalidad o un evento mayor, o es un desenlace intermedio o subrogado?
- Si se utilizan múltiples desenlaces, ¿se definió un desenlace primario *a priori*? Evaluar múltiples desenlaces sin una jerarquía preespecificada aumenta el riesgo de encontrar resultados significativos por azar.

2. **Medición del Desenlace:**

- ¿Cómo se midió el desenlace? ¿Se utilizaron métodos estandarizados y validados?
- ¿La medición del desenlace fue objetiva o subjetiva? Si fue subjetiva, ej. reporte de síntomas, ¿se tomaron medidas para minimizar el sesgo, ej. cegamiento del evaluador?
- ¿Hubo alguna diferencia en la forma en que se midió el desenlace entre los grupos de comparación? Esto podría introducir un sesgo de detección o medición.

3. **Integridad de los Datos del Desenlace:**

- ¿Se obtuvieron datos del desenlace para todos o la gran mayoría de los participantes? ¿Cuál fue la proporción de datos faltantes para el desenlace principal?

- Si hubo datos faltantes, ¿cómo se manejaron? ¿Podría la forma en que se manejaron los datos faltantes introducir sesgo?

4. Consideraciones Específicas del Diseño:

- Si es un ensayo clínico, ¿se realizó el análisis por intención de tratar, ITT, para el desenlace principal? El análisis ITT preserva los beneficios de la aleatorización y proporciona una estimación más realista del efecto en la práctica.
- Si es un estudio observacional, cohorte, casos y controles, ¿se consideró y controló adecuadamente el tiempo de seguimiento hasta el desenlace? Esto es crucial en estudios de sobrevivencia.

5. Interpretación de la Magnitud del Efecto:

- Si se encontró una asociación o un efecto, ¿cuál es su magnitud, ej. diferencia de riesgos, riesgo relativo, diferencia de medias?
- ¿Es esta magnitud clínicamente relevante, independientemente de si es estadísticamente significativa? Un efecto pequeño pero estadísticamente significativo puede no tener un impacto importante en la práctica.

6. Evaluación de Eventos Adversos y Seguridad:

- ¿Se reportaron y evaluaron sistemáticamente los eventos adversos y otros desenlaces de seguridad?
- ¿Se comparó la frecuencia de eventos adversos entre los grupos de comparación? Un tratamiento puede ser efectivo para el desenlace principal pero tener efectos secundarios inaceptables.

Métricas Clave de Interpretación en Estudios de Desenlace:

Al interpretar los resultados de un estudio de desenlace, más allá del valor p y los intervalos de confianza, es fundamental comprender las métricas que cuantifican la magnitud del efecto en términos absolutos y relativos, y su impacto clínico:

- **Diferencia de Riesgos (DR) o Reducción Absoluta del Riesgo (RAR):** Mide la diferencia en la incidencia del evento entre el grupo expuesto (o de intervención) y el grupo no expuesto (o control). Es una medida del efecto absoluto.

- Fórmula: DR= Ie−Io (Incidencia en expuestos - Incidencia en no expuestos)
- *Ejemplo*: En un ensayo clínico, la incidencia de infarto fue 5% en el grupo de tratamiento y 8% en el grupo placebo. DR=5. Esto significa que el tratamiento redujo el riesgo absoluto de infarto en 3 puntos porcentuales. Es una medida muy útil para la toma de decisiones individuales.

- **Reducción Relativa del Riesgo (RRR):** Mide la proporción en que se reduce el riesgo del evento en el grupo expuesto (o de intervención) en relación con el riesgo en el grupo no expuesto (o control). Es una medida del efecto relativo.
  - Fórmula: RRR=(Io−Ie)/Io=1−RR
  - *Ejemplo*: Siguiendo el ejemplo anterior, RRR=(8. Esto significa que el tratamiento redujo el riesgo relativo de infarto en un 37.5%. Aunque el 37.5% suena impresionante, es crucial considerarlo junto con la DR, que nos dice que la reducción absoluta fue solo de 3 puntos porcentuales.
- **Número Necesario para Tratar (NNT):** El número promedio de pacientes que necesitan recibir la intervención para prevenir un evento adicional en comparación con el grupo control. Es una medida del impacto clínico en la práctica.
  - Fórmula: NNT=1/DR (si la DR es positiva, indicando beneficio)
  - *Ejemplo*: Con una DR de -3% (o 3% de reducción absoluta), NNT=1/0.03≈33. Esto significa que se necesita tratar a aproximadamente 33 pacientes con el nuevo fármaco para prevenir un infarto adicional en comparación con el placebo. Este número ayuda a sopesar los beneficios frente a los costos y riesgos del tratamiento.
- **Diferencia de Medias (DM) o Diferencia de Medias Estandarizada (DME):** Se utiliza para desenlaces continuos (ej. cambio en la presión arterial, puntuación en una escala de calidad de vida). La DM es la diferencia promedio en el valor del desenlace entre los grupos. La DME estandariza esta diferencia dividiéndola por la DS, lo que permite comparar efectos entre diferentes escalas.
  - *Ejemplo*: Un estudio encuentra que un programa de ejercicio reduce la presión arterial sistólica en un promedio de 5 mmHg más que el cuidado

habitual (DM = -5 mmHg). Si la DS de la presión arterial en la población es 10 mmHg, la DME sería -0.5 (un efecto moderado).

- **RR o HR:** Como vimos en las secciones anteriores, estas medidas comparan la incidencia o el riesgo instantáneo del evento entre grupos. En estudios de desenlace, nos dicen cuántas veces más (o menos) probable es el desenlace en un grupo comparado con otro.

  - *Ejemplo*: Un estudio de cohorte encuentra un HR de 1.5 para la asociación entre un factor de riesgo y la mortalidad. Esto significa que los individuos con el factor de riesgo tienen 1.5 veces el riesgo instantáneo de morir en cualquier momento dado en comparación con aquellos sin el factor de riesgo.

La evaluación crítica de los estudios de desenlace nos permite ir más allá de la simple verificación estadística para juzgar si los hallazgos son fiables, relevantes y aplicables a los pacientes que buscamos cuidar. Nos recuerda que, en última instancia, la investigación debe centrarse en lo que realmente importa para la salud y el bienestar de las personas.

5.3. Análisis Crítico de Reportes de Caso y Series de Casos

Los reportes de caso y las series de casos son los cimientos de la evidencia médica, a menudo el primer indicio de algo nuevo o inusual. Son como anécdotas clínicas detalladas que, al ser compartidas, pueden alertar a la comunidad médica sobre fenómenos raros, presentaciones atípicas de enfermedades conocidas o posibles efectos adversos de tratamientos. Sin embargo, su naturaleza descriptiva y la falta de un grupo de comparación limitan severamente su capacidad para establecer causalidad o generalizar hallazgos.

Paso a Paso: Evaluación Crítica de Reportes de Caso y Series de Casos

Al evaluar un reporte de caso o una serie de casos, considera:

1. **Claridad y Detalle de la Descripción:** ¿Está la historia clínica del paciente, o de los pacientes, descrita de manera clara y detallada? ¿Se presentan los síntomas, hallazgos del examen físico, resultados de pruebas diagnósticas, tratamiento y evolución de forma comprensible?

2. **Documentación del Diagnóstico:** ¿Está el diagnóstico bien documentado y justificado con evidencia suficiente, ej. resultados de laboratorio, imágenes, histopatología?

3. **Novedad y Relevancia:** ¿Es el caso, o la serie, realmente novedoso o inusual? ¿Aporta información que no está ampliamente disponible en la literatura? ¿Tiene relevancia para la práctica clínica o la generación de hipótesis?

4. **Posibles Sesgos:** Aunque son descriptivos, pueden existir sesgos. ¿Los casos seleccionados para la serie son representativos de todos los casos que los autores podrían haber reportado, sesgo de selección? ¿La forma en que se presenta la información podría estar sesgada para enfatizar ciertos aspectos?

5. **Generación de Hipótesis:** ¿El reporte o la serie sugieren claramente una hipótesis para futuras investigaciones? Este es uno de sus principales valores.

6. **Limitaciones Reconocidas:** ¿Reconocen los autores explícitamente las limitaciones inherentes de un reporte de caso o serie de casos, particularmente su incapacidad para establecer causalidad o generalizar resultados?

Métricas Clave de Interpretación:

En reportes de caso y series de casos, las métricas son principalmente descriptivas:

- **Frecuencia de características:** En una serie de casos, se pueden reportar las frecuencias o porcentajes de ciertas características entre los pacientes (ej. “el 60% de los pacientes presentaron fiebre”).

- **Medidas descriptivas:** Para variables numéricas, se pueden reportar medias, medianas, rangos (ej. “la edad promedio fue de 55 años, rango 30-70”).

Ejemplo: Un reporte de caso describe a un paciente que desarrolló una reacción adversa inusual después de tomar un medicamento. La evaluación crítica se centraría en cuán detallada y documentada está la reacción, si se

descartaron otras posibles causas y si la reacción es realmente novedosa en la literatura. No se buscan métricas de asociación, sino una descripción precisa del evento.

5.4. Análisis Crítico de Estudios Descriptivos

Los estudios descriptivos, como las encuestas de prevalencia o la descripción de la distribución de una enfermedad en el tiempo y el espacio, nos proporcionan una "fotografía" de la salud de una población. Responden a las preguntas de quién, dónde y cuándo, pero no al porqué. Son esenciales para la planificación de salud pública, la identificación de tendencias y la generación de hipótesis.

Paso a Paso: Evaluación Crítica de Estudios Descriptivos

Al evaluar un estudio descriptivo, considera:

1. **Definición Clara de la Población y el Período de Estudio:** ¿A quiénes se estudió y durante qué período de tiempo? ¿Está claramente definida la población de referencia?
2. **Método de Muestreo:** ¿Cómo se seleccionó la muestra de la población? ¿Fue un muestreo probabilístico, aleatorio, para asegurar la representatividad? ¿Podría haber sesgo de selección?
3. **Definición y Medición de Variables:** ¿Cómo se definieron y midieron las variables de interés, ej. la enfermedad, las características demográficas? ¿Fueron las mediciones fiables y válidas? ¿Podría haber sesgo de medición?
4. **Tasa de Respuesta, en encuestas:** Si se basó en una encuesta, ¿cuál fue la tasa de respuesta? Una tasa de respuesta baja puede introducir sesgo de no respuesta si quienes respondieron difieren de quienes no lo hicieron.
5. **Análisis Descriptivo Apropiado:** ¿Se utilizaron las estadísticas descriptivas y las visualizaciones adecuadas para resumir los datos, ej. frecuencias, porcentajes, medias, medianas, gráficos de barras, histogramas?

Métricas Clave de Interpretación:

Las métricas principales en estudios descriptivos son las de ocurrencia y resumen:

- **Prevalencia:** La proporción de individuos con la condición en un momento dado.
  - *Ejemplo:* Un estudio transversal reporta una prevalencia de diabetes del 10% en adultos mayores de 60 años en Cali. Esto significa que, en el momento del estudio, 10 de cada 100 adultos mayores de 60 años en esa población tenían diabetes.
- **Incidencia:** La tasa de casos nuevos que aparecen en una población en riesgo durante un período de tiempo.
  - *Ejemplo:* Un estudio descriptivo de cohorte reporta una incidencia acumulada de influenza del 15% durante la temporada invernal en estudiantes universitarios. Esto significa que el 15% de los estudiantes que no tenían influenza al inicio de la temporada la desarrollaron durante ese período.
- **Medidas de tendencia central y dispersión:** Media, mediana, moda, DS, IQR para variables numéricas.

5.5. Análisis Crítico de Estudios Ecológicos

Los estudios ecológicos examinan la relación entre la exposición y el desenlace a nivel de grupo o población, ej. países, ciudades, no a nivel individual. Comparan, por ejemplo, la tasa promedio de una enfermedad en diferentes regiones con el nivel promedio de exposición a un factor en esas mismas regiones. Son útiles para generar hipótesis y evaluar el impacto de intervenciones a gran escala, pero sufren de una limitación crucial: la falacia ecológica.

Paso a Paso: Evaluación Crítica de Estudios Ecológicos

Al evaluar un estudio ecológico, considera:

1. **Unidad de Análisis:** ¿Cuál es la unidad de análisis, país, región, etc.? ¿Es apropiada para la pregunta?
2. **Calidad y Comparabilidad de los Datos Agregados:** ¿De dónde provienen los datos agregados? ¿Son fiables? ¿Se recolectaron de manera consistente entre los diferentes grupos o a lo largo del tiempo?
3. **Definición y Medición de Variables a Nivel de Grupo:** ¿Cómo se definieron y midieron la exposición y el desenlace a nivel de grupo, ej. promedio de exposición, tasa de incidencia/mortalidad?

4. **Control de Confusores a Nivel de Grupo:** ¿Se consideraron y controlaron posibles variables confusoras a nivel de grupo, ej. nivel socioeconómico promedio de la región? El control de confusores individuales es imposible.

5. **El Riesgo de la Falacia Ecológica:** ¿Intentan los autores inferir asociaciones a nivel individual basándose en hallazgos a nivel de grupo? Este es el error más común y peligroso en este tipo de estudios.

Métricas Clave de Interpretación:

Las métricas en estudios ecológicos suelen ser medidas de asociación o correlación a nivel de grupo:

- **Coeficiente de Correlación Ecológica:** Mide la fuerza y dirección de la asociación lineal entre la exposición promedio y el desenlace promedio a nivel de grupo.
  - *Ejemplo:* Un estudio ecológico encuentra un coeficiente de correlación de Pearson de 0.75 entre el consumo per cápita de azúcar y la tasa de obesidad en diferentes países. Esto indica una fuerte asociación positiva a nivel de país: los países con mayor consumo de azúcar tienden a tener mayores tasas de obesidad. Sin embargo, no podemos concluir que los individuos que consumen más azúcar dentro de esos países son los que son obesos, falacia ecológica.
- **Coeficientes de Regresión Ecológica:** Estiman el cambio en el desenlace promedio asociado a un cambio en la exposición promedio a nivel de grupo.

5.6. Análisis Crítico de Estudios Transversales Analíticos

Los estudios transversales analíticos toman una "fotografía" de una población en un momento dado para explorar la asociación entre una exposición y un desenlace. Miden la prevalencia de ambos simultáneamente. Son útiles para generar hipótesis y describir asociaciones, pero no pueden establecer causalidad porque no determinan si la exposición precedió al desenlace.

Paso a Paso: Evaluación Crítica de Estudios Transversales Analíticos

Al evaluar un estudio transversal analítico, considera:

1. **Representatividad de la Muestra:** ¿Es la muestra representativa de la población objetivo? ¿Podría haber sesgo de selección, ej. sesgo de no respuesta?
2. **Medición Simultánea:** ¿Se midieron la exposición y el desenlace en el mismo momento? Esto es inherente al diseño, pero limita la inferencia causal.
3. **Sesgo de Recuerdo:** Si la exposición se mide preguntando sobre eventos pasados, ¿podría haber sesgo de recuerdo?
4. **Sesgo de Incidencia-Prevalencia:** Si se estudia una enfermedad crónica, ¿los casos prevalentes incluidos son representativos de todos los casos que surgen, o solo de los que sobreviven más tiempo con la enfermedad?
5. **Control de Confusores:** ¿Se midieron y controlaron adecuadamente las variables confusoras?

Métricas Clave de Interpretación:

Las métricas principales son medidas de asociación basadas en la prevalencia:

- **Razón de Prevalencias, PR - *Prevalence Ratio*:** La prevalencia del desenlace en el grupo expuesto dividida por la prevalencia en el grupo no expuesto. Similar al Riesgo Relativo, pero basado en prevalencia, no en incidencia.
  - *Ejemplo*: Un estudio transversal encuentra un PR de 1.8 para la asociación entre el sedentarismo y la obesidad. Esto significa que la prevalencia de obesidad es 1.8 veces mayor en las personas sedentarias que en las activas en el momento del estudio. No podemos decir que el sedentarismo *causa* la obesidad basándonos solo en esto.
- **Odds Ratio de Prevalencia, POR - *Prevalence Odds Ratio*:** La odds de tener el desenlace en el grupo expuesto dividida por la odds de tener el desenlace en el grupo no expuesto. Similar al Odds Ratio.

5.7. Análisis Crítico de Estudios de Casos y Controles

Los estudios de casos y controles son eficientes para investigar factores de riesgo de enfermedades raras o con largos períodos de latencia. Parten de individuos que ya tienen la enfermedad, casos, y los comparan con un grupo sin la enfermedad, controles, investigando retrospectivamente la exposición a factores de interés.

Paso a Paso: Evaluación Crítica de Estudios de Casos y Controles

Al evaluar un estudio de casos y controles, considera:

1. **Definición Clara de Casos y Controles:** ¿Están los criterios para ser un caso y para ser un control claramente definidos y aplicados de manera consistente?
2. **Selección de Controles:** ¿Los controles provienen de la misma población fuente que los casos? ¿Habrían sido identificados como casos si hubieran desarrollado la enfermedad? ¿Podría haber sesgo de selección en la elección de controles, ej. sesgo de Berkson en controles hospitalarios?
3. **Medición Retrospectiva de la Exposición:** ¿Cómo se midió la exposición pasada? ¿Podría haber sesgo de recuerdo? ¿Se midió la exposición de la misma manera en casos y controles?
4. **Secuencia Temporal:** ¿Se aseguró, en la medida de lo posible, que la exposición ocurrió *antes* del inicio de la enfermedad?
5. **Control de Confusores:** ¿Se midieron y controlaron adecuadamente las variables confusoras? ¿Se utilizó apareamiento o ajuste en el análisis?

Métricas Clave de Interpretación:

La métrica principal de asociación es la Razón de Momios u Odds Ratio:

- **OR:** La razón de las odds de exposición en los casos comparada con las odds de exposición en los controles. En enfermedades raras, la OR es una buena aproximación del Riesgo Relativo.
    - *Ejemplo:* Un estudio de casos y controles encuentra una OR de 3.5 para la asociación entre el tabaquismo y el cáncer de pulmón. Esto significa que las odds de haber sido fumador son 3.5 veces mayores en los pacientes con cáncer de pulmón que en los controles sin cáncer de pulmón. Si el cáncer

de pulmón es una enfermedad rara, esto sugiere que los fumadores tienen aproximadamente 3.5 veces el riesgo de desarrollar cáncer de pulmón que los no fumadores.

- **Intervalo de Confianza, IC 95%, para la OR:** Un rango de valores donde es probable que se encuentre la verdadera OR poblacional. Si el IC 95% no incluye el 1, la asociación es estadísticamente significativa.
  - *Ejemplo*: OR = 3.5 (IC 95%: 2.8 - 4.3). Como el intervalo no incluye el 1, la asociación es estadísticamente significativa.

5.8. Análisis Crítico de Estudios de Cohortes

Los estudios de cohortes siguen a grupos de personas, cohortes, a lo largo del tiempo para ver quién desarrolla un desenlace de interés. Clasifican a los participantes según su exposición al inicio del estudio y miden la incidencia del desenlace en los grupos expuestos y no expuestos. Son excelentes para establecer la secuencia temporal, la exposición precede al desenlace, y calcular la incidencia y el riesgo relativo.

Paso a Paso: Evaluación Crítica de Estudios de Cohorte

Al evaluar un estudio de cohorte, considera:

1. **Definición Clara de la Cohorte y la Exposición:** ¿Quiénes fueron incluidos en la cohorte? ¿Cómo se definió y midió la exposición al inicio del estudio?
2. **Libres del Desenlace al Inicio:** ¿Se aseguró que todos los participantes estaban libres del desenlace de interés al inicio del seguimiento?
3. **Seguimiento:** ¿Cómo se realizó el seguimiento de los participantes? ¿Fue completo? ¿Hubo una pérdida de seguimiento significativa? ¿Podría la pérdida de seguimiento ser diferencial entre los grupos de exposición, introduciendo sesgo?
4. **Medición del Desenlace:** ¿Cómo se definió y midió el desenlace durante el seguimiento? ¿Fue la medición consistente y libre de sesgos a lo largo del tiempo y entre grupos?
5. **Control de Confusores:** ¿Se midieron y controlaron adecuadamente las variables confusoras a lo largo del seguimiento?

Métricas Clave de Interpretación:

Las métricas principales son medidas de incidencia y asociación:

- **Incidencia:** La tasa o proporción de casos nuevos en el grupo expuesto y en el grupo no expuesto.
  - *Ejemplo*: Un estudio de cohorte de 10 años encuentra una incidencia de enfermedad cardiovascular del 10% en fumadores y 5% en no fumadores. Esto significa que 10 de cada 100 fumadores y 5 de cada 100 no fumadores desarrollaron la enfermedad en ese período.
- **RR:** La razón de la incidencia en el grupo expuesto dividida por la incidencia en el grupo no expuesto.
  - *Ejemplo*: RR = 10% / 5% = 2. Esto significa que los fumadores tienen el doble de riesgo de desarrollar enfermedad cardiovascular en comparación con los no fumadores.
- **HR:** En análisis de sobrevivencia, el HR compara el riesgo instantáneo del evento entre grupos a lo largo del tiempo. Similar al RR, pero utilizado cuando se analiza el tiempo hasta el evento y hay censura.
  - *Ejemplo*: Un estudio de cohorte encuentra un HR de 0.75 para un tratamiento en la sobrevida libre de progresión del cáncer. Esto significa que los pacientes tratados tienen un riesgo instantáneo de progresión un 25% menor en cualquier momento dado en comparación con el grupo control.
- **Diferencia de Riesgos (DR) o Riesgo Atribuible (RA):** La diferencia absoluta en la incidencia entre el grupo expuesto y el no expuesto.
  - *Ejemplo*: DR = 10% - 5% = 5%. Esto significa que, en este estudio, 5 de cada 100 casos de enfermedad cardiovascular en fumadores son atribuibles al tabaquismo.
- **Intervalo de Confianza (IC 95%) para RR o HR:** Si el IC 95% no incluye el 1, la asociación es estadísticamente significativa.
  - *Ejemplo*: RR = 2 (IC 95%: 1.5 - 2.7). Como el intervalo no incluye el 1, la asociación es estadísticamente significativa.

5.9. Análisis Crítico de Estudios de Sobrevivencia (Tiempo al Evento)

Los estudios de sobrevivencia se centran en el tiempo que transcurre hasta que ocurre un evento de interés, como la muerte, la recurrencia de una enfermedad o la falla de un dispositivo médico. La característica clave de estos datos es la presencia de censura, donde no se observa el evento para todos los participantes durante el período de estudio.

Paso a Paso: Evaluación Crítica de Estudios de Sobrevivencia

Al evaluar un estudio de sobrevivencia, considera:

1. **Definición Clara del Evento y el Tiempo Cero:** ¿Qué evento se está midiendo y cómo se define el inicio del seguimiento (tiempo cero)?
2. **Manejo de la Censura:** ¿Cómo se manejó la censura? ¿Se reportó la proporción de participantes censurados? ¿Podría la censura ser informativa (relacionada con el pronóstico), introduciendo sesgo?
3. **Métodos de Análisis Apropiados:** ¿Se utilizaron métodos adecuados para analizar datos de sobrevivencia (ej. curvas de Kaplan-Meier, prueba de Log-Rank, modelo de regresión de Cox)?
4. **Evaluación de la Asunción de Riesgos Proporcionales (en modelo de Cox):** Si se utilizó un modelo de Cox, ¿se evaluó si el efecto de las covariables sobre el riesgo instantáneo fue constante a lo largo del tiempo? Si no se cumple, ¿se utilizaron métodos alternativos?
5. **Control de Confusores:** ¿Se incluyeron y controlaron adecuadamente las variables confusoras en el modelo de regresión?

Métricas Clave de Interpretación:

Las métricas principales son el tiempo hasta el evento y las medidas de riesgo instantáneo:

- **Mediana de Sobrevida:** El tiempo en el que el 50% de los individuos en un grupo han experimentado el evento.
  - *Ejemplo*: La mediana de sobrevida libre de progresión en el grupo de tratamiento fue de 18 meses, comparado con 12 meses en el grupo control.

Esto significa que la mitad de los pacientes tratados no tuvieron progresión de la enfermedad hasta los 18 meses, mientras que en el grupo control la mitad ya había progresado a los 12 meses.

- **Probabilidad de Sobrevida en un Tiempo Específico:** La proporción de individuos que aún no han experimentado el evento en un punto de tiempo determinado, estimada a partir de curvas de Kaplan-Meier.
  - *Ejemplo*: La probabilidad de sobrevida a 5 años fue del 70% en el grupo de tratamiento y 50% en el grupo control. Esto significa que el 70% de los pacientes tratados y el 50% de los pacientes control estaban vivos a los 5 años.
- **HR:** La medida de efecto más común en la regresión de Cox. Compara el riesgo instantáneo del evento entre grupos.
  - *Ejemplo*: HR = 0.60 para un tratamiento en la mortalidad. Esto significa que el riesgo instantáneo de morir en cualquier momento dado es un 40% menor en el grupo de tratamiento en comparación con el grupo control.
- **Intervalo de Confianza (IC 95%) para el HR:** Si el IC 95% no incluye el 1, el resultado es estadísticamente significativo.

5.10. Análisis Crítico de Ensayos Clínicos Aleatorizados

Los ECA son considerados el estándar de oro para evaluar la eficacia y seguridad de las intervenciones terapéuticas o preventivas. La aleatorización es su característica definitoria, ya que distribuye las características de los participantes, incluyendo los confusores conocidos y desconocidos, de manera aleatoria entre los grupos de comparación, minimizando el sesgo de selección y la confusión.

Paso a Paso: Evaluación Crítica de ECA

Al evaluar un ECA, considera:

1. **Aleatorización:** ¿Fue verdaderamente aleatoria la asignación a los grupos? ¿Se describió el método de aleatorización?
2. **Ocultación de la Asignación (*Allocation Concealment*):** ¿Se ocultó la secuencia de asignación a quienes reclutaban a los participantes? Esto es crucial para prevenir el sesgo de selección.

3. **Cegamiento (*Blinding*):** ¿Estuvieron cegados los participantes, el personal que administró la intervención y quienes evaluaron los desenlaces? El cegamiento reduce el sesgo de performance y el sesgo de detección. Un ECA "doble ciego" (participantes y personal) o "triple ciego" (participantes, personal y evaluadores) es ideal.
4. **Grupo Control/Comparador:** ¿Fue apropiado el grupo control (placebo, tratamiento estándar, no tratamiento)?
5. **Manejo de la Pérdida de Seguimiento y Abandonos:** ¿Cuál fue la proporción de participantes que abandonaron el estudio o se perdieron durante el seguimiento? ¿Fue similar entre los grupos? ¿Cómo se manejaron en el análisis? Una alta pérdida de seguimiento, especialmente si es diferencial, puede introducir sesgo.
6. **Análisis por Intención de Tratar (ITT):** ¿Se realizó el análisis primario por ITT, incluyendo a todos los participantes en el grupo al que fueron asignados aleatoriamente, independientemente de si recibieron o completaron la intervención? El análisis ITT preserva los beneficios de la aleatorización.
7. **Definición y Medición de Desenlaces:** ¿Están los desenlaces claramente definidos y medidos de manera fiable y consistente en todos los grupos? ¿Son los desenlaces clínicamente relevantes?
8. **Consideraciones Éticas:** ¿Se obtuvo consentimiento informado? ¿Fue aprobado por un comité de ética? ¿Existía *equipoise clínico*?

Métricas Clave de Interpretación:

Las métricas dependen del tipo de desenlace (binario, continuo, tiempo al evento), pero a menudo se reportan medidas de efecto absoluto y relativo, y resultados de análisis de sobrevivencia:

- **RR, OR, o HR:** Para desenlaces binarios o tiempo al evento, comparando la frecuencia o riesgo entre los grupos de intervención y control.
  - *Ejemplo:* Un ECA encuentra un RR de 0.70 para la incidencia de un evento adverso grave con el nuevo fármaco vs. placebo. Esto significa que el riesgo de experimentar el evento adverso grave es un 30% menor con el nuevo fármaco.
- **Diferencia de Riesgos (DR) o Reducción Absoluta del Riesgo (RAR):** La diferencia absoluta en la incidencia de un desenlace binario entre grupos.

- *Ejemplo:* La DR para el desenlace primario fue -5%. Esto significa que la intervención redujo el riesgo absoluto del desenlace primario en 5 puntos porcentuales.

- **Número Necesario para Tratar (NNT):** El número de pacientes que necesitan ser tratados para prevenir un evento adicional.

- *Ejemplo:* NNT = 20 para prevenir el desenlace primario.

- **Diferencia de Medias (DM) o Diferencia de Medias Estandarizada (DME):** Para desenlaces continuos.

- *Ejemplo:* La DM en la puntuación de calidad de vida fue +10 puntos con la intervención vs. control.

- **Curvas de Kaplan-Meier y Prueba de Log-Rank:** Para comparar la sobrevida o el tiempo hasta el evento entre grupos.

- **Intervalos de Confianza (IC 95%):** Cruciales para evaluar la precisión de las estimaciones del efecto. Si el IC 95% no cruza el valor nulo (1 para RR/OR/HR, 0 para DM), el resultado es estadísticamente significativo.

5.11. Análisis Crítico de Revisiones Narrativas

Las revisiones narrativas resumen la literatura sobre un tema, a menudo escritas por expertos en el campo. Son útiles para obtener una visión general y contextualizar un tema, pero no siguen una metodología sistemática para identificar y sintetizar estudios. Esto las hace susceptibles a sesgos, como la selección selectiva de estudios que apoyan la perspectiva del autor.

Paso a Paso: Evaluación Crítica de Revisiones Narrativas

Al evaluar una revisión narrativa, considera:

1. **Experiencia del Autor:** ¿Quién escribió la revisión? ¿Son reconocidos expertos en el campo?

2. **Alcance y Pregunta (Implícita):** ¿Cuál es el tema de la revisión? ¿Está claramente delimitado?

3. **Método de Búsqueda:** ¿Se describe cómo se buscaron los estudios? (Generalmente no es sistemático). ¿Podría haber sesgo de selección de estudios?

4. **Selección de Estudios:** ¿Parece que los estudios incluidos son representativos de la literatura existente, o solo se seleccionaron aquellos que apoyan una cierta perspectiva?
5. **Síntesis de la Evidencia:** ¿Cómo se combinaron los hallazgos de los estudios? (Generalmente es un resumen cualitativo basado en la opinión del autor). ¿Se presentan los resultados de manera equilibrada?
6. **Sesgos Potenciales:** ¿Podría la opinión o perspectiva del autor haber influido en la selección o interpretación de los estudios? ¿Hay conflictos de interés?

Métricas Clave de Interpretación:

Las revisiones narrativas no suelen presentar métricas de síntesis cuantitativa. La interpretación se basa en el resumen cualitativo de los hallazgos de los estudios primarios.

- **Resumen de Hallazgos:** Descripción de las principales conclusiones de los estudios individuales citados.
  - *Ejemplo*: La revisión narrativa describe varios estudios que sugieren que la dieta mediterránea se asocia con un menor riesgo de enfermedad cardiovascular, mencionando algunos resultados clave de esos estudios.

5.12. Análisis Crítico de Revisiones Sistemáticas

Las revisiones sistemáticas utilizan métodos rigurosos y transparentes para identificar, seleccionar, evaluar críticamente y sintetizar la evidencia de estudios primarios relevantes sobre una pregunta específica. Su objetivo es proporcionar un resumen imparcial y completo de la mejor evidencia disponible. Son menos susceptibles a sesgos que las revisiones narrativas.

Paso a Paso: Evaluación Crítica de Revisiones Sistemáticas

Al evaluar una revisión sistemática, considera:

1. **Formulación de la Pregunta:** ¿Está la pregunta de investigación claramente definida, idealmente en formato PICOS?
2. **Protocolo y Registro:** ¿Existió un protocolo preespecificado? ¿Fue registrado, ej. en PROSPERO?

3. **Estrategia de Búsqueda:** ¿Fue la búsqueda exhaustiva y sistemática? ¿Se buscaron estudios en múltiples bases de datos y otras fuentes? ¿Se describió la estrategia de búsqueda completa?

4. **Criterios de Elegibilidad:** ¿Están los criterios para incluir o excluir estudios claramente definidos y aplicados de manera consistente?

5. **Selección de Estudios:** ¿Se realizó la selección de estudios por al menos dos revisores independientes para reducir el sesgo?

6. **Extracción de Datos:** ¿Se extrajeron los datos de manera estandarizada y por al menos dos revisores independientes?

7. **Evaluación del Riesgo de Sesgo de los Estudios Primarios:** ¿Se evaluó sistemáticamente la calidad metodológica y el riesgo de sesgo de cada estudio incluido utilizando una herramienta apropiada para el diseño del estudio?

8. **Síntesis de la Evidencia:** ¿Se realizó un metaanálisis? Si no, ¿la síntesis cualitativa fue apropiada y bien realizada?

9. **Evaluación de la Heterogeneidad:** Si se realizó metaanálisis, ¿se evaluó la heterogeneidad entre los estudios? ¿Se exploraron las posibles fuentes de heterogeneidad?

10. **Evaluación del Sesgo de Publicación:** ¿Se evaluó el sesgo de publicación?

Métricas Clave de Interpretación:

Las métricas dependen de si se realizó un metaanálisis.

- **Resultados de la Síntesis Cualitativa:** Resumen de los hallazgos de los estudios incluidos, identificando patrones y diferencias.
- **Resultados del Metaanálisis (si aplica):**
  - **Medida Resumen del Efecto (ej. OR, RR, HR, DM combinados):** La estimación global del efecto a través de todos los estudios.
    - *Ejemplo:* Un metaanálisis encuentra un RR combinado de 0.85 para la incidencia de un evento con una intervención vs. control. Esto sugiere que

la intervención reduce el riesgo relativo del evento en un 15% en promedio a través de los estudios.

- **Intervalo de Confianza (IC 95%) para la Medida Resumen:** Evalúa la precisión de la estimación global.
  - *Ejemplo*: RR combinado = 0.85 (IC 95%: 0.78 - 0.93). Como el intervalo no incluye el 1, el resultado es estadísticamente significativo.
- **Estadísticas de Heterogeneidad (Prueba Q, I^2):** Indican si los resultados de los estudios individuales son consistentes.
  - *Ejemplo*: I^2 = 60%. Esto sugiere una heterogeneidad moderada entre los estudios, lo que indica que el efecto puede variar entre los estudios por razones más allá del azar.
- **Gráficos de Bosque (*Forest Plots*):** Visualización de los resultados de cada estudio y la medida resumen.
- **Gráficos de Embudo (*Funnel Plots*):** Evaluación visual del sesgo de publicación.

5.13. Análisis Crítico de Metaanálisis

Un metaanálisis es la síntesis estadística de los resultados de múltiples estudios primarios. Se realiza como parte de una revisión sistemática cuando los estudios son suficientemente homogéneos. La evaluación crítica de un metaanálisis se superpone con la de la revisión sistemática, pero se enfoca en los aspectos estadísticos de la combinación de datos.

Paso a Paso: Evaluación Crítica de Metaanálisis

Al evaluar un metaanálisis, considera:

1. **Adecuación del Metaanálisis:** ¿Fue apropiado realizar un metaanálisis dados los estudios primarios (homogeneidad clínica y metodológica suficiente)?
2. **Métodos Estadísticos:** ¿Se utilizaron métodos estadísticos apropiados para combinar los resultados (ej. modelo de efectos fijos vs. aleatorios)? ¿Fue la elección del modelo adecuada dada la heterogeneidad?

3. **Evaluación y Manejo de la Heterogeneidad:** ¿Se evaluó la heterogeneidad (Prueba Q, I^2)? Si la heterogeneidad fue alta, ¿se exploraron sus fuentes (análisis de subgrupos, metaregresión)? ¿Se manejó adecuadamente en el análisis (ej. modelo de efectos aleatorios)?
4. **Evaluación del Sesgo de Publicación:** ¿Se evaluó el sesgo de publicación (ej. gráfico de embudo, prueba de Egger)?
5. **Interpretación de la Medida Resumen:** ¿Se interpretó correctamente la medida resumen del efecto (OR, RR, HR, DM combinados)? ¿Se consideró su precisión (IC 95%) y significancia estadística (valor p)?
6. **Análisis de Sensibilidad:** ¿Se realizaron análisis de sensibilidad para evaluar la robustez de los resultados?

Métricas Clave de Interpretación:

Las métricas son las mismas que las reportadas en la sección de Revisiones Sistemáticas con Metaanálisis, pero la evaluación crítica se centra en su validez estadística y la solidez de la síntesis.

- **Medida Resumen del Efecto Combinado (OR, RR, HR, DM):** La estimación global del efecto.
- **Intervalo de Confianza (IC 95%) para la Medida Resumen:** Precisión de la estimación global.
- **Estadísticas de Heterogeneidad (Prueba Q, I^2):** Grado de inconsistencia entre los estudios.
- **Resultados de Análisis de Subgrupos y Metaregresión:** Si se realizó, para explorar fuentes de heterogeneidad.
- **Resultados de la Evaluación del Sesgo de Publicación:** Evidencia de publicación selectiva.
- **Resultados de Análisis de Sensibilidad:** Evaluación de la robustez de los hallazgos.

La evaluación crítica de un metaanálisis nos permite juzgar la solidez de la evidencia sintetizada, comprender la variabilidad entre los estudios y confiar en las conclusiones globales que guiarán la práctica clínica y las futuras investigaciones.

5.14. Análisis Crítico de Contenido de Divulgación en Salud (YouTubers y Expertos)

En la era digital, la información sobre salud fluye a través de múltiples canales, y los videos de YouTube, los podcasts y las publicaciones en redes sociales se han convertido en fuentes populares para el público, sin que la sola autoría garantice su calidad. Conviene distinguir perfiles epistemológicos claramente diferentes: ser médico no convierte automáticamente a alguien en experto en el tema específico que divulga; un especialista con formación y producción científica verificable no equivale a un divulgador sin credenciales; y la popularidad o el tamaño de audiencia de un influencer no constituye evidencia de competencia científica ni sustituye la revisión por pares. Este contenido, aunque a menudo accesible y atractivo, varía enormemente en su calidad, precisión y base científica, independientemente de quién lo produzca. Para el profesional de la salud, es crucial no solo evaluar la literatura académica formal, sino también desarrollar una capacidad crítica para discernir la fiabilidad de la información de divulgación que consumen los pacientes o que puede influir en la percepción pública de la salud, verificando siempre si el contenido se basa en fuentes primarias sólidas, si distingue evidencia de opinión, y si ha sido sometido a algún tipo de contraste o revisión experta, en lugar de asumir que la credencial profesional, la popularidad o el formato mediático garantizan por sí mismos su validez científica.

