

# Aplicación de métodos bioestadísticos en investigaciones en ciencias de la salud: revisión de alcance de literatura científica y metodológica clásica y contemporánea.

## Application of biostatistical methods in health sciences research: scope review of classical and contemporary scientific and methodological literature.

Lisette Beatriz Villavicencio Cedeño<sup>1</sup>, [liso\\_villa@hotmail.com](mailto:liso_villa@hotmail.com), <https://orcid.org/0000-0001-8475-3947>; Marieta del Jesús Azúa Menéndez<sup>2</sup>, [marieta.azua@unesum.edu.ec](mailto:marieta.azua@unesum.edu.ec), <https://orcid.org/0000-0002-5601-6621>; Caleb Isaac Chilán Santana<sup>2</sup>, [caleb.chilan@unesum.edu.ec](mailto:caleb.chilan@unesum.edu.ec), <https://orcid.org/0000-0002-2832-8759>; Yoiler Batista Garcet<sup>3</sup>, [ybatista@sangregorio.edu.ec](mailto:ybatista@sangregorio.edu.ec), <https://orcid.org/0000-0002-7851-5763>

<sup>1</sup> Universidad Nacional de Tumbes, Perú, Unidad Educativa Cristo Rey, Portoviejo. Manabí, Ecuador.

<sup>2</sup> Universidad Estatal del Sur de Manabí, docente de la carrera de Laboratorio Clínico, Jipijapa. Manabí, Ecuador.

<sup>3</sup> Universidad San Gregorio de Portoviejo, Carrera de Medicina. Manabí, Ecuador.

\* Correspondencia: [liso\\_villa@hotmail.com](mailto:liso_villa@hotmail.com)

Recibido: 28/4/26; Aceptado: 19/5/26; Publicado: 21/5/26

### Resumen

**Introducción:** A pesar de la importancia de la bioestadística en la investigación en ciencias de la salud, en la literatura biomédica persisten múltiples deficiencias. **Objetivo:** Analizar la aplicación de métodos bioestadísticos en investigaciones en ciencias de la salud mediante una revisión de alcance con enfoque sistemático de la literatura científica reciente, con énfasis en fundamentos inferenciales, problemas de aplicación y alternativas metodológicas contemporáneas. **Métodos:** Se realizó una revisión de alcance con enfoque sistemático, orientada por PRISMA-ScR. La búsqueda se efectuó principalmente en Scopus y Web of Science, con exploración complementaria en PubMed, ERIC, CINAHL y Google Scholar. Se utilizaron ecuaciones de búsqueda relacionadas con bioestadística, métodos estadísticos, investigación médica, epidemiología, salud pública e investigación clínica. Se identificaron 245 registros; luego de eliminar duplicados y aplicar criterios de elegibilidad, se incluyeron 32 estudios. La información fue organizada en tres categorías: fundamentos teóricos de la inferencia estadística, problemas y limitaciones en la aplicación de métodos bioestadísticos, y enfoques metodológicos contemporáneos. **Resultados:** Los estudios revisados evidenciaron predominio del enfoque frecuentista, especialmente del uso de pruebas de hipótesis y del valor p. Se identificaron errores recurrentes en la interpretación de la significación estadística, uso inapropiado de pruebas, escaso reporte de supuestos, deficiente manejo de comparaciones múltiples, limitada consideración de variables de confusión e insuficiente integración de relevancia clínica. Asimismo, se observaron propuestas emergentes orientadas al uso de intervalos de confianza, tamaños del efecto, modelos multivariados, análisis bayesianos, metaciencia, prácticas de ciencia abierta y mejora del reporte estadístico. **Conclusiones:** La bioestadística sigue siendo un eje decisivo para la calidad de la investigación en ciencias de la salud; no obstante, su aplicación práctica presenta brechas importantes. Superar estas limitaciones exige fortalecer la formación estadística, promover una interpretación inferencial más crítica, abandonar la dependencia exclusiva del valor p, integrar medidas de magnitud y precisión, mejorar la transparencia del reporte y fomentar enfoques analíticos complementarios orientados a la reproducibilidad de la evidencia.

**Palabras clave:** bioestadística; ciencias de la salud; inferencia estadística; valor p; revisión de alcance; investigación biomédica; inferencia bayesiana; reproducibilidad.

## Abstract

**Introduction:** Despite the importance of biostatistics in health sciences research, numerous shortcomings persist in the biomedical literature. **Objective:** To analyze the application of biostatistical methods in health sciences research through a comprehensive, systematic review of recent scientific literature, with an emphasis on inferential foundations, application problems, and contemporary methodological alternatives. **Methods:** A comprehensive, systematic review was conducted, guided by PRISMA-ScR. The search was primarily performed in Scopus and Web of Science, with supplementary searches in PubMed, ERIC, CINAHL, and Google Scholar. Search terms related to biostatistics, statistical methods, medical research, epidemiology, public health, and clinical research were used. 245 records were identified; after removing duplicates and applying eligibility criteria, 32 studies were included. The information was organized into three categories: theoretical foundations of statistical inference, problems and limitations in the application of biostatistical methods, and contemporary methodological approaches. **Results:** The reviewed studies showed a predominance of the frequentist approach, especially regarding the use of hypothesis tests and the p-value. Recurring errors were identified in the interpretation of statistical significance, inappropriate use of tests, poor reporting of assumptions, inadequate handling of multiple comparisons, limited consideration of confounding variables, and insufficient integration of clinically relevant data. Emerging proposals were also observed, focusing on the use of confidence intervals, effect sizes, multivariate models, Bayesian analyses, metascience, open science practices, and improvements in statistical reporting. **Conclusions:** Biostatistics remains a crucial element for the quality of health sciences research; however, its practical application presents significant gaps. Overcoming these limitations requires strengthening statistical training, promoting a more critical inferential interpretation, abandoning the exclusive dependence on the p-value, integrating measures of magnitude and precision, improving the transparency of reporting, and fostering complementary analytical approaches oriented towards the reproducibility of evidence.

**Keywords:** biostatistics; health sciences; statistical inference; p-value; scoping review; biomedical research; Bayesian inference; reproducibility.

---

## 1. Introducción

La investigación en ciencias de la salud se desarrolla en un escenario marcado por la complejidad biológica, la variabilidad individual, la incertidumbre diagnóstica, la heterogeneidad poblacional y la necesidad permanente de tomar decisiones basadas en evidencia. En este contexto, la bioestadística no debe entenderse como un recurso accesorio ni como una etapa mecánica ubicada al final del estudio, sino como un componente transversal del razonamiento científico. Su función inicia desde la formulación de la pregunta de investigación, continúa en la definición del diseño metodológico, orienta la selección de variables y procedimientos de medición, permite calcular tamaños muestrales, organiza la recolección de datos, estructura el análisis e interviene de manera decisiva en la interpretación de los resultados (1).

Pero el uso de la bioestadística en investigación en la literatura científica biomédica tiene sus limitaciones. Uno de los elementos más debatidos en la investigación en salud es el uso del valor p. Históricamente, la inferencia estadística frecuentista consolidó el valor p y las pruebas de hipótesis como instrumentos centrales para evaluar la compatibilidad de los datos con una hipótesis nula (2).