Paso a Paso: Evaluación Crítica de Contenido de Divulgación en Salud

Al evaluar un video de YouTube, un podcast o una publicación de un experto en redes sociales sobre un tema de salud, considera los siguientes puntos:

1. **Identifica la Fuente y la Credibilidad del Creador:** ¿Quién es la persona que presenta la información? ¿Cuál es su formación académica y experiencia relevante en el tema? ¿Están afiliados a instituciones reconocidas? ¿Son sus credenciales apropiadas para hablar sobre este tema? Ten cuidado con los títulos o afirmaciones exageradas.

2. **Define el Propósito y la Audiencia:** ¿Cuál es el objetivo principal del contenido (informar, entretener, vender un producto)? ¿A quién está dirigido (público general, pacientes, otros profesionales)? El propósito puede influir en cómo se presenta la información.

3. **Evalúa las Fuentes de Información Citadas (o la Falta de Ellas):** ¿El creador cita las fuentes de donde obtuvo la información? ¿Son estas fuentes estudios científicos, guías de práctica clínica de organizaciones reconocidas, o son solo opiniones personales o anécdotas? La ausencia de fuentes verificables es una señal de alerta importante.

4. **Busca Sesgos Potenciales y Conflictos de Interés:** ¿El creador tiene algún interés financiero o personal en promover una visión particular (ej. promociona un producto, tiene vínculos con una industria)? ¿El contenido parece sesgado hacia una perspectiva particular sin presentar otras visiones o limitaciones?

5. **Evalúa la Claridad y Precisión del Mensaje:** ¿Es la información presentada de manera clara y fácil de entender? ¿Se simplifican en exceso conceptos complejos? ¿Se utilizan términos médicos de forma correcta? ¿Se evitan afirmaciones exageradas o promesas milagrosas?

6. **Considera la Evidencia Subyacente (Sí se Cita):** Si se citan estudios, ¿se presentan sus hallazgos de manera precisa? ¿Se mencionan las limitaciones de esos estudios (ej. tamaño de muestra pequeño, diseño observacional)? ¿Se generalizan los hallazgos de manera inapropiada?

7. **Evalúa el Tono y el Lenguaje:** ¿El tono es sensacionalista, alarmista o excesivamente optimista? ¿Se utiliza lenguaje que podría generar miedo o falsas esperanzas?

8. **Verifica la Información en Fuentes Fiables:** Si la información parece importante o cuestionable, contrástala con fuentes de evidencia reconocidas (ej. revisiones sistemáticas en Cochrane, guías de práctica clínica de sociedades médicas, bases de datos de investigación como PubMed).

Métricas Clave de Interpretación:

Aunque no hay métricas estadísticas formales como en los estudios primarios, podemos pensar en métricas cualitativas para evaluar este tipo de contenido:

- **Nivel de Credibilidad de la Fuente:** Basado en las credenciales del creador y la reputación de las plataformas o instituciones asociadas (ej. Alto, Medio, Bajo).

- **Transparencia de las Fuentes Citadas:** ¿Se citan fuentes? ¿Son accesibles y verificables? (Sí/No, o Grado: Alto, Medio, Bajo).
- **Evidencia de Sesgo Potencial:** ¿Se detectan posibles conflictos de interés o una presentación unilateral de la información? (Sí/No).
- **Coherencia con la Evidencia Científica Establecida:** ¿La información presentada se alinea con lo que la investigación rigurosa ha demostrado? (Sí/No, o Grado de Coherencia).
- **Calidad de la Comunicación:** ¿Es el mensaje claro, preciso y adaptado a la audiencia sin sacrificar la exactitud? (Grado: Alto, Medio, Bajo).

Ejemplo: Estás viendo un video de YouTube donde un médico discute los beneficios de un suplemento dietético para mejorar la memoria. Al aplicar el análisis crítico, notas que el médico tiene vínculos comerciales con la empresa que produce el suplemento (sesgo potencial). Menciona un estudio que apoya sus afirmaciones, pero no proporciona la referencia. La información parece muy optimista y no menciona posibles efectos secundarios o interacciones. Al verificar en PubMed, encuentras que solo hay un estudio pequeño sobre el suplemento, financiado por la misma empresa, y que otras revisiones sistemáticas no han encontrado evidencia concluyente de su beneficio. Tu evaluación crítica te lleva a concluir que, a pesar de las credenciales del médico, el contenido está sesgado y no se basa en evidencia robusta.

La capacidad de evaluar críticamente el contenido de divulgación en salud es una habilidad esencial para los profesionales de la salud en la era digital, permitiéndonos guiar a nuestros pacientes hacia fuentes de información fiables y protegernos a nosotros mismos y a otros de la desinformación.

5.15. Uso de Herramientas de Inteligencia Artificial para el Análisis Crítico de Contenido de Divulgación

En la era de la sobrecarga de información, discernir la calidad y fiabilidad del contenido de salud en línea, especialmente el proveniente de fuentes no académicas formales como YouTubers o expertos en redes sociales, es un desafío creciente. Afortunadamente, la Inteligencia Artificial está empezando a ofrecer herramientas que pueden asistirnos en este proceso de análisis crítico, ayudándonos a verificar afirmaciones, encontrar evidencia de respaldo

o identificar información contradictoria de manera más eficiente. Piensen en estas herramientas como lupas digitales avanzadas que nos ayudan a examinar las afirmaciones de salud con el respaldo de la literatura científica.

Herramientas como **Consensus** y **Open Evidence** son ejemplos de cómo la IA está siendo aplicada para facilitar el acceso y la síntesis de la evidencia científica. Estas plataformas utilizan procesamiento de lenguaje natural y otras técnicas de IA para buscar en vastas bases de datos de literatura académica (principalmente estudios revisados por pares) y encontrar artículos relevantes que apoyen o refuten afirmaciones específicas.

Cómo la IA puede ayudar en el Análisis Crítico de Contenido de Divulgación:

1. **Verificación Rápida de Afirmaciones:** Si un YouTuber o experto hace una afirmación específica sobre un tratamiento, un factor de riesgo o un beneficio para la salud, puedes usar estas herramientas de IA para buscar rápidamente si existe evidencia científica que respalde o contradiga esa afirmación en la literatura revisada por pares.
2. **Identificación de Estudios Clave:** Las herramientas pueden identificar los estudios más relevantes o las revisiones sistemáticas que abordan la afirmación, proporcionando un punto de partida para una evaluación más profunda.
3. **Resumen de la Evidencia Existente:** Algunas herramientas pueden generar resúmenes concisos de la evidencia encontrada sobre un tema, ayudándote a comprender rápidamente el estado actual del conocimiento científico.
4. **Exploración de Consistencia:** Puedes usar estas herramientas para ver si la evidencia es consistente o si existen estudios con resultados contradictorios, lo cual es una señal importante al evaluar la fiabilidad de una afirmación.

Paso a Paso: Evaluación Crítica de Contenido de Divulgación Asistida por IA

Al evaluar un contenido de divulgación en salud utilizando herramientas de

IA, puedes seguir estos pasos, integrando tu juicio crítico con la asistencia de la tecnología:

1. **Identifica la Afirmación Clave:** ¿Cuál es la afirmación principal sobre salud que el creador está haciendo? Sé lo más específico posible.

- *Ejemplo*: "Tomar suplementos de vitamina D previene los resfriados comunes."

2. **Evalúa la Fuente Inicial (Creador):** Aplica los pasos del análisis crítico de contenido de divulgación (Sección 5.14) para evaluar la credibilidad del creador, el propósito del contenido y los posibles sesgos. Esto te dará un contexto inicial.

3. **Formula la Pregunta para la Herramienta de IA:** Reformula la afirmación clave como una pregunta de investigación clara y concisa que puedas ingresar en una herramienta como Consensus o Open Evidence.

- *Ejemplo*: "¿Los suplementos de vitamina D previenen las infecciones del tracto respiratorio superior (resfriados)?"

4. **Utiliza la Herramienta de IA para Buscar Evidencia:** Ingresa tu pregunta en la herramienta de IA.

- *Ejemplo*: Busca en Consensus o Open Evidence la pregunta formulada.

5. **Evalúa los Resultados de la Búsqueda de IA:** Revisa los estudios o resúmenes que la herramienta de IA te proporciona.

- ¿La IA encuentra estudios directamente relevantes a tu pregunta?

- ¿Qué tipo de estudios son (ensayos clínicos, estudios observacionales, revisiones sistemáticas)? Recuerda la jerarquía de la evidencia.

- ¿Los resúmenes o fragmentos de texto proporcionados por la IA apoyan o contradicen la afirmación original?

- ¿La IA identifica revisiones sistemáticas o metaanálisis sobre el tema? Estos son a menudo la mejor fuente de evidencia.

6. **Profundiza en la Evidencia Clave (Sí es Necesario):** Si la IA identifica estudios o revisiones clave, considera buscar el texto completo para una

evaluación más profunda (aplicando principios de análisis crítico para el diseño específico del estudio).

- *Ejemplo*: Si la IA identifica una revisión sistemática de Cochrane sobre vitamina D y resfriados, busca esa revisión para ver sus métodos, resultados y conclusión.

7. **Integra la Evidencia Encontrada con la Afirmación Original:** Compara lo que dice el creador del contenido de divulgación con lo que encontraste en la literatura científica a través de la herramienta de IA.

- ¿Está la afirmación del creador bien respaldada por evidencia robusta (ej. metaanálisis de ECA)?
- ¿El creador simplificó demasiado los hallazgos (ej. ignoró la heterogeneidad, las limitaciones)?
- ¿La evidencia científica contradice la afirmación?

8. **Formula tu Conclusión Crítica:** Basado en la credibilidad de la fuente original y la evidencia científica encontrada (con la ayuda de la IA), ¿cuán fiable es la afirmación? ¿Qué tan bien se alinea con el conocimiento científico actual?

Métricas Clave de Interpretación (Asistidas por IA):

Las métricas aquí son una combinación de la evaluación cualitativa del contenido original y la evidencia encontrada con la IA:

- **Número de Estudios Relevantes Encontrados por IA:** Un indicador inicial de cuánta investigación existe sobre el tema.
- **Tipo de Estudios Encontrados por IA:** ¿Predominan los ECA, cohortes, estudios *in vitro*? (Relacionado con la jerarquía de la evidencia).
- **Consistencia de los Hallazgos de los Estudios Encontrados:** ¿Los estudios tienden a mostrar resultados similares o son contradictorios? (Evaluado a través de resúmenes de IA o metaanálisis si están disponibles).
- **Medida de Efecto Reportada en Estudios Clave (si aplica):** Si la IA extrae medidas de efecto (ej. RR, OR, HR), ¿cuál es su magnitud y significancia?

- **Calidad de la Evidencia (Evaluación Rápida Asistida por IA):** Algunas herramientas pueden intentar dar una indicación del riesgo de sesgo o la calidad de los estudios, aunque esto requiere validación humana.

Ejemplo: Usas Consensus y encuentras varios estudios, incluyendo una revisión sistemática reciente. La revisión concluye que la suplementación con vitamina D no reduce significativamente el riesgo de infecciones respiratorias agudas en la población general, pero podría tener un pequeño beneficio en individuos con deficiencia severa de vitamina D. Tu evaluación crítica, asistida por la IA, te lleva a concluir que la afirmación del YouTuber es una simplificación excesiva y no se alinea con la evidencia más robusta, que sugiere un beneficio muy limitado y condicionado a la deficiencia basal.

El uso de herramientas de IA como Consensus y Open Evidence no reemplaza el pensamiento crítico, sino que lo potencia, permitiéndonos acceder y evaluar la evidencia científica de manera más eficiente para contrastar las afirmaciones de salud que encontramos en el vasto y a menudo confuso panorama digital.

Conclusiones

1. **Necesidad de Evaluación Crítica:** No toda la literatura científica publicada es de igual calidad; es esencial evaluar críticamente los estudios para determinar su fiabilidad y aplicabilidad.

2. **Conceptos Centrales:** Comprender y buscar activamente el **sesgo** (errores sistemáticos) y la **confusión** (distorsión por terceros factores) es fundamental para evaluar la **validez interna** (grado en que los resultados son correctos para la muestra estudiada).

3. **Generalizabilidad:** La **validez externa** determina si los resultados de un estudio (incluso si tiene alta validez interna) pueden aplicarse a otras poblaciones o entornos.

4. **Proceso Sistemático:** La evaluación crítica sigue pasos lógicos: entender la pregunta, identificar el diseño, evaluar la selección de participantes, la medición de exposición/desenlace, el control de confusores, el análisis estadístico, otros sesgos y la interpretación contextual.

5. **Aplicación Específica por Diseño:** Los principios de evaluación crítica se aplican de forma específica a cada diseño de estudio (ECA, cohorte, casos-controles, etc.), prestando atención a sus fortalezas y debilidades inherentes y sesgos típicos.

6. **Evaluación de Desenlaces:** En estudios de desenlace, es crucial evaluar la relevancia clínica del desenlace medido y la validez de su medición, interpretando métricas como DR, RRR y NNT.

7. **Evaluación de Síntesis:** Las revisiones sistemáticas y metanálisis también requieren evaluación crítica, enfocándose en la rigurosidad de la búsqueda, selección, evaluación de calidad de estudios primarios y métodos de síntesis (incluyendo manejo de heterogeneidad y sesgo de publicación).

8. **Evaluación de Divulgación:** El contenido de divulgación en salud (ej. de YouTubers) debe evaluarse críticamente considerando la credibilidad de la fuente, la citación de evidencia, posibles sesgos/conflictos de interés y la coherencia con el conocimiento científico establecido, pudiendo usar herramientas de IA como apoyo.

9. **Interpretación Integral:** La interpretación final de cualquier estudio debe considerar la validez interna, la magnitud y precisión del efecto (intervalos de confianza), la relevancia clínica y la validez externa.

CAPÍTULO 17

ESTRUCTURA Y REDACCIÓN DEL ARTÍCULO CIENTÍFICO

Structure and Writing of a Scientific Article

Resumen

Este capítulo aborda la comunicación científica escrita como un pilar fundamental de la investigación, enfatizando los principios de claridad, precisión y objetividad. Describe la estructura canónica IMRaD (Introducción, Materiales y Métodos, Resultados y Discusión) como el formato estándar para artículos de investigación original, detallando el propósito y contenido específico de cada sección, incluyendo Título, Autores, Filiación, Resumen y Palabras Clave. Profundiza en aspectos clave del estilo y lenguaje científico: uso de vocabulario técnico preciso y consistente, nomenclatura estándar, y la importancia de la gramática, puntuación y ortografía correctas, así como el uso adecuado de los tiempos verbales. Explica el manejo efectivo de Tablas y Figuras como complementos visuales del texto, que deben ser autocontenidas y referenciadas adecuadamente. Subraya la importancia ética y práctica de la Citación y Referencias Bibliográficas para dar crédito, fundamentar afirmaciones y evitar el plagio. Menciona secciones complementarias como Agradecimientos y Apéndices. Finalmente, destaca la etapa crucial de Revisión y Edición rigurosa (autorevisión, revisión por pares informal) como garantía de calidad para mejorar la claridad, corregir errores y asegurar la coherencia antes de la publicación.

Palabras clave: Redacción científica, artículo científico, estructura IMRaD, comunicación científica, estilo científico, revisión de manuscritos.

Abstract

This chapter addresses written scientific communication as a fundamental pillar of research, emphasizing the principles of clarity, precision, and objectivity. It describes the canonical IMRaD structure (Introduction, Materials and Methods, Results, and Discussion) as the standard format for original research arti-

cles, detailing the specific purpose and content of each section, including Title, Authors, Affiliation, Abstract, and Keywords. It delves into key aspects of scientific style and language: use of precise and consistent technical vocabulary, standard nomenclature, and the importance of correct grammar, punctuation, and spelling, as well as the appropriate use of verb tenses. It explains the effective handling of Tables and Figures as visual complements to the text, which should be self-contained and properly referenced. It underscores the ethical and practical importance of Citation and Bibliographic References for giving credit, substantiating claims, and avoiding plagiarism. Complementary sections like Acknowledgements and Appendices are mentioned. Finally, it highlights the crucial stage of rigorous Revision and Editing (self-review, informal peer review) as a quality assurance step to improve clarity, correct errors, and ensure coherence before publication.

Keywords: *Scientific writing, scientific article, IMRaD structure, scientific communication, scientific style, manuscript revision.*

6.1. Principios Generales de la Comunicación Científica Escrita

Piensen por un momento en la ciencia no solo como un vasto conjunto de conocimientos acumulados, sino como una inmensa y continua conversación global. Desde los primeros naturalistas que garabateaban sus observaciones en cuadernos, pasando por los filósofos que debatían las leyes del universo, hasta los equipos de investigación de hoy en día analizando *big data* con herramientas sofisticadas, el impulso de entender y, crucialmente, de **compartir** ese entendimiento ha sido el motor de todo progreso. Un descubrimiento, por brillante que sea, permanece como una chispa solitaria hasta que se comparte, verificado y utilizado por otros. Aquí es donde la comunicación científica escrita entra en escena, no como una tarea burocrática, sino como el acto fundamental de convertir una idea o un hallazgo en una contribución al edificio del conocimiento humano.

El artículo científico, en particular, es la moneda de cambio en este gran mercado global de ideas. Es el medio primario a través del cual un investigador presenta sus hallazgos, métodos y conclusiones a sus pares y al mundo. No se trata solo de registrar lo que se hizo, sino de persuadir al lector de la solidez y relevancia de la investigación, proporcionando la información

necesaria para que pueda evaluar el trabajo, replicar los experimentos si lo desea o construir sobre esos cimientos. Diversas fuentes, incluidas las seis proporcionadas para esta guía, así como investigaciones complementarias en el campo de la comunicación académica, enfatizan unánimemente que la efectividad de esta comunicación descansa sobre pilares inquebrantables: **claridad, precisión y objetividad**.

La **claridad** es su linterna en la oscuridad. Un texto científico claro, *aquel que se comprende sin esfuerzo intelectual desproporcionado*, guía al lector directamente a través de su razonamiento y sus hallazgos. Pensemos en los escritos de Jared Diamond, quien logra explicar complejas interacciones históricas y geográficas con una fluidez que hace que sus ideas, *profundamente fundamentadas en investigación rigurosa*, sean accesibles a un público amplio sin sacrificar el rigor. En el contexto de un artículo médico o epidemiológico, la claridad asegura que un colega, *quizás al otro lado del mundo y con poco tiempo*, pueda captar rápidamente la pregunta de investigación, cómo se abordó y qué se encontró. La ambigüedad, *en cambio*, es como una niebla que oculta el camino, frustra al lector y, peor aún, puede llevar a malentendidos críticos en la interpretación de resultados que tienen implicaciones para la salud.

La **precisión**, por su parte, es la exigencia de la realidad. En ciencia, y especialmente en campos cuantitativos como la epidemiología, los detalles importan enormemente. No basta con decir "se observó un aumento"; es crucial especificar "se observó un aumento del 15% (IC 95%: 12-18%) en la incidencia". Esta meticulosidad, que se extiende a la descripción exacta de los métodos, los materiales, la población de estudio y las herramientas estadísticas utilizadas, es lo que permite a otros investigadores evaluar la validez interna y externa de su estudio y, si lo desean, intentar replicarlo. Es el equivalente a proporcionar las especificaciones exactas de un engranaje en una máquina compleja; una mínima variación puede hacer que todo el mecanismo falle. La precisión, aunque a veces pueda parecer tediosa, es la salvaguarda contra la mala interpretación y el fundamento de la replicabilidad, uno de los pilares del método científico.

Finalmente, la **objetividad** es el compromiso con la verdad de los datos, independientemente de sus expectativas o preferencias personales. Aunque

la investigación es un proceso impulsado por la curiosidad y la pasión, su comunicación debe esforzarse por presentar los hallazgos de manera imparcial. Esto significa describir los resultados tal como son, incluso si no confirman su hipótesis inicial, discutir las limitaciones del estudio honestamente y separar su interpretación (basada en la evidencia) de los hechos crudos. Ed Yong, por ejemplo, al explorar el fascinante mundo microbiano, presenta la evidencia con una objetividad admirable, permitiendo que la maravilla surja de los hechos mismos, no de una presentación sesgada. En la redacción científica, la objetividad genera confianza en el lector, un elemento esencial para que su trabajo sea aceptado y citado por la comunidad.

Dominar la comunicación científica escrita, por tanto, es mucho más que seguir una serie de reglas gramaticales o de formato; es aprender a ser claros en su exposición, precisos en su descripción y objetivos en su presentación. Es adquirir la habilidad de transformar su laboriosa investigación, su viaje de descubrimiento, en un relato legible y convincente que pueda ser entendido, evaluado y utilizado por otros para seguir construyendo ese vasto edificio del conocimiento. Comprender y aplicar estos principios fundamentales es el primer paso en el camino hacia la publicación efectiva, una travesía que requiere conocer la estructura formal de estos documentos y dominar las herramientas lingüísticas necesarias.

Excelente. Prosigamos con la sección que desvela el esqueleto sobre el cual se construye la mayoría de los artículos científicos, el ubicuo formato IMRaD.

6.2. La Estructura Canónica: El Formato IMRaD

Si los principios de claridad, precisión y objetividad son los cimientos de una comunicación científica sólida, la estructura del artículo es el andamiaje que permite construir el edificio de su estudio de manera coherente y accesible. A diferencia de un ensayo libre o una obra literaria, el artículo científico, especialmente el que presenta investigación original, se adhiere a un formato notablemente estandarizado. Esta uniformidad no es arbitraria; es una convención desarrollada a lo largo del tiempo para optimizar la transmisión de información científica.

En el vasto paisaje de la literatura científica, un formato emerge como el pilar fundamental para los artículos de investigación original: el formato

IMRaD. Este acrónimo, que resuena en los pasillos de la academia y las redacciones de revistas de todo el mundo, proviene de sus secciones principales en inglés: **I***ntroduction*, **M***ethods*, **R***esults*, *and* **D***iscussion*. Introducción, Materiales y Métodos, Resultados y Discusión. Es un lenguaje universal, una especie de mapa para que cualquier científico, sin importar su disciplina exacta, pueda orientarse en un nuevo estudio.

La genialidad del formato IMRaD radica en su lógica implacable, que espejea el propio viaje de la investigación. Piensen en ello:

1. **Introducción (I):** Presenta el **qué** del problema. ¿Cuál es la pregunta de investigación? ¿Por qué es importante abordarla? ¿Qué se sabe ya sobre el tema (y qué no)? Es el preludio que establece el escenario y el propósito de su estudio.
2. **Materiales y Métodos (M):** Explica el **cómo** se abordó el problema. ¿Cómo diseñaron el estudio? ¿Quiénes fueron los participantes o qué materiales se utilizaron? ¿Cómo se recogieron los datos y qué análisis estadísticos se aplicaron? Es la hoja de ruta detallada que permite a otros entender (y, si es posible, replicar) su camino.
3. **Resultados (R):** Presenta el **qué se encontró**. Aquí se exponen los hallazgos de su investigación de forma clara y concisa, a menudo apoyados por tablas y figuras. Es la evidencia empírica que obtuvieron siguiendo sus métodos.
4. **Discusión (D):** Interpreta el **qué significa** lo que se encontró. ¿Qué responden sus resultados a la pregunta de investigación? ¿Cómo se comparan con lo que ya se sabía? ¿Cuáles son las implicaciones de sus hallazgos y cuáles son las limitaciones del estudio? Es el espacio para dialogar con la literatura existente y proyectar futuras investigaciones.

Expertos en publicación científica y directrices de revistas de alto impacto, información que se extrae tanto de las seis fuentes base proporcionadas para esta guía como de otros manuales de estilo ampliamente reconocidos, insisten en que esta estructura facilita enormemente la revisión por pares y la rápida asimilación del contenido por parte de los lectores. Para el lector, familiarizado con este patrón, navegar por un artículo IMRaD es intuitivo; sabe exactamente dónde buscar la información de fondo, cómo se realizó el estudio, qué se halló y cómo interpretarlo. Para el escritor, proporciona un

marco claro, evitando perderse en la organización del vasto material generado por una investigación.

Aunque el IMRaD es el rey indiscutible para la investigación original y es el formato al que nos referiremos principalmente, existen otros formatos para tipos específicos de artículos, como revisiones sistemáticas, reportes de caso o artículos teóricos, que adaptan o reconfiguran estas secciones. Sin embargo, comprender la lógica y el propósito de cada componente del IMRaD es el paso previo fundamental para abordar la tarea de redactarlas efectivamente, que exploraremos a fondo, sección por sección, en la siguiente parte.

6.3. Redacción Detallada de Cada Sección del Artículo

Con el andamiaje del IMRaD firme en nuestra mente, es hora de adentrarnos en la minuciosa labor de dar vida a cada una de sus partes. Cada sección tiene un propósito distinto y exige un enfoque de redacción particular, como si cada sala de un edificio tuviera una función única y, por tanto, requiriera un diseño y mobiliario específicos. Dominar la escritura de cada una es fundamental para que el artículo funcione como un todo coherente y persuasivo.

6.3.1. Título, Autores y Filiación

Aunque parezca obvio, el título es su primera y a menudo única oportunidad de captar la atención del lector. Es la puerta de entrada a su trabajo, a menudo lo único que un lector ve en una lista de resultados de búsqueda o en una tabla de contenido. Un buen título debe ser informativo y conciso, capturando la esencia del estudio sin ser demasiado largo ni demasiado vago. Debe incluir las palabras clave principales, aquellas que un colega usaría para buscar información sobre su tema, facilitando así que su artículo sea descubierto. Comparen: “Estudio sobre diabetes” (demasiado vago) versus “Impacto de la telemedicina en el control glucémico de pacientes diabéticos tipo 2 en zonas rurales” (informativo, conciso y con palabras clave).

Los autores son quienes realizaron el trabajo y se hacen responsables de él. Los criterios de autoría, ampliamente aceptados y detallados en diversas guías éticas y manuales de publicación, suelen basarse en contribuciones significativas al diseño, la adquisición, el análisis o la interpretación de los datos, la redacción del manuscrito y la aprobación final de la versión a publicar. El orden de los autores es crucial y debe ser acordado por todos, reflejando generalmente la magnitud de la contribución de cada persona, con

el primer autor liderando el trabajo y el último (a menudo el investigador principal o supervisor) representando la responsabilidad general o la dirección del proyecto. La filiación indica la institución o instituciones donde se realizó la investigación, conectando el estudio con el entorno académico o de investigación que lo hizo posible.

6.3.2. Resumen y Palabras Clave

El resumen es su 'elevador *pitch*', su oportunidad de vender la esencia de su estudio en unas pocas líneas a un editor, un revisor o un lector apresurado. Es, con frecuencia, la única parte que se lee antes de decidir si vale la pena invertir tiempo en el artículo completo. Debe ser un microrrelato del estudio, condensando la introducción (el problema), los métodos (cómo se estudió), los resultados (lo que se encontró) y la discusión (lo que significa) en un solo párrafo o, en el caso de resúmenes estructurados, en secciones cortas y etiquetadas. Debe ser conciso, informativo y capaz de entenderse por sí solo, sin recurrir al resto del artículo. Eviten jerga excesiva o abreviaturas no estándar, y no incluyan información que no aparezca en el cuerpo principal del texto.

Las palabras clave son etiquetas para su artículo, fundamentales para la indexación en bases de datos como PubMed o Scopus. Deben ser términos específicos que identifiquen el contenido principal del estudio y faciliten que otros investigadores lo encuentren cuando busquen información relevante. Piensen en ellas como los hashtags más importantes para su investigación.

6.3.3. Introducción

Piensen en la introducción como un embudo: comienza amplio, estableciendo el contexto general y la relevancia del tema, para luego estrecharse progresivamente hasta presentar la pregunta específica o la hipótesis que su estudio busca responder. Inician presentando el conocimiento general sobre el tema, situando su investigación en el mapa del conocimiento existente. Luego, identifican el problema o la brecha en el conocimiento actual que su estudio aborda. Es crucial aquí citar la literatura relevante, mostrando que conocen el terreno y que su trabajo se construye sobre lo que ya se ha hecho. Finalmente, justifican por qué era importante realizar este estudio (su relevancia) y exponen claramente los objetivos o la hipótesis que guiaron su investigación. La introducción debe ser cautivadora, enganchando al lector y convenciéndolo de la importancia de seguir leyendo.

6.3.4. Materiales y Métodos

Esta sección es el corazón metodológico, el manual de instrucciones detallado que permitiría a otro investigador replicar su camino o, al menos, evaluar rigurosamente su validez. La precisión aquí no es negociable; cada detalle sobre la población de estudio, el diseño utilizado, las intervenciones (si las hubo), los procedimientos de recolección de datos, las mediciones, los instrumentos y los métodos de análisis estadístico es crucial. Deben describir con suficiente detalle "quién", "qué", "dónde", "cuándo" y "cómo" se realizó la investigación. Generalmente se redacta en tiempo pasado, ya que se describen acciones que ya se llevaron a cabo. Cuando se utilizan métodos estándar o previamente publicados, es suficiente con citar la referencia, pero si hay modificaciones, deben describirlas detalladamente. Piensen en esta sección como si estuvieran creando una '*checklist*' para que alguien más siga exactamente sus pasos.

6.3.5. Resultados

Aquí, los datos hablan por sí mismos. La sección de Resultados es la presentación objetiva de lo que encontraron, libre de interpretación, discusión o especulación (eso viene después). Deben presentar sus hallazgos de forma clara, lógica y secuencial, siguiendo el orden de los métodos descritos. Utilicen el texto para narrar los hallazgos principales, pero apóyense fuertemente en tablas y figuras (con leyendas claras) para mostrar datos detallados o tendencias visualmente impactantes. No repitan en el texto toda la información de tablas y figuras, sino que destaquen los hallazgos más relevantes y remitan al lector a los elementos visuales para el detalle. Esta sección también se redacta generalmente en tiempo pasado, describiendo lo que se observó o calculó.

6.3.6. Discusión

Si la introducción era un embudo que iba de lo general a lo específico, la discusión es un embudo invertido, comenzando con sus resultados específicos y expandiéndose para considerar sus implicaciones en el contexto más amplio del conocimiento. Inician resumiendo sus hallazgos clave, recordando al lector lo más importante que encontraron. Luego, interpretan esos resultados: ¿qué significan realmente? Comparan sus hallazgos con los de otros estudios (la literatura que mencionaron en la introducción), explicando si concuerdan o difieren y por qué. Es vital discutir las limitaciones de su pro-

pio estudio, mostrando honestidad intelectual y ayudando a los lectores a evaluar la robustez de sus conclusiones. Finalmente, discuten las implicaciones de sus hallazgos (¿qué aportan? ¿quién se beneficia?) y sugieren líneas de investigación futuras. La discusión es el espacio para el pensamiento crítico, donde conectan sus datos con el panorama general. Aquí, el uso de los tiempos verbales puede variar, usando el pasado para referirse a sus resultados específicos y el presente para referirse a hechos establecidos, conocimientos generales o las conclusiones que se derivan.

Dominar la redacción de cada una de estas secciones, comprendiendo su propósito único y sus convenciones, es la clave para transformar su investigación en un artículo publicable. Sin embargo, la estructura y el contenido por sección son solo una parte del desafío; la forma en que utilizamos el lenguaje dentro de ellas es igualmente vital, un aspecto que exploraremos a continuación al abordar el estilo, la gramática, la puntuación y otros detalles cruciales que pulen su mensaje hasta hacerlo impecable.

6.4. Estilo y Lenguaje en la Redacción Científica

Si el formato IMRaD es el andamiaje que da forma a su artículo, el estilo y el lenguaje son los materiales de construcción finos: el cemento bien mezclado, los ladrillos alineados con precisión, el acabado pulido. No basta con tener un gran descubrimiento o un análisis impecable; la forma en que se comunica ese conocimiento es tan vital como el conocimiento mismo. Un estudio revolucionario mal redactado puede pasar desapercibido o, peor aún, ser malentendido. Esta sección profundiza en el arte (y la ciencia) de usar las palabras con efectividad en el contexto académico.

6.4.1. Claridad, Concisión y Precisión (De Nuevo, pero Ahora en la Práctica)

Retomamos estos principios fundamentales de la sección 6.1, pero ahora enfocados en su aplicación directa en la escritura. La **claridad**, siempre su faro guía, implica que su texto se lea de manera fluida y lógica. Eviten las construcciones de frases enrevesadas, esas que obligan al lector a releer para desentrañar el significado. Usen transiciones suaves entre ideas y párrafos. Definir los términos técnicos o las abreviaturas en su primera aparición es un gesto de cortesía hacia el lector, asegurando que todos hablan el mismo idioma.

La **concisión**, un sello distintivo de la buena escritura científica, respeta el tiempo del lector y la necesidad de comunicar información de forma eficien-

te. Eliminen las palabras innecesarias. Comparen: "Debido al hecho de que la prevalencia de la enfermedad ha aumentado, es necesario investigar..." versus "Dado el aumento en la prevalencia de la enfermedad, es necesario investigar...". Ambas dicen lo mismo, pero la segunda es más directa y profesional. Eviten redundancias y circunloquios; vayan al grano.

La **precisión**, particularmente crítica en campos como la medicina y la epidemiología, exige un uso exacto del vocabulario y una descripción sin ambigüedades. No digan "un gran número de pacientes" si pueden decir "150 pacientes". Especificar el tipo de diseño de estudio, la técnica estadística exacta utilizada o la concentración de un reactivo, son ejemplos de precisión indispensable. Como señalan diversas guías de redacción científica, información que se subraya consistentemente en las fuentes consultadas, la falta de precisión puede impedir la replicación del estudio o llevar a interpretaciones erróneas con consecuencias potencialmente serias.

6.4.2. Vocabulario y Nomenclatura

El vocabulario científico es, por naturaleza, técnico. Utilicen los términos específicos de su disciplina de forma correcta y coherente. La **nomenclatura** estándar (por ejemplo, nombres de enfermedades según clasificaciones internacionales, nombres genéricos de fármacos en lugar de marcas comerciales) es fundamental para la comprensión global de su trabajo. Sin embargo, eviten la jerga innecesaria o excesivamente esotérica si existe un término más accesible que comunica lo mismo sin sacrificar la precisión. Crucialmente, una vez que eligen un término para referirse a un concepto, úsenlo consistentemente a lo largo de todo el manuscrito. Eviten la tentación de usar sinónimos solo por variedad; en ciencia, la consistencia es reina. Las abreviaturas deben ser definidas la primera vez que aparecen en el texto (salvo las muy comunes y estándar en el campo, como unidades de medida).

6.4.3. Aspectos Lingüísticos Fundamentales: Gramática, Puntuación y Ortografía

Estos son los engranajes finos que hacen que la máquina de su texto funcione sin problemas. Una gramática deficiente, más allá de ser un error formal, puede alterar el significado de una oración y generar confusión. Asegúrense de que los sujetos y los verbos concuerden, de que las frases estén bien construidas y de evitar errores comunes en español que pueden minar la credibi-

lidad, como el uso incorrecto de preposiciones o la construcción de frases demasiado largas y complejas.

La **puntuación** es la señalización de su texto. Un punto, una coma, un punto y coma, cada signo tiene una función específica para organizar las ideas y guiar al lector a través de la estructura sintáctica de la oración. Un uso incorrecto de la puntuación puede cambiar drásticamente el sentido. Por ejemplo, consideren la diferencia entre "Los pacientes que recibieron el tratamiento, mejoraron rápidamente" (implica que todos los pacientes tratados mejoraron) y "Los pacientes que recibieron el tratamiento mejoraron rápidamente" (implica que solo algunos de los pacientes que recibieron el tratamiento mejoraron). **Recuerden siempre usar comas para separar incisos**, esas pausas informativas que se insertan en la oración principal y añaden un detalle o una aclaración, en lugar de guiones, que tienen otros usos. Revisen las reglas de puntuación para asegurar que su texto fluya de manera clara y precisa.

La **ortografía** correcta es simplemente no negociable. Un texto con errores ortográficos, por brillante que sea la investigación subyacente, pierde instantáneamente profesionalismo y credibilidad. Usen correctores automáticos, pero no confíen ciegamente en ellos; revisen manualmente y, si es posible, pidan a alguien más que lea su texto.

6.4.4. Uso Correcto de los Tiempos Verbales en Textos Científicos

El uso de los tiempos verbales en un artículo científico sigue convenciones específicas que ayudan al lector a distinguir entre lo que se hizo (su estudio), lo que se encontró (sus resultados) y lo que se sabe (hechos establecidos o conclusiones).

- **Pasado simple (Pretérito Perfecto Simple):** Es el tiempo predominante en las secciones de **Materiales y Métodos** y **Resultados**. Se utiliza para describir acciones específicas que se realizaron en su estudio ("Administramos el fármaco", "Analizamos los datos", "La prevalencia fue del 10%").
- **Presente simple:** Se utiliza para describir hechos establecidos universalmente ("El agua hierve a 100°C"), referencias a tablas o figuras ("La Tabla 1 muestra..."), conclusiones generales que se derivan de su trabajo o de la literatura ("Estos resultados sugieren que...", "La diabetes es una

enfermedad crónica"). También para referirse a la hipótesis o los objetivos en la Introducción ("El objetivo de este estudio es determinar...").

- **Futuro simple o Presente:** Ocasionalmente se usan para referirse a trabajo futuro o implicaciones ("Será necesario investigar...", "Esta línea de investigación continuará...").

La clave es la **consistencia** dentro de cada sección y el uso apropiado que guíe al lector a través del tiempo verbal correcto para cada tipo de información que presenta.

Dominar estos aspectos del estilo y el lenguaje no es meramente una cuestión de formalidad, sino de efectividad comunicativa. Un texto pulido, libre de ambigüedades y errores, permite que la fortaleza de su investigación brille sin distracciones, preparando el terreno para la correcta presentación de sus datos visuales y la gestión de las fuentes que dan sustento a sus afirmaciones.

6.5. Manejo de Tablas y Figuras en el Texto

Mientras que las palabras construyen la narrativa y el argumento de su estudio, hay información que se comunica de manera más poderosa y eficiente a través de representaciones visuales. Las tablas y figuras no son adornos decorativos; son herramientas esenciales para presentar datos complejos, tendencias, patrones o estructuras de forma concisa y comprensible. Piensen en ellas como los mapas, gráficos o diagramas que acompañan un buen reportaje de investigación, ofreciendo al lector una visión rápida y clara de la evidencia que sustenta el relato escrito.

El propósito principal de las tablas y figuras es complementar (no duplicar) la información presentada en el texto principal. Una tabla es excelente para mostrar datos exactos y comparaciones puntuales, especialmente cuando hay muchos valores. Una figura, como un gráfico de barras, una curva de supervivencia o un diagrama de flujo, es ideal para ilustrar tendencias, relaciones o procesos de forma visualmente intuitiva.

Es fundamental referenciar cada tabla y figura en el texto, y hacerlo en orden secuencial la primera vez que se mencionan. Esto guía al lector desde la narrativa escrita hacia la visualización de los datos específicos que se están discutiendo. Al referenciar una figura, por ejemplo, no se limiten a decir "Ver

Figura 1". Utilicen la referencia para resaltar el hallazgo clave o la tendencia que la figura ilustra: "La prevalencia mostró un aumento constante a lo largo del periodo de estudio (Figura 1), incrementándose significativamente entre 2010 y 2020". De esta manera, integran la visualización en su argumento.

Aunque los principios detallados para diseñar tablas claras y figuras informativas, incluyendo la elección del tipo de gráfico adecuado y la presentación visual efectiva de los datos, se abordaron extensamente en la Parte II de esta guía (especialmente en los capítulos sobre descripción y visualización de datos clínicos), es crucial recordar aquí que deben ser comprensibles por sí solas. Cada tabla debe tener un título claro y conciso en la parte superior, explicando de qué trata. Cada figura debe tener una leyenda (o pie de figura) en la parte inferior, describiendo lo que representa, las variables mostradas, las unidades y cualquier otra información necesaria para su interpretación sin tener que buscar en el texto principal. Las notas al pie de tabla, por ejemplo, son útiles para explicar abreviaturas, símbolos estadísticos o detalles metodológicos específicos de esa tabla.

En la sección de Resultados, el texto debe guiar al lector a través de las tablas y figuras principales, destacando los hallazgos clave que presentan sin repetir cada dato numérico. En la Discusión, pueden referenciar tablas y figuras nuevamente para interpretar los resultados en un contexto más amplio o compararlos con otros estudios. El manejo efectivo de estos elementos visuales, combinado con una redacción clara en el texto, maximiza el impacto y la comprensión de sus hallazgos. Las directrices de las revistas y los manuales de estilo, basados en la experiencia para optimizar la comunicación de datos complejos, ofrecen recomendaciones específicas para el formato y la presentación de estos elementos visuales, información que es vital consultar antes de la sumisión.

Presentar sus datos de manera efectiva a través de tablas y figuras es un componente esencial de la comunicación científica. Igualmente, es esencial dar el crédito adecuado a las fuentes de información e ideas que han nutrido su propia investigación y sustentan sus argumentos, un aspecto crucial que abordaremos a continuación.

6.6. Citación y Referencias Bibliográficas (Principios Básicos)

Después de pulir el lenguaje y organizar los datos visuales, nos encontramos ante otro pilar ético y práctico de la investigación: la correcta atribución de

las fuentes de información. Ningún estudio científico nace de la nada; siempre se construye sobre el conocimiento acumulado por generaciones de investigadores. La citación y la elaboración de la lista de referencias bibliográficas son los mecanismos formales para reconocer esa deuda intelectual y, al mismo tiempo, permitir que otros sigan el rastro de sus fuentes.

¿Por Qué Citar? Los Pilares de la Integridad Académica

Citar correctamente es mucho más que una formalidad académica; es un acto con múltiples propósitos esenciales:

1. **Dar Crédito:** Es el reconocimiento justo a los autores originales de las ideas, datos, métodos o palabras que han utilizado. Es un pilar de la honestidad y la ética en la investigación.
2. **Fundamentar Sus Afirmaciones:** Las citas proporcionan la evidencia que respalda sus aseveraciones en el texto. Al decir "Estudios previos han demostrado que...", deben respaldar esa afirmación con referencias específicas.
3. **Permitir la Verificación:** La lista de referencias permite que los lectores interesados localicen las fuentes originales para verificar la información, profundizar en el tema o entender el contexto completo.
4. **Situar su Trabajo:** Las citas muestran cómo su estudio se relaciona con la literatura existente, identificando si sus hallazgos apoyan, contradicen o amplían el conocimiento previo.
5. **Evitar el Plagio:** Crucialmente, citar es la forma de evitar el plagio, el uso de ideas o palabras ajenas como si fueran propias.

¿Qué Citar? Reconociendo la Influencia

Deben citar cualquier idea, dato, método, teoría o argumento que hayan tomado de otra fuente, ya sea que lo parafraseen (lo expresen con sus propias palabras), lo resuman o lo citen textualmente. Esto incluye artículos de revistas, libros, capítulos de libros, ponencias de conferencias, tesis, reportes, páginas web (con cautela y evaluando su fiabilidad) y cualquier otra fuente de información publicable o accesible públicamente. La única excepción es el conocimiento general o ampliamente conocido en el campo.

Principios Básicos de la Citación

La forma exacta de citar varía según el **estilo de citación** requerido por la revista o institución (los veremos brevemente más adelante). Sin embargo, hay principios universales:

- **Consistencia:** Una vez que eligen un estilo, deben aplicarlo rigurosamente y sin variaciones a lo largo de todo el manuscrito.
- **Precisión:** Cada fuente citada en el texto debe tener una entrada correspondiente y precisa en la lista de referencias, y cada entrada en la lista debe haber sido citada en el texto.
- **Ubicación:** La cita en el texto debe aparecer inmediatamente después de la información que están tomando de la fuente, o en un lugar donde quede claro a qué información se refiere.

La Citación en el Texto

La forma más común de citar en el texto es usando un sistema numérico (ej. [1], [2-4]) o un sistema de autor-año (ej. (Smith, 2020), (Smith & Jones, 2021)). Cuando citan, deben integrar la referencia de manera que no interrumpa la fluidez de la lectura. Pueden mencionarla al final de la frase o el párrafo donde se usa la información, o incluso mencionar a los autores como parte de la oración si es relevante para el argumento (ej. "Según Smith (2020)...").

La Lista de Referencias Bibliográficas

La lista de referencias, generalmente ubicada al final del manuscrito, proporciona la información completa sobre cada fuente citada en el texto, permitiendo que el lector la localice. El formato de cada entrada en esta lista (cómo se presenta el autor, año, título, revista, volumen, páginas, etc.) depende completamente del estilo de citación elegido.

Existen numerosos estilos de citación, cada uno con sus propias reglas minuciosas. En medicina y ciencias de la salud, estilos como Vancouver o APA (Asociación Americana de Psicología) son muy comunes, aunque existen muchos otros. Las directrices de las revistas, información que se extrae de las fuentes consultadas y es un aspecto crucial a verificar antes de redactar, especificarán cuál deben seguir. Usualmente, hay softwares de gestión de referencias (como Zotero, Mendeley o EndNote) que pueden ayudar a ges-

tionar sus bases de datos de referencias y formatearlas automáticamente en el estilo deseado, facilitando enormemente esta tarea.

La correcta citación y referenciación cierra un ciclo de construcción académica, asegurando que nuestro trabajo se integre ética y sólidamente en la conversación científica global. Con el cuerpo principal del manuscrito y sus referencias listo, nos acercamos a las etapas finales de preparación del manuscrito, incluyendo posibles secciones complementarias y el crucial proceso de revisión antes de que su investigación emprenda su viaje hacia la publicación.

6.7. Secciones Complementarias (Agradecimientos, Apéndices)

Más allá del cuerpo principal del artículo y su lista de referencias, hay ocasiones en que es necesario incluir información adicional o expresar gratitudes formales. Estas secciones complementarias, ubicadas usualmente al final del manuscrito, cumplen roles específicos y contribuyen a la transparencia y completitud de su publicación.

La sección de **Agradecimientos** es un espacio formal para reconocer a todas aquellas personas o instituciones que contribuyeron a su estudio pero cuyas contribuciones, siendo valiosas y necesarias, no cumplen con los criterios formales de autoría definidos en la sección 6.3.1. Aquí pueden mencionar a técnicos de laboratorio que realizaron análisis específicos, estadísticos que ofrecieron consultoría sin participar en la redacción principal, personal administrativo que facilitó la logística, o colegas que revisaron versiones tempranas del manuscrito ofreciendo sugerencias valiosas (siempre con su permiso). Es fundamental agradecer explícitamente a las **fuentes de financiación** que hicieron posible la investigación, indicando los nombres de las instituciones o becas recibidas. Este reconocimiento es un estándar ético y un requisito común de las revistas.

Por otro lado, los **Apéndices** (o Material Suplementario en las publicaciones online) son para detalles que, aunque relevantes para algunos lectores muy interesados, cuya inclusión en el cuerpo principal del artículo interrumpiría el flujo de la narrativa o lo haría excesivamente largo. Piensen en tablas de datos crudos muy extensas, protocolos detallados de laboratorio, cuestionarios completos utilizados en el estudio, figuras adicionales que ilustran análisis secundarios, o incluso código de programación si es relevante para

la reproducibilidad de los análisis. Las directrices de las revistas suelen especificar qué tipo de contenido es apropiado para material suplementario y cómo debe ser formateado, dado que frecuentemente se publica aparte del principal del artículo. Mantener esta información separada permite que el lector promedio se concentre en los hallazgos y la discusión principal, mientras que el lector que necesita detalles metodológicos precisos o datos extendidos puede encontrarlos fácilmente en los apéndices.

En esencia, estas secciones adicionales son una oportunidad para ser exhaustivos y transparentes sin sacrificar la concisión y claridad del cuerpo principal del artículo. Cumplen una función importante en la comunicación científica, asegurando que se dé el crédito adecuado y que los datos de respaldo estén disponibles para quienes los necesiten.

Una vez que estas secciones complementarias están listas y el manuscrito principal ha sido cuidadosamente elaborado, el proceso de preparación se acerca a su etapa final. Lo que sigue es un paso indispensable, una etapa crítica que garantiza la calidad final antes de que su trabajo sea compartido con el mundo: la revisión y edición rigurosa.

De acuerdo. Finalicemos este capítulo fundamental sobre la redacción científica abordando la etapa crucial del pulido final: la revisión.

6.8. La Importancia de la Revisión y Edición

Hemos recorrido el camino desde los principios generales hasta la estructura IMRaD, pulido el lenguaje y organizado los elementos visuales y complementarios. Ahora, con lo que podría considerarse la "primera versión completa" de su manuscrito en mano, llegamos a una de las etapas más críticas, un paso que separa un buen manuscrito de uno excelente, o incluso un manuscrito aceptado de uno rechazado: la revisión y edición rigurosa. Consideren la redacción como la construcción de una casa detallada; la revisión y edición son la inspección final y los últimos retoques que garantizan que es sólida, segura y estéticamente agradable antes de abrirla al público.

Ningún primer borrador, por muy cuidadoso y experimentado que sea el autor, está realmente listo para ser publicado. El proceso de revisión no es un mero trámite; es una oportunidad indispensable para perfeccionar su mensaje, corregir errores y asegurar que su investigación se presenta con la máxima claridad y profesionalismo.

¿Por Qué Revisar y Editar? La Garantía de Calidad

Las razones para invertir tiempo y esfuerzo considerable en esta etapa son múltiples y convincentes:

- **Mejorar la Claridad y Precisión:** Una relectura crítica ayuda a identificar frases ambiguas, construcciones confusas o detalles faltantes que pueden ser evidentes para ustedes, pero no para un lector nuevo.
- **Corregir Errores:** Desde erratas tipográficas y faltas de ortografía hasta errores gramaticales o de puntuación (¡esos usos de las comas!), cada error, por pequeño que parezca, puede distraer al lector y minar su credibilidad.
- **Asegurar la Coherencia:** La revisión permite verificar que su argumento fluye lógicamente de principio a fin, que las secciones se conectan sin problemas y que no hay contradicciones internas.
- **Fortalecer Argumentos:** Al revisar, pueden identificar puntos débiles en su argumentación o interpretación y pulir la forma en que presentan sus hallazgos para que su impacto sea máximo.
- **Cumplir con las Normas de la Revista:** La revisión final es esencial para asegurar que el manuscrito cumple con todas las instrucciones específicas de la revista (estilo de citación, formato de tablas/figuras, extensión, etc.).

Etapas y Tipos de Revisión

La revisión es un proceso en capas, y es útil abordarlo desde diferentes perspectivas:

1. **Autorevisión:** El primer revisor siempre son ustedes mismos. Pero para que sea efectiva, necesitan distanciarse del texto. Dejar reposar el manuscrito por un tiempo, incluso unos pocos días si es posible, permite verlo con ojos frescos, como si fuera el trabajo de otra persona. Lean lentamente, en voz alta si ayuda, para detectar frases extrañas o errores que el cerebro pasa por alto en una lectura rápida. Revisen específicamente la lógica del argumento, la precisión de los datos presentados, la consistencia en la terminología y el cumplimiento de las instrucciones de la revista.

2. **Revisión por Pares (Informal):** Pedir a colegas de confianza, idealmente aquellos con experiencia en el tema o en redacción científica, que lean su manuscrito antes de enviarlo a una revista es invaluable. Ojos frescos y expertos pueden identificar problemas científicos, metodológicos o de claridad que ustedes no vieron. Aborden la revisión no como una crítica a su trabajo, sino como una colaboración esencial para mejorarlo.

3. **Edición Profesional (Opcional):** Para autores no nativos del idioma de publicación, o para manuscritos de alta complejidad donde la precisión lingüística es crítica, considerar una corrección de estilo o una corrección de pruebas profesional puede marcar una gran diferencia. Un corrector de estilo se enfoca en la fluidez, la claridad, la concisión y el cumplimiento de las normas de estilo, mientras que un corrector de pruebas se centra en errores tipográficos, gramaticales, de puntuación y formato de último minuto.

La importancia de una revisión exhaustiva es un mantra en el mundo editorial científico, una práctica universalmente recomendada en las guías de publicación y destacada en las fuentes consultadas. Un manuscrito bien revisado y editado refleja profesionalismo, incrementa sus posibilidades de ser aceptado y asegura que su valiosa investigación sea comunicada de manera efectiva a la comunidad científica.

Una vez superada esta etapa de revisión y pulido, el manuscrito está listo para emprender su camino formal hacia la comunidad científica. La preparación final para la sumisión, la interacción con los editores y revisores, y los procesos editoriales son los pasos siguientes en la vida de su publicación.

Conclusiones

1. **Comunicación Esencial:** La comunicación científica escrita eficaz, basada en la claridad, precisión y objetividad, es fundamental para compartir hallazgos y contribuir al conocimiento.

2. **Estructura Estándar (IMRaD):** El formato IMRaD (Introducción, Métodos, Resultados y Discusión) es la estructura estándar para artículos de investigación original, facilitando la lectura y comprensión al seguir la lógica del proceso de investigación.

3. **Propósito de Cada Sección:** Cada parte del artículo (título, resumen, palabras clave, introducción, métodos, resultados, discusión, referencias, tablas/figuras) tiene un propósito específico y requiere un enfoque de redacción particular.

4. **Importancia del Estilo y Lenguaje:** La precisión terminológica, la gramática correcta, la puntuación adecuada, la ortografía impecable y el uso apropiado de tiempos verbales son cruciales para la credibilidad y claridad del manuscrito.

5. **Visualización y Referenciación:** Las tablas y figuras deben usarse para complementar el texto, ser autoexplicativas y estar correctamente referenciadas. La citación rigurosa y una lista de referencias completa son indispensables por razones éticas y para la verificabilidad.

6. **Revisión Indispensable:** Ningún primer borrador es perfecto. La revisión y edición exhaustivas (por el propio autor y, preferiblemente, por colegas) son pasos críticos para corregir errores, mejorar la claridad, asegurar la coherencia y aumentar las posibilidades de publicación y el impacto del trabajo.

Parte IV

El Viaje de los Datos en Salud: de la Recolección al Conocimiento

CAPÍTULO 18

FUNDAMENTOS DEL CICLO DE VIDA DE LOS DATOS

Data Lifecycle Fundamentals

Resumen

Este capítulo introduce los fundamentos del manejo de datos en salud, comenzando con el modelo DIKW que describe la transformación de Datos brutos en Información contextualizada, luego en Conocimiento interpretado y finalmente en Sabiduría para la toma de decisiones. Explora los tipos de análisis aplicables: Descriptivo (qué pasó), Predictivo (qué podría pasar) y Prescriptivo (qué hacer). Subraya la importancia crítica de los datos para mejorar la atención al paciente, la investigación, la salud pública y la gestión operativa, destacando su rol central en la Mejora de la Calidad (QI). Detalla el modelo de QI, incluyendo las tres preguntas fundamentales (objetivo, mediciones, cambios), el ciclo PDSA (*Plan-Do-Study-Act*) para probar cambios, la importancia de un *Aim Statement* efectivo, los tipos de mediciones (resultado, proceso, balance) y el benchmarking. Clasifica los tipos de datos (cualitativos y cuantitativos discretos/continuos), enfatiza la necesidad de asegurar la Calidad de los Datos (precisión, completitud, consistencia, etc.) y aborda la importancia de la Gobernanza de Datos y la Ética (privacidad, consentimiento). Finalmente, describe las etapas del Ciclo de Vida de los Datos, desde la planificación hasta su disposición final.

Palabras Clave: Ciclo de vida de los datos, modelo DIKW, análisis de datos en salud, calidad de datos, gobernanza de datos, mejora de la calidad.

Abstract

This chapter introduces the fundamentals of data management in health, starting with the DIKW model which describes the transformation of raw Data into contextualized Information, then into interpreted Knowledge, and finally into Wisdom for decision-making. It explores the types of applicable analyses: Descriptive (what happened), Predictive (what might happen), and Prescriptive

(what to do). It underscores the critical importance of data for improving patient care, research, public health, and operational management, highlighting its central role in Quality Improvement (QI). It details the QI model, including the three fundamental questions (aim, measures, changes), the PDSA (Plan-Do-Study-Act) cycle for testing changes, the importance of an effective Aim Statement, the types of measures (outcome, process, balancing), and benchmarking. It classifies data types (qualitative and quantitative discrete/continuous), emphasizes the need to ensure Data Quality (accuracy, completeness, consistency, etc.), and addresses the importance of Data Governance and Ethics (privacy, consent). Finally, it describes the stages of the Data Lifecycle, from planning to final disposition.

Keywords: Data lifecycle, DIKW model, health data analytics, data quality, data governance, quality improvement.

Introducción

Hemos explorado los métodos para diseñar estudios, analizar la evidencia de estudios individuales y sintetizarla a través de revisiones sistemáticas y metaanálisis. Todas estas actividades, en la era moderna, se basan fundamentalmente en los **datos**. Para aprovechar el poder de la información en salud, es esencial comprender qué son los datos, de dónde provienen, cómo se transforman y cómo se utilizan para generar conocimiento y guiar decisiones.

Este capítulo introduce los conceptos fundamentales del ciclo de vida de los datos en salud, desde su origen hasta su aplicación en la práctica clínica y la investigación.

En esencia, los **datos** son hechos brutos o mediciones sin procesar. Por sí solos, pueden tener poco significado. Sin embargo, a través de un proceso de transformación y análisis, los datos pueden convertirse en información valiosa, luego en conocimiento aplicable y, en última instancia, guiar la sabiduría en la toma de decisiones.

1.1. El Modelo DIKW (Datos → Información → Conocimiento → Sabiduría)

Este modelo describe una jerarquía conceptual que ilustra cómo los datos se convierten en insumos útiles para la toma de decisiones:

- **Datos (*Data*):** Símbolos o hechos brutos, discretos y objetivos, sin contexto ni interpretación. Son las unidades más básicas (ej. Una lectura de presión arterial: "120/80", un resultado de laboratorio: "Glucosa: 105 mg/dL", un registro de fecha: "2023-10-27").
- **Información (*Information*):** Datos que han sido procesados, organizados y estructurados dentro de un contexto. La información responde a preguntas básicas como "¿quién?", "¿qué?", "¿dónde?" y "¿cuándo?". (Ej. "La presión arterial del paciente X el 27 de octubre de 2023 fue 120/80 mmHg", "El nivel de glucosa en ayunas del paciente Y fue 105 mg/dL").
- **Conocimiento (*Knowledge*):** Información que ha sido interpretada, relacionada y comprendida. El conocimiento responde a la pregunta "¿cómo?". Implica la aplicación de la información, a menudo a través de la identificación de patrones, relaciones y tendencias. (Ej. "Un nivel de glucosa en ayunas de 105 mg/dL en el paciente Y sugiere prediabetes, según las guías clínicas").
- **Sabiduría (*Wisdom*):** La aplicación del conocimiento y la experiencia para tomar decisiones informadas y éticas. La sabiduría responde a la pregunta "¿por qué?" y "¿cuál es el mejor curso de acción?". Implica juicio, intuición y la consideración de valores. (Ej. "Basado en el conocimiento de que el paciente Y tiene prediabetes y considerando sus factores de riesgo y preferencias, el mejor curso de acción es recomendar cambios en el estilo de vida y monitoreo regular, en lugar de iniciar medicación de inmediato").

El trabajo con datos en salud implica ascender en esta jerarquía, transformando hechos brutos en acciones que mejoran la salud.

1.2. Tipología de Análisis

El tipo de análisis aplicado a los datos determina el tipo de ideas clave que podemos obtener y su utilidad para la toma de decisiones. Los análisis se pueden clasificar en general en:

- **Análisis Descriptivo:** Responde a la pregunta "¿Qué pasó?". Se centra en resumir y describir las características principales de un conjunto de datos. Incluye estadísticas descriptivas (medias, medianas, desviaciones estándar, frecuencias) y visualizaciones (gráficos, tablas). (Ej. "El 30% de los pacientes en nuestro hospital fueron readmitidos dentro de los 30 días", "La edad promedio de los participantes en el estudio fue de 65 años").

- **Análisis Predictivo:** Responde a la pregunta "¿Qué podría pasar?". Utiliza datos históricos para hacer predicciones sobre eventos futuros. Emplea modelos estadísticos y de aprendizaje automático para identificar patrones y construir modelos predictivos. (Ej. "Este paciente tiene un 70% de probabilidad de ser readmitido en los próximos 30 días basado en sus características", "Predecir qué pacientes responderán mejor a un tratamiento particular").
- **Análisis Prescriptivo:** Responde a la pregunta "¿Qué deberíamos hacer?". Va más allá de la predicción para recomendar cursos de acción óptimos. Considera posibles acciones, sus resultados predichos y sus implicaciones, a menudo utilizando técnicas de optimización y simulación. (Ej. "Dado el riesgo de readmisión de este paciente, la mejor intervención para reducir ese riesgo es una llamada de seguimiento telefónico dentro de las 48 horas post-alta").

En salud/medicina, trabajar con datos es fundamental para:

- Mejorar la atención al paciente: Personalizar tratamientos, predecir riesgos (ej. De enfermedades, de complicaciones), optimizar la asignación de recursos.
- Avanzar en la investigación: Identificar factores de riesgo, evaluar la efectividad de intervenciones, descubrir nuevos biomarcadores.
- Optimizar la gestión de la salud pública: Monitorear brotes de enfermedades, planificar campañas de vacunación, evaluar el impacto de políticas de salud.
- Mejorar la eficiencia operativa: Optimizar flujos de trabajo en hospitales, reducir costos, mejorar la seguridad del paciente.
- Mejora de la Calidad (*Quality Improvement*): Utilizar datos (mediciones) para identificar áreas de mejora en los sistemas y procesos de atención médica, monitorear el impacto de los cambios implementados y asegurar que la atención sea segura, efectiva, centrada en el paciente, oportuna, eficiente y equitativa.

1.3. Datos y Mejora de la Calidad

La mejora continua de la calidad de la atención médica es un pilar fundamental en los sistemas de salud. Este proceso se basa intrínsecamente en el

uso de datos para entender el rendimiento actual, identificar oportunidades de mejora y evaluar el impacto de los cambios implementados.

Objetivos de la Mejora de la Calidad:

Los esfuerzos de mejora de la calidad buscan responder a preguntas fundamentales y lograr metas específicas, como:

- Formular las tres preguntas fundamentales del modelo de mejora de la calidad.
- Conocer los pasos para redactar un *Aim Statement* efectivo.
- Identificar *process*, *outcome* y *balancing measures*.
- Comprender la relevancia del *benchmarking* en salud.
- Explicar el ciclo PDSA y su utilidad para testar cambios.

Mejora de la Calidad vs. Auditoría Clínica:

Aunque ambos buscan mejorar la atención, difieren en su enfoque. La **Auditoría Clínica** compara la práctica actual con estándares predefinidos (a menudo basados en guías clínicas) en un punto fijo en el tiempo, identificando brechas. La **Mejora de la Calidad** es un proceso continuo que implica probar cambios a pequeña escala, medir su impacto y refinar el proceso, buscando la mejora incremental en los sistemas y procesos. Si bien la auditoría puede identificar la necesidad de mejora, la mejora de la calidad proporciona el marco para lograrla. Ambos enfoques son complementarios y utilizan datos.

El Modelo para la Mejora de la Calidad:

Un modelo común para la mejora de la calidad comienza con tres preguntas fundamentales que guían el proceso y resaltan el papel central de los datos:

1. **Objetivo (*Aim*):** ¿Qué estamos tratando de lograr? (Define la meta específica de mejora).
2. **Mediciones (*Measures*):** ¿Cómo sabremos si un cambio es una mejora? (Aquí es donde los **datos** son esenciales. Necesitamos medir el estado actual y el impacto de los cambios).
3. **Cambios (*Changes*):** ¿Qué cambios podemos hacer que resultarán en una mejora? (Implica idear intervenciones basadas en el conocimiento y la comprensión del sistema).

Una vez que se han respondido estas preguntas, el proceso avanza a través de ciclos de prueba y aprendizaje.

1.4. El Ciclo PDSA (Plan-Do-Study-Act)

El ciclo Planificar-Hacer-Estudiar-Actuar (PDSA) es la herramienta fundamental en la mejora de la calidad para probar cambios a pequeña escala. Es un ciclo iterativo que se basa en la recolección y análisis de datos:

- **Planificar (*Plan*):** Planificar el cambio a probar, predecir el resultado y planificar la recolección de datos.
- **Hacer (*Do*):** Implementar el cambio a pequeña escala y **recolectar los datos** según lo planeado.
- **Estudiar (*Study*): Analizar los datos** recolectados, comparar los resultados con las predicciones e identificar aprendizajes.
- **Actuar (Act):** Basado en los aprendizajes, decidir el siguiente paso: adoptar el cambio, adaptarlo o abandonarlo, y planificar el próximo ciclo PDSA.

Este ciclo demuestra claramente cómo los datos son el motor del aprendizaje y la adaptación en el proceso de mejora.

1.5. Redactando un Aim Statement Efectivo

Un *Aim Statement* (Declaración de Objetivo) define claramente lo que se busca lograr. Debe ser específico, medible, alcanzable, relevante y con un plazo definido (SMART, por sus siglas en inglés). Incluye:

- El grado de mejora esperado.
- El plazo para lograr la mejora.
- La población o el sistema afectado.

1.6. Tipos de Mediciones en Mejora de la Calidad

Para saber si un cambio ha llevado a una mejora, se utilizan diferentes tipos de mediciones, que son ejemplos concretos de los **datos** recolectados en proyectos de mejora:

- **Mediciones de Resultado (*Outcome Measures*):** Miden el resultado final que se busca mejorar (ej. Tasa de infecciones asociadas a la atención, porcentaje de pacientes con presión arterial controlada, tiempo de espera promedio).
- **Mediciones de Proceso (*Process Measures*):** Miden si se están siguiendo los pasos del proceso que se espera que conduzcan al resultado deseado (ej. Porcentaje de pacientes a los que se les midió la presión arterial dos veces al año, porcentaje de prescripciones que cumplen un protocolo, tiempo para administrar un antibiótico).
- **Mediciones de Balance (*Balancing Measures*):** Miden si los cambios realizados para mejorar un área están causando problemas no deseados en otras partes del sistema (ej. Si reducir el tiempo de espera en una clínica aumenta la carga de trabajo en el laboratorio, si un nuevo protocolo de alta aumenta las llamadas de pacientes al día siguiente).

1.7. Benchmarking en Salud

El *benchmarking* es el proceso de comparar el rendimiento de un proceso o servicio con el de organizaciones líderes o con estándares de la industria. En salud, implica comparar métricas de calidad o eficiencia con otras instituciones similares o con los mejores en su clase. Esto requiere la recolección y análisis de **datos** de rendimiento para identificar brechas y áreas de oportunidad para la mejora.

En conclusión, la Mejora de la Calidad es una aplicación fundamental de los datos en salud. Utiliza mediciones (datos) para entender el rendimiento de los sistemas, guiar la prueba de cambios a través de ciclos PDSA y asegurar que los esfuerzos de mejora conduzcan a resultados positivos para los pacientes y las organizaciones de salud.

1.8. Tipos de Datos

Los datos pueden clasificarse de diversas maneras, lo que influye en los métodos de análisis apropiados:

- **Datos Cualitativos:** Representan características o atributos que no pueden medirse numéricamente. Describen cualidades o categorías.

- *Ejemplos*: Sexo (masculino/femenino), grupo sanguíneo (A, B, AB, O), estado civil, raza, opiniones de un paciente.

- **Datos Cuantitativos:** Representan mediciones numéricas.

- **Datos Cuantitativos Discretos:** Toman valores enteros y generalmente representan conteos.

- *Ejemplos*: Número de hijos, número de hospitalizaciones, número de camas en un hospital.

- **Datos Cuantitativos Continuos:** Pueden tomar cualquier valor dentro de un rango dado.

- *Ejemplos*: Presión arterial, peso, altura, temperatura corporal, niveles de glucosa en sangre.

Comprender el tipo de datos es crucial para seleccionar las técnicas estadísticas y de visualización adecuadas.

1.9. Calidad de los Datos

La utilidad de los datos para generar información y conocimiento depende críticamente de su calidad. Los datos de baja calidad pueden llevar a análisis erróneos y decisiones incorrectas. Aspectos clave de la calidad de los datos incluyen:

- **Precisión:** Los datos reflejan la realidad de forma correcta.
- **Completitud:** No faltan valores importantes.
- **Consistencia:** Los datos son coherentes entre sí y a lo largo del tiempo.
- **Actualidad (*Timeliness*):** Los datos están disponibles cuando se necesitan y son relevantes para el período de tiempo actual.
- **Validez:** Los datos miden lo que se supone que deben medir (distinto de la validez de una prueba diagnóstica, que se refiere a su capacidad para clasificar correctamente).

Asegurar la calidad de los datos requiere procesos rigurosos de recolección, almacenamiento y limpieza.

1.10. Gobernanza de Datos y Ética

Trabajar con datos de salud implica una gran responsabilidad. La gobernanza de datos establece políticas y procedimientos para gestionar los datos a lo largo de su ciclo de vida, asegurando su disponibilidad, usabilidad, integridad y seguridad. La ética es fundamental, especialmente en lo que respecta a la privacidad del paciente, la confidencialidad de la información y el consentimiento informado para el uso de datos. Cumplir con regulaciones como HIPAA (en EE. UU.) o GDPR (en Europa) es obligatorio.

1.11. El Ciclo de Vida de los Datos

Los datos en salud pasan por varias etapas a lo largo de su existencia:

1. **Planificación:** Definir qué datos se necesitan, cómo se recolectarán, almacenarán y utilizarán.
2. **Recolección:** Adquirir los datos de diversas fuentes.
3. **Almacenamiento:** Guardar los datos de forma segura y organizada (ej. Bases de datos).
4. **Procesamiento:** Limpiar, transformar y preparar los datos para el análisis.
5. **Análisis:** Aplicar métodos estadísticos y computacionales para extraer información y conocimiento.
6. **Archivado/Disposición:** Retener los datos según las políticas y regulaciones, o eliminarlos de forma segura cuando ya no son necesarios.

Comprender este ciclo ayuda a gestionar los datos de manera efectiva y a garantizar que estén disponibles y sean utilizables para la investigación y la toma de decisiones.

Conclusiones

1. **Jerarquía del Valor:** Los datos crudos se transforman secuencialmente en información, conocimiento y sabiduría (Modelo DIKW), incrementando su valor para la toma de decisiones en cada paso.
2. **Tipos de Análisis:** El análisis de datos puede ser descriptivo (resumir el pasado), predictivo (anticipar el futuro) o prescriptivo (recomendar acciones), cada uno con un propósito distinto.

3. **Datos en Salud:** Los datos son cruciales en múltiples ámbitos de la salud, incluyendo la mejora de la atención, la investigación, la salud pública y la gestión.
4. **Mejora de la Calidad (QI) Basada en Datos:** La QI depende intrínsecamente de los datos para definir objetivos (*Aim Statement*), medir el rendimiento (*Outcome*, *Process*, *Balancing Measures*), comparar con los mejores (Benchmarking) y probar cambios de forma iterativa (Ciclo PDSA).
5. **Tipología y Calidad de Datos:** Comprender si los datos son cualitativos o cuantitativos (discretos/continuos) es clave para el análisis. La calidad de los datos (precisión, completitud, consistencia, actualidad, validez) es esencial para la fiabilidad de los resultados.
6. **Gobernanza y Ética:** El manejo de datos de salud requiere una gobernanza robusta y un estricto cumplimiento de principios éticos, asegurando la privacidad, confidencialidad y uso adecuado de la información.
7. **Ciclo de Vida:** Los datos siguen un ciclo de vida que incluye planificación, recolección, almacenamiento, procesamiento, análisis y disposición final, requiriendo gestión en cada etapa.

CAPÍTULO 19

RECOLECCIÓN Y DISEÑO DE DATOS EN SALUD

Health Data Collection and Design

Resumen

Este capítulo enfatiza que la calidad de la recolección y diseño de datos es crucial para la validez de cualquier análisis posterior en salud. Explora las diversas fuentes de datos disponibles, incluyendo Historias Clínicas Electrónicas (HCE), registros administrativos, encuestas/cuestionarios, sensores/wearables, registros de enfermedades, datos de salud pública y datos genómicos/moleculares. Distingue entre datos primarios (recolectados específicamente para el estudio, mayor control y relevancia) y datos secundarios (reutilización de datos existentes, más rápidos y económicos pero con posibles limitaciones de calidad y adecuación). Discute la importancia de seleccionar el instrumento de recolección adecuado (formularios, escalas, HCE) y la necesidad fundamental de la validación psicométrica/clinimétrica para asegurar que los instrumentos midan de forma válida y fiable, especialmente para constructos abstractos. Detalla las buenas prácticas de recolección de datos, como protocolos estandarizados, capacitación del personal, control de calidad, documentación rigurosa, manejo de datos faltantes y protección de la privacidad/seguridad. Se ilustra la aplicación de estos principios con un ejemplo práctico sobre la investigación del tabaquismo en estudiantes universitarios mediante encuestas (datos primarios).

Palabras Clave: Recolección de datos, fuentes de datos en salud, datos primarios, datos secundarios, validación de instrumentos, calidad de datos.

Abstract

This chapter emphasizes that the quality of data collection and design is crucial for the validity of any subsequent health analysis. It explores diverse available data sources, including Electronic Health Records (EHR), administrative records, surveys/questionnaires, sensors/wearables, disease registries, public health data,

and genomic/molecular data. It distinguishes between primary data (collected specifically for the study, offering greater control and relevance) and secondary data (reusing existing data, faster and cheaper but with potential limitations in quality and suitability). The importance of selecting the appropriate data collection instrument (CRFs, scales, EHR interface) and the fundamental need for psychometric/clinimetric validation to ensure instruments measure validly and reliably, especially for abstract constructs, is discussed. Good data collection practices are detailed, such as standardized protocols, staff training, quality control, rigorous documentation, handling missing data, and protecting privacy/security. The application of these principles is illustrated with a practical example of researching smoking habits in university students using surveys (primary data).

Keywords: *Data collection, health data sources, primary data, secondary data, instrument validation, data quality.*

Introducción

Hemos sentado las bases conceptuales sobre qué son los datos y cómo se transforman en conocimiento. Ahora, nos adentramos en el origen de todo: la **recolección de datos**. La forma en que se diseñan y recolectan los datos en salud es un paso crítico que determina la calidad y la utilidad de cualquier análisis posterior. Un análisis sofisticado sobre datos deficientes solo producirá resultados deficientes.

La recolección de datos en el ámbito de la salud es un proceso complejo y multifacético. Implica identificar las fuentes de información relevantes, seleccionar los métodos e instrumentos adecuados, y asegurar que el proceso se lleve a cabo de manera rigurosa para garantizar la precisión e integridad de los datos.

2.1. Fuentes de Datos en Salud

Los datos que impulsan la investigación y la toma de decisiones en salud provienen de una diversidad de fuentes, cada una con sus propias características, fortalezas y limitaciones:

- **Historia Clínica Electrónica (HCE) / *Electronic Health Records* (EHR):** Son quizás la fuente más rica de datos clínicos detallados a nivel individual. Capturan información sobre visitas médicas, diagnósticos (codificados y en texto libre), procedimientos, medicamentos recetados y administrados, resultados de laboratorio, informes de radiología, notas de progreso, datos demográficos del paciente y más. Las HCE son

valiosas para la investigación retrospectiva y para obtener una visión longitudinal de la atención del paciente.