Esta tradición tiene raíces en los aportes de Fisher y en el enfoque de Neyman-Pearson, que estructuró la toma de decisiones mediante el control de errores tipo I y tipo II (3). Sin embargo, el valor  $p$  suele interpretarse erróneamente como la probabilidad de que la hipótesis nula sea verdadera, como la probabilidad de que el resultado sea producto del azar o incluso como una medida directa de importancia clínica(4). Estas interpretaciones han sido cuestionadas por declaraciones metodológicas que recomiendan utilizar el valor  $p$  con prudencia y siempre dentro del contexto del diseño, los datos y la evidencia disponible (5). Además, se ha señalado que los valores  $p$ , los intervalos de confianza y la potencia estadística son con frecuencia malinterpretados en la investigación biomédica (6).

La dependencia excesiva del umbral  $p < 0,05$  ha favorecido una cultura de decisión dicotómica: significativo/no significativo, útil/no útil, publicable/no publicable. Esta lógica puede empobrecer la interpretación científica al reducir fenómenos complejos a una frontera arbitraria. Además, puede favorecer prácticas como la repetición de análisis hasta obtener significación, la selección selectiva de resultados y el subreporte de análisis negativos. En ciencias de la salud, donde las decisiones pueden impactar diagnósticos, tratamientos y políticas sanitarias, esta simplificación representa un desafío metodológico y ético (7).

Otro problema frecuente es la omisión de medidas complementarias como los intervalos de confianza y los tamaños del efecto. Mientras el valor  $p$  informa sobre la compatibilidad entre los datos y una hipótesis bajo supuestos específicos, los intervalos de confianza aportan información sobre la precisión de la estimación y los tamaños del efecto permiten valorar la magnitud de la diferencia o asociación observada (8-9). En investigación clínica, una diferencia puede ser estadísticamente significativa y, al mismo tiempo, irrelevante desde el punto de vista clínico. También puede ocurrir lo contrario: un resultado no significativo puede ser potencialmente importante si el estudio carece de potencia estadística suficiente. Por ello, la interpretación rigurosa exige integrar significación estadística, magnitud del efecto, precisión, plausibilidad biológica, consistencia con la evidencia previa y relevancia clínica (10).

La selección inadecuada de pruebas estadísticas constituye otra limitación persistente. En investigaciones en salud se observan errores como aplicar pruebas paramétricas sin verificar supuestos, usar pruebas para muestras independientes cuando los datos son pareados, analizar variables ordinales como si fueran continuas sin justificación, emplear múltiples comparaciones sin ajustes, ignorar la distribución de los datos, no controlar variables de confusión, no evaluar la bondad de ajuste de los modelos estadísticos planteados o interpretar asociaciones bivariadas como si fueran relaciones causales. Estos errores no siempre son visibles para el lector no especializado, pero pueden modificar sustancialmente las conclusiones (1).

La reproducibilidad científica también se ha convertido en un eje crítico. La falta de transparencia en el análisis, la escasa disponibilidad de datos, la ausencia de protocolos, el uso flexible de decisiones analíticas y el sesgo de publicación han contribuido a que muchos resultados no puedan ser replicados (11-13). En este contexto, la bioestadística se relaciona directamente con la metaciencia, la ciencia abierta y los estándares de reporte. No se trata únicamente de aplicar métodos más avanzados, sino de reportarlos con claridad, justificar las decisiones analíticas y permitir que otros investigadores puedan comprender, evaluar y, cuando sea posible, reproducir el análisis (14-15).

Frente a estas limitaciones, han surgido propuestas metodológicas orientadas a mejorar la interpretación de los datos. Entre ellas se encuentran el mayor uso de intervalos de confianza, tamaños del efecto, análisis de sensibilidad, modelos multivariados, enfoques bayesianos, estimaciones basadas en verosimilitud, correcciones por comparaciones múltiples, preregistro de

estudios, publicación de protocolos, disponibilidad de bases de datos y códigos analíticos en abierto, así como guías de reporte específicas para diferentes diseños. La inferencia bayesiana, en particular, ha ganado interés porque permite integrar conocimiento previo con datos observados y expresar la evidencia de manera probabilística (16); el factor de Bayes se ha propuesto como una alternativa o complemento al valor  $p$ , al permitir comparar directamente la evidencia relativa entre hipótesis (17). Los modelos bayesianos ofrecen herramientas para modelar incertidumbre y analizar datos complejos en investigación biomédica y clínica (18-19).

En consecuencia, resulta necesario revisar la evidencia disponible sobre la aplicación de métodos bioestadísticos en investigaciones en ciencias de la salud, identificando patrones de uso, limitaciones recurrentes y alternativas metodológicas contemporáneas para reconocer deficiencias metodológicas, orientar procesos formativos, fortalecer estándares de reporte y contribuir a una investigación más rigurosa, transparente y útil para la toma de decisiones. El presente estudio tuvo como objetivo analizar la aplicación de métodos bioestadísticos en investigaciones en ciencias de la salud mediante una revisión de alcance orientada por procedimientos sistemáticos de búsqueda, selección y síntesis de literatura científica y metodológica clásica y contemporánea, con el propósito de contribuir al fortalecimiento del rigor metodológico, la calidad de la evidencia científica y la toma de decisiones basada en evidencia.

## 2. Métodos

### 2.1 Diseño del estudio

Se realizó una revisión de alcance orientada por procedimientos sistemáticos de búsqueda, selección, extracción y síntesis de literatura científica y metodológica clásica y contemporánea (20-21). El proceso metodológico se orientó por la extensión PRISMA para revisiones de alcance, conocida como PRISMA-ScR (22). Se utilizó este marco porque proporciona criterios para reportar de manera transparente la identificación, selección, extracción y síntesis de fuentes de evidencia en revisiones de alcance. Aunque el estudio empleó procedimientos sistemáticos de búsqueda, cribado y síntesis, se evitó presentarlo como una revisión sistemática clásica de intervención, debido a que el objetivo fue cartográfico y descriptivo más que evaluativo de efectividad (22)

### 2.2 Pregunta de revisión

La pregunta que orientó la revisión fue: ¿cómo se aplican, interpretan y reportan los métodos bioestadísticos en investigaciones en ciencias de la salud según la literatura científica y metodológica clásica y contemporánea? A partir de esta pregunta general se derivaron tres interrogantes específicos, ¿cuáles son los principales fundamentos teóricos e inferenciales que sustentan la aplicación de la bioestadística en investigaciones en salud? ¿qué problemas, errores o limitaciones se reportan con mayor frecuencia en el uso de métodos bioestadísticos? ¿qué enfoques metodológicos contemporáneos se proponen para mejorar la interpretación, transparencia y reproducibilidad de los análisis estadísticos?

### 2.3 Fuentes de información

La búsqueda bibliográfica se realizó principalmente en Scopus y Web of Science, por tratarse de bases de datos multidisciplinarias con amplia cobertura de literatura científica revisada por pares. Con el fin de ampliar la sensibilidad de la búsqueda, se exploraron adicionalmente PubMed, ERIC, CINAHL y Google Scholar en el mes de abril de 2026. Estas fuentes complementarias permitieron verificar la existencia de publicaciones relevantes en salud pública, educación en ciencias de la salud, investigación clínica, epidemiología y metodología científica.

#### 2.4 Estrategia de búsqueda

La estrategia de búsqueda combinó términos relacionados con bioestadística, métodos estadísticos e investigación en salud. La ecuación principal fue:

ALL = ("biostatistics" OR "statistical methods") AND ("health sciences" OR "medical research" OR "epidemiology").

Posteriormente, se realizó una segunda iteración con términos ampliados:

("biostatistics" OR "statistical analysis" OR "epidemiological methods") AND ("health sciences" OR "clinical research" OR "public health").

La combinación de términos se ajustó según las particularidades de cada base de datos. Se consideraron publicaciones en español e inglés, con énfasis en literatura científica y metodológica relevante, clásica y contemporánea y documentos metodológicos de alta relevancia conceptual. La búsqueda no se limitó exclusivamente a estudios empíricos, debido a que el objetivo de la revisión incluía fundamentos teóricos, debates inferenciales y propuestas metodológicas.

#### 2.5 Criterios de inclusión

Se incluyeron estudios y documentos científicos que cumplieron los siguientes criterios:

- Publicaciones metodológicas relacionadas con la aplicación, interpretación o fundamentación de métodos bioestadísticos en ciencias de la salud.
- Estudios empíricos, revisiones, ensayos metodológicos, artículos teóricos o documentos de consenso vinculados con inferencia estadística, valor p, pruebas de hipótesis, modelos estadísticos, reproducibilidad, métodos bayesianos o calidad del reporte estadístico.
- Trabajos publicados en revistas científicas, documentos académicos reconocidos o fuentes metodológicas de amplia aceptación.
- Estudios con pertinencia directa para investigación clínica, epidemiológica, biomédica, de salud pública o educación en ciencias de la salud.
- Publicaciones en español o inglés.

#### 2.6 Criterios de exclusión

Se excluyeron:

- Publicaciones sin relación explícita con métodos bioestadísticos o inferencia estadística.
- Estudios centrados exclusivamente en desarrollos matemáticos sin aplicabilidad en ciencias de la salud.
- Documentos de opinión sin sustento metodológico suficiente.
- Trabajos duplicados.
- Publicaciones cuyo texto completo no estuvo disponible.
- Estudios con información insuficiente para categorizar sus aportes.
- Cartas al editor, editoriales sin desarrollo metodológico o documentos sin aporte conceptual pertinente para los objetivos de la revisión.

La búsqueda inicial identificó 245 registros. Luego de eliminar 35 duplicados, quedaron 210 registros para revisión de títulos y resúmenes. En esta fase se excluyeron 130 estudios por no cumplir los criterios de inclusión, principalmente por falta de pertinencia con ciencias de la salud, ausencia de discusión bioestadística o enfoque exclusivamente técnico sin aplicación biomédica. Posteriormente se evaluaron 80 artículos a texto completo. De ellos, 48 fueron excluidos por presentar deficiencias metodológicas para los fines de la revisión, ausencia de análisis bioestadístico sustantivo o falta de correspondencia con el objetivo del estudio. Finalmente, se incluyeron 32 estudios en la síntesis (figura 1).

### 2.7 Extracción y organización de la información

La información de los estudios incluidos fue extraída mediante una matriz diseñada para registrar: autor, año, tipo de publicación, objetivo, área de aplicación, método estadístico o problema analizado, principales hallazgos y aporte a la revisión. Posteriormente, los estudios fueron agrupados en tres categorías temáticas:

- Fundamentos teóricos de la inferencia estadística.
- Problemas y limitaciones en la aplicación de métodos bioestadísticos.
- Enfoques metodológicos contemporáneos y alternativas analíticas.

Esta categorización permitió organizar la evidencia de manera interpretativa y facilitar la integración de hallazgos heterogéneos. No se realizó metaanálisis debido a la diversidad conceptual, metodológica y documental de las fuentes incluidas mediante la utilización de zotero.

### 2.8 Síntesis de la evidencia

La síntesis se realizó de forma narrativa y temática. Se identificaron patrones recurrentes, coincidencias conceptuales, tensiones metodológicas y propuestas emergentes. La información fue presentada mediante tablas de resumen que agrupan los estudios según su contribución principal. Este procedimiento permitió integrar literatura histórica, metodológica y aplicada, sin forzar una homogeneidad inexistente entre documentos de distinta naturaleza.

### 2.9 Consideraciones éticas

Por tratarse de una revisión de literatura basada en fuentes publicadas, el estudio no requirió consentimiento informado ni aprobación de un comité de ética. No se trabajó con datos individuales de participantes humanos. Se respetaron los principios de integridad académica, citación de fuentes y uso responsable de la evidencia científica.

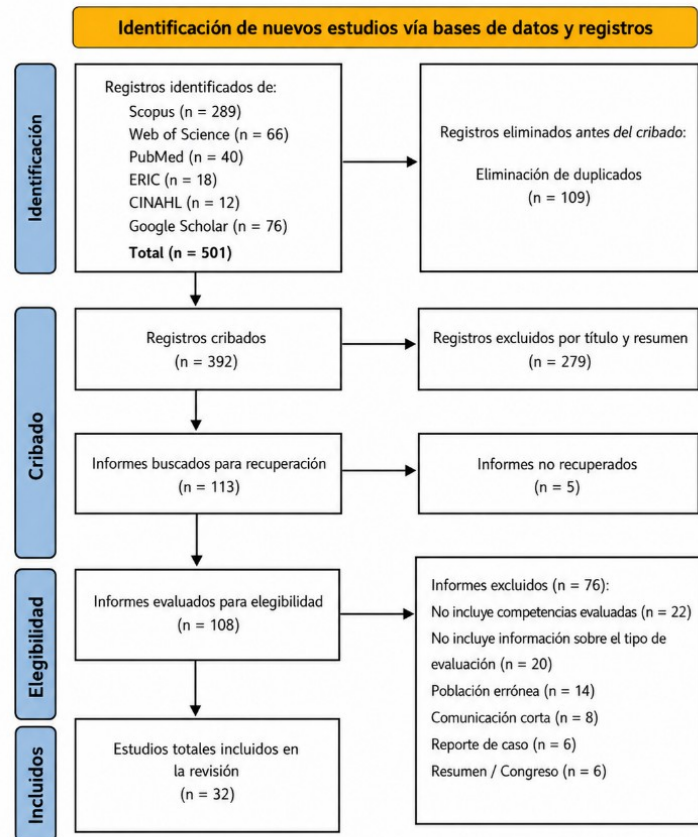


Figura 1. Diagrama de flujo del proceso de selección de estudios según PRISMA-ScR.

### 3. Resultados

#### 3.1 Características generales de los estudios incluidos

Los 32 estudios incluidos abordaron la aplicación de métodos bioestadísticos en ciencias de la salud desde perspectivas teóricas, metodológicas y críticas. De acuerdo con el contenido principal de cada documento, los estudios fueron organizados en tres categorías temáticas: fundamentos teóricos de la inferencia estadística, problemas y limitaciones en la aplicación de métodos bioestadísticos, y enfoques metodológicos contemporáneos o alternativas analíticas. La primera categoría agrupó 11 estudios relacionados con los fundamentos teóricos de la inferencia estadística. La segunda categoría incluyó 13 estudios centrados en errores, limitaciones o problemas frecuentes en la aplicación de métodos bioestadísticos. La tercera categoría integró 8 estudios orientados a propuestas metodológicas contemporáneas, entre ellas inferencia bayesiana, factor de Bayes, ciencia abierta, reproducibilidad y mejora del reporte estadístico. Todos los artículos seleccionados se muestran en la tabla 1.