- **Registros Administrativos y de Facturación:** Estos datos se generan como parte de los procesos de gestión de los sistemas de salud y el reembolso. Incluyen información sobre admisiones hospitalarias, altas, transferencias, procedimientos facturados y servicios utilizados. Son útiles para análisis de utilización de servicios, costos, acceso a la atención y patrones de derivación, a menudo a gran escala poblacional.
- **Encuestas y Cuestionarios:** Permiten recolectar datos directamente de los individuos, incluyendo información que no suele estar disponible en registros clínicos o administrativos, como síntomas reportados por el paciente (PRO), calidad de vida, comportamientos de salud (dieta, ejercicio, tabaquismo), percepciones, actitudes y satisfacción con la atención. Pueden administrarse en persona, por teléfono, correo o electrónicamente.
- **Sensores y Dispositivos Portátiles (*Wearables*):** Una fuente de datos cada vez más importante, que proporciona mediciones fisiológicas continuas o frecuentes fuera del entorno clínico tradicional. Incluyen datos de actividad física, patrones de sueño, ritmo cardíaco, variabilidad de la frecuencia cardíaca, niveles de glucosa (monitores continuos), e incluso parámetros respiratorios. Son valiosos para estudios de comportamiento en salud y monitoreo remoto.
- **Registros de Enfermedades (*Disease Registries*):** Bases de datos especializadas que recopilan información detallada sobre pacientes con una condición específica (ej. Cáncer, enfermedades raras, enfermedades cardiovasculares). Suelen tener un enfoque más profundo en la enfermedad de interés que las HCE generales y son fundamentales para estudios de historia natural de la enfermedad, pronóstico y evaluación de tratamientos a largo plazo.
- **Datos de Salud Pública y Vigilancia:** Incluyen estadísticas vitales (nacimientos, defunciones), datos de enfermedades notificables, resultados de programas de detección poblacional y datos ambientales. Son esenciales para monitorear la salud de la población, identificar brotes y evaluar el impacto de políticas de salud pública.

- **Datos Genómicos y Moleculares:** Información sobre la composición genética de un individuo, la expresión génica, perfiles proteómicos o metabolómicos. Cruciales para la medicina de precisión, la identificación de factores de riesgo genéticos y la comprensión de los mecanismos de la enfermedad.

2.2. Datos Primarios vs. Secundarios

Una distinción fundamental en la recolección de datos es si los datos se recopilan específicamente para el propósito de la investigación actual o si se reutilizan datos existentes:

- **Datos Primarios:** Son aquellos datos que el investigador o equipo de investigación recolecta directamente para responder a una pregunta de investigación específica. El investigador tiene control sobre el diseño del estudio, los instrumentos de recolección, la población de estudio y el proceso de recolección.
  - *Ventajas:* Alta relevancia para la pregunta de investigación, mayor control sobre la calidad de los datos, posibilidad de recolectar datos específicos no disponibles en otras fuentes.
  - *Desventajas:* Generalmente más costosos y requieren más tiempo y recursos para recolectar.
  - *Ejemplos:* Datos de un ensayo clínico, datos de una encuesta diseñada específicamente para un estudio, mediciones de laboratorio realizadas para un proyecto de investigación.
- **Datos Secundarios:** Son datos que ya han sido recolectados para un propósito diferente (ej. Atención clínica, administración, facturación) y que son reutilizados para la investigación. El investigador no tuvo control sobre el diseño de la recolección original.
  - *Ventajas:* Generalmente más rápidos y económicos de obtener, permiten acceder a grandes volúmenes de datos y estudiar poblaciones amplias o eventos raros.
  - *Desventajas:* Pueden no ser perfectamente alineados con la pregunta de investigación, la calidad de los datos puede ser variable y difícil de verificar, pueden faltar datos importantes, y la información puede estar sesgada por el propósito original de la recolección.
  - *Ejemplos:* Datos de HCE, registros administrativos, datos de registros de enfermedades, datos de reclamaciones de seguros.

La elección entre datos primarios y secundarios depende de la pregunta de investigación, los recursos disponibles y la factibilidad. A menudo, se utilizan ambos tipos en enfoques de investigación mixtos.

2.3. Selección del Instrumento Adecuado

El "instrumento" de recolección de datos es la herramienta o método utilizado para capturar la información. La selección del instrumento adecuado es crucial para obtener datos precisos y relevantes. Los instrumentos deben ser apropiados para el tipo de datos, la fuente y la población de estudio.

- **Formularios de Recolección de Datos (CRF):** Estructuras estandarizadas utilizadas en investigación (especialmente en ensayos clínicos) para documentar la información de cada participante. Deben ser claros, concisos y diseñados para capturar todos los datos necesarios de manera uniforme.
- **Escalas de Medición y Cuestionarios:** Utilizados para recolectar datos reportados por pacientes, evaluaciones de síntomas, calidad de vida, etc. La elección del cuestionario debe basarse en su relevancia para la pregunta de investigación y, fundamentalmente, en su **validación**.
- **Sistemas de Historia Clínica Electrónica (HCE):** Aunque son una fuente de datos, la interfaz de entrada de datos dentro de la HCE actúa como un instrumento. El diseño de los campos de entrada y la estructura de la base de datos subyacente influyen en la calidad y la estructura de los datos secundarios disponibles.
- **Interfaces de Dispositivos/Sensores:** El software o hardware que captura datos de dispositivos médicos o wearables.

Validación Psicométrica/Clinimétrica

Cuando se utilizan cuestionarios o escalas para medir constructos abstractos (como dolor, fatiga, ansiedad, calidad de vida), es esencial que estos instrumentos hayan sido **validados**. La validación es el proceso de asegurar que un instrumento mide lo que pretende medir (validez) de manera consistente (fiabilidad/confiabilidad).

- **Validación Psicométrica:** Se aplica a instrumentos que miden constructos psicológicos.

- **Validación Clinimétrica:** Se aplica a instrumentos que miden constructos relacionados con la salud o la clínica.

Los aspectos clave de la validación incluyen:

- **Validez:** ¿El instrumento mide el constructo correcto? Tipos de validez incluyen validez de contenido (¿cubre todos los aspectos relevantes?), validez de constructo (¿se relaciona con otros constructos teóricos como se espera?) y validez relacionada con criterios (¿predice un resultado futuro o se correlaciona con un estándar de oro?).
- **Fiabilidad (Confiabilidad):** ¿El instrumento produce resultados consistentes? Tipos de fiabilidad incluyen consistencia interna (¿los ítems dentro de una escala miden lo mismo?), fiabilidad test-retest (¿los resultados son consistentes a lo largo del tiempo en ausencia de cambios reales?) y fiabilidad inter-observador (¿diferentes evaluadores obtienen resultados similares?).

Utilizar instrumentos validados es crucial para la calidad y la interpretabilidad de los datos recolectados, especialmente para datos reportados por pacientes o evaluaciones clínicas subjetivas.

2.4. Buenas Prácticas de Recolección

Independientemente de la fuente o el instrumento, la implementación de buenas prácticas de recolección es fundamental para minimizar errores y sesgos, y asegurar la integridad de los datos:

- **Protocolos Estandarizados:** Desarrollar y seguir protocolos detallados que especifiquen exactamente cómo se deben recolectar los datos, quién es responsable, cuándo y dónde. Esto asegura la uniformidad en el proceso de recolección entre diferentes recolectores y a lo largo del tiempo.
- **Capacitación Rigurosa del Personal:** Todo el personal involucrado en la recolección de datos debe recibir una capacitación exhaustiva sobre el protocolo del estudio, el uso de los instrumentos y los procedimientos de control de calidad.
- **Control de Calidad y Aseguramiento de la Integridad:** Implementar mecanismos para verificar la precisión, completitud y consistencia de los datos en el momento de la recolección y posteriormente. Esto puede incluir:

- **Validaciones en la Entrada de Datos:** Utilizar sistemas electrónicos con validaciones integradas para prevenir errores de formato, rango o consistencia en el momento en que se ingresan los datos.
- **Doble Entrada de Datos:** Para datos recolectados en papel, ingresar los datos dos veces por personas diferentes y comparar las entradas para identificar errores.
- **Monitoreo de Datos:** Revisar regularmente una muestra de los datos recolectados para verificar su precisión y adherencia al protocolo.
- **Consultas de Datos:** Generar consultas para identificar inconsistencias o valores atípicos que requieran verificación con la fuente original.

- **Documentación Detallada:** Mantener registros claros de todo el proceso de recolección, incluyendo cualquier desviación del protocolo, problemas encontrados y cómo se resolvieron.
- **Manejo de Datos Faltantes:** Establecer procedimientos claros para documentar por qué los datos faltan y cómo se manejarán durante el análisis.
- **Seguridad y Privacidad:** Implementar medidas técnicas y administrativas para proteger los datos de salud sensibles, cumpliendo con las regulaciones de protección de datos.

Un diseño de recolección de datos cuidadoso y la adherencia a buenas prácticas son la piedra angular de cualquier investigación o análisis de datos en salud fiable. Sentar estas bases sólidas es esencial antes de pasar a las etapas de pretratamiento y análisis.

Hemos explorado los conceptos teóricos detrás de la recolección y el diseño de datos en salud. Ahora, para solidificar estos conocimientos, vamos a trabajar a través de un ejemplo práctico. Imagina que eres parte de un equipo de investigación encargado de abordar una pregunta de salud específica. ¿Cómo aplicarías los principios que acabamos de discutir?

2.5. Ejemplo Práctico: Investigación sobre Tabaquismo en Estudiantes Universitarios

Pregunta de Investigación: Queremos determinar la prevalencia actual del tabaquismo entre los estudiantes universitarios de la ciudad de Cali, e investigar las actitudes que tienen estos estudiantes hacia dejar de fumar.

Esta pregunta nos interesa porque el tabaquismo es un factor de riesgo importante para muchas enfermedades, y entender su prevalencia y las barreras para dejarlo en una población joven puede informar programas de prevención y cesación más efectivos.

A. Identificación de Fuentes de Datos

Para responder a nuestra pregunta, ¿dónde podríamos encontrar la información necesaria?

- **Historias Clínicas Electrónicas (HCE):** Las historias clínicas de los servicios de salud universitarios podrían contener información sobre si un estudiante ha reportado ser fumador. Sin embargo, es probable que esta información no esté completa o estandarizada, y no nos daría detalles sobre actitudes hacia dejar de fumar.
- **Registros Administrativos:** No son una fuente probable para este tipo de información de comportamiento y actitud.
- **Encuestas y Cuestionarios:** Esta parece ser una fuente muy prometedora. Podríamos diseñar un cuestionario específico para preguntar sobre hábitos de tabaquismo (frecuencia, cantidad, edad de inicio) y sobre actitudes, creencias y motivaciones relacionadas con dejar de fumar.
- **Sensores/Wearables:** No son adecuados para medir hábitos de tabaquismo o actitudes.
- **Registros de Enfermedades:** No aplicable a esta pregunta.
- **Datos de Salud Pública:** Las encuestas de salud pública a nivel general podrían tener datos sobre tabaquismo en grupos de edad relevantes, pero no serían específicas para la población universitaria de Cali.

Conclusión sobre Fuentes: Las **encuestas y cuestionarios** administrados directamente a los estudiantes universitarios de Cali son la fuente de datos más adecuada y directa para responder a nuestra pregunta.

B. Datos Primarios vs. Secundarios

Dada nuestra pregunta de investigación específica sobre la prevalencia *actual* y las *actitudes* en una población particular (estudiantes universitarios de Cali), es poco probable que encontremos datos secundarios existentes que sean lo suficientemente detallados, recientes y específicos para nuestra población objetivo.

- **Datos Secundarios:** Podríamos encontrar encuestas generales sobre tabaquismo en jóvenes o adultos jóvenes, pero no serían específicas de universidades en Cali. La información sobre actitudes es aún menos probable que esté disponible en bases de datos secundarias.
- **Datos Primarios:** Recolectar datos directamente de los estudiantes nos permitirá obtener la información exacta que necesitamos, definir precisamente nuestra población de estudio y controlar la calidad de la recolección.

Conclusión sobre Tipo de Datos: Optaremos por recolectar **datos primarios** a través de encuestas.

C. Selección y Diseño del Instrumento Adecuado

Dado que vamos a recolectar datos primarios mediante encuestas, necesitamos diseñar un **cuestionario**. Este cuestionario será nuestro instrumento de recolección de datos.

El cuestionario debe incluir preguntas para medir:

- **Variables Demográficas:** Edad, sexo, carrera universitaria, año de estudio.
- **Variables sobre Hábito de Tabaquismo:**
  - ¿Fuma actualmente? (Sí/No) – Si sí:
    - ¿Cuántos cigarrillos fuma al día?
    - ¿A qué edad empezó a fumar?
    - ¿Fuma a diario o solo ocasionalmente?
  - Si no fuma actualmente: ¿Ha fumado alguna vez? ¿Cuándo dejó de fumar?
- **Variables sobre Actitudes hacia Dejar de Fumar:**
  - ¿Ha intentado dejar de fumar alguna vez?
  - ¿Le gustaría dejar de fumar? (Sí/No/No estoy seguro)
  - Si le gustaría dejar de fumar: ¿Qué tan importante es para usted dejar de fumar en este momento? (Escala, ej. 1-10)

- ¿Qué tan seguro(a) está de poder dejar de fumar si lo intentara? (Escala, ej. 1-10)
- ¿Cuáles considera que son las principales barreras para que usted deje de fumar? (Pregunta abierta o lista de opciones: estrés, amigos que fuman, miedo a engordar, falta de apoyo, etc.)
- ¿Qué tipo de apoyo le sería útil si decidiera dejar de fumar? (Lista de opciones: terapia individual, terapia grupal, medicamentos, apps, apoyo de amigos/familia, etc.)

D. Diseño del Cuestionario:

- Las preguntas deben ser claras, concisas y fáciles de entender.
- Evitar preguntas ambiguas o sesgadas.
- Utilizar formatos de respuesta apropiados (opción múltiple, escalas Likert, preguntas abiertas).
- Organizar las preguntas de manera lógica.
- Considerar la longitud del cuestionario para evitar la fatiga del respondedor.
- Decidir el formato de administración (en línea, en papel, asistido por entrevistador). Un cuestionario en línea podría ser eficiente para llegar a muchos estudiantes.

E. Validación del Instrumento

Dado que estamos midiendo constructos como "actitudes hacia dejar de fumar", que no son directamente observables, es crucial que las secciones del cuestionario que miden estas actitudes hayan sido **validadas** previamente en poblaciones similares. Si no existen escalas validadas que se ajusten a nuestras necesidades, tendríamos que embarcarnos en un proceso de validación de nuestra propia escala, lo cual es un proceso complejo (validación psicométrica). Para este ejemplo, asumiremos que buscamos o adaptamos escalas validadas existentes para medir las actitudes.

F. Buenas Prácticas de Recolección

Para asegurar que los datos recolectados a través de nuestro cuestionario sean de alta calidad:

- **Protocolo Estandarizado:** Desarrollar un protocolo detallado que describa cómo se seleccionarán los estudiantes (ej. Muestreo aleatorio en diferentes facultades), cómo se les invitará a participar, cómo se administrará el cuestionario (ej. Enlace en línea, instrucciones claras), cómo se manejarán las preguntas de los participantes y cómo se almacenarán los datos de forma segura.
- **Capacitación:** Si hay personal involucrado en la administración (ej. Ayudantes de investigación que invitan a participar), deben estar capacitados en el protocolo y en cómo responder a preguntas comunes de los participantes.
- **Control de Calidad:**
  - Si es en línea, utilizar validaciones en el formulario para asegurar que las respuestas estén en el formato correcto (ej. Solo números para cantidad de cigarrillos) y dentro de rangos lógicos.
  - Monitorear la tasa de respuesta y la completitud de los cuestionarios.
  - Revisar una muestra de los datos recolectados para identificar patrones de respuesta inconsistentes.
- **Manejo de Datos Faltantes:** El protocolo debe especificar qué hacer si un estudiante no responde ciertas preguntas.
- **Documentación:** Registrar cualquier problema técnico con la plataforma en línea, interrupciones en la recolección o cambios en el protocolo.
- **Seguridad y Privacidad:** Asegurar que la plataforma en línea cumpla con estándares de seguridad, que los datos se almacenen de forma anonimizada o pseudonimizada, y que se obtenga el consentimiento informado de los participantes antes de que completen el cuestionario.

Siguiendo estos pasos, desde la formulación de la pregunta hasta la implementación de buenas prácticas de recolección con un instrumento validado, sentaremos las bases para obtener datos de alta calidad que nos permitan responder de manera fiable a nuestra pregunta de investigación sobre el tabaquismo en estudiantes universitarios. Este proceso es una aplicación directa de los conceptos discutidos en el Capítulo 2.

Conclusiones

1. **Base Fundamental:** La calidad del diseño y la recolección de datos determina la fiabilidad y utilidad de cualquier análisis o investigación en salud posterior.
2. **Diversidad de Fuentes:** Existen múltiples fuentes de datos en salud (HCE, registros, encuestas, wearables, etc.), cada una con ventajas y desventajas que deben considerarse según la pregunta de investigación.
3. **Primario vs. Secundario:** La elección entre recolectar datos primarios (mayor control, más costoso) o usar datos secundarios (más rápido/barato, menor control) depende de los objetivos y recursos del estudio.
4. **Instrumentos y Validación:** Es crucial seleccionar instrumentos de recolección adecuados (formularios, cuestionarios) y, especialmente para medidas subjetivas o constructos, utilizar instrumentos validados (psicométrica/clinimétricamente) que midan con precisión y consistencia.
5. **Buenas Prácticas Esenciales:** La implementación de buenas prácticas (protocolos estandarizados, capacitación, control de calidad, documentación, seguridad) es indispensable para garantizar la integridad, precisión y fiabilidad de los datos recolectados.
6. **Aplicación Práctica:** El diseño de la recolección de datos debe adaptarse a la pregunta específica, seleccionando las fuentes, el tipo de datos (primarios/secundarios) y los instrumentos más apropiados, siempre aplicando buenas prácticas.

CAPÍTULO 20

PRETRATAMIENTO Y PREPARACIÓN DE DATOS

Data Pretreatment and Preparation

Resumen

Este capítulo destaca que los datos recolectados raramente están listos para el análisis y requieren pretratamiento y preparación. Describe este proceso en etapas clave. Primero, la Limpieza de Datos aborda valores faltantes (identificando su tipo y cantidad, y manejándolos mediante eliminación o imputación con media, mediana, moda, interpolación o modelos predictivos como k-NN y MICE), valores atípicos (*outliers*, detectándolos y tratándolos según su causa) y datos ruidosos (suavizándolos con *binning*, regresión o *clustering*). Segundo, las Transformaciones Básicas adaptan los datos para el análisis, incluyendo Escalado y Normalización (Min-Max, Z-score) para variables numéricas, Codificación de Variables Categóricas (numérica/ordinal, *One-Hot Encoding*) para convertirlas a formato numérico, Discretización y Generación de Jerarquía de Conceptos. Tercero, la Preparación según el Modelo ajusta los datos a los requisitos específicos del algoritmo analítico a utilizar (regresión, series temporales, redes neuronales). Se incluye un ejemplo práctico que ilustra la aplicación de estos pasos (manejo de duplicados, imputación de faltantes, codificación y estandarización) en un conjunto de datos hipotético de salud.

Palabras Clave: Pretratamiento de datos, preparación de datos, limpieza de datos, imputación de datos faltantes, transformación de datos, codificación de variables.

Abstract

This chapter highlights that collected data is rarely ready for analysis and requires pretreatment and preparation. It describes this process in key stages. First, Data Cleaning addresses missing values (identifying their type and amount, and handling them through deletion or imputation using mean, median, mode, interpolation, or predictive models like k-NN and MICE), outliers

(detecting and treating them analysis cause), and noisy data (smoothing using binning, regression, or clustering). Second, Basic Transformations adapt data for analysis, including Scaling and Normalization (Min-Max, Z-score) for numerical variables, Encoding Categorical Variables (numerical/ordinal, One-Hot Encoding) to convert them to numerical format, Discretization, and Concept Hierarchy Generation. Third, Preparation according to the Model adjusts data to the specific requirements of the analytical algorithm to be used (regression, time series, neural networks). A practical example illustrating the application of these steps (handling duplicates, imputing missing values, encoding, and standardization) on a hypothetical health dataset is included.

Keywords: *Data preprocessing, data preparation, data cleaning, missing data imputation, data transformation, variable encoding.*

Introducción

Hemos visto cómo los datos brutos son la materia prima y cómo un diseño y recolección cuidadosos son esenciales. Sin embargo, los datos recolectados rara vez están listos para ser analizados directamente. A menudo contienen imperfecciones: valores faltantes, errores, inconsistencias o formatos inadecuados para los métodos analíticos. El **pretratamiento y la preparación de datos** son pasos cruciales para transformar estos datos brutos en un conjunto de datos limpio, estructurado y listo para el análisis. Este proceso, aunque a veces laborioso, es fundamental para garantizar la validez y fiabilidad de los hallazgos.

Este proceso se puede dividir en varias etapas clave:

3.1. Limpieza de Datos

La limpieza de datos se enfoca en manejar los datos "ruidosos" o incompletos, identificando y corrigiendo errores e inconsistencias.

- **Detección y Manejo de Valores Faltantes:** Los valores faltantes son un problema común en los datos de salud (ej. Un paciente no respondió una pregunta en una encuesta, una medición de laboratorio no se realizó). Identificar la cantidad y el patrón de los datos faltantes es el primer paso. ¿Faltan datos completamente al azar (MCAR – *Missing Completely At Random*)? ¿Faltan al azar (MAR – *Missing At Random*), donde la probabilidad de que falte un valor depende de otras variables observadas? ¿O no faltan

al azar (NMAR – *Missing Not At Random*), donde la probabilidad de que falte un valor depende del valor faltante en sí mismo? El manejo adecuado depende de la naturaleza de los datos faltantes y la cantidad.

Según la presentación adjunta, podemos evaluar cuánto es "mucho" en términos de datos faltantes:

- **< 5% de datos faltantes:** Generalmente se pueden imputar sin problema.
- **5% – 20%:** Se puede imputar, pero con cuidado; se deben considerar métodos de imputación más robustos.
- **> 20%:** Se deben considerar otras opciones, como eliminar variables o usar métodos avanzados.

Opciones comunes para manejar valores faltantes (Imputación de datos faltantes):

- **Ignorar la Tupla (Eliminación de Filas):** Eliminar las filas (registros de pacientes) que contienen valores faltantes. Esto es simple pero puede reducir drásticamente el tamaño de la muestra y sesgar los resultados si los datos no faltan completamente al azar.
- **Eliminar la Variable (Columna):** Si una variable tiene un porcentaje muy alto de valores faltantes (ej. > 50%, como se sugiere en la presentación) y no es crítica para el análisis, se puede eliminar.
- **Imputación (Rellenar los Valores Faltantes):** Reemplazar los valores faltantes por valores estimados. Las opciones varían según el tipo de datos:

- **Imputación de datos cuantitativos (numéricos):**
- **Media (Promedio):** Reemplazar los valores faltantes con la media de la variable. Útil si la distribución es simétrica.
- **Mediana:** Reemplazar los valores faltantes con la mediana. Preferible cuando hay datos atípicos (*outliers*) o la distribución es asimétrica.
- **Moda:** Para datos numéricos discretos (ej. Número de hospitalizaciones), aunque menos común que para datos categóricos.
- **Interpolación:** Estimar el valor faltante basándose en valores anteriores y posteriores. Muy útil en series de tiempo (ej. Estimar una lectura de presión arterial faltante entre dos mediciones conocidas).

- **Modelos Predictivos (Regresión, k-NN, etc.):** Predecir el valor faltante utilizando otras variables como predictores en un modelo de regresión o utilizando los valores de vecinos cercanos (*k-Nearest Neighbors – k-NN Imputation*).
- **Imputación de datos categóricos:**
- **Moda (Valor más frecuente):** Reemplazar los valores faltantes con la categoría que aparece con mayor frecuencia. Es el método más simple.
- **Imputación Aleatoria:** Elegir un valor aleatorio de entre las categorías existentes, a menudo ponderado por su frecuencia.
- **Modelos Predictivos (Clasificación):** Usar un modelo de clasificación (ej. Árbol de decisión, k-NN) para predecir la categoría faltante basándose en otras variables.
- **Nueva Categoría ("Desconocido" o "Missing"):** Crear una nueva categoría para representar explícitamente los valores faltantes. Útil si la ausencia de datos es en sí misma informativa.

o **Imputación Avanzada (cuando eliminar no es opción y hay muchas filas incompletas):**

- ***K-Nearest Neighbors* (KNN):** Imputa valores faltantes basándose en los valores de las observaciones más similares (vecinos) en otras variables.
- **Modelos Predictivos:** Imputación más precisa utilizando modelos de regresión o clasificación para predecir los valores faltantes.
- **MICE (*Multiple Imputation by Chained Equations*):** Un método iterativo y robusto que genera múltiples conjuntos de datos imputados y luego combina los resultados del análisis realizado en cada conjunto. Considera la incertidumbre de la imputación.

• **Identificación y Tratamiento de *Outliers* (Valores Atípicos):** Los *outliers* son observaciones que se desvían significativamente de la mayoría de los otros datos. Pueden ser errores de entrada, errores de medición o representar variabilidad real pero extrema.

La identificación de *outliers* se puede hacer visualmente (ej. Diagramas de caja, histogramas, gráficos de dispersión) o utilizando métodos estadísticos (ej. Puntuación Z, IQR). El tratamiento de los *outliers* depende de su causa:

o Si son errores, deben corregirse o eliminarse.

- Si representan variabilidad real, la decisión de incluirlos, transformarlos o analizarlos por separado depende del análisis y su impacto.

- **Manejo de Datos Ruidosos (*Noisy Data*):** Los datos ruidosos contienen errores aleatorios o varianza. La presentación menciona métodos como:

- ***Binning Method*:** Agrupar datos ordenados en "bins" (cubetas) y suavizar los valores dentro de cada bin (ej. Por la media o mediana del bin).

- ***Regression*:** Usar un modelo de regresión para ajustar los datos y suavizar el ruido.

- ***Clustering*:** Agrupar datos similares en clústeres; los valores que caen fuera de los clústeres principales pueden considerarse ruido.

- **Transformaciones Básicas**

- Las transformaciones de datos se aplican para convertir los datos a un formato más útil o apropiado para el análisis.

- **Escalado y Normalización:** Crucial para muchos algoritmos de aprendizaje automático que son sensibles a la escala de las variables (ej. Algoritmos basados en distancias como k-NN, máquinas de vectores de soporte con kernels, redes neuronales).

- **Normalización (Min-Max *Scaling*):** Escala los datos a un rango específico, comúnmente entre 0 y 1. $X' = (x - \min(x)) / (\max(x) - \min(x))$.

- **Estandarización (Z-score *Normalization*):** Transforma los datos para que tengan una media de 0 y una DS de 1. $X' = (x - \mu) / \sigma$.

- **Codificación de Variables Categóricas:** La mayoría de los algoritmos analíticos requieren entrada numérica. Las variables categóricas deben convertirse.

- **Codificación Numérica/Ordinal:** Asignar un número entero a cada categoría. Apropiado si las categorías tienen un orden intrínseco (ej. Leve=1, moderado=2, severo=3).

- ***One-Hot Encoding*:** Crear una nueva columna binaria (0 o 1) para cada categoría única. Apropiado para categorías nominales sin orden (ej. Raza, Grupo Sanguíneo) para evitar que el algoritmo asuma una relación de orden.

- **Discretización:** Convertir variables continuas o cuantitativas discretas en variables categóricas (ej. Agrupar edades en rangos: 18-25, 26-40, etc.).

- **Generación de Jerarquía de Conceptos:** Agrupar categorías de bajo nivel en categorías de alto nivel (ej. Agrupar códigos postales en ciudades, o tipos específicos de cáncer en categorías más amplias).

3.2. Preparación según el Modelo

La etapa final de preparación de datos a menudo depende de los requisitos específicos del modelo o algoritmo analítico que se utilizará.

- **Requisitos de Entrada Específicos:** Diferentes modelos tienen diferentes sensibilidades a la escala de los datos, la distribución, la presencia de *outliers*, la multicolinealidad (alta correlación entre variables predictoras), etc. La preparación debe asegurar que se cumplan estas asunciones o requisitos.

- **Formatos de Datos:**

  - **Regresión y Clasificación:** Generalmente requieren datos en formato tabular (filas = observaciones, columnas = variables).

  - **Series Temporales:** Requieren datos ordenados cronológicamente con un índice de tiempo claro. A menudo implican la creación de variables rezagadas (valores pasados de la variable de interés) o características basadas en el tiempo.

  - **Redes Neuronales:** Suelen requerir datos numéricos estandarizados o normalizados. Para variables categóricas con muchas categorías o texto, a menudo se utilizan ***Embeddings***, que son representaciones vectoriales densas y de baja dimensión que capturan relaciones semánticas.

El pretratamiento y la preparación de datos son un ciclo iterativo que a menudo implica regresar a etapas anteriores a medida que se descubren nuevos problemas o se prueban diferentes enfoques analíticos. Es una fase intensiva en tiempo pero indispensable para garantizar que el análisis posterior sea significativo y produzca resultados fiables.

Hemos aprendido la importancia del pretratamiento y la preparación de datos para transformar los datos brutos en un formato adecuado para el análisis. Ahora, veamos un ejemplo práctico de cómo aplicar estos pasos a un pequeño conjunto de datos hipotético en salud.

3.3. Ejemplo Práctico: Preparando Datos de un Estudio Piloto

Imagina que hemos realizado un pequeño estudio piloto para investigar la relación entre algunas características de los pacientes y un biomarcador sanguíneo. Hemos recolectado datos de 10 pacientes.

Datos Brutos Iniciales:

Aquí está nuestro conjunto de datos inicial. Observa que hay algunos problemas (ver tabla 22).

Tabla 22. Conjunto de datos de ejemplo.

ID Paciente	Edad	Sexo	Nivel Biomarcador (ng/mL)	Historial Familiar Diabetes	IMC (kg/m²)
1	45	Femenino	12.5	Sí	28.1
5	62	Masculino	18.9	No	31.5
3	38	Femenino	9.1	No	25.3
4	55	Masculino	15.2	Sí	29.9
5	70	Masculino	250.0	No	35.2
6	51	Femenino	14.0	Sí	27.5
7	49	Masculino	NaN	No	30.1
8	65	Femenino	20.5	Sí	NaN
9	42	Masculino	11.8	No	26.8
10	58	Femenino	16.1	Sí	32.0

Nota. NaN significa "Not a Number", indicando un valor faltante

Problemas en los Datos Brutos:

1. **Valores Faltantes:** El Paciente 7 tiene un valor faltante para "Nivel Biomarcador", y el Paciente 8 tiene un valor faltante para "IMC".
2. **Outlier Potencial:** El Paciente 5 tiene un "Nivel Biomarcador" de 250.0, que parece muy alto en comparación con los otros valores (12.5 a 20.5). Podría ser un error de entrada o un valor real pero extremo.

3. **Duplicado:** El "ID Paciente" 5 aparece dos veces. Uno de ellos es probablemente un error.

4. **Variables Categóricas:** "Sexo" y "Historial Familiar Diabetes" son variables categóricas que necesitarán ser codificadas para muchos modelos.

5. **Escala Variable:** Las variables numéricas ("Edad", "Nivel Biomarcador", "IMC") tienen diferentes rangos de valores.

Proceso de Pretratamiento y Preparación:

Aplicaremos los pasos de limpieza y transformación:

Paso 1: Limpieza de Datos

- **Manejo de Duplicados:** Identificamos que el ID Paciente 5 está duplicado. Al revisar los datos originales (si fuera un estudio real), determinaríamos cuál es el registro correcto o si ambos son válidos (lo cual sería inusual para un ID de paciente). Asumamos que el registro del Paciente 5 con Nivel Biomarcador 250.0 es un error de entrada del ID y deberíamos haber ingresado un ID diferente o que es simplemente un registro duplicado erróneo del paciente 5. Para este ejemplo, eliminaremos el segundo registro del Paciente 5 (el que tiene el Nivel Biomarcador de 250.0) asumiendo que fue un error de entrada del ID o un duplicado.

Datos después de eliminar duplicado (asumiendo que el segundo registro del paciente 5 fue el error):

Tabla 23. Conjunto de datos después de corrección del ID.

ID Paciente	Edad	Sexo	Nivel Biomarcador (ng/mL)	Historial Familiar Diabetes	IMC (kg/m²)
1	45	Femenino	12.5	Sí	28.1
2	62	Masculino	18.9	No	31.5
3	38	Femenino	9.1	No	25.3
4	55	Masculino	15.2	Sí	29.9
6	51	Femenino	14.0	Sí	27.5

ID Paciente	Edad	Sexo	Nivel Biomarcador (ng/mL)	Historial Familiar Diabetes	IMC (kg/m²)
7	49	Masculino	NaN	No	30.1
8	65	Femenino	20.5	Sí	NaN
9	42	Masculino	11.8	No	26.8
10	58	Femenino	16.1	Sí	32.0

Nota. Hemos corregido el ID del segundo paciente 5 a 2 para el ejemplo

- **Manejo de Valores Faltantes:**

  - **Nivel Biomarcador (Paciente 7):** Es un dato cuantitativo. Tenemos 1 valor faltante de 9 registros (aproximadamente 11%). Según la guía (<20%), la imputación es una opción razonable. La distribución de los niveles de biomarcador restantes (9.1 a 20.5) parece relativamente simétrica, por lo que podríamos imputar la **media** de los valores observados. Media = (12.5 + 18.9 + 9.1 + 15.2 + 14.0 + 20.5 + 11.8 + 16.1) / 8 = 14.76. Imputamos 14.76 al Paciente 7.

  - **IMC (Paciente 8):** Es un dato cuantitativo. Tenemos 1 valor faltante de 9 registros (aproximadamente 11%). La distribución de IMC (25.3 a 32.0) parece razonablemente simétrica. Podríamos imputar la **media** (o la mediana si hubiera *outliers* evidentes). Media = (28.1 + 31.5 + 25.3 + 29.9 + 27.5 + 30.1 + 26.8 + 32.0) / 8 = 28.9. Imputamos 28.9 al Paciente 8.

Tabla 24. Datos después de imputar valores faltantes.

ID Paciente	Edad	Sexo	Nivel Biomarcador (ng/mL)	Historial Familiar Diabetes	IMC (kg/m²)
1	45	Femenino	12.50	Sí	28.10
2	62	Masculino	18.90	No	31.50
3	38	Femenino	9.10	No	25.30
4	55	Masculino	15.20	Sí	29.90
6	51	Femenino	14.00	Sí	27.50

ID Paciente	Edad	Sexo	Nivel Biomarcador (ng/mL)	Historial Familiar Diabetes	IMC (kg/m²)
7	49	Masculino	14.76	No	30.10
8	65	Femenino	20.50	Sí	28.90
9	42	Masculino	11.80	No	26.80
10	58	Femenino	16.10	Sí	32.00

Manejo de *Outliers*: El valor 250.0 en "Nivel Biomarcador" del registro duplicado ya fue eliminado al manejar los duplicados. Si hubiera sido un valor real pero extremo en un registro válido, podríamos haber considerado transformarlo (ej. aplicar logaritmo) o usar métodos de análisis robustos menos sensibles a *outliers*.

Paso 2: Transformaciones Básicas

- **Codificación de Variables Categóricas:**

  - **Sexo:** Es nominal (sin orden). Usaremos *One-Hot Encoding*. Crearemos dos nuevas columnas: "Sexo_Femenino" (1 si es Femenino, 0 si es Masculino) y "Sexo_Masculino" (1 si es Masculino, 0 si es Femenino). Para evitar multicolinealidad perfecta, a menudo se elimina una de las columnas (ej. solo mantener "Sexo_Femenino").

  - **Historial Familiar Diabetes:** Es binaria (Sí/No). Podemos codificarla numéricamente: Sí=1, No=0.

- **Escalado de Variables Numéricas:** "Edad", "Nivel Biomarcador" e "IMC" tienen diferentes rangos. Si planeamos usar un modelo sensible a la escala (ej. SVM, k-NN, redes neuronales), deberíamos escalarlas. Usemos Estandarización (Z-score *Normalization*). Para cada variable, restaremos la media y dividiremos por la DS.

La Tabla 25 presenta los datos de ejemplo una vez aplicadas las transformaciones necesarias: la codificación de variables categóricas y la estandarización (Z-score) de variables numéricas. Para comprender mejor qué representan estos nuevos valores, analicemos los ejemplos de los primeros tres pacientes:

- **Paciente 1:**
  - Tiene una **edad** ligeramente superior a la media (valor estandarizado: 0.45).
  - Es **mujer** (Sexo_Femenino = 1).
  - Tiene **historial familiar** de diabetes (Historial Familiar Diabetes (Codificado) = 1).
  - Su **nivel de biomarcador** está notablemente por encima del promedio (valor estandarizado: 1.15).
  - Su **IMC** también es superior a la media (valor estandarizado: 0.82).
- **Paciente 2:**
  - Presenta una **edad** inferior a la media (valor estandarizado: -1.20).
  - Es **hombre** (Sexo_Femenino = 0).
  - No tiene **historial familiar** de diabetes (Historial Familiar Diabetes (Codificado) = 0).
  - Su **nivel de biomarcador** se encuentra ligeramente por debajo del promedio (valor estandarizado: -0.30).
  - Su **IMC** es muy cercano a la media (valor estandarizado: -0.15).
- **Paciente 3:**
  - Su **edad** está bastante por encima de la media (valor estandarizado: 1.55).
  - Es **mujer** (Sexo_Femenino = 1).
  - No tiene **historial familiar** de diabetes (Historial Familiar Diabetes (Codificado) = 0).
  - El **nivel de biomarcador** es muy cercano al promedio (valor estandarizado: 0.05).
  - Sin embargo, su **IMC** es significativamente elevado (valor estandarizado: 2.10).

Esta interpretación muestra cómo los valores numéricos estandarizados indican la posición relativa de cada paciente respecto a la media del grupo (valores positivos por encima, negativos por debajo) y cómo las variables categóricas se representan ahora numéricamente (0 o 1).

Tabla 25. *Datos después de codificación y estandarización (ejemplo de las primeras filas y variables transformadas).*

ID Paciente	Edad (Estandarizada)	Sexo Femenino	Historial Familiar Diabetes (Codificado)	Nivel Biomarcador (Estandarizado)	IMC (Estandarizado)
1	0.45	1	1	1.15	0.82
2	-1.20	0	0	-0.30	-0.15
3	1.55	1	0	0.05	2.10
...	...	...	...	...	...

Nota: Los valores en las columnas estandarizadas son ejemplos numéricos. Representan cuántas desviaciones estándar por encima (+) o por debajo (-) de la media se encuentra el valor original de cada paciente para esa variable. Tienen una media cercana a 0 y una DS cercana a 1 en todo el conjunto de datos.