#### 3.2 Categoría 1. Fundamentos teóricos de la inferencia estadística

En esta categoría se incluyeron 11 documentos que aportaron bases conceptuales, históricas y epistemológicas para comprender la inferencia estadística aplicada a la investigación en salud. Los estudios seleccionados abordaron el desarrollo de la inferencia frecuentista, las pruebas de significación, el contraste de hipótesis, el control de errores tipo I y II, la probabilidad, la lógica de la investigación científica, la epidemiología moderna y la interpretación basada en estimaciones. Los documentos clásicos de Fisher y Neyman-Pearson se relacionaron con la construcción histórica del enfoque frecuentista y de las pruebas de hipótesis (2-3). Otros textos, como los de Jeffreys, Popper y Hacking, aportaron una perspectiva probabilística y epistemológica sobre la construcción del conocimiento científico (16, 23-24). Asimismo, los textos de Altman, Rothman, Greenland y Lash, y Szklo y Nieto contribuyeron desde la estadística médica aplicada y la epidemiología (1, 25-26). Finalmente, los aportes de Cumming, Cohen y Gigerenzer permitieron identificar críticas al uso ritualizado de la significación estadística y propuestas de interpretación basadas en estimaciones (8, 27-28).

#### 3.3 Categoría 2. Problemas y limitaciones en la aplicación de métodos bioestadísticos

En esta categoría se incluyeron 13 estudios que abordaron errores y limitaciones en el uso de métodos estadísticos en investigación biomédica, clínica, epidemiológica y de salud pública. Los problemas identificados se relacionaron principalmente con la interpretación incorrecta del valor  $p$ , la dependencia del umbral  $p < 0,05$ , la baja potencia estadística, el  $p$ -hacking, los falsos positivos, las comparaciones múltiples, la diferencia entre significación estadística y significación clínica, y el uso inadecuado de procedimientos de medición o análisis. Los estudios revisados describieron problemas vinculados con el uso de pruebas de significación como criterio aislado de decisión (29), la interpretación incorrecta del valor  $p$  y de otros indicadores inferenciales (4-6), el aumento del riesgo de hallazgos falsos positivos y baja reproducibilidad (11, 13, 30-31), la multiplicidad analítica y los múltiples desenlaces (32-33), los errores en estudios de comparación de métodos de medición (34) y la necesidad de diferenciar entre resultados estadísticamente significativos y resultados clínicamente relevantes (10).

#### 3.3 Categoría 3. Enfoques metodológicos contemporáneos y alternativas analíticas

En esta categoría se incluyeron 8 estudios que abordaron alternativas analíticas y propuestas metodológicas orientadas a mejorar la interpretación, transparencia y reproducibilidad de los resultados en ciencias de la salud. Los enfoques identificados incluyeron el factor de Bayes, la inferencia bayesiana, el análisis bayesiano de datos, la superación del uso dicotómico del valor  $p$ , el reporte del tamaño del efecto, la cultura de investigación abierta, la ciencia reproducible y los

lineamientos PRISMA-ScR. Estos estudios aportaron estrategias para complementar o superar la interpretación basada exclusivamente en el valor  $p$  (7). También destacaron la importancia de reportar la magnitud de los efectos (9), promover prácticas de ciencia abierta, transparentar los procedimientos analíticos y fortalecer la reproducibilidad (14-15), así como utilizar guías de reporte metodológico cuando se desarrollan revisiones de alcance o estudios de síntesis de evidencia (22).

#### 4. Discusión

Este diseño fue seleccionado porque el propósito principal no fue estimar un efecto combinado ni responder una pregunta clínica estrecha mediante metaanálisis, sino mapear la producción científica relacionada con la aplicación de métodos bioestadísticos en investigaciones en ciencias de la salud, identificar categorías temáticas, describir tendencias metodológicas y reconocer vacíos de conocimiento. Las revisiones de alcance son útiles cuando el área de estudio es amplia, heterogénea o conceptualmente diversa, como ocurre con la bioestadística aplicada, donde convergen fundamentos teóricos, prácticas analíticas, problemas de interpretación, estándares de reporte y enfoques emergentes (20-21). Los resultados de esta revisión muestran que la bioestadística ocupa una posición central en la investigación en ciencias de la salud, pero su aplicación práctica continúa marcada por tensiones importantes entre disponibilidad técnica, comprensión conceptual y calidad interpretativa (1, 6). Si bien existen documentos metodológicos y revisiones previas sobre inferencia estadística, valor  $p$ , reproducibilidad y enfoques bayesianos, esta revisión integra de forma conjunta fundamentos teóricos, problemas recurrentes de aplicación y alternativas metodológicas contemporáneas orientadas al fortalecimiento del rigor metodológico. En este sentido, se cumplió el objetivo de analizar la aplicación de métodos bioestadísticos en investigaciones en ciencias de la salud, al identificar patrones de uso, limitaciones persistentes y propuestas dirigidas a mejorar la interpretación, transparencia y reproducibilidad de la evidencia científica.

Los estudios incluidos en esta revisión evidenciaron que el valor  $p$  continúa ocupando un lugar central como criterio de interpretación estadística en investigaciones en ciencias de la salud (5). No obstante, diversos autores advierten que su uso aislado puede favorecer interpretaciones reduccionistas, especialmente cuando se asume como indicador directo de magnitud del efecto, probabilidad de verdad de la hipótesis o relevancia clínica del hallazgo (4, 6). En consecuencia, la literatura revisada recomienda complementar la significación estadística con intervalos de confianza, medidas de tamaño del efecto y una interpretación contextualizada según el diseño del estudio, la plausibilidad biológica y la pertinencia clínica de los resultados (8-10).

Los hallazgos coinciden con debates internacionales que han cuestionado el uso simplista de la significación estadística. El movimiento hacia una interpretación más rica de los resultados propone considerar estimaciones, intervalos de confianza, plausibilidad, calidad del diseño, sesgos y consecuencias prácticas (8). Esta perspectiva resulta especialmente pertinente en ciencias de la salud, donde una diferencia estadísticamente significativa puede carecer de valor clínico, mientras que un resultado no significativo puede ser relevante si el estudio tiene limitaciones de potencia o tamaño muestral (10).

La selección inadecuada de pruebas estadísticas observada en la literatura revisada revela otra brecha crítica. Muchos errores surgen de aplicar pruebas sin considerar el tipo de variable, la distribución de los datos, la independencia de las observaciones, el tamaño muestral o el diseño del estudio. Esta situación se agrava por el uso acrítico de software estadístico. Los programas informáticos facilitan el análisis, pero no garantizan pertinencia metodológica. El software ejecuta órdenes; el criterio científico lo pone el investigador (1).

El manejo de comparaciones múltiples requiere atención particular según la literatura consultada. En investigaciones con numerosos desenlaces o múltiples subgrupos, la probabilidad

de encontrar resultados aparentemente significativos aumenta por simple azar. Si no se aplican ajustes o no se define una estrategia analítica previa, los falsos positivos pueden multiplicarse. Este problema es especialmente relevante en estudios exploratorios, análisis secundarios y contextos de alta dimensionalidad. La solución no consiste en prohibir la exploración, sino en distinguir claramente entre análisis exploratorios y confirmatorios, aplicar controles adecuados y reportar los procedimientos con transparencia (32-33).

La confusión entre asociación y causalidad constituye una de las limitaciones más delicadas en ciencias de la salud. Los estudios observacionales son indispensables, pero sus resultados deben interpretarse con prudencia. Una asociación estadística puede deberse a causalidad, confusión, sesgo, azar o combinación de estos elementos. Para acercarse a una inferencia causal se requiere algo más que una prueba significativa: se necesitan diseños sólidos, temporalidad clara, medición adecuada de variables, control de confusores, plausibilidad biológica y, cuando sea pertinente, modelos causales explícitos. La bioestadística proporciona herramientas, pero la causalidad exige pensamiento epidemiológico (25, 35).

La revisión también evidencia que el reporte estadístico incompleto sigue siendo una barrera para la evaluación crítica. La falta de información sobre supuestos, manejo de datos perdidos, criterios de exclusión, selección de pruebas, software utilizado o análisis de sensibilidad limita la evaluación crítica de los estudios clínicos y observacionales (36-37). En revisiones de evidencia, PRISMA 2020 contribuye a transparentar los procesos de búsqueda, selección y síntesis documental (38). En ensayos clínicos aleatorizados, CONSORT 2010 actualiza los criterios para comunicar de forma completa los procedimientos y resultados (39). El estándar PRISMA original también constituye una guía de base para el reporte de revisiones sistemáticas (40). En una época donde la ciencia exige transparencia, el reporte insuficiente ya no es un detalle menor; es una falla estructural para la reproducibilidad (15). Un lector debe poder comprender qué se hizo, por qué se hizo y cómo se llegó a las conclusiones.

Los enfoques contemporáneos identificados ofrecen una ruta de mejora. La inferencia bayesiana, por ejemplo, permite integrar información previa con datos nuevos y expresar conclusiones en términos probabilísticos (16, 18-19). Esto puede resultar especialmente útil en contextos donde existe evidencia acumulada o cuando se requiere actualizar conocimiento de manera progresiva. El factor de Bayes permite comparar evidencia relativa entre hipótesis y puede complementar la interpretación basada en valores  $p$  (17). No obstante, su uso no debe idealizarse: los métodos bayesianos también requieren supuestos, decisiones y transparencia, especialmente en la elección de distribuciones previas (18-19).

La adopción de análisis bayesianos en ciencias de la salud sigue siendo limitada, según la literatura metodológica revisada sobre inferencia estadística, análisis bayesiano y crítica al uso exclusivo del valor  $p$ . Los trabajos de Jeffreys, Gelman et al., Goodman y Spiegelhalter et al. señalan que los enfoques bayesianos permiten integrar información previa, actualizar la evidencia disponible y expresar los resultados en términos probabilísticos (16-19); sin embargo, también advierten que su aplicación requiere mayor formación conceptual, definición transparente de distribuciones previas y comprensión del razonamiento probabilístico (18-19). Asimismo, las recomendaciones de Wasserstein, Schirm y Lazar, junto con los planteamientos de Cumming, evidencian la necesidad de superar la dependencia exclusiva de la estadística frecuentista tradicional y avanzar hacia interpretaciones más integradoras de la evidencia (7-8). En este sentido, el predominio de la enseñanza centrada en pruebas de hipótesis y valores  $p$  puede limitar la incorporación de enfoques bayesianos y otros modelos estadísticos contemporáneos en la investigación en salud.

La reproducibilidad científica y la ciencia abierta aparecen como dimensiones clave. El preregistro de estudios, la publicación de protocolos, la disponibilidad de bases de datos anonimizadas y el acceso a códigos analíticos contribuyen a reducir sesgos y fortalecer la reproducibilidad (15). Asimismo, las prácticas de ciencia abierta promueven la transparencia metodológica y el acceso verificable a los productos de investigación (14). Estas prácticas no son adornos editoriales; son mecanismos de control de calidad científica. En investigaciones en salud, donde los resultados pueden orientar decisiones clínicas y políticas públicas, la transparencia metodológica es una responsabilidad ética.

El fortalecimiento de la formación bioestadística debe ser una prioridad. No se trata de convertir a todos los profesionales de la salud en estadísticos puros, sino de formar investigadores capaces de dialogar críticamente con los datos (1). Esto implica comprender diseños de estudio, reconocer tipos de variables, seleccionar análisis pertinentes, interpretar estimaciones, identificar sesgos, leer críticamente resultados y comunicar hallazgos con prudencia (8). En la formación médica y de posgrado, la bioestadística debería enseñarse desde problemas reales, no como una colección de fórmulas desconectadas de la práctica investigativa.

Una implicación importante de esta revisión es la necesidad de integrar bioestadística, epidemiología y metodología de la investigación (1, 25). Estas áreas no deben enseñarse ni aplicarse como compartimentos aislados. La pregunta de investigación determina el diseño; el diseño condiciona el análisis; el análisis delimita la interpretación; y la interpretación debe volver al problema clínico o sanitario (25, 37). Cuando esa cadena se rompe, aparecen conclusiones débiles, aunque el análisis parezca sofisticado.

En el contexto latinoamericano y ecuatoriano, estas reflexiones tienen especial relevancia. La producción científica en salud ha crecido, pero aún enfrenta desafíos relacionados con formación metodológica, acceso a asesoría estadística, cultura de publicación, financiamiento y disponibilidad de bases de datos robustas. Fortalecer la bioestadística aplicada puede mejorar la calidad de tesis, artículos, proyectos institucionales y decisiones sanitarias. La mejora no depende únicamente de usar métodos más complejos, sino de aplicar correctamente los métodos adecuados para cada pregunta.

Los resultados de esta revisión permiten proponer varias recomendaciones prácticas. Primero, los estudios deberían justificar explícitamente la selección de pruebas estadísticas. Segundo, todo análisis inferencial debería acompañarse de medidas del efecto y precisión. Tercero, los investigadores deben evitar interpretar el valor  $p$  como criterio absoluto de verdad. Cuarto, los estudios observacionales deben controlar confusión y evitar lenguaje causal cuando el diseño no lo permite. Quinto, los artículos deben reportar con claridad el manejo de datos perdidos, supuestos de aplicabilidad del análisis realizado, control de la confusión, medición de la exposición, posibles sesgos, software, criterios analíticos y análisis de sensibilidad. Sexto, las instituciones académicas deberían fortalecer la asesoría bioestadística desde la fase de diseño, no solo al final del estudio.

Esta revisión presenta limitaciones. En primer lugar, al tratarse de una revisión de alcance con enfoque sistemático, no se realizó evaluación formal del riesgo de sesgo como en revisiones sistemáticas de intervención. En segundo lugar, la heterogeneidad de los documentos incluidos impidió realizar síntesis cuantitativa. En tercer lugar, aunque se exploraron varias bases de datos, no se consultó EMBASE, por lo que es posible que algunos estudios relevantes no hayan sido identificados. No obstante, el enfoque empleado permitió mapear un campo amplio y ofrecer una visión integradora sobre la aplicación de métodos bioestadísticos en ciencias de la salud.

Como fortaleza, el estudio recopila e integra fundamentos teóricos, problemas prácticos y alternativas contemporáneas, lo que permite comprender la bioestadística no solo como técnica de

análisis, sino como parte del razonamiento científico. Esta mirada resulta útil para investigadores, docentes, estudiantes de posgrado, revisores de revistas y comités académicos interesados en mejorar la calidad metodológica de la producción científica.

En síntesis, la bioestadística seguirá siendo indispensable para la investigación en salud, pero su valor depende de cómo se utilice. No es suficiente reportar un valor  $p$  ni aplicar una prueba por costumbre. La investigación biomédica necesita una cultura estadística más sólida, capaz de interpretar incertidumbre, reconocer limitaciones, integrar evidencia y comunicar resultados con responsabilidad. Estas reflexiones también deben trasladarse al campo de la formación en ciencias médicas. La enseñanza de la bioestadística no debería limitarse al aprendizaje mecánico de fórmulas o pruebas estadísticas aisladas, sino orientarse al desarrollo del razonamiento crítico, la interpretación clínica de los datos y la toma de decisiones basada en evidencia. En la formación de estudiantes y profesionales de la salud resulta esencial fortalecer competencias para analizar resultados científicos, identificar errores metodológicos, comprender la incertidumbre inherente a la investigación biomédica y utilizar la evidencia de manera ética y responsable en la práctica clínica, la salud pública y la investigación.