Paso 3: Preparación según el Modelo

Ahora que los datos están limpios y transformados, están listos para ser utilizados en un modelo analítico. Si nuestro objetivo es, por ejemplo, construir un modelo de **regresión lineal** para predecir el "Nivel Biomarcador" basándose en las otras variables, nuestro conjunto de datos preparado tendría las columnas "Edad (Estandarizada)", "Sexo_Femenino", "Historial Familiar Diabetes (Codificado)" e "IMC (Estandarizado)" como variables predictoras (variables independientes) y "Nivel Biomarcador (Estandarizado)" como la variable objetivo (variable dependiente).

Si estuviéramos preparando datos para una **red neuronal**, probablemente mantendríamos todas las variables numéricas estandarizadas y usaríamos *one-hot encoding* para "Sexo" (manteniendo ambas columnas "Sexo_Femenino" y "Sexo_Masculino" si el modelo lo requiere, o usando *embeddings* si tuviéramos más categorías).

Este ejemplo, aunque simplificado, ilustra los pasos comunes involucrados en el pretratamiento y la preparación de datos, destacando la importancia de la limpieza, el manejo de valores faltantes y *outliers*, y las transformaciones necesarias para preparar los datos para el análisis.

Conclusiones

1. **Paso Obligatorio:** El pretratamiento y la preparación de datos son etapas esenciales después de la recolección y antes del análisis para asegurar la calidad y adecuación de los datos.
2. **Limpieza Fundamental:** La limpieza de datos es crucial e incluye:

- **Manejo de Valores Faltantes:** Identificar y tratar adecuadamente los datos ausentes (según su patrón y cantidad) mediante eliminación o métodos de imputación (simple o avanzada como MICE/k-NN).
- **Manejo de *Outliers*:** Detectar y gestionar valores atípicos, corrigiéndolos si son errores o decidiendo cómo tratarlos si son valores reales extremos.
- **Manejo de Ruido:** Aplicar técnicas de suavizado si los datos presentan errores aleatorios.

3. **Transformación Necesaria:** Los datos a menudo requieren transformaciones para adecuarse a los métodos de análisis:

- **Escalado/Normalización:** Ajustar el rango o distribución de variables numéricas, especialmente para algoritmos sensibles a la escala.
- **Codificación Categórica:** Convertir variables no numéricas a un formato numérico (ej. *One-Hot Encoding*).
- **Discretización:** Agrupar valores continuos en categorías si es necesario.

4. **Dependencia del Modelo:** La preparación final de los datos a menudo depende de los requisitos específicos del modelo de análisis o aprendizaje automático que se vaya a utilizar.
5. **Proceso Iterativo:** La preparación de datos no es estrictamente lineal; a menudo requiere volver a etapas anteriores a medida que se entiende mejor los datos o se refinan los objetivos del análisis.

6. **Impacto en Resultados:** La calidad del pretratamiento y la preparación impacta directamente en la fiabilidad, validez y precisión de los resultados del análisis y los modelos generados.

CAPÍTULO 21

DE LOS DATOS AL CONOCIMIENTO APLICABLE

From Data to Actionable Knowledge

Resumen

Este capítulo se enfoca en transformar los resultados del análisis de datos en conocimiento accionable para la toma de decisiones en salud. Detalla la etapa de Integración de Resultados, que implica interpretar hallazgos en contexto, considerar la incertidumbre, evaluar la relevancia clínica (más allá de la significancia estadística) y sintetizar con evidencia existente para guiar decisiones clínicas (ej. guías, CDSS). Aborda la Presentación de Hallazgos, enfatizando la necesidad de claridad y adaptación a la audiencia. Describe el diseño de *Dashboards* efectivos, incluyendo la selección de métricas clave (KPI relevantes, medibles, accionables) y la elección adecuada de gráficos (barras, líneas, dispersión, etc.) para diferentes propósitos (comparación, distribución, tendencia), así como su organización lógica. Introduce el concepto de *Storytelling* con Datos como el arte de narrar la historia inherente en los datos de forma convincente para comunicar *ideas clave* y motivar la acción. Finalmente, cubre la Divulgación de la Investigación a través de diversos canales (publicaciones, conferencias, informes, comunicación pública) para asegurar que el conocimiento generado sea accesible y utilizado. Se ilustra con un ejemplo práctico de análisis y presentación de datos para una clínica.

Palabras Clave: Conocimiento aplicable, integración de resultados, presentación de datos, visualización de datos, *storytelling* con datos, divulgación científica.

Abstract

This chapter focuses on transforming data analysis results into actionable knowledge for health decision-making. It details the Results Integration stage, involving interpreting findings in context, considering uncertainty, evaluating

clinical relevance (beyond statistical significance), and synthesizing with existing evidence to guide clinical decisions (e.g., guidelines, CDSS). It addresses the Presentation of Findings, emphasizing clarity and audience adaptation. Effective Dashboard design is described, including selecting relevant Key Performance Indicators (KPIs - relevant, measurable, actionable) and choosing appropriate charts (bar, line, scatter, etc.) for different purposes (comparison, distribution, trend), along with logical layout. The concept of Data Storytelling is introduced as the art of narrating the inherent story in data convincingly to communicate insights and motivate action. Finally, it covers Research Dissemination through various channels (publications, conferences, reports, public communication) to ensure generated knowledge is accessible and utilized. A practical example of analyzing and presenting data for a clinic illustrates these points.

Keywords: *Actionable knowledge, results integration, data presentation, data visualization, data storytelling, research dissemination.*

Introducción

El objetivo final de trabajar con datos en salud no es simplemente realizar análisis complejos, sino generar *ideas clave* que tengan un impacto positivo en la atención al paciente, la salud pública o la gestión de los sistemas de salud. Esta transformación de resultados analíticos en conocimiento accionable requiere una cuidadosa **integración** de los hallazgos y una **presentación** clara y adaptada al público objetivo, culminando en la **divulgación** de la investigación.

4.1. Integración de Resultados: Traduciendo Hallazgos en Decisiones Clínicas

La etapa de análisis nos proporciona resultados: estadísticas descriptivas, parámetros de modelos predictivos, medidas de asociación, estimaciones de efectos de tratamiento, etc. Sin embargo, estos resultados por sí solos no son conocimiento aplicable. La integración de resultados implica:

- **Interpretación en Contexto:** Los hallazgos deben interpretarse dentro del contexto clínico, epidemiológico o de gestión relevante. ¿Qué significan estos números para un paciente individual? ¿Cómo impactan en la práctica clínica habitual? ¿Qué implican para una población específica?

- **Consideración de la Incertidumbre:** Reconocer y comunicar la incertidumbre asociada a los resultados (ej. intervalos de confianza). Ningún hallazgo es 100% seguro.
- **Evaluación de la Relevancia Clínica o de Gestión:** Un resultado estadísticamente significativo no siempre es clínicamente relevante o útil desde una perspectiva de gestión. Se debe evaluar la magnitud del efecto y su importancia práctica.
- **Síntesis con Evidencia Existente:** Los hallazgos de un análisis de datos particular deben integrarse con el cuerpo de conocimiento existente (ej. resultados de estudios previos, guías de práctica clínica). ¿Confirman, contradicen o complementan lo que ya se sabe?
- **Consideraciones Éticas y de Viabilidad:** Antes de traducir un hallazgo en una recomendación o decisión, se deben considerar las implicaciones éticas, la viabilidad de la implementación en el entorno real y los posibles costos y beneficios.

Cómo traducir hallazgos en decisiones clínicas

Este es quizás el paso más crítico y desafiante. Un hallazgo analítico (ej. "nuestro modelo predice que los pacientes con característica X tienen un mayor riesgo de complicación Y") debe transformarse en una acción concreta a nivel clínico (ej. "¿deberíamos monitorear más de cerca a estos pacientes?", "¿necesitan un tratamiento preventivo diferente?"). Esto implica un proceso de toma de decisiones basado en evidencia, donde los hallazgos del análisis de datos se combinan con la experiencia clínica, los valores del paciente y los recursos disponibles.

- **Desarrollo de Guías o Protocolos:** Los hallazgos robustos pueden informar el desarrollo o la actualización de guías de práctica clínica o protocolos de manejo de pacientes.
- **Sistemas de Soporte a la Decisión Clínica (CDSS):** Los modelos predictivos pueden integrarse en sistemas de HCE para proporcionar alertas o recomendaciones a los clínicos en el punto de atención.
- **Herramientas de Comunicación con el Paciente:** Los hallazgos sobre pronóstico o riesgo pueden traducirse en herramientas visuales o explicaciones sencillas para discutir opciones con los pacientes.

- **Capacitación del Personal:** Educar a los clínicos y al personal de salud sobre los nuevos hallazgos y cómo aplicarlos en su práctica.

La traducción efectiva de hallazgos en decisiones clínicas requiere una estrecha colaboración entre analistas de datos, clínicos, gestores y pacientes.

4.2. Presentación de Hallazgos: Comunicando Ideas clave Efectivamente

La forma en que se presentan los hallazgos es tan importante como la calidad del análisis mismo. La presentación debe ser clara, concisa, precisa y adaptada a la audiencia para maximizar el impacto y la comprensión.

Diseñando *Dashboards* Efectivos:

Los dashboards son herramientas poderosas para presentar información clave de manera visual e interactiva, especialmente útiles para monitoreo y toma de decisiones operativas o tácticas.

- **Escogiendo Buenas Métricas (KPI - *Key Performance Indicators*):**
  - **Relevancia:** Las métricas deben estar directamente relacionadas con los objetivos del análisis o del área que se monitorea (ej. tasa de reingreso hospitalario, tiempo promedio de estancia, adherencia a guías de tratamiento, prevalencia de una condición).
  - **Medibilidad:** Deben ser cuantificables y los datos para calcularlas deben estar disponibles y ser fiables.
  - **Accionabilidad:** Las métricas deben proporcionar información que permita tomar decisiones o emprender acciones.
  - **Claridad:** Deben ser fáciles de entender por la audiencia objetivo.
  - **Contexto:** Presentar la métrica junto con su valor objetivo, valor de referencia (ej. promedio histórico, benchmark) o tendencia a lo largo del tiempo ayuda a su interpretación.
- **Cómo Escoger Métricas Adecuadamente:**
  - Comprender profundamente la pregunta de negocio o investigación y los objetivos de la audiencia.

- Identificar los procesos clave que influyen en los resultados.
- Realizar sesiones con los stakeholders (clínicos, gestores) para definir qué información necesitan para tomar decisiones.
- Comenzar con un número limitado de métricas clave y añadir más gradualmente si es necesario.

- **Escogiendo Adecuadamente el Gráfico:** La elección del tipo de gráfico es crucial para visualizar los datos de manera efectiva.
  - **Comparación:** Gráficos de barras (para comparar categorías), gráficos de líneas (para comparar tendencias a lo largo del tiempo), gráficos de dispersión (para comparar dos variables numéricas).
  - **Distribución:** Histogramas (para la distribución de una variable numérica), diagramas de caja (para comparar la distribución entre grupos).
  - **Composición:** Gráficos circulares o de pastel (para mostrar partes de un todo, usar con precaución y pocas categorías), gráficos de barras apiladas.
  - **Relación:** Gráficos de dispersión (para mostrar la relación entre dos variables numéricas), mapas de calor.
  - **Tendencia a lo Largo del Tiempo:** Gráficos de líneas.
- **Cómo Organizar los Datos en un *Dashboard*:**
  - **Layout Lógico:** Organizar los elementos visuales de manera que sigan un flujo lógico, a menudo de arriba a abajo o de izquierda a derecha, comenzando con la información más importante.
  - **Agrupación:** Agrupar gráficos y métricas relacionadas temáticamente.
  - **Uso del Espacio:** Utilizar el espacio de manera eficiente sin sobrecargar el *dashboard*. Dejar espacios en blanco ayuda a la legibilidad.
  - **Interactividad:** Incorporar filtros, segmentación y opciones de desglose (*drill-down*) para permitir a los usuarios explorar los datos según sus necesidades.
  - **Consistencia Visual:** Mantener una paleta de colores, fuentes y estilos consistentes en todo el *dashboard*.

Otros Formatos de Presentación:

- **Informes Ejecutivos:** Resumen conciso para audiencias de alto nivel. Se centran en los hallazgos clave y las recomendaciones.
- **Informes Técnicos/Analíticos:** Documento detallado para otros analistas o investigadores, incluyendo metodología y resultados completos.
- **Publicaciones Científicas:** Artículos formales para difundir nuevos conocimientos a la comunidad científica.
- **Presentaciones Orales/Pósteres:** Comunicación concisa y visual en eventos académicos o internos.

4.3. Storytelling con Datos: Narrando la Historia Detrás de los Números

Los datos y los análisis proporcionan los hechos, pero el ***storytelling* con datos** es el arte de comunicar estos hechos de una manera convincente y memorable que resuene con la audiencia y los motive a la acción. No se trata de inventar una historia, sino de estructurar la presentación de los hallazgos de manera que revele la narrativa inherente en los datos.

- **Comprende a tu Audiencia:** ¿Cuáles son sus intereses, conocimientos y motivaciones? ¿Qué preguntas necesitan responder?
- **Define el Mensaje Principal:** ¿Cuál es el hallazgo más importante o la conclusión clave que quieres que la audiencia recuerde?
- **Establece el Contexto:** Comienza explicando el problema o la pregunta de investigación y por qué es importante.
- **Presenta los Datos de Forma Lógica:** Guía a la audiencia a través de los hallazgos paso a paso, construyendo la narrativa. Utiliza visualizaciones para ilustrar los puntos clave.
- **Enfócate en las *Ideas clave*, no Solo en los Datos Brutos:** Explica qué significan los hallazgos, no solo cuáles son los números.
- **Utiliza Visualizaciones Efectivas para Contar la Historia:** Un gráfico bien diseñado puede comunicar un mensaje complejo de manera instantánea.

- **Concluye con una Llamada a la Acción (Sí Aplica):** ¿Qué deberían hacer la audiencia con esta información?

El *storytelling* con datos es crucial para cerrar la brecha entre el análisis técnico y la comprensión y acción por parte de quienes toman las decisiones.

4.4. Divulgación de la Investigación: Compartiendo el Conocimiento Generado

La divulgación es el proceso de comunicar los hallazgos de la investigación a la comunidad científica, a los profesionales de la salud, a los responsables de políticas y al público en general. Va más allá de la simple presentación de resultados y busca asegurar que el conocimiento generado sea accesible y utilizado.

- **Publicaciones en Revistas Científicas:** El método tradicional y fundamental para difundir nuevos conocimientos a la comunidad académica. La publicación en revistas revisadas por pares asegura la calidad y credibilidad de la investigación.
- **Presentaciones en Conferencias:** Permiten compartir hallazgos con colegas, obtener retroalimentación y establecer colaboraciones.
- **Informes Técnicos y White Papers:** Documentos detallados o resúmenes para audiencias específicas (ej. agencias gubernamentales, organizaciones de atención médica).
- **Comunicación a Responsables de Políticas:** Presentar hallazgos y sus implicaciones a quienes toman decisiones sobre políticas de salud para informar la práctica basada en evidencia.
- **Comunicación a Profesionales de la Salud:** Compartir hallazgos relevantes con médicos, enfermeras y otros profesionales a través de presentaciones, boletines informativos, o integración en guías de práctica clínica.
- **Comunicación al Público General:** Traducir hallazgos complejos a un lenguaje comprensible para el público a través de comunicados de prensa, artículos de divulgación, blogs o redes sociales. Esto es particularmente importante para temas de salud pública.

- **Repositorios de Datos y Código Abierto:** Compartir los datos (anonimizados) y el código utilizado en la investigación permite la replicación, validación y el uso de los datos para futuras investigaciones.

La divulgación efectiva es un proceso continuo que requiere identificar las audiencias clave, adaptar el mensaje y utilizar los canales de comunicación más apropiados para maximizar el impacto del conocimiento generado a partir de los datos.

En resumen, la fase de "De los Datos al Conocimiento Aplicable" es donde el valor real del análisis de datos se materializa. Implica no solo obtener resultados, sino integrarlos, presentarlos de manera efectiva utilizando herramientas como *dashboards* y *storytelling*, y divulgarlos ampliamente para informar la toma de decisiones y avanzar en el campo de la salud.

Hemos aprendido sobre el ciclo de vida de los datos, desde su recolección y preparación hasta la generación de conocimiento. Ahora, un paso crucial es cómo presentamos y comunicamos ese conocimiento de manera efectiva para informar la toma de decisiones. Este ejemplo práctico ilustrará cómo podemos pasar de un conjunto de datos a un *dashboard* visual y luego usar el *storytelling* para transmitir las *ideas clave* clave.

4.5. Ejemplo Práctico: Análisis de Datos en una Clínica de Atención Primaria

Imagina que trabajas en una pequeña clínica de atención primaria y tienes acceso a un conjunto de datos de pacientes recientes. La gerencia de la clínica está interesada en entender mejor a su población de pacientes y optimizar la atención.

4.5.1. *Conjunto de Datos de Ejemplo (Simplificado)*

Hemos extraído y limpiado un pequeño conjunto de datos hipotético de 20 pacientes que visitaron la clínica el mes pasado:

Tabla 26. *Conjunto de datos de ejemplo.*

ID Paciente	Edad	Sexo	Motivo Consulta Principal	Diagnóstico Principal	Tensión Arterial Sistólica (TAS)	Tensión Arterial Diastólica (TAD)	Fumador	Citas Previas (último año)
101	45	Femenino	Chequeo Anual	Saludable	125	80	No	2
102	62	Masculino	Dolor de Pecho	Angina de Pecho	140	90	Sí	5
103	38	Femenino	Síntomas Gripales	Infección Respiratoria Viral	118	75	No	1
104	55	Masculino	Dolor de Espalda	Lumbalgia	130	85	No	3
105	70	Femenino	Chequeo Anual	Hipertensión	155	95	No	4
106	51	Femenino	Erupción Cutánea	Dermatitis de Contacto	120	78	No	1
107	49	Masculino	Dolor de Cabeza	Cefalea Tensional	135	88	Sí	2
108	65	Femenino	Seguimiento Diabetes	Diabetes Mellitus Tipo 2	148	92	No	6
109	42	Masculino	Dolor de Garganta	Faringitis Estreptocócica	122	76	No	1
110	58	Femenino	Chequeo Anual	Hipertensión	150	90	No	3
111	35	Masculino	Lesión Deportiva	Esguince de Tobillo	115	70	No	0

ID Paciente	Edad	Sexo	Motivo Consulta Principal	Diagnóstico Principal	Tensión Arterial Sistólica (TAS)	Tensión Arterial Diastólica (TAD)	Fumador	Citas Previas (último año)
112	75	Femenino	Problemas de Sueño	Insomnio	140	85	No	7
113	29	Masculino	Acné	Acné Vulgar	110	68	No	1
114	68	Masculino	Seguimiento Colesterol	Dislipidemia	138	82	Sí	4
115	53	Femenino	Dolor Articular	Osteoartritis	132	86	No	3
116	47	Masculino	Chequeo Anual	Saludable	128	82	No	2
117	60	Femenino	Fatiga	Hipotiroidismo	130	80	No	5
118	40	Masculino	Control Presión	Hipertensión	145	90	Sí	2
119	72	Femenino	Dolor Abdominal	Síndrome Intestino Irritable	135	85	No	6
120	56	Masculino	Chequeo Anual	Obesidad	140	90	No	3

4.5.2. *Diseño del Dashboard: del Dato a las Métricas y Visualizaciones*

El objetivo del *dashboard* es proporcionar una visión rápida y comprensible de las características de los pacientes y los patrones de consulta.

Escogiendo Buenas Métricas (KPI):

Basándonos en los datos disponibles y lo que podría interesarle a la gerencia de una clínica:

- **Métricas de Demografía:**
  - Distribución por Sexo.
  - Distribución por Grupos de Edad.
- **Métricas de Actividad/Utilización:**
  - Número Total de Citas.
  - Promedio de Citas Previas por Paciente.
  - Distribución de Citas Previas.
- **Métricas Clínicas/de Salud:**
  - Distribución de Diagnósticos Principales.
  - Prevalencia de Tabaquismo.
  - Promedio de Tensión Arterial (Sistólica y Diastólica).
  - Proporción de Pacientes con Hipertensión (TAS ≥140 o TAD ≥90).
- **Métricas de Motivo de Consulta:**
  - Distribución de Motivos de Consulta Principales.

Cómo Escoger Métricas Adecuadamente (Proceso):

1. **Entender a la Audiencia:** ¿Quién usará este *dashboard*? (Gerencia de la clínica). ¿Qué decisiones necesitan tomar? (Optimizar horarios, identificar necesidades de salud en la población, planificar recursos).
2. **Relacionar con Objetivos:** Los objetivos podrían ser: entender la carga de trabajo, identificar condiciones comunes, evaluar factores de riesgo. Las métricas seleccionadas (número de citas, diagnósticos, tabaquismo, hipertensión) se alinean con estos objetivos.

3. **Disponibilidad de Datos:** Todas las métricas seleccionadas pueden calcularse directamente a partir de nuestro conjunto de datos.
4. **Simplicidad:** Las métricas son relativamente fáciles de entender.

Escogiendo Adecuadamente el Gráfico:

Ahora, elegimos el tipo de gráfico apropiado para cada métrica o comparación:

- Distribución por Sexo: **Gráfico de Pastel** o **Gráfico de Barras** (para mostrar proporciones).
- Distribución por Grupos de Edad: **Histograma** o **Gráfico de Barras** (agrupando edades en rangos como <40, 40-60, >60).
- Número Total de Citas: Un simple **Indicador Numérico** grande.
- Promedio de Citas Previas: Un simple **Indicador Numérico**.
- Distribución de Citas Previas: **Gráfico de Barras** (contando cuántos pacientes tuvieron 0 citas, 1 cita, 2, etc.).
- Distribución de Diagnósticos Principales: **Gráfico de Barras** (mostrando la frecuencia de cada diagnóstico).
- Prevalencia de Tabaquismo: **Gráfico de Pastel** o **Gráfico de Barras** (proporción de fumadores vs no fumadores).
- Promedio de Tensión Arterial: **Indicadores Numéricos** o **Gráfico de Barras** comparando TAS y TAD promedio.
- Proporción de Pacientes con Hipertensión: **Gráfico de Pastel** o **Indicador Numérico** con porcentaje.
- Distribución de Motivos de Consulta: **Gráfico de Barras**.

Cómo Organizar los Datos en un *Dashboard* (*Layout* Conceptual):

Organizaríamos estos gráficos y métricas en un *dashboard* de manera lógica y visualmente atractiva.

- **Parte Superior (Resumen Ejecutivo):** Indicadores numéricos clave (Total Citas, Promedio Citas Previas, Prevalencia Tabaquismo, Proporción Hipertensión).
- **Sección de Demografía:** Gráficos de distribución por Sexo y Edad.

- **Sección de Actividad:** Gráfico de distribución de Citas Previas.
- **Sección Clínica:** Gráficos de distribución de Diagnósticos Principales, Tensión Arterial Promedio.
- **Sección de Motivos de Consulta:** Gráfico de distribución de Motivos de Consulta.

Podríamos usar filtros interactivos para ver los datos por sexo o grupo de edad.

Figura 62. *Ejemplo de un dashboard*. Elaborado con GeminiPRO.



4.5.3. *Storytelling con Datos: Narrando la Historia de la Clínica*

Una vez que el *dashboard* está listo, no solo lo presentamos; contamos una historia con los datos que contiene. Imaginemos que presentamos este *dashboard* a la gerencia de la clínica.

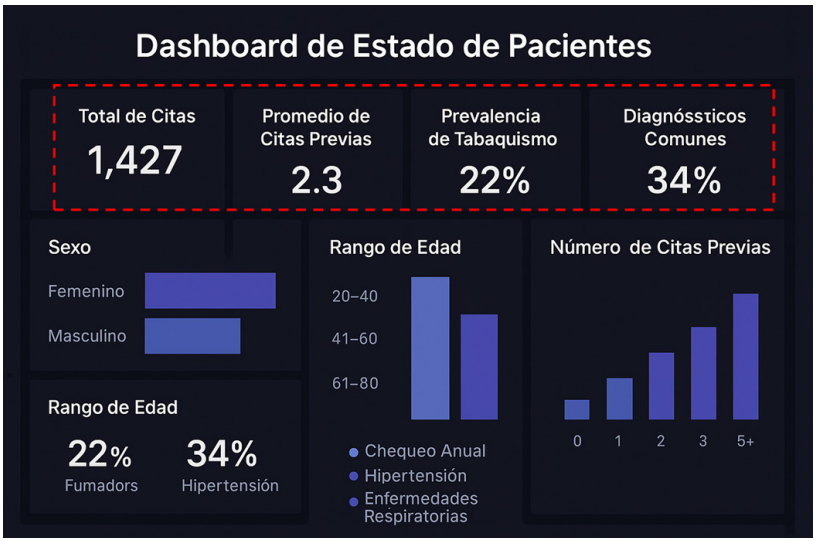
Audiencia: Gerencia de la Clínica.

Mensaje Principal: Comprender las características de nuestros pacientes y los patrones de consulta para mejorar la planificación y la atención.

Narrativa (Ejemplo de Presentación):

"Buenos días, equipo. Hoy vamos a sumergirnos en los datos de nuestros pacientes del mes pasado para entender mejor quiénes son y por qué vienen a vernos (ver figura 63).

Figura 63. *Ejemplo de dashboard para el storytelling*. Elaborado con GeminiPRO.



Primero, vemos que tuvimos un total de [Número Total de Citas] citas. El promedio de citas previas por paciente en el último año es de [Promedio Citas Previas], lo que nos da una idea de la frecuencia con la que nuestros pacientes utilizan nuestros servicios.

[Mostrar Gráficos de Demografía]

Analizando la demografía, observamos que nuestra población de pacientes está bastante equilibrada entre hombres y mujeres [Mostrar Gráfico de Sexo]. Sin embargo, vemos una concentración significativa de pacientes en los grupos de edad de [Rangos de Edad con más pacientes] [Mostrar Gráfico de Edad]. Esto es importante para adaptar nuestros servicios y programas preventivos a estas edades.

[Mostrar Gráfico de Diagnósticos Principales]

En cuanto a los diagnósticos, los motivos de consulta más comunes el mes pasado fueron [Diagnósticos más frecuentes, ej. Infecciones Respiratorias, Hipertensión, Chequeos Anuales] [Mostrar Gráfico de Diagnósticos]. Esto nos indica las principales necesidades de salud que estamos atendiendo.

[Mostrar Prevalencia de Tabaquismo e Hipertensión]

Un hallazgo importante es que el [Porcentaje]% de nuestros pacientes reportaron ser fumadores [Mostrar Métrica de Tabaquismo]. Además, el [Porcenta-

je]% de los pacientes en este conjunto de datos presentaron valores de tensión arterial que cumplen los criterios de hipertensión [Mostrar Métrica de Hipertensión]. Estas son áreas clave donde podríamos enfocarnos en programas de detección y manejo activo.

[Mostrar Gráfico de Citas Previas]

Interesantemente, vemos que hay un grupo de pacientes con un número elevado de citas previas [Mostrar Gráfico de Citas Previas]. Podríamos investigar más a fondo si estos pacientes tienen condiciones crónicas que requieren manejo intensivo.

Conclusión y Llamada a la Acción:

En resumen, nuestros datos nos muestran que atendemos a una población diversa en edad, con una carga significativa de condiciones comunes como infecciones respiratorias e hipertensión, y un porcentaje considerable de fumadores. Esta información nos puede ayudar a:

1. Planificar recursos y horarios basándonos en los motivos de consulta más frecuentes.
2. Desarrollar o fortalecer programas de prevención y manejo para tabaquismo e hipertensión, quizás dirigidos específicamente a los grupos de edad más representados.
3. Identificar a los pacientes con alta utilización de servicios para ofrecerles un manejo más integrado de sus condiciones crónicas.

"Este *dashboard* estará disponible para monitorear estas métricas continuamente."

4.5.4. *Divulgación de la Investigación (en este contexto)*

En el contexto de una clínica pequeña, la "divulgación" podría no ser una publicación científica, sino:

- **Compartir el *Dashboard*:** Poner el *dashboard* a disposición del personal clínico y administrativo para que puedan acceder a la información.
- **Presentaciones Internas:** Realizar presentaciones regulares al personal para discutir los hallazgos y su impacto en la práctica diaria.
- **Informes a la Gerencia:** Utilizar los datos del *dashboard* para informar decisiones de gestión, como la asignación de personal o la inversión en programas de salud.

- **Mejoras en los Procesos:** Implementar cambios en los flujos de trabajo o en la atención al paciente basados en las *ideas clave* obtenidos.

Este ejemplo ilustra cómo los datos, una vez recolectados, limpiados y analizados (aunque sea de forma descriptiva simple en este caso), pueden ser transformados en conocimiento aplicable a través de la visualización en un *dashboard* y la comunicación efectiva mediante *storytelling*, informando así la toma de decisiones en un entorno de atención médica.

Conclusiones

1. **Objetivo Final:** El análisis de datos en salud busca generar conocimiento aplicable que informe y mejore la toma de decisiones clínicas, de gestión o de salud pública.
2. **Integración Contextual:** Los resultados analíticos deben ser interpretados considerando el contexto clínico/epidemiológico, la incertidumbre (IC), la relevancia práctica (no solo estadística) y la evidencia existente.
3. **Traducción a la Práctica:** La transformación de hallazgos en acciones requiere un proceso deliberado que puede incluir la actualización de guías, el desarrollo de sistemas de apoyo a la decisión (CDSS) o la mejora de la comunicación con el paciente.
4. **Comunicación Efectiva:** La presentación de hallazgos debe ser clara, concisa y adaptada a la audiencia. Los ***Dashboards*** son herramientas visuales clave que requieren una cuidadosa selección de KPI y tipos de gráficos apropiados.
5. **Poder de la Narrativa:** El ***Storytelling* con Datos** es esencial para comunicar *ideas clave* de manera memorable y convincente, estructurando los hallazgos en una narrativa lógica centrada en un mensaje clave.
6. **Divulgación Amplia:** Compartir el conocimiento generado a través de publicaciones científicas, conferencias, informes y comunicación pública es crucial para maximizar el impacto de la investigación.
7. **Ejemplo Práctico:** La aplicación de estos principios (selección de KPI, diseño de *dashboard, storytelling*) es fundamental para traducir datos, incluso descriptivos simples, en información útil para la gestión y mejora en entornos clínicos reales.

Parte V.

Inteligencia Artificial en Investigación

CAPÍTULO 22

HISTORIA Y FUNDAMENTOS DE LA IA

History and Fundamentals of AI

Resumen

Este capítulo traza una breve historia de la Inteligencia Artificial (IA), desde sus inicios con sistemas basados en reglas explícitas (años 50-80) hasta el auge del Aprendizaje Automático (*Machine Learning* - ML) (años 80-2010s), donde las máquinas aprenden de los datos. Destaca el posterior despegue del Aprendizaje Profundo (*Deep Learning* - DL) (años 2010s en adelante), un subcampo del ML que utiliza redes neuronales profundas y ha sido clave para avances en el procesamiento de imágenes, audio y texto, incluyendo el reciente desarrollo de Grandes Modelos de Lenguaje (LLM). Clasifica la IA en Débil (especializada en tareas específicas, la actual) y Fuerte (inteligencia general humana, teórica). Además, describe los principales tipos de Aprendizaje Automático: Supervisado (aprende de datos etiquetados para clasificación o regresión), No Supervisado (encuentra patrones en datos no etiquetados mediante *clustering* o reducción de dimensionalidad) y Por Refuerzo (aprende mediante prueba y error interactuando con un entorno y recibiendo recompensas/penalizaciones).

Palabras Clave: Inteligencia artificial, historia de la IA, aprendizaje automático, aprendizaje profundo, aprendizaje supervisado, aprendizaje no supervisado.

Abstract

This chapter outlines a brief history of Artificial Intelligence (AI), from its early days with explicit rule-based systems (1950s-80s) to the rise of Machine Learning (ML) (1980s-2010s), where machines learn from data. It highlights the subsequent takeoff of Deep Learning (DL) (2010s onwards), an ML subfield using deep neural networks, which has been key to advancements in image, audio, and text processing, including the recent development of Large Language Models (LLM). It classifies AI into Weak (specialized in specific tasks, current AI) and Strong (hu-

man-level general intelligence, theoretical). Additionally, it describes the main types of Machine Learning: Supervised (learns from labeled data for classification or regression), Unsupervised (finds patterns in unlabeled data through clustering or dimensionality reduction), and Reinforcement Learning (learns via trial and error by interacting with an environment and receiving rewards/penalties).

Keywords: *Artificial intelligence, history of AI, machine learning, deep learning. supervised learning, unsupervised learning.*

Introducción

La idea de crear máquinas inteligentes no es nueva, pero el campo de la Inteligencia Artificial, tal como lo conocemos hoy, tiene una historia relativamente reciente marcada por periodos de gran optimismo y otros de "invierno de la IA" debido a limitaciones tecnológicas y de financiamiento.

Breve Evolución de la IA:

La historia de la IA se puede trazar a través de diferentes enfoques y paradigmas:

- **Primeros Días y Sistemas Basados en Reglas (Años 50-80):** Los inicios de la IA se centraron en simular el razonamiento humano a través de la lógica simbólica y sistemas basados en reglas. Los programas intentaban seguir reglas explícitas proporcionadas por expertos humanos (ej. "Si el paciente tiene fiebre Y tos, ENTONCES considerar infección respiratoria"). Estos sistemas, conocidos como "sistemas expertos", tuvieron cierto éxito en dominios limitados, pero luchaban con la incertidumbre, la ambigüedad y la necesidad de codificar manualmente un vasto conocimiento del mundo.
- **El Auge del Aprendizaje Automático (*Machine Learning*) (Años 80-2010s):** En lugar de programar reglas explícitas, el foco se trasladó a desarrollar algoritmos que permitieran a las máquinas *aprender* de los datos. El aprendizaje automático se convirtió en el subcampo dominante. Algoritmos como árboles de decisión, máquinas de vectores de soporte (SVM), regresión logística y redes neuronales superficiales ganaron popularidad. La disponibilidad creciente de datos y el aumento del poder computacional impulsaron este progreso.

- **El Despegue del Aprendizaje Profundo (*Deep Learning*) (Años 2010s en adelante):** Un subcampo del aprendizaje automático que utiliza redes neuronales artificiales con múltiples capas (de ahí el término "profundo"). El aprendizaje profundo ha demostrado ser extraordinariamente eficaz en tareas que implican datos no estructurados, como imágenes, audio y texto. Arquitecturas como las redes neuronales convolucionales (CNN) para imágenes y las redes neuronales recurrentes (RNN) y, más recientemente, los Transformers para secuencias (incluido texto), han logrado resultados de vanguardia en muchas áreas.
- **LLM (Principios de los 2020s en adelante):** Un desarrollo reciente y notable dentro del aprendizaje profundo, particularmente utilizando la arquitectura Transformer. Los LLM son modelos entrenados en vastas cantidades de datos de texto y código, capaces de comprender, generar y manipular lenguaje humano a un nivel sin precedentes. Tienen el potencial de transformar tareas relacionadas con la información textual en investigación, como la revisión de literatura, la redacción de borradores y la extracción de información.

Clasificación de la IA:

La IA se puede clasificar de varias maneras, dependiendo de su capacidad y el tipo de aprendizaje que utiliza:

- **IA Débil (Narrow AI) vs. IA Fuerte (General AI):**

  - **IA Débil:** Sistemas diseñados y entrenados para realizar una tarea específica (ej. reconocer imágenes de tumores, traducir texto, jugar ajedrez). La mayoría de la IA que existe hoy en día es IA débil. Es muy buena en su dominio específico, pero no puede realizar tareas fuera de él.
  - **IA Fuerte (o IA General - AGI):** Sistemas con la capacidad intelectual general de un ser humano. Podrían comprender, aprender y aplicar el conocimiento para resolver cualquier problema. La IA fuerte es actualmente un concepto teórico y un objetivo a largo plazo, no una realidad.

- **Tipos de Aprendizaje Automático:** El aprendizaje automático, como motor de gran parte de la IA moderna, se clasifica según el tipo de datos de entrenamiento y el enfoque del aprendizaje:

- **Aprendizaje Supervisado:** El algoritmo aprende de un conjunto de datos "etiquetado", donde cada ejemplo de entrada está asociado con un resultado o "etiqueta" conocido. El objetivo es aprender un mapeo de las entradas a las salidas para poder predecir la etiqueta de nuevas entradas no vistas (ver figura 64).
  - *Tareas Comunes*: **Clasificación** (predecir una categoría, ej. si una imagen muestra un tumor [Sí/No], si un paciente tiene diabetes [Sí/No]) y **Regresión** (predecir un valor numérico continuo, ej. predecir la presión arterial, predecir el tiempo de estancia hospitalaria).

Figura 64. *Aprendizaje Supervisado (Identificación de un Gato)*. Elaborado con GeminiPRO.



Nota. En esta ilustración del aprendizaje supervisado, una maestra primero muestra a una estudiante imágenes etiquetadas de un perro y un gato. Luego, presenta una nueva imagen con ambos animales y le pregunta "¿Cuál es el gato?". La estudiante, basándose en el aprendizaje previo de las etiquetas, señala correctamente al gato. Este proceso ejemplifica cómo un algoritmo de aprendizaje supervisado aprende a clasificar nuevas instancias a partir de datos previamente etiquetados.

- **Aprendizaje No Supervisado:** El algoritmo aprende de un conjunto de datos "no etiquetado", donde solo se proporcionan los datos de entrada

sin resultados conocidos. El objetivo es encontrar patrones, estructuras o relaciones ocultas dentro de los datos (ver figura 65).

- *Tareas Comunes*: **Clustering** (agrupar datos similares, ej. identificar subgrupos de pacientes con características similares), **Reducción de Dimensionalidad** (reducir el número de variables manteniendo la información importante, ej. análisis de componentes principales - PCA), y **Detección de Anomalías** (identificar puntos de datos inusuales).

Figura 65. *Aprendizaje No Supervisado*. Elaborado con GeminiPRO.