## 5. Conclusiones

- La presente revisión permitió identificar tres grandes núcleos de evidencia: los fundamentos teóricos de la inferencia estadística, los problemas recurrentes en la aplicación de métodos bioestadísticos y las alternativas metodológicas contemporáneas orientadas a mejorar la interpretación y reproducibilidad de los resultados.
- Los hallazgos evidencian que el enfoque frecuentista continúa siendo predominante, especialmente mediante el uso de pruebas de hipótesis y valores  $p$ . Aunque estas herramientas siguen siendo útiles, su aplicación aislada y dicotómica limita la calidad interpretativa de la investigación. El valor  $p$  no debe ser utilizado como medida única de evidencia ni como sustituto del razonamiento clínico, epidemiológico y metodológico.
- Entre las principales deficiencias identificadas se encuentran la interpretación incorrecta de la significación estadística, la omisión de intervalos de confianza y tamaños del efecto, el uso inadecuado de pruebas, la falta de control de variables de confusión, las comparaciones múltiples sin ajuste, el reporte incompleto de procedimientos analíticos y la confusión entre asociación y causalidad. Estas limitaciones pueden comprometer la validez de los resultados y reducir la utilidad de la evidencia científica.
- Los enfoques contemporáneos, como la inferencia bayesiana, el uso del factor de Bayes, los análisis de sensibilidad, la interpretación basada en estimaciones, la ciencia abierta y las guías de reporte, ofrecen oportunidades para fortalecer la investigación en salud. Además, la expansión de nuevos campos como la inteligencia artificial y el aprendizaje automático en ciencias médicas exige bases conceptuales y metodológicas sólidas en bioestadística, debido a que el desarrollo, validación e interpretación de modelos predictivos requieren comprender la incertidumbre, el ajuste de modelos, el sesgo, la reproducibilidad y la calidad de los datos.
- Mejorar la aplicación de la bioestadística en ciencias de la salud exige actuar desde la formación, el diseño de los estudios, el análisis, el reporte y la revisión editorial. La calidad estadística no debe corregirse al final del manuscrito; debe construirse desde el inicio del proyecto. Solo así la evidencia científica podrá ser más válida, reproducible y útil para la toma de decisiones clínicas y sanitarias.

## 6. Referencias

1. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991. <https://www.routledge.com/Practical-Statistics-for-Medical-Research/Altman/p/book/9780412276309>.

2. Fisher RA. Statistical methods for research workers. Edinburgh: Oliver and Boyd; 1925. <https://archive.org/details/statisticalmetho00fish>.
3. Neyman J, Pearson ES. On the problem of the most efficient tests of statistical hypotheses. *Philos Trans R Soc Lond A*. 1933, 231, 289-337. <https://doi.org/10.1098/rsta.1933.0009>.
4. Goodman SN. Toward evidence-based medical statistics. 1: The p value fallacy. *Ann Intern Med*. 1999, 130(12), 995-1004. <https://pubmed.ncbi.nlm.nih.gov/10383371/>.
5. Wasserstein RL, Lazar NA. The ASA statement on p-values: context, process, and purpose. *Am Stat*. 2016, 70(2), 129-133. <https://doi.org/10.1080/00031305.2016.1154108>.
6. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, p values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol*. 2016, 31(4), 337-350. <https://doi.org/10.1007/s10654-016-0149-3>.
7. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond  $p < 0.05$ . *Am Stat*. 2019, 73(sup1), 1-19. <https://doi.org/10.1080/00031305.2019.1583913>.
8. Cumming G. The new statistics: why and how. *Psychol Sci*. 2014, 25(1), 7-29. <https://doi.org/10.1177/0956797613504966>.
9. Sullivan GM, Feinn R. Using effect size - or why the p value is not enough. *J Grad Med Educ*. 2012, 4(3), 279-282. <https://doi.org/10.4300/JGME-D-12-00156.1>.
10. Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: clinical versus statistical significance. *Perspect Clin Res*. 2015, 6(3), 169-170. <https://doi.org/10.4103/2229-3485.159943>.
11. Ioannidis JPA. Why most published research findings are false. *PLoS Med*. 2005, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>.
12. Head ML, Holman L, Lanfear R, Kahn AT, Jennions MD. The extent and consequences of p-hacking in science. *PLoS Biol*. 2015, 13(3), e1002106. <https://doi.org/10.1371/journal.pbio.1002106>.
13. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013, 14(5), 365-376. <https://doi.org/10.1038/nrn3475>.
14. Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, et al. Promoting an open research culture. *Science*. 2015, 348(6242), 1422-1425. <https://doi.org/10.1126/science.aab2374>.
15. Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, Percie du Sert N, et al. A manifesto for reproducible science. *Nat Hum Behav*. 2017, 1, 0021. <https://doi.org/10.1038/s41562-016-0021>.
16. Jeffreys H. Theory of probability. 3rd ed. Oxford: Oxford University Press; 1961. <https://archive.org/details/theoryofprobabil00jeff>.
17. Goodman SN. Toward evidence-based medical statistics. 2: The Bayes factor. *Ann Intern Med*. 1999, 130(12), 1005-1013. <https://pubmed.ncbi.nlm.nih.gov/10383350/>.
18. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis. 3rd ed. Boca Raton: CRC Press; 2013. <https://www.routledge.com/Bayesian-Data-Analysis/Gelman-Carlin-Stern-Dunson-Vehtari-Rubin/p/book/9781439840955>.
19. Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and health-care evaluation. Chichester: Wiley; 2004. <https://doi.org/10.1002/0470092602>.
20. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005, 8(1), 19-32. <https://doi.org/10.1080/1364557032000119616>.
21. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci*. 2010, 5, 69. <https://doi.org/10.1186/1748-5908-5-69>.
22. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018, 169(7), 467-473. <https://doi.org/10.7326/M18-0850>.
23. Popper KR. The logic of scientific discovery. London: Routledge; 1959. <https://www.routledge.com/The-Logic-of-Scientific-Discovery/Popper/p/book/9780415278447>.

24. Hacking I. The emergence of probability. Cambridge: Cambridge University Press; 1975. <https://doi.org/10.1017/CBO9780511817557>.
25. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008. <https://shop.lww.com/Modern-Epidemiology/p/9781451190052>.
26. Szklo M, Nieto FJ. Epidemiology: beyond the basics. 4th ed. Burlington: Jones & Bartlett Learning; 2019. <https://www.jblearning.com/catalog/productdetails/9781284116595>.
27. Cohen J. The earth is round ( $p < .05$ ). *Am Psychol*. 1994, 49(12), 997-1003. <https://doi.org/10.1037/0003-066X.49.12.997>.
28. Gigerenzer G. Mindless statistics. *J Socio Econ*. 2004, 33(5), 587-606. <https://doi.org/10.1016/j.socec.2004.09.033>.
29. Sterne JAC, Smith GD. Sifting the evidence - what's wrong with significance tests? *BMJ*. 2001, 322(7280), 226-231. <https://doi.org/10.1136/bmj.322.7280.226>.
30. Colquhoun D. An investigation of the false discovery rate and the misinterpretation of p-values. *R Soc Open Sci*. 2014, 1(3), 140216. <https://doi.org/10.1098/rsos.140216>.
31. Nuzzo R. Scientific method: statistical errors. *Nature*. 2014, 506(7487), 150-152. <https://doi.org/10.1038/506150a>.
32. Streiner DL. Best but oft-forgotten practices: the multiple problems of multiplicity - whether and how to correct for many statistical tests. *Am J Clin Nutr*. 2015, 102(4), 721-728. <https://doi.org/10.3945/ajcn.115.113548>.
33. Feise RJ. Do multiple outcome measures require p-value adjustment? *BMC Med Res Methodol*. 2002, 2, 8. <https://doi.org/10.1186/1471-2288-2-8>.
34. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986, 327(8476), 307-310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8).
35. Vandembroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the reporting of observational studies in epidemiology: explanation and elaboration. *PLoS Med*. 2007, 4(10), e297. <https://doi.org/10.1371/journal.pmed.0040297>.
36. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Lancet*. 2001, 357(9263), 1191-1194. [https://doi.org/10.1016/S0140-6736\(00\)04337-3](https://doi.org/10.1016/S0140-6736(00)04337-3).
37. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandembroucke JP. The STROBE statement: guidelines for reporting observational studies. *Lancet*. 2007, 370(9596), 1453-1457. [https://doi.org/10.1016/S0140-6736\(07\)61602-X](https://doi.org/10.1016/S0140-6736(07)61602-X).
38. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021, 372, n71. <https://doi.org/10.1136/bmj.n71>.
39. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010, 340, c332. <https://doi.org/10.1136/bmj.c332>.
40. Moher D, Liberati A, Tetzlaff A, Altman DG, PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. *PLoS Med*. 2009, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>.



© 2026 Universidad de Murcia. Enviado para su publicación en acceso abierto bajo los términos y condiciones de la licencia Creative Commons Reconocimiento-NoComercial-Sin Obra Derivada 4.0 España (CC BY-NC-ND) (<http://creativecommons.org/licenses/by/4.0/>).

**Tabla 1.** Matriz de extracción y análisis de los estudios incluidos.

<b>Autor, año (referencia)</b>	<b>Tipo de documento</b>	<b>Categoría temática</b>	<b>Objetivo o enfoque principal</b>	<b>Método / problema bioestadístico analizado</b>	<b>Principales hallazgos</b>	<b>Aporte para la revisión</b>
<b>Fisher, 1925 (2)</b>	Libro metodológico clásico	Fundamentos teóricos de la inferencia estadística	Establecer bases para las pruebas de significación y la inferencia frecuentista.	Valor p, pruebas de significación e inferencia frecuentista.	Consolidó el uso de pruebas estadísticas para valorar evidencia frente a la hipótesis nula.	Establece bases históricas del uso del valor p como indicador de evidencia contra la hipótesis nula.
<b>Neyman y Pearson, 1933 (3)</b>	Artículo teórico	Fundamentos teóricos de la inferencia estadística	Desarrollar un marco formal para el contraste de hipótesis.	Contraste de hipótesis, errores tipo I y II.	Propuso un enfoque decisional basado en reglas, potencia y control de errores a largo plazo.	Fundamenta el enfoque decisional de la inferencia estadística y el control de errores.
<b>Jeffreys, 1961 (16)</b>	Libro teórico	Fundamentos teóricos de la inferencia estadística	Desarrollar fundamentos probabilísticos para la inferencia bayesiana.	Probabilidad, comparación de hipótesis e inferencia bayesiana.	Planteó la probabilidad como herramienta para actualizar y comparar hipótesis.	Introduce fundamentos para comparar hipótesis desde una perspectiva bayesiana.
<b>Popper, 1959 (23)</b>	Libro epistemológico	Fundamentos teóricos de la inferencia estadística	Explicar la lógica de la investigación científica y la falsación.	Falsación y contrastación empírica de hipótesis.	Destacó que el conocimiento científico avanza mediante contrastación crítica y posibilidad de refutación.	Aporta marco epistemológico para comprender la contrastación empírica de hipótesis.

<b>Hacking, 1975 (24)</b>	Libro histórico-epistemológico	Fundamentos teóricos de la inferencia estadística	Analizar la emergencia histórica del pensamiento probabilístico.	Probabilidad como base del razonamiento científico moderno.	Explicó la evolución conceptual de la probabilidad y su papel en la ciencia moderna.	Explica la evolución de la probabilidad como base del razonamiento científico.
<b>Altman, 1991 (1)</b>	Libro metodológico aplicado	Fundamentos teóricos de la inferencia estadística	Sistematizar procedimientos estadísticos aplicados a investigación médica.	Estadística médica aplicada.	Organizó métodos estadísticos esenciales para el diseño, análisis e interpretación de estudios clínicos.	Sistematiza procedimientos estadísticos esenciales para investigación clínica y biomédica.
<b>Rothman, Greenland y Lash, 2008 (25)</b>	Libro metodológico	Fundamentos teóricos de la inferencia estadística	Integrar epidemiología moderna, sesgos, confusión e inferencia causal.	Inferencia causal, sesgo, confusión y análisis epidemiológico.	Mostró la necesidad de articular diseño, análisis estadístico y razonamiento causal en estudios observacionales.	Integra razonamiento estadístico, sesgos, confusión y causalidad.
<b>Szklo y Nieto, 2019 (26)</b>	Libro metodológico	Fundamentos teóricos de la inferencia estadística	Desarrollar fundamentos de epidemiología aplicada y análisis de estudios poblacionales.	Diseño epidemiológico, análisis estadístico e interpretación.	Refuerza la relación entre diseño del estudio, análisis y validez de la interpretación epidemiológica.	Refuerza la articulación entre diseño epidemiológico, análisis estadístico e interpretación.
<b>Cohen, 1994 (27)</b>	Artículo conceptual	Fundamentos teóricos de la inferencia estadística	Cuestionar el ritualismo de la significación estadística.	Uso mecánico de $p < 0,05$ y tamaño del efecto.	Argumentó que la dependencia del valor $p$ limita la interpretación científica y debe complementarse con tamaño del efecto.	Cuestiona el uso mecánico de $p < 0,05$ y promueve interpretación basada en tamaño del efecto.
<b>Gigerenzer, 2004 (28)</b>	Artículo teórico-crítico	Fundamentos teóricos de la inferencia estadística	Criticar el uso acrítico de las pruebas estadísticas.	Aplicación automática de pruebas estadísticas.	Advirtió sobre el uso ritualizado de pruebas sin razonamiento estadístico suficiente.	Advierte sobre la aplicación automática de pruebas sin razonamiento