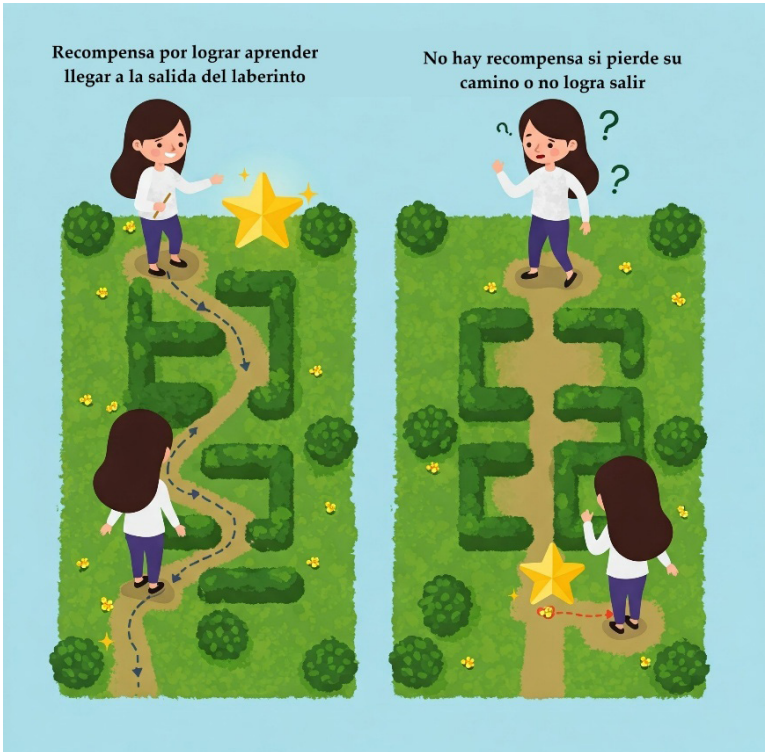


Nota. Agrupación de Animales por Características Similares. Esta ilustración representa el aprendizaje no supervisado. En el panel izquierdo, se muestra una niña en una habitación con una variedad de animales sin ninguna etiqueta o guía previa. En el panel derecho, se observa cómo la niña ha agrupado los animales basándose en sus características similares, como el tamaño, la forma de las orejas o el tipo de pelaje, creando categorías sin que se le hayan proporcionado etiquetas predefinidas. Este proceso refleja cómo un algoritmo de aprendizaje no supervisado identifica patrones y estructuras inherentes en los datos sin la necesidad de información etiquetada.

- **Aprendizaje por Refuerzo:** El algoritmo aprende a tomar decisiones a través de la interacción con un entorno. Recibe una "recompensa" o "penalización" por sus acciones y aprende a maximizar la recompensa a lo largo del tiempo. Es útil para tareas de toma de decisiones secuenciales (ver figura 66).

- *Tareas Comunes*: Robótica, juegos, sistemas de recomendación, optimización de procesos. Aunque menos común en la investigación biomédica directa, tiene potencial en la optimización de protocolos de tratamiento o estrategias de ensayo clínico.

Figura 66. *Aprendizaje por Refuerzo: Aprendiendo a Salir de un Laberinto mediante Recompensas*. Elaborado con GeminiPRO.



Nota. Esta ilustración ejemplifica el aprendizaje por refuerzo. En el panel izquierdo, se muestra a una niña navegando por un laberinto. Cada vez que toma el camino correcto hacia la salida, recibe una estrella como recompensa. En el panel derecho, se observa que cuando toma un camino incorrecto

o se desorienta, no recibe ninguna recompensa y muestra signos de duda. A través de repetidas interacciones y la asociación de acciones con recompensas (o la falta de ellas), la niña aprenderá a identificar el camino óptimo para salir del laberinto. Este proceso ilustra cómo un agente de aprendizaje por refuerzo aprende a tomar decisiones secuenciales para maximizar una recompensa acumulada en un entorno determinado.

Conclusiones

1. **Evolución de la IA:** La IA ha evolucionado desde sistemas basados en reglas codificadas manualmente hacia enfoques donde las máquinas aprenden patrones directamente de los datos (ML y DL).
2. **ML y DL Dominantes:** El Aprendizaje Automático (ML) y su subcampo, el Aprendizaje Profundo (DL), son los motores de la mayoría de las aplicaciones de IA actuales, con DL destacando en tareas con datos no estructurados. Los LLM son un avance reciente significativo dentro de DL.
3. **IA Débil vs. Fuerte:** La IA actual es predominantemente "Débil" o "Estrecha", diseñada para tareas específicas. La IA "Fuerte" o General (AGI), con capacidades cognitivas humanas, sigue siendo un objetivo teórico.
4. **Paradigmas del Aprendizaje Automático:** Existen tres enfoques principales para que las máquinas aprendan:

- **Supervisado:** Aprender un mapeo de entradas a salidas usando datos etiquetados (para clasificación y regresión).
- **No Supervisado:** Descubrir estructura o patrones ocultos en datos no etiquetados (para *clustering*, reducción de dimensionalidad).
- **Por Refuerzo:** Aprender a tomar decisiones óptimas interactuando con un entorno y recibiendo retroalimentación (recompensas/penalizaciones).

CAPÍTULO 23

HERRAMIENTAS CLAVE Y SUS NICHOS

Key Tools and Their Niches

Resumen

Este capítulo explora diversas herramientas de Inteligencia Artificial (IA) que asisten en diferentes etapas del proceso de investigación. Describe herramientas clave y sus nichos específicos: Modelos de Lenguaje Grandes (LLM) como ChatGPT para ideación, redacción de borradores, resumen y edición; herramientas de búsqueda y síntesis de literatura como Elicit y DeepSeek para explorar artículos y extraer información estructurada; herramientas de extracción rápida de ideas clave como Napkin y RapidRabbit para identificar puntos clave en documentos; plataformas de maquetación como Typeset.io para formatear manuscritos y referencias; cuadernos de laboratorio inteligentes como Monica AI y NotebookLM para organizar notas y análisis preliminares; y herramientas de creación visual como Gamma App para generar presentaciones y reportes interactivos. Se mencionan también nichos emergentes como análisis de imágenes médicas y descubrimiento de fármacos. Se enfatiza que estas herramientas son asistentes que aumentan la eficiencia pero requieren supervisión humana y juicio crítico.

Palabras Clave: Herramientas de IA, inteligencia artificial en investigación, modelos de lenguaje grandes (LLM), síntesis de literatura, asistentes de investigación, automatización de la investigación.

Abstract

This chapter explores various Artificial Intelligence (AI) tools that assist in different stages of the research process. It describes key tools and their specific niches: Large Language Models (LLMs) like ChatGPT for ideation, drafting, summarizing, and editing; literature search and synthesis tools like Elicit and DeepSeek for exploring articles and extracting structured information; rapid insight extraction tools like Napkin and RapidRabbit for identifying key points

in documents; layout platforms like Typeset.io for formatting manuscripts and references; smart laboratory notebooks like Monica AI and NotebookLM for organizing notes and preliminary analysis; and visual creation tools like Gamma App for generating interactive presentations and reports. Emerging niches such as medical image analysis and drug discovery are also mentioned. It emphasizes that these tools are assistants that increase efficiency but require human oversight and critical judgment.

Keywords: *AI tools, artificial intelligence in research, large language models (LLMs), literature synthesis, research assistants, research automation.*

La rápida evolución de la IA, particularmente en el campo del procesamiento del lenguaje natural y el aprendizaje automático, ha dado lugar a una proliferación de herramientas diseñadas para asistir en tareas que antes requerían una considerable inversión de tiempo y esfuerzo humano. Desde la ideación inicial hasta la redacción y difusión de resultados, la IA está encontrando aplicaciones valiosas en casi todos los aspectos de la investigación.

A continuación, exploramos algunas de estas herramientas y los roles específicos (nichos) que están desempeñando en el ecosistema de la investigación (ver tabla 27):

- **ChatGPT (y otros Modelos de Lenguaje Grandes - LLM):**
  - **Descripción:** Modelos de lenguaje avanzados capaces de comprender y generar texto similar al humano, basados en arquitecturas de aprendizaje profundo como los Transformers.
  - **Usos en Investigación:**
    - Brainstorming de ideas
    - Redacción de objetivos
    - Generación de borradores de protocolo
    - Resumen de textos
    - Revisión y edición de texto
    - Traducción
- **Elicit y DeepSeek (y otras herramientas de búsqueda y síntesis de literatura basadas en IA):**

- **Descripción:** Plataformas diseñadas específicamente para la exploración y síntesis de literatura académica utilizando IA. Analizan artículos para extraer información clave y presentarla de forma estructurada.
- **Usos en Investigación:**
  - Búsqueda de literatura
  - Síntesis de literatura académica
  - Identificación de lagunas en el conocimiento
  - Organización de referencias

- **Napkin y RapidRabbit (y otras herramientas de extracción rápida de ideas clave):**
  - **Descripción:** Herramientas diseñadas para procesar documentos de texto (como artículos PDF) y extraer rápidamente los puntos clave, argumentos principales o datos relevantes.
  - **Usos en Investigación:**
    - Extracción rápida de ideas clave de papers
    - Identificación de metodologías clave
    - Generación de notas y resúmenes concisos
- **Typeset.io (y otras herramientas de maquetación y formato):**
  - **Descripción:** Plataformas que utilizan IA y automatización para ayudar en la maquetación y el formato de manuscritos de investigación según los requisitos de diferentes revistas o estilos de citación.
  - **Usos en Investigación:**
    - Maquetación de manuscritos con plantillas
    - Gestión de referencias y citas
    - Verificación de formato
- **Monica AI y NotebookLM (y otras herramientas de cuadernos de laboratorio inteligentes):**
  - **Descripción:** Herramientas que actúan como asistentes o cuadernos de

laboratorio digitales, a menudo integrando capacidades de IA para procesar notas, documentos y datos.

- **Usos en Investigación:**

- Organización de notas y experimentos
- Análisis preliminar de datos
- Generación de resúmenes de experimentos
- Conexión de ideas

- **Gamma App (y otras herramientas de creación de presentaciones y reportes interactivos):**

- **Descripción:** Herramientas que utilizan IA para facilitar la creación de presentaciones, documentos o reportes visualmente atractivos e interactivos a partir de texto o esquemas.

- **Usos en Investigación:**

- Creación rápida de presentaciones
- Diseño de reportes interactivos
- Visualización de datos

Tabla 27. *Resumen de Herramientas Clave de IA para Investigación y sus Aplicaciones*

Herramienta/ Categoría Principal	Descripción Principal	Usos Clave en Investigación
ChatGPT (y otros LLM)	Modelos de lenguaje generativos	Ideación, redacción (borradores, objetivos), resumen, edición de texto, traducción.
Elicit, DeepSeek	Búsqueda y síntesis de literatura académica	Exploración de literatura, extracción estructurada de información de artículos, identificación de lagunas.
Napkin, RapidRabbit	Extracción rápida de ideas clave de documentos	Obtención rápida de puntos clave y hallazgos principales de artículos.

Herramienta/ Categoría Principal	Descripción Principal	Usos Clave en Investigación
Typeset.io	Maquetación y formato de manuscritos	Adaptación a formatos de revistas, gestión de referencias, verificación de formato.
Monica AI, NotebookLM	Cuadernos de laboratorio inteligentes / Asistentes de notas	Organización de notas y experimentos, análisis preliminar, resumen de experimentos, conexión de ideas.
Gamma App	Creación de presentaciones y reportes visuales	Generación rápida de presentaciones y reportes interactivos, asistencia en visualización.

Otras Herramientas y Nichos Emergentes: La lista de herramientas de IA en investigación está en constante crecimiento. Otros nichos incluyen:

- Análisis de Imágenes Médicas
- Descubrimiento de Fármacos
- Análisis de Datos Genómicos
- Procesamiento de Lenguaje Natural Clínico (Clinical NLP)
- Diseño Experimental Automatizado
- Revisión por Pares Asistida por IA

La integración de estas herramientas de IA en el flujo de trabajo de investigación tiene el potencial de aumentar la eficiencia, acelerar el descubrimiento y permitir a los investigadores abordar preguntas más complejas. Sin embargo, es crucial recordar que estas herramientas son asistentes; el juicio crítico del investigador, la comprensión de los fundamentos científicos y la supervisión humana siguen siendo indispensables. La elección de la herramienta adecuada dependerá de la tarea específica, los datos disponibles y los objetivos del investigador.

Conclusiones

1. **Diversidad de Herramientas:** Existe una creciente variedad de herramientas de IA diseñadas para apoyar tareas específicas a lo largo del ciclo de investigación.

2. **Nicho de Aplicación:** Cada herramienta o categoría de herramienta (LLM, búsqueda de literatura, extracción de ideas clave, formato, notas, presentaciones) tiene aplicaciones particulares donde puede ser más útil.

3. **Potencial de Eficiencia:** Estas herramientas tienen el potencial de aumentar significativamente la eficiencia en tareas como la revisión de literatura, redacción de borradores, formato y creación de presentaciones.

4. **Asistentes, No Sustitutos:** Es crucial entender que estas herramientas de IA son asistentes y no reemplazan la necesidad de comprensión profunda, pensamiento crítico y juicio experto por parte del investigador.

5. **Supervisión Humana Esencial:** Los resultados generados por IA deben ser revisados, validados y editados por humanos para asegurar su precisión, relevancia y adecuación al contexto científico.

6. **Campo en Evolución:** El panorama de herramientas de IA para investigación está en constante desarrollo, con nuevas aplicaciones y capacidades emergiendo continuamente.

CAPÍTULO 24

DESDE LA IDEA AL PROTOCOLO

From Idea to Protocol

Resumen

Este capítulo detalla cómo las herramientas de Inteligencia Artificial (IA) pueden facilitar la transición desde una idea inicial de investigación hasta un protocolo estructurado. Explica que la IA, especialmente los LLM como ChatGPT, puede asistir en la generación y refinamiento de preguntas de investigación (usando marcos como PCC para *scoping reviews*), la definición de objetivos específicos y variables relevantes. Demuestra, mediante un ejemplo de análisis bibliométrico sobre "Gemelo Digital" en salud, cómo la IA puede ayudar a buscar literatura existente, analizar su alcance, identificar lagunas y validar la factibilidad de un tema novedoso. Describe también cómo los LLM pueden ayudar a estructurar el protocolo de investigación siguiendo guías estándar (ej. PRISMA-ScR). Además, resalta el uso de herramientas de búsqueda y extracción rápida de ideas clave (como Elicit, DeepSeek, Napkin) para validar el alcance y la viabilidad de un proyecto mediante la exploración preliminar eficiente de la literatura. Se subraya que la IA actúa como un asistente, pero el juicio crítico y la experiencia humana son indispensables.

Palabras Clave: Protocolo de investigación, pregunta de investigación, inteligencia artificial, revisión de alcance, análisis bibliométrico, herramientas de IA en investigación.

Abstract

This chapter details how Artificial Intelligence (AI) tools can facilitate the transition from an initial research idea to a structured protocol. It explains that AI, especially Large Language Models (LLM) like ChatGPT, can assist in generating and refining research questions (using frameworks like PCC for scoping reviews), defining specific objectives, and relevant variables. It demonstrates, through

an example of a bibliometric analysis on "Digital Twin" in health, how AI can help search existing literature, analyze its scope, identify gaps, and validate the feasibility of a novel topic. It also describes how LLM can help structure the research protocol following standard guidelines (e.g., PRISMA-ScR). Furthermore, it highlights the use of AI-powered search and rapid insight extraction tools (like Elicit, DeepSeek, Napkin) to validate a project's scope and feasibility through efficient preliminary literature exploration. It underscores that AI acts as an assistant, but human critical judgment and expertise remain indispensable.

Keywords: *Research protocol, research question, artificial intelligence, scoping review, bibliometric analysis, AI research tools.*

Introducción

El inicio de cualquier proyecto de investigación es a menudo una fase creativa pero desafiante. Transformar una idea general en una pregunta de investigación clara y un protocolo riguroso requiere pensamiento crítico, conocimiento de la literatura y atención al detalle. Las herramientas de IA pueden actuar como poderosos copilotos en este proceso, ayudando a refinar conceptos, explorar la viabilidad y estructurar la planificación.

3.1. Generación y Refinamiento de la Pregunta de Investigación

Una pregunta de investigación bien formulada es la base de un estudio exitoso. Debe ser clara, concisa y enfocada. A menudo, comenzamos con un área de interés amplia y necesitamos delimitarla.

- **Cómo la IA puede ayudar:** Herramientas como **ChatGPT** (u otros LLM) son excelentes para el *brainstorming* y la exploración conceptual. Puedes presentarle a la IA una idea general o un área de interés y pedirle que genere diferentes preguntas de investigación potenciales, explore ángulos relacionados o sugiera cómo delimitar el tema.
- **Ejemplo (Contexto: Revisión de alcance):**
  - **Idea Inicial:** "Quiero investigar sobre el uso de la telemedicina en pacientes mayores."
  - **Prompt a ChatGPT:** "Estoy planeando una revisión de alcance sobre telemedicina en pacientes mayores. Mi interés inicial es amplio. ¿Podrías ayudarme a generar diferentes preguntas de investigación más específicas usando el formato PCC, adecuado para revisión de alcance?"

- **Respuesta Potencial de ChatGPT:** ChatGPT podría sugerir preguntas como:
  - "¿Cuál es el *alcance* de la evidencia sobre la *aceptabilidad* de la *telemedicina* en *pacientes mayores*?" (P: Pacientes mayores, C: Aceptabilidad de la telemedicina, C: N/A)
  - "¿Qué *tipos de tecnologías de telemedicina* se han *utilizado* para el *manejo de enfermedades crónicas* en *pacientes mayores en entornos rurales*?" (P: Pacientes mayores con enfermedades crónicas, C: Tipos de tecnologías de telemedicina utilizadas, C: Entornos rurales)
  - "¿Cuáles son las *barreras y facilitadores* para la *implementación de servicios de telemedicina* dirigidos a *pacientes mayores en atención primaria*?" (P: Pacientes mayores, C: Barreras y facilitadores de la implementación de telemedicina, C: Atención primaria)
- **Refinamiento:** Puedes refinar estas sugerencias interactuando con la IA, pidiendo que se enfoque en un aspecto particular (ej. solo monitoreo remoto, solo salud mental) o que ajuste la población/contexto.

3.2. Ejemplo de Búsqueda de Idea de Investigación en Gemini Pro

En este ejemplo, utilizaremos la *IA Gemini Pro Deep Research* para desarrollar un tema novedoso y factible para un análisis bibliométrico sobre "Digital Twin" (Gemelo Digital) en el ámbito de la salud.

Idea Inicial: Realizar un análisis bibliométrico sobre el uso de Gemelos Digitales en Salud.

Objetivo con Asistencia IA o promt: Refinar esta idea inicial, identificar lagunas en la literatura bibliométrica existente y proponer un tema específico, novedoso y factible para un nuevo estudio bibliométrico.

Proceso Guiado por la IA:

1. Búsqueda de Artículos Bibliométricos Existentes

- **Cómo la IA puede ayudar:** La IA puede ejecutar búsquedas complejas y rápidas en bases de datos científicas (o simular el acceso a sus índices) para identificar estudios que cumplan criterios específicos.
- **Prompt/Tarea para la IA (Implícita en la solicitud del usuario):** "Busca artículos científicos que sean específicamente análisis bibliomé-

tricos sobre 'digital twin' OR 'gemelo digital' AND ('health' OR 'healthcare' OR 'medicine')."

- **Simulación de Resultados de Búsqueda:** Se esperaría encontrar varios estudios bibliométricos publicados en los últimos años, dada la creciente popularidad del tema. Títulos potenciales podrían ser:
  - "A Bibliometric Review of Digital Twin in Healthcare: Trends and Future Directions"
  - "Mapping the Research Landscape of Digital Twin in Medicine: A Bibliometric Study"
  - "Gemelos Digitales en la Gestión Hospitalaria: Un Análisis Bibliométrico"
  - "Research Hotspots and Trends of Digital Twin in Health from 20XX to 20YY: A Bibliometric Analysis"

2. Análisis de los Artículos Bibliométricos Encontrados

- **Cómo la IA puede ayudar:** La IA puede procesar resúmenes y textos completos de los artículos encontrados para extraer y sintetizar información clave sobre su alcance, metodología y hallazgos.
- **Prompt/Tarea para la IA:** "Analiza los artículos bibliométricos identificados sobre 'digital twin' en salud y extrae: alcance temático, bases de datos usadas, periodos cubiertos, metodologías y principales conclusiones/tendencias."
- **Simulación de Análisis (Basado en la naturaleza esperada de estos estudios):**
  - **(a) Alcance Temático:** La mayoría de los estudios bibliométricos iniciales en campos emergentes tienden a ser amplios, cubriendo aplicaciones diversas: diagnóstico, tratamiento, gestión hospitalaria, medicina personalizada, simulación quirúrgica, monitoreo de pacientes, etc. Algunos pueden tener un ligero sesgo hacia la gestión o la aplicación general, pero pocos se centrarían *exclusivamente* en una subdisciplina médica muy específica.
  - **(b) Bases de Datos Consultadas:** Predominantemente Scopus y/o Web of Science, que son estándar para análisis bibliométricos amplios y cu-

bren una vasta gama de revistas científico-técnicas. PubMed podría ser incluido en algunos, pero no siempre es la fuente principal para temas con fuerte componente tecnológico como DT.

- **(c) Periodos de Tiempo Analizados:** Desde los primeros artículos relevantes (quizás 2018-2019, aunque el concepto es anterior, su aplicación explícita en salud es más reciente) hasta el año de la búsqueda o el año anterior a la publicación del estudio (ej. 2018-2022, 2019-2023).
- **(d) Metodologías Bibliométricas Utilizadas:** Análisis de concitación (identificar autores influyentes, revistas clave), análisis de coocurrencia de palabras clave (identificar temas emergentes y clústeres), análisis de rendimiento (producción por país, institución, autor; evolución temporal), visualización de redes (VOSviewer, CiteSpace). Son metodologías estándar para análisis bibliométricos en campos técnicos.
- **(e) Principales Conclusiones y Tendencias:** Alto crecimiento en el número de publicaciones, identificación de clústeres temáticos (ej. enfermedades crónicas, simulación, IA), países e instituciones líderes, revistas más relevantes, y mención de desafíos y oportunidades (datos, validación, ética).

3. Identificación de Lagunas o Áreas No Cubiertas

- **Cómo la IA puede ayudar:** La IA puede comparar el alcance de los estudios existentes con el amplio espectro de aplicaciones potenciales del gemelo digital en salud, señalando áreas que parecen subrepresentadas en los *análisis bibliométricos* específicos.
- **Prompt/Tarea para la IA:** "Basado en el análisis, ¿qué subáreas de la salud o enfoques de aplicación del gemelo digital no han sido el foco *principal* de los análisis bibliométricos existentes?"
- **Simulación de Identificación de Lagunas:** A pesar de que los estudios amplios *mencionan* diversas aplicaciones, es probable que falten análisis bibliométricos *profundamente enfocados* en:
  - DT aplicado **exclusivamente** a **diagnóstico por imagen y radiología**.
  - DT aplicado **exclusivamente** a **planificación y simulación quirúrgica**.

- DT aplicado **exclusivamente** a **farmacología personalizada y predicción de respuesta a fármacos** (modelos farmacocinéticos/farmacodinámicos - PK/PD basados en DT).
- DT aplicado a la gestión de **ensayos clínicos**.
- DT aplicado a la **salud mental**.

Estas son áreas muy específicas donde el modelado (núcleo del DT) es crucial, pero que podrían no haber sido desglosadas en análisis bibliométricos separados.

4. Verificación de Factibilidad (Búsqueda de Literatura Primaria)

- **Cómo la IA puede ayudar:** La IA puede realizar búsquedas rápidas de artículos *no bibliométricos* en las áreas identificadas como lagunas para estimar el volumen de investigación primaria existente. Esto valida si hay suficiente material para un análisis bibliométrico enfocado.
- **Prompt/Tarea para la IA:** "Para las lagunas identificadas (ej. DT en radiología, DT en cirugía, DT en farmacología), busca artículos de investigación primaria (no bibliométricos) para ver si hay un número significativo de publicaciones."
- **Simulación de Resultados de Búsquedas Primarias:**
  - Búsqueda: "digital twin" radiology imaging -> Resultaría en artículos sobre creación de gemelos digitales de órganos a partir de imágenes médicas, simulación de procedimientos radiológicos, etc. *Resultado: Sí, hay artículos.*
  - Búsqueda: "digital twin" surgical planning simulation -> Resultaría en artículos sobre gemelos digitales para planificar cirugías complejas, entrenar cirujanos, simular resultados. *Resultado: Sí, hay artículos.*
  - Búsqueda: "digital twin" personalized pharmacology "drug response" -> Resultaría en artículos sobre modelos basados en datos de pacientes para predecir cómo responderán a tratamientos específicos, optimización de dosis, etc. *Resultado: Sí, hay artículos, y esta área (PK/PD modeling + patient data) se presta bien al concepto de DT.*

La simulación de estas búsquedas primarias confirma que existe un cuerpo de literatura no bibliométrica suficiente en varias de las áreas identificadas como posibles lagunas para justificar un análisis bibliométrico enfocado.

5. Síntesis y Propuesta de Tema Novedoso y Factible

- **Cómo la IA puede ayudar:** La IA sintetiza los hallazgos de las etapas anteriores (qué se ha hecho bibliométricamente, qué falta, qué es factible) para formular una propuesta de tema específica.
- **Prompt/Tarea para la IA:** "Basado en el análisis de la literatura bibliométrica existente y la verificación de literatura primaria, propone un tema para un nuevo análisis bibliométrico sobre 'digital twin' en salud que sea novedoso y factible."
- **Síntesis y Propuesta de Tema:**
  - Los análisis bibliométricos existentes sobre "Digital Twin en Salud" tienden a ser amplios.
  - Áreas específicas como la **farmacología personalizada** y la **predicción de respuesta a fármacos** mediante gemelos digitales son subcampos activos pero no suelen ser el foco *principal* de los análisis bibliométricos generales.
  - Existe suficiente literatura primaria que combina los conceptos de "digital twin", "pharmacology", "personalized medicine", y "drug response" para sustentar un análisis bibliométrico detallado.
  - Este enfoque es relevante porque el gemelo digital tiene un potencial significativo para revolucionar cómo se diseñan y aplican los tratamientos farmacológicos a nivel individual.
- **Tema Propuesto: "Análisis Bibliométrico del Uso de Gemelos Digitales en Farmacología Personalizada y Predicción de Respuesta a Fármacos"**

Justificación del Tema Propuesto:

- **Novedoso:** Si bien los análisis amplios mencionan la medicina personalizada, un estudio dedicado *exclusivamente* al análisis bibliométrico de la intersección entre Gemelos Digitales y Farmacología Personalizada/Predicción de Respuesta a Fármacos es probable que no se haya publica-

do como tema *central* en los estudios bibliométricos generales encontrados. Permite una inmersión profunda en autores, revistas, temas emergentes y colaboraciones específicas de esta subdisciplina.

- **Factible:** La verificación mediante búsqueda de literatura primaria (Paso 4) confirma que hay un número creciente de publicaciones en esta área específica, proporcionando una base de datos suficiente para realizar un análisis bibliométrico significativo.
- **Relevante:** La farmacología personalizada es un campo de gran interés, y los gemelos digitales representan una tecnología de vanguardia con potencial transformador en este ámbito.

Este proceso, simulando la interacción con una IA capaz de buscar, analizar y sintetizar información de la literatura científica, nos ha permitido pasar de una idea general a un tema de investigación específico, justificado como novedoso y factible para un análisis bibliométrico. Los siguientes pasos serían refinar la pregunta de investigación exacta y desarrollar el protocolo detallado para la búsqueda y el análisis bibliométrico.

3.3. Definición de Objetivos y Variables

Una vez que la pregunta de investigación está tomando forma, es necesario definir los objetivos específicos del estudio y las variables clave que se medirán o extraerán.

- **Cómo la IA puede ayudar:** Los LLM pueden ayudar a desglosar la pregunta de investigación en objetivos específicos y a pensar en las variables relevantes que deben ser consideradas para responder a esos objetivos.
- **Ejemplo (Continuando con la revisión de alcance):**
  - **Pregunta Refinada:** «¿Cuál es el alcance de la evidencia sobre la *aceptabilidad* de la *telemedicina* en *pacientes mayores*?"
  - **Prompt a ChatGPT:** "Para esta revisión de alcance sobre la aceptabilidad de la telemedicina en pacientes mayores, ¿cuáles serían los objetivos específicos que debería plantearme? ¿Y qué variables o tipos de datos sobre la aceptabilidad debería buscar en los estudios?"

- **Respuesta Potencial de ChatGPT:**

- **Objetivos:**

- Identificar los estudios que han evaluado la aceptabilidad de la telemedicina en pacientes mayores.
- Describir las metodologías utilizadas para evaluar la aceptabilidad en estos estudios.
- Sintetizar los factores reportados que influyen en la aceptabilidad (positivos y negativos).
- Identificar las lagunas en la investigación sobre la aceptabilidad de la telemedicina en esta población.
- **Variables/Datos a Extraer:**
- Diseño del estudio
- Población de pacientes mayores (edad promedio, comorbilidades, nivel educativo)
- Tipo de tecnología de telemedicina utilizada
- Aspectos de la aceptabilidad evaluados (ej. facilidad de uso, satisfacción, confianza, disposición a usar)
- Métodos de recolección de datos sobre aceptabilidad (ej. encuestas, entrevistas, datos de uso)
- Factores que influyen en la aceptabilidad (ej. apoyo técnico, costo, usabilidad de la interfaz, relación con el clínico)

3.4. Estructuración del Protocolo/Metodología

Un protocolo detallado es esencial para garantizar la transparencia, reproducibilidad y rigor metodológico de la investigación. Para revisión de alcance, existen guías y checklists (ej. PRISMA-ScR).

- **Cómo la IA puede ayudar:** Los LLM pueden ayudar a generar la estructura inicial de un protocolo y a redactar secciones. Las herramientas de búsqueda de literatura basadas en IA pueden ayudar a encontrar las guías metodológicas y checklists relevantes.

- **Ejemplo (Continuando con la *Scoping Review*):**

  - **Prompt a ChatGPT:** "Necesito estructurar el protocolo para mi revisión de alcance. ¿Podrías proporcionarme un esquema típico de protocolo para una revisión de alcance, incluyendo las secciones clave según las guías metodológicas?"

  - **Respuesta Potencial de ChatGPT:** ChatGPT podría ofrecer un esquema con secciones como: Introducción (justificación, pregunta), Objetivos, Criterios de Elegibilidad (PCC, Tipos de Fuentes), Estrategia de Búsqueda (fuentes de datos, términos de búsqueda), Proceso de Selección de Estudios, Extracción de Datos, Síntesis de Resultados, y Discusión.

  - **Uso de Elicit/DeepSeek:** Puedes usar Elicit o DeepSeek para buscar artículos sobre "cómo realizar una revisión de alcance" o "guías metodológicas revisión de alcance" para encontrar ejemplos de protocolos publicados o las checklists como PRISMA-ScR.

  - **Redacción de Secciones:** Puedes usar ChatGPT para redactar borradores de secciones específicas, como la justificación o la descripción de los criterios de elegibilidad, basándote en la información que le proporciones.

3.5. Validación de Alcance y Factibilidad

Antes de embarcarse completamente en la revisión, es crucial validar que la pregunta y el alcance son apropiados y que es factible realizar la revisión con los recursos disponibles.

- **Cómo la IA puede ayudar:** Las herramientas de búsqueda y extracción rápida de ideas clave son invaluables aquí.

- **Ejemplo (Continuando con la revisión de alcance):**

  - **Uso de Elicit/DeepSeek:** Realiza búsquedas preliminares con tu pregunta de investigación propuesta. ¿Cuántos estudios relevantes aparecen? ¿Parece que hay suficiente literatura para una revisión de alcance, o es un tema demasiado nuevo o demasiado amplio?

  - **Uso de Napkin/RapidRabbit:** Descarga algunos de los artículos más relevantes encontrados por Elicit/DeepSeek y utiliza Napkin o RapidRabbit

para extraer rápidamente sus resúmenes, métodos y hallazgos principales. ¿Están los estudios abordando tu concepto y población de la manera que esperabas? ¿Hay mucha heterogeneidad en cómo se define o mide la aceptabilidad? Además, Napkin te permite elaborar gráficos tipo esquema para facilitar la transmisión o comunicación de la información.

- **Evaluación del Alcance:** Si Elicit/DeepSeek encuentran miles de estudios, tu alcance podría ser demasiado amplio. Si encuentran muy pocos, podría ser demasiado específico o el tema es incipiente. RapidRabbit te dan una idea rápida del contenido de los estudios para ver si se alinean con tu concepto.
- **Evaluación de la Factibilidad:** La cantidad de estudios y la complejidad de la extracción de datos (evaluada rápidamente con RapidRabbit o un LLM) te dan una idea de cuánto tiempo y esfuerzo requerirá la revisión, ayudando a evaluar si es factible con tu equipo y plazos.

La integración estratégica de herramientas de IA en estas etapas iniciales puede acelerar significativamente el proceso, mejorar la calidad de la planificación y ayudar a los investigadores a tomar decisiones más informadas sobre la dirección de su estudio. Sin embargo, el juicio humano y la experiencia siguen siendo insustituibles para la interpretación final y la toma de decisiones críticas.

Conclusiones

1. **Asistencia en Fases Iniciales:** Las herramientas de IA pueden ser valiosos asistentes en las etapas tempranas de la investigación, ayudando a transformar ideas generales en planes de estudio concretos.
2. **Refinamiento de Preguntas y Objetivos:** Los LLM (ej. ChatGPT) son útiles para brainstorming, generar preguntas de investigación específicas (utilizando marcos como PCC) y definir objetivos claros y variables relevantes.
3. **Exploración de Literatura y Lagunas:** La IA (ej. Elicit, DeepSeek, Gemini Pro) puede analizar eficientemente la literatura existente (incluyendo estudios bibliométricos previos) para identificar el estado del arte, detectar lagunas en el conocimiento y proponer temas de investigación novedosos.

4. **Validación de Factibilidad:** Herramientas de IA permiten realizar búsquedas preliminares rápidas para estimar el volumen de literatura disponible sobre un tema específico, ayudando a evaluar la factibilidad de un proyecto (ej. una revisión sistemática o bibliométrica).
5. **Estructuración de Protocolos:** Los LLM pueden generar esquemas iniciales de protocolos de investigación basados en guías metodológicas estándar (ej. PRISMA-ScR), que luego el investigador refina.
6. **Rol de Co-piloto:** La IA actúa como un co-piloto en este proceso, acelerando tareas y ofreciendo nuevas perspectivas, pero la dirección estratégica, el juicio crítico y la validación final dependen del investigador humano.

CAPÍTULO 25

ANÁLISIS Y VISUALIZACIÓN ASISTIDOS POR IA

AI-Assisted Analytics and Visualization

Resumen

Este capítulo explora cómo la IA puede asistir en las fases de análisis y visualización de datos en investigación. Describe que, aunque el pretratamiento detallado se vio antes, la IA puede sugerir pasos automáticos de limpieza y transformación basados en un análisis preliminar de los datos. Detalla cómo la IA, especialmente los LLM, puede generar código para crear visualizaciones efectivas (diagramas de caja, dispersión, curvas ROC, etc.) en lenguajes como Python o R, o directamente generar gráficos en plataformas integradas a partir de descripciones en lenguaje natural (*prompts*). Además, explica cómo la IA puede ayudar en la interpretación de resultados estadísticos complejos o salidas de modelos (ej. coeficientes de regresión, importancia de variables), generando resúmenes automáticos y explicaciones en lenguaje más accesible. Se enfatiza que la IA actúa como un asistente que potencia la eficiencia, pero no reemplaza la comprensión estadística fundamental ni el juicio crítico del investigador.

Palabras Clave: Análisis de datos con IA, visualización de datos con IA, interpretación de resultados, generación de código, inteligencia artificial, asistentes de análisis.

Abstract

This chapter explores how AI can assist in the data analysis and visualization phases of research. It describes that, although detailed preprocessing was covered earlier, AI can suggest automated cleaning and transformation steps based on a preliminary data analysis. It details how AI, especially LLM, can generate code to create effective visualizations (box plots, scatter plots, ROC curves, etc.) in languages like Python or R, or directly generate graphs in integrated platforms from natural language descriptions (prompts). Furthermore, it explains

how AI can aid in interpreting complex statistical results or model outputs (e.g., regression coefficients, variable importance) by generating automatic summaries and explanations in more accessible language. It emphasizes that AI acts as an assistant that enhances efficiency but does not replace the researcher's fundamental statistical understanding or critical judgment.

Keywords: *AI data analysis, AI data visualization, results interpretation, code generation, artificial intelligence, analysis assistants.*

Introducción

La fase de análisis de datos es donde transformamos los datos procesados en información y conocimiento. Implica aplicar métodos estadísticos y computacionales para explorar los datos, probar hipótesis, construir modelos y extraer ideas clave. La visualización es una herramienta indispensable en este proceso, permitiendo comprender distribuciones, identificar relaciones y comunicar hallazgos de manera efectiva. Las herramientas de IA están emergiendo como asistentes valiosos en ambas áreas.

4.1. Preparación de Datos: Sugerencias Automáticas de Limpieza y Transformación

Aunque ya se cubrió en detalle el pretratamiento de datos, algunas herramientas de IA (o plataformas de análisis con capacidades de IA integradas) pueden ofrecer asistencia en esta etapa, particularmente al sugerir pasos de limpieza o transformación basados en el análisis preliminar de los datos.

- **Cómo la IA puede ayudar:** Algunas plataformas de análisis de datos con IA pueden escanear un conjunto de datos e identificar problemas potenciales como valores faltantes, *outliers* o distribuciones sesgadas. Luego, pueden sugerir métodos apropiados para manejar estos problemas. Los LLM también pueden ayudar a entender por qué ciertos pasos de pretratamiento son necesarios para tipos de análisis específicos.
- **Ejemplo:**
    - **Escenario:** Tienes un conjunto de datos con una variable de ingresos que está fuertemente sesgada hacia la derecha y planeas usar un modelo de regresión lineal.