						estadístico.
<b>Cumming, 2014 (8)</b>	Artículo metodológico	Fundamentos teóricos de la inferencia estadística	Proponer una interpretación estadística basada en estimaciones.	Nueva estadística, intervalos de confianza y tamaño del efecto.	Propuso priorizar estimaciones, intervalos de confianza y visualización frente a decisiones dicotómicas.	Propone priorizar intervalos de confianza, tamaños del efecto y pensamiento estadístico visual.
<b>Sterne y Smith, 2001 (29)</b>	Artículo metodológico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Examinar limitaciones de las pruebas de significación en investigación biomédica.	Limitaciones de las pruebas de significación.	Mostró que la significación estadística usada de forma aislada puede inducir interpretaciones erróneas.	Evidencia riesgos interpretativos cuando la significación estadística se usa aisladamente.
<b>Ioannidis, 2005 (11)</b>	Artículo metacientífico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Analizar causas de falsos positivos y baja reproducibilidad en investigación biomédica.	Falsos positivos, sesgos y reproducibilidad.	Argumentó que múltiples factores metodológicos elevan la probabilidad de hallazgos falsos.	Sustenta la necesidad de rigor metodológico y transparencia analítica.
<b>Goodman, 1999a (4)</b>	Artículo metodológico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Explicar falacias comunes en la interpretación del valor p.	Falacia del valor p.	Mostró que el valor p suele interpretarse erróneamente como probabilidad de verdad de la hipótesis o del azar.	Explica consecuencias de interpretar incorrectamente el valor p en medicina basada en evidencia.
<b>Greenland et al., 2016 (6)</b>	Artículo metodológico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Sistematizar malinterpretaciones de pruebas estadísticas, p, intervalos y potencia.	P, intervalos de confianza, potencia e inferencia estadística.	Identificó errores frecuentes en la lectura de significación estadística, precisión y potencia.	Orienta una interpretación más rigurosa de pruebas, intervalos y potencia.
<b>Wasserstein y Lazar, 2016 (5)</b>	Declaración metodológica	Problemas y limitaciones en la aplicación de	Presentar recomendaciones de la ASA sobre el uso del	Uso e interpretación del valor p.	Advirtió que el valor p no debe usarse como medida única	Presenta posición institucional sobre el uso responsable del

		métodos bioestadísticos	valor p.		de evidencia ni como criterio absoluto de decisión.	valor p.
<b>Head et al., 2015 (12)</b>	Estudio metacientífico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Evaluar extensión y consecuencias del p-hacking en ciencia.	P-hacking y sesgo analítico.	Identificó prácticas analíticas flexibles asociadas a búsqueda de significación estadística.	Aporta evidencia sobre sesgos derivados de decisiones analíticas flexibles.
<b>Nuzzo, 2014 (31)</b>	Artículo de análisis científico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Describir errores estadísticos frecuentes en publicaciones científicas.	Errores estadísticos en investigación.	Señaló fallas frecuentes en interpretación y reporte de resultados estadísticos.	Refuerza la necesidad de formación estadística y reporte transparente.
<b>Colquhoun, 2014 (30)</b>	Artículo metodológico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Analizar la tasa de falsos descubrimientos y la interpretación del valor p.	Falsos descubrimientos y mala interpretación del valor p.	Mostró que la interpretación inadecuada del valor p puede inflar la evidencia aparente.	Sustenta la cautela ante resultados significativos aislados.
<b>Button et al., 2013 (13)</b>	Artículo metodológico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Analizar el impacto de bajo tamaño muestral sobre la confiabilidad de hallazgos.	Baja potencia estadística.	Evidenció que muestras pequeñas reducen la confiabilidad y reproducibilidad de los resultados.	Respalda la importancia de potencia estadística y tamaño muestral adecuado.
<b>Ranganathan, Pramesh y Buyse, 2015 (10)</b>	Artículo educativo	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Diferenciar significación clínica y significación estadística.	Significación clínica frente a significación estadística.	Explicó que un hallazgo significativo puede carecer de relevancia clínica y viceversa.	Apoya la integración de relevancia clínica en la interpretación estadística.

<b>Streiner, 2015 (32)</b>	Artículo metodológico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Explicar problemas derivados de múltiples comparaciones.	Multiplicidad y ajuste por comparaciones múltiples.	Mostró cuándo y cómo aplicar correcciones para reducir falsos positivos.	Sustenta la necesidad de controlar multiplicidad en análisis inferenciales.
<b>Feise, 2002 (33)</b>	Artículo metodológico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Discutir la necesidad de ajustar valores p ante múltiples desenlaces.	Ajuste por múltiples desenlaces.	Analizó escenarios en los que múltiples medidas de resultado requieren ajuste estadístico.	Aporta criterios para el manejo de comparaciones múltiples.
<b>Bland y Altman, 1986 (34)</b>	Artículo metodológico	Problemas y limitaciones en la aplicación de métodos bioestadísticos	Proponer métodos adecuados para comparar mediciones clínicas.	Comparación entre métodos de medición.	Planteó procedimientos de concordancia más adecuados que la correlación simple.	Contribuye a evitar análisis inadecuados en estudios de medición.
<b>Goodman, 1999b (17)</b>	Artículo metodológico	Enfoques metodológicos contemporáneos y alternativas analíticas	Proponer el factor de Bayes como alternativa para interpretar evidencia.	Factor de Bayes.	Mostró cómo comparar evidencia relativa entre hipótesis competidoras.	Propone el factor de Bayes como complemento o alternativa al valor p.
<b>Gelman et al., 2013 (18)</b>	Libro metodológico	Enfoques metodológicos contemporáneos y alternativas analíticas	Desarrollar fundamentos y aplicaciones del análisis bayesiano de datos.	Modelos bayesianos y análisis de datos complejos.	Expuso métodos bayesianos para modelar incertidumbre e integrar información previa.	Aporta bases para enfoques bayesianos aplicados a salud.
<b>Spiegelhalter, Abrams y Myles, 2004 (19)</b>	Libro metodológico	Enfoques metodológicos contemporáneos y alternativas analíticas	Aplicar enfoques bayesianos a ensayos clínicos y evaluación sanitaria.	Inferencia bayesiana en investigación clínica.	Mostró la utilidad de modelos bayesianos en evaluación de intervenciones y servicios sanitarios.	Evidencia aplicabilidad clínica y sanitaria de la inferencia bayesiana.

<b>Wasserstein, Schirm y Lazar, 2019 (7)</b>	Artículo de consenso	Enfoques metodológicos contemporáneos y alternativas analíticas	Promover una interpretación científica más allá del umbral $p < 0,05$ .	Superación del uso dicotómico del valor p.	Recomendó abandonar decisiones rígidas basadas en significación estadística.	Apoya interpretaciones integrales de la evidencia científica.
<b>Sullivan y Feinn, 2012 (9)</b>	Artículo educativo	Enfoques metodológicos contemporáneos y alternativas analíticas	Destacar la importancia del tamaño del efecto en educación médica e investigación.	Tamaño del efecto.	Explicó por qué el valor p no es suficiente para interpretar importancia práctica.	Refuerza la necesidad de reportar magnitud del efecto.
<b>Nosek et al., 2015 (14)</b>	Artículo de política científica	Enfoques metodológicos contemporáneos y alternativas analíticas	Promover prácticas de ciencia abierta para mejorar reproducibilidad.	Cultura de investigación abierta.	Propuso apertura de datos, materiales y protocolos como estrategias de transparencia.	Apoya prácticas abiertas para mejorar calidad y reproducibilidad.
<b>Munafò et al., 2017 (15)</b>	Manifiesto metodológico	Enfoques metodológicos contemporáneos y alternativas analíticas	Plantear acciones para fortalecer la ciencia reproducible.	Ciencia reproducible, preregistro y reducción de sesgos.	Propuso reformas metodológicas para aumentar robustez, transparencia y reproducibilidad.	Aporta recomendaciones para mejorar la robustez de la investigación.
<b>Tricco et al., 2018 (22)</b>	Guía metodológica	Enfoques metodológicos contemporáneos y alternativas analíticas	Establecer criterios de reporte para revisiones de alcance.	PRISMA-ScR.	Definió elementos para reportar identificación, selección, extracción y síntesis en scoping reviews.	Fortalece la transparencia metodológica del presente artículo.