- **Sugerencia de IA (Plataforma con IA):** Una plataforma podría detectar el sesgo en la variable de ingresos y sugerir automáticamente aplicar una transformación logarítmica para hacer su distribución más simétrica, lo cual es una asunción común en la regresión lineal.
- **Uso de LLM (ChatGPT):** Puedes describir tu variable y el modelo que planeas usar a un LLM: "Tengo una variable 'Ingresos del Paciente' que está muy sesgada. Voy a usar regresión lineal. ¿Qué transformaciones de datos me recomiendas aplicar a esta variable y por qué?" La IA podría explicar el problema del sesgo para la regresión lineal y sugerir transformaciones como el logaritmo, la raíz cuadrada o la transformación de Box-Cox, explicando la justificación estadística detrás de cada una.
- **Sugerencia de Limpieza:** Si la IA detecta un alto porcentaje de valores faltantes en una columna, podría sugerir opciones como imputación por la media/mediana o explorar si la eliminación de la columna es viable basándose en la cantidad de datos faltantes.

4.2. Generación de Gráficas: Prompts para Visualizaciones Efectivas

La visualización de datos es crucial para el análisis exploratorio y la comunicación de resultados. La IA puede ayudar a generar código para gráficos o a crear visualizaciones directamente a partir de descripciones en lenguaje natural.

- **Cómo la IA puede ayudar:** Los LLM pueden generar código (ej. en Python con bibliotecas como Matplotlib o Seaborn, o en R con ggplot2) para crear gráficos específicos basándose en tu descripción de los datos y el tipo de visualización deseada. Algunas herramientas de visualización con IA integrada permiten crear gráficos simplemente describiendo lo que quieres ver.
- **Ejemplo:**
- **Escenario:** Tienes un conjunto de datos de pacientes con variables como 'Edad', 'IMC', 'Nivel de Colesterol' y 'Género'. Quieres visualizar la distribución de IMC por género y la relación entre Edad y Nivel de Colesterol.
- **Prompt a ChatGPT (Generación de Código Python):** "Tengo un DataFrame de pandas con columnas 'Edad', 'IMC', 'Nivel_Colesterol' y 'Genero'. Quiero generar dos gráficos: 1) Un diagrama de caja del IMC

para cada género. 2) Un gráfico de dispersión del Nivel_Colesterol versus Edad. Por favor, proporciona el código Python usando Seaborn y Matplotlib."

- **Respuesta Potencial de ChatGPT:** ChatGPT generaría el código Python completo para cargar datos de ejemplo (o usar tus datos si los proporcionas en un formato adecuado) y generar los dos gráficos solicitados, incluyendo etiquetas de ejes y títulos.
- **Prompt a Herramienta de Visualización con IA:** Si usas una herramienta como Tableau o Power BI con capacidades de lenguaje natural, podrías simplemente escribir: "Mostrar diagrama de caja de IMC por Género" o "Graficar Nivel de Colesterol vs Edad como dispersión". La herramienta interpretaría el *prompt* y generaría la visualización.
- **Generación de Curvas ROC:** Para evaluar el rendimiento de un modelo de clasificación (como vimos en el capítulo de diagnóstico), necesitas una curva ROC.
- **Prompt a ChatGPT (Generación de Código Python para Curva ROC):** "Tengo las predicciones de probabilidad de mi modelo ('probabilidades_predichas') y las etiquetas reales ('etiquetas_reales') para un problema de clasificación binaria. Genera el código Python usando scikit-learn y Matplotlib para calcular y graficar la curva ROC y mostrar el valor del AUC."

4.3. Interpretación de Resultados: Resúmenes Automáticos y Explicaciones de Modelos

Entender lo que significan los resultados estadísticos o los parámetros de un modelo complejo puede ser desafiante. La IA puede ayudar a resumir hallazgos clave y proporcionar explicaciones en lenguaje más accesible.

- **Cómo la IA puede ayudar:** Los LLM pueden analizar salidas de software estadístico o descripciones de modelos y generar resúmenes concisos, explicar términos técnicos o interpretar el significado práctico de los resultados (ej. qué significa un coeficiente de regresión o un Hazard Ratio).
- **Ejemplo:**
- **Escenario:** Has ejecutado un modelo de regresión logística para predecir la probabilidad de desarrollar diabetes y obtienes una tabla de coeficientes, errores estándar, valores p e intervalos de confianza.

- **Prompt a ChatGPT (Interpretación de Regresión Logística):** "Aquí está la salida de mi modelo de regresión logística [pegar la tabla de salida]. La variable dependiente es 'Diabetes' (1=Sí, 0=No). Las variables independientes son 'Edad', 'IMC' y 'Historial_Familiar_Diabetes' (1=Sí, 0=No). Por favor, resume los hallazgos clave, explicando qué variables son predictores significativos de diabetes y cómo se interpreta el Odds Ratio para cada una."

- **Respuesta Potencial de ChatGPT:** La IA identificaría las variables con valores $p < 0.05$ (o el umbral de significancia especificado), indicando que son predictores significativos. Explicaría que el coeficiente para 'Edad' indica cómo cambia el log-odds de diabetes por cada año adicional, y convertiría el coeficiente a un Odds Ratio (OR), interpretando el OR (ej. "Por cada año adicional de edad, los odds de desarrollar diabetes aumentan en X%, manteniendo otras variables constantes"). Haría lo mismo para IMC e Historial Familiar de Diabetes.

- **Escenario:** Has entrenado un modelo de aprendizaje automático complejo (ej. un bosque aleatorio) para predecir la respuesta a un tratamiento y quieres entender qué variables fueron más importantes para la predicción.

- **Prompt a ChatGPT (Explicación de Importancia de Variables):** "He entrenado un modelo de bosque aleatorio para predecir la respuesta al tratamiento. El modelo me dio una lista de 'importancia de las variables' [pegar la lista]. ¿Cómo interpreto esta lista? ¿Qué significa que una variable tenga una puntuación de importancia alta?"

- **Respuesta Potencial de ChatGPT:** La IA explicaría que la importancia de la variable indica cuánto contribuyó esa variable a la precisión predictiva del modelo. Una puntuación alta significa que la variable fue muy útil para el modelo al tomar sus decisiones de predicción. Sin embargo, también podría añadir una advertencia importante: la importancia de la variable en el modelo no necesariamente implica causalidad.

La IA en el análisis y la visualización no reemplaza la necesidad de comprender los principios estadísticos y la interpretación clínica experta. Más bien, actúa como un asistente inteligente que puede automatizar tareas repetiti-

vas, generar borradores de código, sugerir enfoques y proporcionar explicaciones iniciales, permitiendo a los investigadores enfocarse en el pensamiento crítico y la interpretación profunda de los resultados.

Conclusiones

1. **Asistencia en Análisis y Visualización:** Las herramientas de IA están emergiendo como asistentes valiosos para tareas de análisis y visualización de datos en investigación.
2. **Sugerencias de Preparación:** La IA puede analizar datos preliminarmente y sugerir pasos adecuados de limpieza (ej. manejo de faltantes) o transformación (ej. para corregir sesgos) basados en las características de los datos o los requisitos del modelo.
3. **Generación de Visualizaciones:** Se puede usar IA (LLM o herramientas integradas) para generar código o crear directamente gráficos (desde básicos hasta especializados como curvas ROC) a partir de descripciones en lenguaje natural, agilizando el proceso de exploración visual.
4. **Apoyo en Interpretación:** La IA puede ayudar a interpretar resultados estadísticos complejos o salidas de modelos de machine learning, generando resúmenes, explicando términos técnicos o el significado práctico de los parámetros (ej. OR, HR, coeficientes, importancia de variables).
5. **Rol Complementario:** La IA actúa como un asistente que puede automatizar tareas, generar código y ofrecer explicaciones iniciales, pero no sustituye la necesidad de una sólida comprensión estadística, el pensamiento crítico y la interpretación experta por parte del investigador humano.

CAPÍTULO 26

BÚSQUEDA Y GESTIÓN DE REFERENCIAS

Searching and Reference Management

Resumen

Este capítulo aborda cómo las herramientas de IA están transformando la búsqueda y gestión de la literatura científica. Explica que herramientas basadas en IA como Elicit y DeepSeek utilizan procesamiento de lenguaje natural para ir más allá de las búsquedas por palabras clave, comprendiendo preguntas de investigación y encontrando artículos relevantes, además de extraer información clave y generar resúmenes estructurados. Menciona cómo herramientas como NotebookLM permiten analizar y obtener ideas clave de colecciones de documentos ya recopilados. Discute el potencial de la IA para integrarse con gestores de referencias tradicionales (Zotero, Mendeley), automatizando tareas como la extracción de metadatos y el etiquetado temático. Si bien estas herramientas pueden aumentar significativamente la eficiencia, se enfatiza la necesidad de verificar la precisión de la información generada por IA, ser consciente de posibles sesgos algorítmicos y considerar aspectos de privacidad, seguridad y costo. La supervisión y el juicio crítico del investigador siguen siendo indispensables.

Palabras clave: Búsqueda de literatura, gestión de referencias, inteligencia artificial, síntesis de evidencia, herramientas de IA, procesamiento de lenguaje natural (PLN).

Abstract

This chapter addresses how AI tools are transforming scientific literature search and management. It explains that AI-powered tools like Elicit and DeepSeek use natural language processing to go beyond keyword searches, understanding research questions and finding relevant articles, as well as extracting key information and generating structured summaries. It mentions how tools like NotebookLM allow analysis and insight generation from already collected document collections. It discusses AI's potential integration with traditional reference man-

agers (Zotero, Mendeley) to automate tasks like metadata extraction and thematic tagging. While these tools can significantly increase efficiency, the chapter emphasizes the need to verify the accuracy of AI-generated information, be aware of potential algorithmic biases, and consider privacy, security, and cost aspects. Researcher supervision and critical judgment remain indispensable.

Keywords: *Literature search, reference management, artificial intelligence, evidence synthesis, AI tools, natural language processing* (NLP).

Introducción

La investigación se construye sobre el trabajo previo. Realizar una búsqueda de literatura exhaustiva y sistemática es esencial para contextualizar un estudio, identificar lagunas en el conocimiento, informar el diseño metodológico y discutir los hallazgos. Sin embargo, la cantidad de publicaciones científicas es abrumadora. Las herramientas de IA están emergiendo para hacer este proceso más inteligente, rápido y efectivo.

5.1. Identificación de Artículos Clave y Generación de Resúmenes

Las bases de datos tradicionales (como PubMed, Scopus, Web of Science) son fundamentales, pero encontrar los artículos *más* relevantes y comprender rápidamente su contenido puede ser un desafío. Las herramientas basadas en IA van más allá de la simple búsqueda por palabras clave, utilizando técnicas de procesamiento de lenguaje natural (PLN) para comprender el significado y el contexto de los artículos.

- **Cómo la IA puede ayudar:** Herramientas como **Elicit** y **DeepSeek** están diseñadas para entender tu pregunta de investigación y encontrar artículos que la aborden directamente, incluso si no utilizan las palabras clave exactas que esperas. Pueden analizar los artículos encontrados y extraer automáticamente información estructurada (ej. objetivos, métodos, resultados clave, limitaciones), presentándola en formatos que facilitan la comparación y la síntesis. **NotebookLM**, al permitir cargar y "chatear" con tus documentos, también puede ayudarte a obtener resúmenes e ideas clave de colecciones de artículos que ya has identificado.

- **Ejemplos Prácticos:**

  - **Escenario 1: Encontrar artículos que respondan a una pregunta específica.**

    - **Pregunta:** «¿Es efectiva la terapia cognitivo-conductual para reducir la ansiedad en adultos mayores con enfermedades crónicas?»

    - **Uso de Elicit/DeepSeek:** En lugar de solo buscar por "terapia cognitivo-conductual", "ansiedad", "adultos mayores" y "enfermedades crónicas", puedes ingresar la pregunta completa en Elicit o DeepSeek. Estas herramientas buscarán artículos relevantes y te presentarán una tabla resumiendo aspectos clave de cada estudio (ej. número de participantes, tipo de enfermedad crónica, medida de ansiedad utilizada, resultado principal). Esto te permite identificar rápidamente los estudios más pertinentes sin tener que leer el resumen de cientos de artículos.

  - **Escenario 2: Sintetizar información de un conjunto de artículos.**

    - **Uso de Elicit/DeepSeek:** Después de una búsqueda, puedes pedirle a la herramienta que sintetice los hallazgos sobre un aspecto particular (ej. "¿Cuáles son los principales hallazgos sobre la magnitud del efecto de la TCC en la ansiedad en estos estudios?"). La IA puede extraer y resumir esta información de los artículos identificados.

  - **Escenario 3: Obtener resúmenes y ideas clave de artículos ya descargados.**

    - **Uso de NotebookLM:** Carga una carpeta con PDF de artículos que has recopilado. Puedes pedirle a NotebookLM que genere un resumen de cada artículo, un resumen general de la colección, o hacerle preguntas específicas sobre el contenido de los documentos (ej. "¿Cuáles fueron las principales limitaciones reportadas en estos estudios sobre telemedicina?").

5.2. Integración con Gestores de Referencias y Etiquetado Automático

Una vez que se han identificado los artículos relevantes, el siguiente paso es organizarlos y gestionarlos. Los gestores de referencias como Zotero y Mendeley son herramientas estándar para esto, pero la IA puede mejorar su funcionalidad.

- **Cómo la IA puede ayudar:** La integración de IA con gestores de referencias puede automatizar tareas como la extracción y el etiquetado de metadatos, y potencialmente ayudar en la organización temática.
- **Ejemplos Prácticos:**
  - **Escenario 1: Extracción automática de metadatos.**
    - **Uso:** Al añadir un archivo PDF a tu gestor de referencias (Zotero, Mendeley), la mayoría ya utiliza alguna forma de automatización para extraer el título, autores, revista, año, etc. La IA puede mejorar esta capacidad, siendo más precisa con formatos de PDF variados o extrayendo información más allá de los metadatos básicos, como el tipo de estudio o el número de pacientes.
  - **Escenario 2: Etiquetado automático de metadatos (etiquetas/tags).**
    - **Uso:** Algunas herramientas o plugins con IA pueden leer el contenido de un artículo y sugerir automáticamente etiquetas relevantes basadas en los temas principales, métodos o población. Por ejemplo, un artículo sobre un ECA de un nuevo fármaco para la diabetes en adultos mayores podría ser etiquetado automáticamente con "ECA", "Fármaco", "Diabetes", "Adultos Mayores". Esto facilita la organización y búsqueda dentro de tu propia biblioteca de referencias.
  - **Escenario 3: Organización temática sugerida.**
    - **Uso Potencial:** Herramientas futuras podrían analizar tu colección de referencias y sugerir cómo agruparlas temáticamente o identificar clústeres de investigación relacionados, ayudándote a estructurar tu revisión de literatura o identificar subtemas para explorar.

5.3. Otras Herramientas y Consideraciones

Además de las herramientas mencionadas, otras plataformas (a menudo integrando capacidades de LLM) están surgiendo para asistir en la gestión de referencias y el flujo de trabajo de escritura, como herramientas que ayudan a insertar citas y generar bibliografías mientras escribes, o que verifican la consistencia de tus citas con la lista de referencias.

Consideraciones Importantes:

- **Precisión:** Aunque la IA es potente, no es infalible. Siempre verifica la información extraída automáticamente (metadatos, resúmenes) con el artículo original.
- **Sesgo:** Los algoritmos de IA pueden reflejar sesgos presentes en los datos con los que fueron entrenados. Sé consciente de esto al interpretar los resultados de las búsquedas o los resúmenes generados.
- **Privacidad y Seguridad:** Al utilizar herramientas en línea, considera cómo se manejan tus datos y los documentos que cargas, especialmente si contienen información sensible.
- **Costo:** Algunas de las herramientas más avanzadas requieren suscripciones de pago.

La integración de IA en la búsqueda y gestión de referencias tiene el potencial de liberar un tiempo valioso para los investigadores, permitiéndoles dedicar más energía al pensamiento crítico y la síntesis de la información, en lugar de a las tareas manuales de organización. A medida que estas herramientas continúan evolucionando, se espera que se conviertan en componentes aún más integrales del flujo de trabajo de investigación.

Conclusiones

1. **Optimización de Búsqueda:** Las herramientas de IA (ej. Elicit, DeepSeek) mejoran la búsqueda de literatura al comprender preguntas complejas y encontrar artículos relevantes más allá de las palabras clave.
2. **Extracción y Síntesis Automatizada:** La IA puede extraer información estructurada clave (objetivos, métodos, resultados) de artículos y generar resúmenes, acelerando la fase de revisión y síntesis inicial.
3. **Gestión Inteligente:** Existe potencial para que la IA mejore los gestores de referencias (ej. Zotero, Mendeley) mediante la extracción automática y precisa de metadatos y el etiquetado temático sugerido.
4. **Eficiencia Mejorada:** El uso de IA en estas tareas puede liberar tiempo significativo para que los investigadores se enfoquen en el análisis crítico y la interpretación.

5. **Precaución Necesaria:** Los resultados generados por IA (resúmenes, extracciones) no son infalibles y requieren verificación humana contra las fuentes originales.
6. **Conciencia de Limitaciones:** Se debe ser consciente de los posibles sesgos en los algoritmos de IA, así como de las consideraciones de privacidad, seguridad de datos y costos asociados a estas herramientas.
7. **Asistencia, No Reemplazo:** La IA actúa como un asistente poderoso para la búsqueda y gestión de referencias, pero no reemplaza el juicio crítico y la experiencia del investigador.

CAPÍTULO 27

REDACCIÓN Y MAQUETACIÓN AUTOMATIZADA

Automated Writing and Layout

Resumen

Este capítulo explora el uso de la Inteligencia Artificial (IA) para asistir en la redacción y maquetación de artículos científicos. Detalla cómo los Modelos de Lenguaje Grandes (LLM), como ChatGPT, pueden utilizarse para generar borradores iniciales de secciones, refinar el estilo del texto (mejorando claridad, concisión y fluidez), revisar gramática y ortografía, e incluso ayudar con traducciones. Describe también herramientas especializadas, como Typeset.io, que emplean IA para automatizar la maquetación de manuscritos según las directrices específicas de las revistas, incluyendo el formato de texto, tablas, figuras y, crucialmente, la gestión y formateo de citas y referencias bibliográficas. Además, menciona el potencial de la IA para revisar la consistencia terminológica a lo largo del documento y validar la correspondencia entre citas y referencias. Se concluye que, si bien estas herramientas aumentan la eficiencia y reducen errores manuales, la revisión humana y el control editorial final por parte del investigador son indispensables para garantizar la precisión científica y la calidad del manuscrito.

Palabras clave: Redacción científica con IA, maquetación automatizada, formato de manuscritos, asistentes de escritura, LLM en investigación, validación de citas.

Abstract

This chapter explores the use of Artificial Intelligence (AI) to assist in the writing and layout (formatting) of scientific articles. It details how Large Language Models (LLM), like ChatGPT, can be used to generate initial drafts of sections, refine text style (improving clarity, conciseness, and flow), check grammar and spelling, and even aid with translations. It also describes specialized tools, such as Typeset.io, that employ AI to automate manuscript formatting according to

specific journal guidelines, including text layout, tables, figures, and, crucially, citation and bibliography management and formatting. Furthermore, it mentions AI's potential role in checking terminological consistency throughout the document and validating the correspondence between citations and references. It concludes that while these tools enhance efficiency and reduce manual errors, human review and final editorial control by the researcher are indispensable to ensure scientific accuracy and manuscript quality.

Keywords: *AI scientific writing, automated formatting, manuscript formatting, writing assistants, LLMs in research, citation validation.*

Introducción

La redacción científica exige precisión, claridad y adherencia a estructuras y estilos específicos. La maquetación, por su parte, implica formatear el manuscrito según las estrictas directrices de las revistas o las normas de presentación. La Inteligencia Artificial está emergiendo como un asistente valioso en estas etapas, ayudando a generar borradores, refinar el lenguaje, gestionar referencias y adaptar el formato.

6.1. Uso de LLM para Borradores y Revisión de Estilo

Los Modelos de Lenguaje Grandes (LLM) como ChatGPT, Gemini, u otros similares, son herramientas potentes para generar y refinar texto. Pueden ayudar a superar el bloqueo del escritor y mejorar la calidad de la prosa científica.

- **Cómo la IA puede ayudar:** Los LLM pueden generar borradores de secciones del manuscrito, reformular oraciones para mayor claridad, sugerir sinónimos, verificar la gramática y la ortografía, y adaptar el tono y el estilo.
- **Ejemplos Prácticos:**
  - **Generación de Borradores:**
    - **Escenario:** Tienes los resultados de tu análisis y necesitas escribir la sección de 'Resultados'.
    - **Prompt a LLM:** "Aquí están los hallazgos clave de mi estudio sobre [tema]. [Proporciona un resumen de los resultados, quizás en viñetas o

una tabla simple]. Por favor, redacta un borrador inicial para la sección de Resultados de un manuscrito científico basado en esta información. Asegúrate de mencionar [variables clave, análisis realizados] y presenta los resultados de forma clara y concisa."

- **Refinamiento y Mejora de Estilo:**
  - **Escenario:** Has escrito un párrafo pero sientes que podría ser más claro o tener un mejor flujo.
  - **Prompt a LLM:** "Por favor, revisa el siguiente párrafo para mejorar la claridad, concisión y fluidez. [Pega el párrafo]."
  - **Escenario:** Quieres asegurar que el tono de tu sección de 'Discusión' sea apropiado para una revista académica.
  - **Prompt a LLM:** "He escrito la sección de Discusión de mi manuscrito. ¿Podrías leerla y sugerir mejoras para asegurar que el tono sea formal y académico, y que las interpretaciones se presenten con la debida cautela?"
- **Verificación Lingüística:**
  - **Escenario:** Has escrito el manuscrito en un idioma y quieres una revisión exhaustiva de gramática y ortografía, o incluso una traducción inicial.
  - **Prompt a LLM:** "Por favor, revisa la gramática y ortografía de este texto. [Pega el texto]." o "Traduce este texto al inglés académico. [Pega el texto]."

6.2. Formateo según Guías de Revistas (Maquetación)

Una vez que el contenido está listo, formatearlo según las directrices específicas de una revista puede ser un proceso tedioso y propenso a errores. Herramientas especializadas, a menudo con capacidades de IA, pueden automatizar gran parte de este trabajo.

- **Cómo la IA puede ayudar:** Herramientas como **Typeset.io** o plataformas como Overleaf con funcionalidades de IA integrada pueden aplicar automáticamente las plantillas y requisitos de formato de miles de revistas científicas, ajustando estilos de encabezado, formato de tablas y figuras, y, crucialmente, el estilo de citación y bibliografía.

- **Ejemplos Prácticos:**
  - **Escenario:** Tienes tu manuscrito en un formato genérico y quieres formatearlo para la revista 'Journal of Health Data Science'.
  - **Uso de Overleaf + IA:** Si usas LaTeX con Overleaf, ciertas extensiones o herramientas de terceros con IA podrían ayudar a verificar el cumplimiento de las directrices de formato de la revista o a automatizar tareas de formato complejas.
  - **Adaptación Rápida para Múltiples Revistas:** Si tu manuscrito es rechazado por una revista y necesitas enviarlo a otra con directrices diferentes, estas herramientas pueden aplicar rápidamente el nuevo formato, ahorrando mucho tiempo en comparación con el ajuste manual.

6.3. Revisión de Consistencia Terminológica y Validación de Citas

Mantener la consistencia en la terminología a lo largo de un manuscrito y asegurar que todas las citas en el texto correspondan a entradas correctas en la lista de referencias son tareas críticas para la calidad y la credibilidad. La IA puede ayudar a identificar y corregir errores en estas áreas.

- **Cómo la IA puede ayudar:** Los LLM y herramientas especializadas pueden escanear el manuscrito para identificar variaciones en la forma en que se nombran conceptos o variables (ej. usar "HCE", "EHR", "Historia Clínica Electrónica" indistintamente) y sugerir unificación. También pueden verificar que cada cita en el texto tenga una entrada correspondiente en la lista de referencias y viceversa, y que el formato de las referencias sea consistente.
- **Ejemplos Prácticos:**
  - **Revisión de Consistencia Terminológica:**
    - **Prompt a LLM:** "Por favor, revisa este manuscrito para detectar inconsistencias en la terminología. Identifica diferentes formas de referirse al mismo concepto o variable y sugiere una terminología unificada. [Pega el texto]."
  - **Validación de Citas y Referencias:**
    - **Uso de Herramientas Integradas (ej. en Gestores de Referencias o Maquetación):** Herramientas como Zotero, Mendeley, o las funcionali-

dades de Typeset.io/Overleaf a menudo tienen verificadores integrados que usan algoritmos (potencialmente mejorados por IA) para:

- Verificar que cada cita en el texto tiene una entrada en la bibliografía.
- Verificar que cada entrada en la bibliografía se cita al menos una vez en el texto.
- Verificar el formato de cada entrada bibliográfica según el estilo seleccionado.
- **Uso de LLM (para verificación de formato o detalles):** Aunque menos robusto que las herramientas especializadas, un LLM podría ayudar a verificar el formato de una referencia específica si le proporcionas las reglas del estilo (ej. "Según el estilo Vancouver, ¿está bien formateada esta referencia? [Pega la referencia]").

La automatización de la redacción y la maquetación con herramientas de IA tiene el potencial de acelerar significativamente el proceso de preparación de manuscritos, reducir errores manuales y permitir a los investigadores dedicar más tiempo a la investigación en sí. Sin embargo, es fundamental que el investigador revise cuidadosamente todo el contenido generado por la IA, asegure su precisión científica y mantenga el control editorial total sobre el manuscrito final. La IA es un asistente, no un reemplazo del autor.

Conclusiones

1. **Asistencia en Redacción:** Los LLM (ej. ChatGPT) pueden ser herramientas útiles para generar borradores iniciales, superar bloqueos, refinar el estilo (claridad, concisión), y realizar revisiones lingüísticas (gramática, ortografía).
2. **Automatización de Formato:** Herramientas específicas (ej. Typeset.io) utilizan IA para automatizar la tediosa tarea de formatear manuscritos según las guías de estilo de diferentes revistas, incluyendo la gestión de referencias.
3. **Mejora de la Consistencia:** La IA puede ayudar a identificar y corregir inconsistencias terminológicas y a validar la correcta correspondencia entre las citas en el texto y la lista de referencias.

4. **Aumento de Eficiencia:** El uso de estas herramientas puede acelerar significativamente el proceso de preparación del manuscrito y reducir errores manuales de formato y estilo.

5. **Revisión Humana Indispensable:** La IA actúa como un asistente; es crucial que el investigador revise, edite y valide todo el contenido y formato generado o modificado por IA para asegurar la precisión científica, la coherencia argumental y la calidad final.

6. **Control Editorial del Autor:** El investigador mantiene siempre el control y la responsabilidad final sobre el contenido y la presentación del manuscrito.

CAPÍTULO 28

PRESENTACIÓN DE RESULTADOS

Presentation of Results

Resumen

Este capítulo se enfoca en cómo la IA puede asistir en la presentación efectiva de resultados de investigación. Describe el uso de IA para generar imágenes como diagramas conceptuales, esquemas de flujo (incluyendo la generación de código para herramientas como Mermaid para diagramas PRISMA) e infografías, utilizando herramientas de generación de imágenes o diseño asistido (Midjourney, DALL-E, Canva + IA, Gamma App). Explica cómo la IA puede ayudar a generar y formatear tablas automáticamente a partir de datos o resumir su contenido. Discute la creación de dashboards interactivos asistidos por IA (usando Gamma App, Streamlit + IA), que pueden incluir funcionalidades como resúmenes automáticos de tendencias o consultas en lenguaje natural. Se subraya que estas herramientas pueden automatizar y mejorar la presentación, pero la revisión humana es esencial para asegurar la precisión y adecuación de los elementos generados.

Palabras clave: Presentación de resultados, inteligencia artificial, visualización de datos, generación de imágenes con IA, dashboards interactivos, herramientas de IA.

Abstract

This chapter focuses on how AI can assist in the effective presentation of research results. It describes the use of AI to generate images such as conceptual diagrams, flowcharts (including code generation for tools like Mermaid for PRISMA diagrams), and infographics, using image generation or AI-assisted design tools (Midjourney, DALL-E, Canva + AI, Gamma App). It explains how AI can help automatically generate and format tables from data or summarize their content. It discusses the creation of AI-assisted interactive dashboards (using Gamma App,

Streamlit + AI), which can include functionalities like automated trend summaries or natural language querying. It underscores that these tools can automate and enhance presentation, but human review is essential to ensure the accuracy and appropriateness of the generated elements.

Keywords: *Results presentation, artificial intelligence, data visualization, AI image generation, interactive dashboards, AI tools.*

Introducción

Presentar los resultados de la investigación de manera efectiva es tanto un arte como una ciencia. Implica seleccionar las visualizaciones y tablas adecuadas, organizarlas lógicamente y añadir contexto y narrativa. Las herramientas de IA pueden automatizar aspectos de este proceso, generar elementos visuales complejos y ayudar a crear presentaciones dinámicas.

7.1. Imágenes Generadas por IA: Diagramas Conceptuales, Esquemas de Flujo, Infografías

Las imágenes son herramientas poderosas para comunicar ideas complejas de manera rápida y clara. En investigación, utilizamos diagramas para ilustrar modelos conceptuales, esquemas de flujo para describir procesos metodológicos e infografías para resumir hallazgos clave de forma visualmente atractiva.

- **Cómo la IA puede ayudar:** Los LLM avanzados con capacidades de generación de imágenes o herramientas especializadas de diseño asistidas por IA pueden crear borradores o versiones completas de estos elementos visuales basándose en descripciones de texto.
- **Ejemplos Prácticos:**
  - **Diagramas Conceptuales:**
    - **Escenario:** Quieres ilustrar el modelo conceptual de tu estudio que relaciona la exposición, varios mediadores y el desenlace.
    - **Prompt a Herramienta de Generación de Imágenes (ej. Midjourney, DALL-E, o LLM con esta función):** "Crea un diagrama conceptual simple que muestre una flecha de 'Exposición' a 'Desenlace', con tres cajas intermedias etiquetadas 'Mediador 1', 'Mediador 2', 'Mediador 3' con

flechas que van de 'Exposición' a cada mediador y de cada mediador a 'Desenlace'. Usa un estilo limpio y académico."

- **Esquemas de Flujo (ej. para Diagramas PRISMA):**

- **Escenario:** Necesitas crear un diagrama de flujo PRISMA para tu revisión sistemática (como vimos en capítulos anteriores).
- **Uso de LLM (Generación de Código para Diagramas):** Aunque no generan la imagen directamente, algunos LLM pueden generar código para herramientas de diagramación (ej. Mermaid, Graphviz).
- **Prompt a LLM:** "Genera el código Mermaid para un diagrama de flujo PRISMA simplificado que muestre los pasos de Identificación, Cribado, Elegibilidad e Inclusión, con el número de registros en cada etapa." (Luego usarías un editor Mermaid para renderizar el código).

- **Infografías:**

- **Escenario:** Quieres crear una infografía simple para resumir los 3 hallazgos principales de tu estudio para una presentación.
- **Uso de Herramientas de Diseño con IA (ej. Canva con funciones de IA, Gamma App):** Algunas plataformas de diseño gráfico integran IA para sugerir diseños, iconos o elementos visuales basándose en el contenido que proporcionas. Puedes ingresar texto clave o un resumen de tus hallazgos y la IA puede ayudar a estructurar una infografía inicial.
- **Uso de Herramientas de Extracción de Ideas clave (ej. Napkin AI):** Aunque no diseñan la infografía directamente, herramientas como Napkin pueden ser útiles en la fase previa para extraer rápidamente los puntos clave, datos o citas de tus fuentes de investigación, que luego utilizarás como contenido para la infografía que diseñes con otras herramientas.

7.2. Tablas Inteligentes: Generación Automática y Resumen

Las tablas son esenciales para presentar datos numéricos de forma organizada (ej. características basales de los participantes, resultados de análisis estadísticos, parámetros de modelos). La IA puede simplificar la creación y el formato de estas tablas.

- **Cómo la IA puede ayudar:** Los LLM pueden generar código para crear tablas a partir de estructuras de datos (como DataFrames en Python). También pueden ayudar a formatear tablas existentes o a resumir su contenido.
- **Ejemplos Prácticos:**
  - **Generación de Tablas a partir de DataFrames:**
    - **Escenario:** Tienes un DataFrame de pandas con estadísticas descriptivas (media, DS, N) para varias variables y quieres presentarlo como una tabla en formato Markdown o LaTeX para un informe.
    - **Prompt a LLM (Generación de Código Python):** "Tengo un DataFrame de pandas llamado 'stats_df' con columnas 'Variable', 'Media', 'Desviacion_Estandar', 'N'. Genera código Python para imprimir este DataFrame como una tabla en formato Markdown."
  - **Formato de Tablas:**
    - **Escenario:** Tienes una tabla de resultados de regresión y quieres asegurarte de que el formato (número de decimales, negritas para resultados significativos) cumpla con un estilo específico.
    - **Prompt a LLM:** "Aquí está mi tabla de resultados de regresión [pega la tabla]. Por favor, reformatea la columna 'Valor p' a 3 decimales y pon en negrita los coeficientes donde el valor p sea menor a 0.05."
  - **Generación Automática de Tablas (Plataformas de Análisis con IA):** Algunas plataformas de análisis de datos con IA integrada pueden generar automáticamente tablas de estadísticas descriptivas o resultados de modelos con un solo clic.

7.3. Dashboards Interactivos con IA

Como vimos en el Capítulo 4 de la Parte IV, los *dashboards* son excelentes para presentar datos de manera interactiva. La IA puede llevar los *dashboards* al siguiente nivel, haciéndolos más fáciles de crear y más inteligentes.

- **Cómo la IA puede ayudar:** Herramientas como **Gamma App** (que genera presentaciones/reportes interactivos) o *frameworks* como **Streamlit**

(para crear aplicaciones web de datos en Python) integrados con capacidades de IA pueden facilitar la creación de *dashboards*. La IA también puede añadir funcionalidades inteligentes *dentro* del *dashboard*, como resúmenes automáticos de tendencias, detección de anomalías o la capacidad de hacer preguntas sobre los datos en lenguaje natural.

- **Ejemplos Prácticos:**

  - **Creación Rápida de *Dashboards*/Reportes Interactivos (Gamma App):**

    - **Escenario:** Quieres crear un reporte interactivo rápido de los hallazgos clave de tu estudio para compartir con colaboradores.

    - **Uso de Gamma App:** Puedes empezar con un esquema de tu reporte o pegar texto y datos. Gamma App puede sugerir diseños, formatear el contenido y crear elementos interactivos (como gráficos).

  - ***Dashboards* con Funcionalidades de IA (Streamlit + IA):**

    - **Escenario:** Has construido un modelo predictivo y quieres crear un *dashboard* interactivo donde los usuarios puedan ingresar características de un paciente y obtener una predicción, junto con una explicación generada por IA.

    - **Uso de Streamlit + LLM API:** Puedes usar Streamlit para construir la interfaz del *dashboard* (campos de entrada, botones, visualizaciones). Integrar una API de LLM (como OpenAI, Gemini) te permitiría enviar los resultados del modelo y las características del paciente a la IA para generar una explicación personalizada en lenguaje natural (ej. "Basado en la edad, IMC e historial familiar, el modelo predice un riesgo alto/bajo de X. Los factores que más contribuyeron a esta predicción fueron...").

  - **Ideas clave Generados por IA en *Dashboards*:** Algunas plataformas de BI (*Business Intelligence*) están incorporando IA para analizar los datos que se muestran en un *dashboard* y generar automáticamente ideas clave textuales (ej. "La tasa de reingreso aumentó significativamente en el último trimestre, particularmente en la población mayor de 65 años").

La IA está transformando la presentación de resultados, haciendo que sea más fácil crear visualizaciones impactantes, tablas precisas y reportes interactivos. Esto permite a los investigadores comunicar sus hallazgos de ma-

nera más efectiva a una audiencia más amplia, acelerando la traducción del conocimiento científico a la práctica. Sin embargo, la revisión humana y la curación de los elementos generados por IA son esenciales para garantizar la precisión y la adecuación al contexto.

Conclusiones

1. **Asistencia Visual:** La IA ofrece herramientas para generar elementos visuales clave (diagramas conceptuales, esquemas de flujo, infografías) a partir de descripciones textuales, facilitando la comunicación de ideas complejas.
2. **Automatización de Tablas:** Se puede usar IA para generar código para tablas a partir de datos estructurados, formatear tablas según estilos específicos y potencialmente resumir su contenido clave.
3. ***Dashboards* Inteligentes:** La IA facilita la creación de *dashboards* interactivos y puede añadir funcionalidades avanzadas como resúmenes automáticos o la capacidad de interactuar con los datos mediante lenguaje natural.
4. **Eficiencia en Presentación:** Estas herramientas pueden ahorrar tiempo y esfuerzo en la creación de materiales de presentación de alta calidad y visualmente atractivos.
5. **Revisión Humana Clave:** Es fundamental que el investigador revise cuidadosamente todos los elementos visuales y textuales generados por IA para garantizar la precisión científica, la claridad y la adecuación al mensaje y la audiencia.
6. **Potencial Comunicativo:** La IA amplía las posibilidades para comunicar hallazgos de investigación de manera más efectiva e interactiva a diversas audiencias.

CAPÍTULO 29

DIVULGACIÓN CIENTÍFICA CON IA

Presentation of Results

Resumen

Este capítulo explora cómo la IA facilita la divulgación científica más allá de los canales tradicionales. Describe el uso de herramientas de IA para la creación de materiales de comunicación diversos. Para posters (tradicionales o digitales), herramientas como Canva + IA o Gamma App ayudan a sugerir diseños, formatear contenido y crear presentaciones interactivas. Para videos y animaciones, los LLM pueden generar guiones y storyboards, mientras que herramientas como Pictory o Synthesia pueden automatizar la creación de videos con voz en off o avatares a partir de texto. Para redes sociales, la IA asiste en la generación de microcontenidos concisos (tweets, posts) y visuales (infografías simples, carruseles automáticos). Se concluye que la IA democratiza la creación de contenido de divulgación de alta calidad, ampliando el alcance de la investigación, pero siempre requiriendo supervisión humana para garantizar la precisión científica y la integridad del mensaje.

Palabras clave: Divulgación científica, inteligencia artificial, comunicación científica, visualización de datos, herramientas de IA, contenido digital.

Abstract

This chapter explores how AI facilitates scientific dissemination beyond traditional channels. It describes the use of AI tools for creating diverse communication materials. For posters (traditional or digital), tools like Canva + AI or Gamma App help suggest designs, format content, and create interactive presentations. For videos and animations, LLMs can generate scripts and storyboards, while tools like Pictory or Synthesia can automate video creation with voice-overs or avatars from text. For social media, AI assists in generating concise microcontent (tweets, posts) and visuals (simple infographics, automated carousels). It concludes that AI democratizes the creation of high-quality dissemination con-

tent, broadening research reach, but always requires human oversight to ensure scientific accuracy and message integrity.

Keywords: *Science communication, artificial intelligence, scientific dissemination, data visualization, AI tools, digital content.*

Introducción

Tradicionalmente, la divulgación científica se ha centrado en publicaciones en revistas y presentaciones en conferencias. Si bien estos siguen siendo pilares fundamentales, la necesidad de comunicar la ciencia de manera más amplia y comprensible ha llevado a la exploración de nuevos formatos. La IA ofrece herramientas que pueden simplificar la creación de materiales de divulgación para diferentes audiencias y canales.

8.1. Pósteres Digitales: Plantillas Adaptadas con IA

Los pósteres son un formato común en conferencias para presentar un resumen visual de la investigación. Los pósteres digitales y las presentaciones interactivas están ganando popularidad. La IA puede asistir en el diseño y la adaptación de estos materiales.

- **Cómo la IA puede ayudar:** Herramientas de diseño gráfico con funciones de IA (como **Canva + IA**) o plataformas diseñadas para presentaciones interactivas (como **Gamma App**) pueden sugerir diseños, formatear texto, recomendar paletas de colores y organizar elementos visuales basándose en el contenido que proporcionas.
- **Ejemplos Prácticos:**
  - **Creación Rápida de Borradores de Pósteres:**
    - **Escenario:** Tienes el resumen, los puntos clave de resultados y algunas figuras de tu investigación y necesitas crear un póster para una conferencia.
    - **Uso de Canva + IA:** Puedes usar las funciones de "Diseño Mágico" o "Magic Write" de Canva. Ingresa un resumen de tu investigación o pega secciones de tu manuscrito. La IA puede sugerir diseños de póster, organizar el texto en secciones (introducción, métodos, resultados, conclu-

sión) y recomendar elementos visuales. Luego, puedes personalizarlo fácilmente.

- **Adaptación de Contenido a Formato de Póster:**

  - **Prompt a LLM:** "He escrito un resumen ejecutivo de mi estudio [pega el resumen]. Ayúdame a extraer los puntos clave y a reformularlos en frases cortas y concisas, adecuadas para las secciones de un póster científico (Objetivo, Métodos Breves, Resultados Clave, Conclusión)."

- **Creación de Pósteres Digitales Interactivos:**

  - **Uso de Gamma App:** En lugar de un póster estático, puedes usar Gamma App para crear una presentación interactiva tipo póster. Puedes pegar tu contenido y la IA te ayudará a darle formato en tarjetas visuales, añadir interactividad y asegurar un diseño atractivo para visualización en pantalla.

8.2. Videos y Animaciones: Generación de Guiones y Storyboards con IA

Los videos y las animaciones son formatos muy efectivos para explicar conceptos complejos o resumir hallazgos de investigación de manera atractiva para audiencias más amplias (incluido el público no especializado).

- **Cómo la IA puede ayudar:** Los LLM pueden generar guiones para videos explicativos o resúmenes animados. Herramientas especializadas (como **Pictory** o **Synthesia**) pueden automatizar la creación del video a partir del guion, incluso generando voces en off o avatares que "hablan".
- **Ejemplos Prácticos:**
  - **Generación de Guion para Video Resumen:**
    - **Escenario:** Quieres crear un video corto (2-3 minutos) explicando los hallazgos principales de tu revisión sistemática sobre [tema].
    - **Prompt a LLM:** "Genera un guion para un video explicativo de 3 minutos sobre los hallazgos de mi revisión sistemática [proporciona un resumen de la revisión y los hallazgos clave]. El guion debe estar dirigido a profesionales de la salud y debe incluir una introducción, los resultados más importantes y una conclusión sobre las implicaciones clínicas."

- **Creación de Storyboards Sencillos:**

- **Prompt a LLM:** "Basado en el guion anterior, sugiere ideas visuales o un storyboard simple para cada sección del video (ej. qué imágenes o animaciones podrían acompañar el texto)."

- **Creación Automatizada de Video:**

- **Uso de Pictory:** Puedes pegar tu guion en Pictory, y la herramienta utilizará IA para seleccionar clips de video o imágenes de archivo relevantes, añadir voz en off (generada por IA o tu propia grabación) y música de fondo para crear un borrador de video.

- **Uso de Synthesia:** Si necesitas un video con un presentador, Synthesia te permite crear avatares de IA que narran tu guion en diferentes idiomas y estilos.

8.3. Microcontenidos para Redes Sociales: Resúmenes Visuales y Carruseles Automáticos

Las redes sociales son plataformas importantes para la divulgación rápida y directa de la investigación. El contenido debe ser conciso, visualmente atractivo y fácil de consumir en formatos como publicaciones cortas, imágenes o carruseles.

- **Cómo la IA puede ayudar:** Los LLM pueden ayudar a condensar hallazgos en mensajes cortos para redes sociales. Herramientas de diseño gráfico con IA o plataformas especializadas pueden generar automáticamente resúmenes visuales o carruseles a partir de texto o contenido más largo.

- **Ejemplos Prácticos:**

- **Redacción de Tweets o Posts Cortos:**

- **Prompt a LLM:** "Formula 5 tweets diferentes resumiendo el hallazgo principal de mi estudio: '[Hallazgo principal]'. Cada tweet debe ser conciso y usar hashtags relevantes (ej. #investigación #salud #ciencia)."

- **Generación de Resúmenes Visuales/Infografías Pequeñas:**

- **Uso de Canva + IA:** Utiliza las funciones de diseño de Canva para crear

imágenes simples que resuman un punto clave de tu investigación, quizás usando iconos o gráficos sencillos. La IA puede ayudar a organizar el texto y sugerir elementos de diseño.

- **Creación de Carruseles Automáticos:**

- **Uso de Herramientas Especializadas:** Algunas herramientas (a menudo integradas en plataformas de gestión de redes sociales o diseño) pueden tomar un texto más largo (ej. un resumen de tu artículo) y generar automáticamente una serie de imágenes o diapositivas cortas, formateadas como un carrusel para plataformas como Instagram o LinkedIn.

La IA está democratizando la creación de materiales de divulgación científica, permitiendo a los investigadores con recursos o habilidades de diseño limitadas producir contenido de alta calidad. Al aprovechar estas herramientas, los científicos pueden ampliar el alcance de su investigación, involucrar a diversas audiencias y acelerar la traducción del conocimiento científico a la sociedad. Sin embargo, es crucial mantener la precisión científica, la transparencia y la supervisión humana en todo el contenido generado por IA.

Conclusiones

1. **Nuevos Formatos:** La IA permite a los investigadores crear fácilmente materiales de divulgación en formatos más allá de los tradicionales, como pósters interactivos, videos y contenido para redes sociales.
2. **Asistencia en Diseño:** Herramientas con IA (ej. Canva, Gamma App) pueden sugerir diseños, organizar contenido y formatear texto para crear pósters y presentaciones visualmente atractivos con menor esfuerzo.
3. **Creación de Contenido Multimedia:** La IA (LLM, Pictory, Synthesia) puede generar guiones, storyboards e incluso borradores de videos explicativos o animaciones, facilitando la comunicación de conceptos complejos a audiencias amplias.
4. **Adaptación a Microformatos:** Se puede usar IA para condensar hallazgos de investigación en mensajes cortos y visuales adecuados para la rápida diseminación en redes sociales.

5. **Democratización:** Estas herramientas bajan la barrera de entrada para la creación de contenido de divulgación de alta calidad, independientemente de las habilidades de diseño o recursos del investigador.

6. **Supervisión Humana Crítica:** A pesar de la asistencia de la IA, la precisión científica, la claridad del mensaje y la integridad del contenido deben ser siempre validadas y controladas por el investigador humano.

REFERENCIAS

References

Aguilar, L. J. (2023). *Ciencias de datos*. Marcombo. https://books.google.com.co/books?id=B4jnEAAAQBAJ

Altman, D. G. (1990). *Practical Statistics for Medical Research*. Chapman & Hall/CRC.

Angulo Medina, A. S., Aguilar Bonilla, M. I., Rodríguez Giraldo, I. D., Montenegro Palacios, J. F., Cáceres Gutiérrez, D. A., & Liscano, Y. (2024). Electroencephalography-Based Brain-Computer Interfaces in Rehabilitation: A Bibliometric Analysis (2013–2023). *Sensors*, *24*(22), 7125. https://doi.org/10.3390/s24227125

Aragón-Muriel, A., Liscano, Y., Upegui, Y., Robledo, S. M., Ramírez-Apan, M. T., et al. (2021). In vitro evaluation of the potential pharmacological activity and molecular targets of new benzimidazole-based Schiff base metal complexes. *Antibiotics*, *10*(6), 728.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*(7604), 452–454. https://doi.org/10.1038/533452a

Bencardino, C. M., & Ediciones, E. (2019). *Estadística básica aplicada*. Ecoe Ediciones. https://books.google.com.co/books?id=WlckEAAAQBAJ

Berktas, M., Hiziroglu, A., Erbaycu, A. E., Er, O., & Kahyaoglu, S. B. (2024). *The Impact of Artificial Intelligence on Healthcare Industry: Volume 1: Non-Clinical Applications*. CRC Press. https://books.google.com.co/books?id=v-YwEQAAQBAJ

Blackard, J. T., Knapke, J. M., Ryan, P. H., Schuckman, S., Veevers, J., Hardie, W. D., & Heubi, J. E. (2022). *Research 101*: An online course introducing medical students to research. *Journal of Clinical and Translational Science*, *6*(1), e102. https://doi.org/10.1017/cts.2022.435

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., et al. (2015). STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*, 351, h5527. https://doi.org/10.1136/bmj.h5527

Caicedo, N., Liscano, Y., & Oñate-Garzón, J. (2025). Bioactive peptides from quinoa (Chenopodium quinoa Willd.) as modulators of the gut microbiome: a scoping review of preclinical evidence. *Nutrients*, 17(20), 3215.

Chan, A. W., Tetzlaff, J. M., Altman, D. G., Laupacis, A., Gøtzsche, P. C., Krleža-Jerić, K., et al. (2013). SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Annals of Internal Medicine*, 158(3), 200–207. https://doi.org/10.7326/0003-4819-158-3-201302050-00583

Chen, J. H., & Asch, S. M. (2017). Machine learning and prediction in medicine — beyond the peak of inflated expectations. *New England Journal of Medicine*, 376(26), 2507–2509. https://doi.org/10.1056/NEJMp1702071

Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology*, 110, 12–22. https://doi.org/10.1016/j.jclinepi.2019.02.004

Ciro, M. B., & Ediciones, E. (2012). *Estadística básica aplicada*. Ecoe Ediciones. https://books.google.com.co/books?id=tKQwDgAAQBAJ

Cobo, E. (2007). *Bioestadística para no estadísticos: Bases para interpretar artículos científicos*. Elsevier Health Sciences Spain. https://books.google.com.co/books?id=IXvZYjtYxkYC

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates.

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Annals of Internal Medicine*, 162(1), 55–63. https://doi.org/10.7326/M14-0697

Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., et al. (2024). TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine

learning methods. *BMJ*, 385, e078378. https://doi.org/10.1136/bmj-2023-078378

Cruz Mosquera, F. E., Lizcano Martinez, S., & Liscano, Y. (2024). Effectiveness of psychobiotics in the treatment of psychiatric and cognitive disorders: a systematic review of randomized clinical trials. *Nutrients*, *16*(9), 1352.

Cruz Mosquera, F. E., Guevara-Montoya, M. C., Serna-Ramirez, V., & Liscano, Y. (2024). Neuroinflammation and schizophrenia: new therapeutic strategies through psychobiotics, nanotechnology, and artificial intelligence (AI). *Journal of Personalized Medicine*, 14(4), 391.

Cruz Mosquera, F. E., Murillo, S. R., Naranjo Rojas, A., Perlaza, C. L., et al. (2024). Effect of exercise and pulmonary rehabilitation in pre- and post-surgical patients with lung cancer: systematic review and meta-analysis. *Medicina*, *60*(11), 1725.

Cruz Mosquera, F. E., Benítez, E. C., Benavides, M. C. C., Muñoz, J. E. C., Castañeda, C. A., et al. (2025). Fluid resuscitation with lactated Ringer vs. normal saline in acute pancreatitis: a systematic review and meta-analysis of clinical trials. *Diseases*, *13*(9), 300.

Cruz Mosquera, F. E., de la Rosa Caldas, M., Naranjo Rojas, A., Perlaza, C. L., et al. (2025). Probiotic, prebiotic, and synbiotic supplementation for the prevention and treatment of acute otitis media: a systematic review and meta-analysis. *Children*, *12*(5), 591.

Cruz Mosquera, F. E., Perlaza, C. L., Naranjo Rojas, A., Murillo Rios, S., et al. (2025). Effectiveness of probiotics, prebiotics, and symbiotic supplementation in cystic fibrosis patients: a systematic review and meta-analysis of clinical trials. *Medicina*, *61*(3), 489.

Cruz Rivera, S., Liu, X., Chan, A. W., Denniston, A. K., Calvert, M. J., & SPIRIT-AI and CONSORT-AI Working Group. (2020). Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nature Medicine*, *26*(9), 1351–1363. https://doi.org/10.1038/s41591-020-1037-7

Cáceres-Gutiérrez, D. A., Bonilla-Bonilla, D. M., Liscano, Y., & Díaz Vallejo, J. A. (2025). From architecture to outcomes: mapping the landscape of digital twins for personalized diabetes care—a scoping review. *Journal of Personalized Medicine*, 15(11), 504.

Daniel, W. W. (2002). *Bioestadistica: Base para el analisis de las ciencias de la salud*. Editorial Limusa S.A. De C.V. https://books.google.com.co/books?id=hT2YPQAACAAJ

de Jesús Celis de la Rosa, A., & Martagón, V. L. (2014). *Bioestadística*. Editorial El Manual Moderno. https://books.google.com.co/books?id=-jPEWCQAAQBAJ

Dueñas, D., Daza, J., & Liscano, Y. (2023). Coinfections and Superinfections Associated with COVID-19 in Colombia: A Narrative Review. *Medicina*, *59*(7), 1336. https://doi.org/10.3390/medicina59071336

Díaz-Vallejo, J. A., Liscano, Y., Hernández, M. del M., Cuji-Galarza, W. D., et al. (2024). Scientific output by Latin American women in pediatric surgical sciences over the past 11 years: a bibliometric and visual analysis. *Journal of Pediatric Surgery*, *59*(9), 1680–1686.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical Methods for Rates and Proportions* (3rd ed.). John Wiley & Sons.

Forister, J. G., & Blessing, J. D. (Eds.). (2020). *Introduction to research and medical literature for health professionals* (Fifth edition). Jones & Bartlett Learning.

Gagnier, J. J., Kienle, G., Altman, D. G., Moher, D., Sox, H., Riley, D., & CARE Group. (2013). The CARE guidelines: consensus-based clinical case reporting guideline development. *BMJ Case Reports*, 2013, bcr2013201554. https://doi.org/10.1136/bcr-2013-201554

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence in-

tervals, and power: a guide to misinterpretations. *European Journal of Epidemiology*, 31(4), 337–350. https://doi.org/10.1007/s10654-016-0149-3

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, *336*(7650), 924–926. https://doi.org/10.1136/bmj.39489.470347.AD

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS Biology*, 13(3), e1002106. https://doi.org/10.1371/journal.pbio.1002106

Horwitz, R. I., Hayes-Conroy, A., Caricchio, R., & Singer, B. H. (2017). From Evidence Based Medicine to Medicine Based Evidence. *The American Journal of Medicine*, 130(11), 1246-1250. https://doi.org/10.1016/j.amjmed.2017.06.012

Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D. G., & Newman, T. B. (2013). *Designing Clinical Research* (4th ed.). Lippincott Williams & Wilkins.

Husereau, D., Drummond, M., Augustovski, F., de Bekker-Grob, E., Briggs, A. H., Carswell, C., et al. (2022). Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement. *Value in Health*, 25(1), 3–9. https://doi.org/10.1016/j.jval.2021.11.1351

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124. https://doi.org/10.1371/journal.pmed.0020124

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: with Applications in R* (2nd ed.). Springer. https://doi.org/10.1007/978-1-0716-1418-1

Jones, H. (2019). *Ciencia de los datos: La guía definitiva sobre análisis de datos, minería de datos, almacenamiento de datos, visualización de datos, Big Data para ... Para principiantes (Spanish Edition)*. Bravex Publications. https://books.google.com.co/books?id=KcfHzAEACAAJ

Kaura, A. (2013). *Evidence-based medicine: Reading and writing medical papers*. Elsevier - Health Sciences Division.

Kelleher, J. D., & Tierney, B. (2021). *Ciencia de datos: La serie de conocimientos esenciales de MIT Press*. Ediciones UC. https://books.google.com.co/books?id=dkpQEAAAQBAJ

Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis: A Self-Learning Text* (3rd ed.). Springer.

Kumar, R. (2012). *Research methodology: A step-by-step guide for beginners* (3. ed., reprinted). Sage Publ.

Kurup, V., & Sakai, T. (2020). Demystifying research in medical education: A novel framework, resources, and ethical challenges. *International Anesthesiology Clinics*, *58*(4), 46-51. https://doi.org/10.1097/AIA.0000000000000289

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied Linear Statistical Models* (5th ed.). McGraw-Hill/Irwin.

Lee, G. S. J., Chin, Y. H., Jiang, A. A., Mg, C. H., Nistala, K. R. Y., Iyer, S. G., Lee, S. S., Chong, C. S., & Samarasekera, D. D. (2021). Teaching Medical Research to Medical Students: A Systematic Review. *Medical Science Educator*, *31*(2), 945-962. https://doi.org/10.1007/s40670-020-01183-w

Liscano, Y., & Sanchez-Palacio, N. (2023). A Critical Look at Omega-3 Supplementation: A Thematic Review. *Healthcare*, *11*(23), 3065. https://doi.org/10.3390/healthcare11233065

Liscano, Y., Oñate-Garzón, J., & Delgado, J. P. (2020). Peptides with dual antimicrobial–anticancer activity: strategies to overcome peptide limitations and rational design of anticancer peptides. *Molecules*, *25*(18), 4245. https://doi.org/10.3390/molecules25184245

Liscano, Y., Anillo Arrieta, L. A., Montenegro, J. F., Prieto-Alvarado, D., et al. (2025). Early warning of infectious disease outbreaks using social media and digital data: a scoping review. *International Journal of Environmental Research and Public Health*, *22*(7), 1104.

Liscano, Y., Bernal, L. M., & Díaz Vallejo, J. A. (2025). Effectiveness of AI-assisted digital therapies for post-stroke aphasia rehabilitation: a systematic review. *Brain Sciences*, *15*(9), 1007.

Liscano, Y., Muñoz Morales, D., Suarez Daza, F., Vidal Cañas, S., et al. (2025). Microbial interventions for inflammatory skin diseases: a systematic review and meta-analysis of atopic dermatitis and psoriasis. *Microorganisms*, *13*(11), 2416.

Liscano, Y., Caicedo, N., & Oñate-Garzón, J. (2025). Unlocking new bioactive peptides from coffee pulp: a metagenomics and AI-driven discovery paradigm. *Foods*, *14*(21), 3682.

Liscano, Y., Martinez Guevara, D., Urriago-Osorio, G. A., & Quintana, J. (2025). Beyond conventional meta-analysis: a meta-learning model to predict cohort-level mortality after transcatheter aortic valve replacement (TAVR). *Journal of Cardiovascular Development and Disease*, *12*(10), 376.

Liscano, Y., Arias Coronel, F., & Martínez, D. (2025). Effects of virtual reality on motor function and balance in incomplete spinal cord injury: a systematic review and meta-analysis of controlled trials. *Brain Sciences*, *15*(10), 1071.

Liscano, Y., Álvarez-Caballero, J. M., & Aragón-Muriel, A. (2026). Spider venom peptides as potential allosteric inhibitors of undecaprenyl diphosphatase (UppP) from Acinetobacter baumannii: in silico identification and characterization. *Toxins*, *18*(5), 210.

Liscano, Y., Salamanca, C. H., Vargas, L., Cantor, S., Laverde-Rojas, V., et al. (2019). Increases in hydrophilicity and charge on the polar face of alyteserin 1c helix change its selectivity towards gram-positive bacteria. *Antibiotics*, *8*(4), 238.

Liscano, Y., Oñate-Garzón, J., & Ocampo-Ibáñez, I. D. (2020). In silico discovery of antimicrobial peptides as an alternative to control SARS-CoV-2. *Molecules*, *25*(23), 5535.

Liscano, Y. (2026). *R para la investigación clínica y epidemiológica: del dato crudo a la inferencia causal, un pipeline reproducible en R*. https://www.amazon.in/dp/B0H9B27KKG

Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J., Denniston, A. K., & SPIRIT-AI and CONSORT-AI Working Group. (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nature Medicine*, *26*(9), 1364–1374. https://doi.org/10.1038/s41591-020-1034-x

Lozada-Martinez, I. D., Neira-Rodado, D., Martinez-Guevara, D., Cruz-Soto, H. S., et al., & Liscano, Y. (2025). Why is it important to implement meta-research in universities and institutes with medical research activities? *Frontiers in Research Metrics and Analytics*, 10, 1497280.

Lozada-Martinez, I. D., Hernandez-Paez, D. A., Palacios Velasco, I., et al. (2024). Meta-research in geriatric surgery: improving the quality of surgical evidence for older persons in a multidimensional-scale research field. *Journal of Clinical Medicine*, *13*(18), 5441.

López-Kleine, L. (2017). *Bioestadística*. Universidad Nacional de Colombia. https://books.google.com.co/books?id=dd_0DwAAQBAJ

Marta, M. M. (2015). A BRIEF HISTORY OF THE EVOLUTION OF THE MEDICAL RESEARCH ARTICLE. *Medicine and Pharmacy Reports*, *88*(4), 567-570. https://doi.org/10.15386/cjmed-560

Martinez Guevara, D., Vidal Cañas, S., Palacios, I., Gómez, A., Estrada, M., Gallego, J., & Liscano, Y. (2024). Effectiveness of Probiotics, Prebiotics, and Synbiotics in Managing Insulin Resistance and Hormonal Imbalance in Women with Polycystic Ovary Syndrome (PCOS): A Systematic Review of Randomized Clinical Trials. *Nutrients*, *16*(22), 3916. https://doi.org/10.3390/nu16223916

Masic, I., Miokovic, M., & Muhamedagic, B. (2008). Evidence Based Medicine—New Approaches and Challenges. *Acta Informatica Medica*, *16*(4), 219. https://doi.org/10.5455/aim.2008.16.219-225

Montenegro-Palacios, J. F., Vidal-Cañas, S., Murillo-Benítez, N. E., et al. (2025). Common risk factors for atrial fibrillation after transcatheter aortic valve implantation: a systematic review from 2009 to 2024. *Journal of Cardiovascular Development and Disease*, 12(3), 90.

Montoya-Quintero, K. F., Galván-Barrios, J., Martinez-Guevara, D., Dueñas, D., et al. (2025). Bridging the gap: cancer scientific equity, global child health, and distribution of CAR T-cell therapy clinical trials in childhood cancer. *Frontiers in Pediatrics*, 13, 1611187.

Moons, K. G. M., Damen, J. A. A., Kaul, T., Hooft, L., Navarro, C. A., Dhiman, P., et al. (2025). PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*, 388, e082505. https://doi.org/10.1136/bmj-2024-082505

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., et al. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1, 0021. https://doi.org/10.1038/s41562-016-0021

Murillo, S. R., Hinestroza Mancilla, A. M., Pérez Ordoñez, L. M., et al. (2025). Physical training and pulmonary rehabilitation in patients with cystic fibrosis: a systematic review and meta-analysis of clinical trials. *Healthcare*, 13(16), 2017.

Noland, D., Drisko, J. A., & Wagner, L. (Eds.). (2020). *Integrative and Functional Medical Nutrition Therapy: Principles and Practices*. Springer International Publishing. https://doi.org/10.1007/978-3-030-30730-1

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future — big data, machine learning, and clinical medicine. *New England Journal of Medicine*, 375(13), 1216–1219. https://doi.org/10.1056/NEJMp1606181

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372, n71. https://doi.org/10.1136/bmj.n71

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, *49*(12), 1373–1379. https://doi.org/10.1016/s0895-4356(96)00236-3

Perakakis, N., Yazdani, A., Karniadakis, G. E., & Mantzoros, C. (2018). Omics, big data and machine learning as tools to propel understanding of biological mechanisms and to discover novel diagnostics and therapeutics. *Metabolism*, *87*, A1-A9. https://doi.org/10.1016/j.metabol.2018.08.002

Perlaza, C. L., Rojas, S. M., Choco, L. D., González, M. P. P., Cruz Mosquera, F. E., et al. (2025). Effectiveness of digital health tools for asthma self-management: a systematic review and meta-analysis of clinical trials. *Applied Sciences*, *15*(23), 12471.

Polgar, S., & Thomas, S. A. (2021). *Introducción a la investigación en ciencias de la salud*. Elsevier. https://books.google.com.co/books?id=jzwoEAAAQBAJ

Prieto-Alvarado, D. E., Parada-Gereda, H. M., Molano, D., Martinez, Y. L., et al. (2025). Risk factors and outcomes of ventilator-associated pneumonia in patients with traumatic brain injury: a systematic review and meta-analysis. *Journal of Critical Care*, 85, 154922.

Pérez, V. H. E., Cruz Mosquera, F. E., de la Rosa Caldas, M., Rodríguez, O. A. P., et al. (2025). Effectiveness of intravenous and nebulized MgSO4 in children with asthma exacerbation: a systematic review and meta-analysis of clinical trials. *Children*, *12*(8), 1064.

Rahman, M. M., Ghoshal, U. C., Ragunath, K., Jenkins, G., Rahman, M., Edwards, C., Hasan, M., & Taylor-Robinson, S. D. (2020). Biomedical research in developing countries: Opportunities, methods, and challenges. *Indian Journal of Gastroenterology*, *39*(3), 292-302. https://doi.org/10.1007/s12664-020-01056-5

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, *380*(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

Ratnani, I., Fatima, S., Abid, M. M., Surani, Z., & Surani, S. (2023). Evidence-Based Medicine: History, Review, Criticisms, and Pitfalls. *Cureus*. https://doi.org/10.7759/cureus.35266

Restrepo, J. C., Martínez Guevara, D., Pareja López, A., Montenegro Palacios, J. F., & Liscano, Y. (2024). Identification and Application of Emerging Biomarkers in Treatment of Non-Small-Cell Lung Cancer: Systematic Review. *Cancers*, *16*(13), 2338. https://doi.org/10.3390/cancers16132338

Restrepo, J. C., Dueñas, D., Corredor, Z., & Liscano, Y. (2023). Advances in genomic data and biomarkers: revolutionizing NSCLC diagnosis and treatment. *Cancers*, *15*(13), 3474.

Riley, R. D., Ensor, J., Snell, K. I. E., Harrell, F. E., Martin, G. P., Reitsma, J. B., et al. (2019). Minimum sample size for developing a multivariable prediction model. *Statistics in Medicine*, *38*(7), 1276–1296. https://doi.org/10.1002/sim.7992

Romero, A. R. V., México, U. N. A. de, de la Torre, G. S. G., & Altamirano, L. M. (2011). *Epidemiología y estadística en salud pública*. McGraw-Hill Interamericana Editores. https://books.google.com.co/books?id=kfsyLgEACAAJ

Rosa, J. M., & Frutos, E. L. (2022). Ciencia de datos en salud: Desafíos y oportunidades en América Latina. *Revista Médica Clínica Las Condes*. https://api.semanticscholar.org/CorpusID:254306274

Rosner, B. (2015). *Fundamentals of Biostatistics* (8th ed.). Cengage Learning.

Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern Epidemiology* (3rd ed.). Lippincott Williams & Wilkins.

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ*, *312*(7023), 71–72. https://doi.org/10.1136/bmj.312.7023.71

Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-Based Medicine: How to Practice and Teach EBM* (2nd ed.). Churchill Livingstone.

Schulz, K. F., Altman, D. G., Moher, D., & CONSORT Group. (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*, 340, c332. https://doi.org/10.1136/bmj.c332

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Sterne, J. A. C., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., et al. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, 355, i4919. https://doi.org/10.1136/bmj.i4919

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., et al. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*, 366, l4898. https://doi.org/10.1136/bmj.l4898

Steyerberg, E. W. (2019). *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating* (2nd ed.). Springer. https://doi.org/10.1007/978-3-030-16399-0

Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., et al. (2000). Meta-analysis of observational studies in epidemiology: a proposal for reporting (MOOSE). *JAMA*, 283(15), 2008–2012. https://doi.org/10.1001/jama.283.15.2008

Szklo, M., & Nieto, F. J. (2019). *Epidemiology: Beyond the Basics* (4th ed.). Jones & Bartlett Learning.

Tang, C.-M., Qureshi, Z., & Fischbacher, C. (Eds.). (2015). *The unofficial guide to medical research, audit and teaching* (First edition). Zeshan Qureshi.

Ter Wee, M. M., & Lissenberg-Witte, B. I. (2019). *A Quick Guide on How to Conduct Medical Research: From Set-Up to Publication*. Bohn Stafleu van Loghum. https://doi.org/10.1007/978-90-368-2248-0

Thomas, C. G. (2021). *Research Methodology and Scientific Writing*. Springer International Publishing. https://doi.org/10.1007/978-3-030-64865-7

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Trejos, M., Aristizabal, Y., Aragón-Muriel, A., Oñate-Garzón, J., & Liscano, Y. (2023). Characterization and Classification In Silico of Peptides with Dual Activity (Antimicrobial and Wound Healing). *International Journal of Molecular Sciences*, 24(17), 13091. https://doi.org/10.3390/ijms241713091

Urbano-Cano, A. L., Álvarez-Rosero, R. E., & Liscano, Y. (2025). Association of CETP, APOA5, IL6, and PON1 gene variants with dyslipidemia and cardiovascular risk in a population from Cauca Department, Colombia. *Genes*, *16*(5), 545.

Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gøtzsche, P. C., Mulrow, C. D., Pocock, S. J., et al. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Medicine*, 4(10), e297. https://doi.org/10.1371/journal.pmed.0040297

Villanueva Parra, I., Muñoz Diaz, V., Martinez Guevara, D., Cruz Mosquera, F. E., Prieto-Alvarado, D. E., & Liscano, Y. (2024). A Scoping Review of Angiostrongyliasis and Other Diseases Associated with Terrestrial Mollusks, Including Lissachatina fulica: An Overview of Case Reports and Series. *Pathogens*, *13*(10), 862. https://doi.org/10.3390/pathogens13100862

Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2012). *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models* (2nd ed.). Springer.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *The Lancet*, *370*(9596), 1453–1457. https://doi.org/10.1016/S0140-6736(07)61602-X

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: context, process, and purpose. *The American Statistician*, 70(2), 129–133. https://doi.org/10.1080/00031305.2016.1154108

Wheelan, C. (2013). *Naked Statistics: Stripping the Dread from the Data*. W. W. Norton. https://books.google.com.co/books?id=BgFJfC_CrTAC

Whitehead, A. J., Woodring, T., & Klein, B. S. (2024). Immunity to fungi and vaccine considerations. *Cell Host & Microbe*, 32(10), 1681-1690. https://doi.org/10.1016/j.chom.2024.09.011

Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., et al. (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529–536. https://doi.org/10.7326/0003-4819-155-8-201110180-00009

Williams, N. H. (2024). *Artificial Intelligence and Healthcare: The Impact of Algorithmic Bias on Health Disparities*. Springer International Publishing. https://books.google.com.co/books?id=E7jrEAAAQBAJ

Wolff, R. F., Moons, K. G. M., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., et al. (2019). PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine*, 170(1), 51–58. https://doi.org/10.7326/M18-1376

Zacarías, H., & Supo, J. (2020). *MetodologÍa de la InvestigaciÓn CientÍfica: Para Las Ciencias de la Salud y Las Ciencias Sociales*. Bioestadistico. https://books.google.com.co/books?id=WruXzQEACAAJ

ACERCA DEL AUTOR

About the Authors

Yamil Liscano Martínez

Filiación institucional: Universidad Saniago de Cali.
yamil.liscano00@usc.edu.co; liscano2@gmail.com
https://orcid.org/0000-0002-2674-8725

Doctor en Biotecnología de la Universidad de Antioquia, donde se centró en la identificación y caracterización de péptidos con potencial actividad antibacteriana y anticancerígena. También posee una maestría en Ciencias de la Biotecnología de la Universidad Nacional de Colombia, una especialización en Epidemiología de la Universidad de Caldas, y otra en Desarrollo Farmacéutico y Farmacología. Su formación de pregrado en Bacteriología y Laboratorio Clínico la realizó en la Universidad del Valle.

Destacado investigador y académico en el campo de las ciencias biológicas y la biotecnología con más de 20 publicaciones indexadas en revistas internacionales en cuartiles Q1 y Q2. Actualmente, se desempeña como docente titular en la Universidad Santiago de Cali, donde imparte sus conocimientos en áreas como la inmunología, enfermedades infecciosas, modelamiento molecular, bioinformática, dinámica molecular y redes de interacción. En cuanto a su área de especialización se enfoca en varias ramas de las ciencias naturales y médicas, incluyendo la biología celular y microbiología, bioquímica y biología molecular, biología teórica, biotecnología en salud, epidemiología y bioinformática.

PARES EVALUADORES

Peer reviewers

Wilfred Fabian Rivera Martinez
Centro de Desarrollo Tecnológico Creatic. Popayán, Colombia
Investigador Asociado (I)
https://orcid.org/0000-0003-2888-7929

Adriana Orejuela Upegui
Institución Universitaria Escuela Nacional del Deporte. Cali, Colombia
Investigador Junior (IJ)
https://orcid.org/0000-0002-2544-3567

Angie Milena Cárdenas Silva
Fundación Hospital Infantil Los Ángeles. Pasto, Colombia
Investigador Junior (IJ)
https://orcid.org/0000-0002-6190-8341

Rodrigo Andrés Sarria Villa
Universidad del Cauca. Popayán, Colombia
Investigador Junior (IJ)
https://orcid.org/0000-0003-1295-7865

Anibal Vicente Arteaga Noriega
Corporación Universitaria Remington. Medellín, Colombia
Investigador Senior (IS)
https://orcid.org/0000-0002-6612-1169

Pedro Antonio Calero Saa
Institución Universitaria Escuela Nacional del Deporte. Cali, Colombia
Investigador Asociado (I)
https://orcid.org/0000-0002-9978-7944

Distribución y Comercialización /
Distribution and Marketing:

Universidad Santiago de Cali
Publicaciones / Editorial USC
Bloque 7 - Piso 5
Calle 5 No. 62 - 00
Tel: (57+) (2+) 518 3000
Ext. 323 - 324 - 414
editor@usc.edu.co
publica@usc.edu.co
Cali, Valle del Cauca
Colombia

Diagramación / Design & Layout by:

Diego Pablo Guerra Gonzalez
diagramacioneditorialusc@usc.edu.co
Tel: (57+) (2+) 518 3000 Ext. 9131

Este libro se diagramó utilizando fuentes tipográficas Literata en sus respectivas variaciones a 11 puntos en el contenido y Fira Sans, para los capitulares 17 puntos.

Impreso en el mes de Noviembre.
Se imprimieron 50 ejemplares en la
Editorial Diké S.A.S
Tel: (+57) 301 242 7399
Bogotá - Colombia
2026

Fue publicado por la Facultad de Salud
y de la Universidad Santiago de Cali.

Querido lector:

Si alguna vez has sentido que la investigación es un mundo ajeno a la práctica clínica, este libro se escribió pensando en ti. Durante años he acompañado a estudiantes, residentes y profesionales de la salud que llegan a la investigación por obligación: una tesis por terminar, un artículo por publicar, un protocolo que nadie les enseñó a escribir. Muchos llegan con miedo a la estadística y con la sensación de que la metodología es un laberinto sin salida. Este manual es la brújula que me hubiera gustado tener cuando empecé.

En estas páginas encontrarás un recorrido completo y ordenado, desde la pregunta clínica hasta la publicación. Empezamos por los fundamentos epistemológicos y la medicina basada en la evidencia. Después vienen la descripción y visualización de datos clínicos, la probabilidad, la inferencia y la elección de la prueba estadística adecuada. Luego llegan la epidemiología, los diseños de estudio, la lectura crítica de la literatura y la redacción del artículo científico. Por último, recorremos el ciclo de vida de los datos en salud y el uso responsable de la inteligencia artificial para diseñar protocolos, analizar resultados, gestionar referencias y divulgar la ciencia.

No necesitas ser matemático ni programador para aprovecharlo. Cada concepto se explica en un lenguaje claro y con ejemplos tomados de la práctica médica. La idea es que entiendas por qué se hace cada cosa, no solo cómo hacerla. Tampoco pretendo convertirte en un investigador de laboratorio. Lo que quiero es darte criterio para leer un artículo y saber si sus conclusiones se sostienen, para diseñar un estudio que responda una pregunta real y para usar las nuevas herramientas digitales sin perder el rigor.

La inteligencia artificial está cambiando la manera de hacer ciencia, pero solo potencia a quien domina los fundamentos. Por eso este libro une lo clásico con lo emergente.

Te invito a leerlo con lápiz en mano, a volver a él cuando lo necesites y a hacer de la investigación una extensión natural de la curiosidad que te trajo a las ciencias de la salud.

Yamil Liscano Martínez

VIGILADA MINEDUCACIÓN
USC UNIVERSIDAD SANTIAGO DE CALI
EDITORIAL

ISBN: 978-628-7914-26-1
9 786287 914261