

Revista del Hospital Italiano de Buenos Aires

Volumen 46 | Número 3 | Año 2026

El precio de revisar gratis: el trabajo invisible que sostiene la ciencia [Alfredo Eymann, et al.](#)

Otoesclerosis en el adulto mayor: la indicación más allá de la edad cronológica [Graciela L. Tiberti](#)

Evaluación clínica del bocio y riesgo de malignidad [Marcelo F. Figari](#)

Factores clínicos y demográficos asociados a carcinoma diferenciado de tiroides en pacientes con bocio de larga data en Bolivia [Álvaro B. Daza Mier](#)

Cirugía estapediovestibular en pacientes con otoesclerosis de 65 años o más: resultados auditivos posoperatorios. Experiencia de 20 años [Micaela Redivo, et al.](#)

Síndrome de seno silente: resultados quirúrgicos y propuesta de algoritmo para la reconstrucción orbitaria [Antonella Giorgi, et al.](#)

Sinusitis odontógena: ¿por dónde empezar? La importancia del abordaje multidisciplinario [Candelaria Serrano, et al.](#)

El papel del *benchmarking* en la gestión sanitaria: más allá de la medición, una estrategia para aprender y mejorar [María F. Grande Ratti, et al.](#)

Nuevas terapias para enfermedad hepática por depósito de grasa/metabólica (MASLD/MASH): resultados clínicos y retos en implementación [Javier J. Flores, et al.](#)

Aneurisma aórtico micótico por enterobacterias multirresistentes en paciente anciano: desafío diagnóstico [Anabela Games, et al.](#)

Lesión hepática grave asociada al consumo de productos fitoterapéuticos no regulados: Reporte de caso y revisión de la literatura [María A. De la Rosa, et al.](#)

En búsqueda de una $p < 0,05$: el problema de las comparaciones múltiples [María E. Knorre, et al.](#)

Desenlaces subrogados en investigación clínica [Mariana A. Burgos, et al.](#)

Decisiones médicas compartidas: una perspectiva desde la ciencia de la complejidad [Alicia Rodríguez López](#)

¿Puede un antidepresivo cambiar una vida? Evidencia, experiencia y relevancia clínica [Nahuel E. Baca](#)

ISSN 2314-3312
<https://ojs.hospitalitaliano.org.ar>



25%
DE DESCUENTO

DURANTE
MESES **3**
EN TU CUOTA

Afiliate al Italiano.
Hay equipo.

Hoy es muy importante tener tranquilidad, respaldo
y la mejor calidad médica.



PLAN
Vita

CONOCÉ MÁS



PLAN
Magna

CONOCÉ MÁS

PLAN DE SALUD

 **HOSPITAL ITALIANO**
de Buenos Aires

TODO EL ITALIANO DETRÁS DE TU PLAN

**TRATAMIENTOS
DE ALTA PRECISIÓN
ADAPTADOS A
CADA PACIENTE**



3D Radioterapia Tridimensional Conformada

IMRT-VMAT Radioterapia de Intensidad Modulada

IGRT Radioterapia Guiada por Imágenes

SBRT Radioterapia Estereotáxica Extracraneal

SRS Radiocirugía Estereotáxica Craneal

TBI Irradiación Corporal Total 3D Conformada

BQT 3D HDR Braquiterapia Tridimensional de Alta Tasa de Dosis

TSEI/RTSEI Baño de Electrones

IORT Radioterapia Intraoperatoria
(Primer Equipo INTRABEAM móvil en Argentina)



innovación

Brindamos atención médica de excelencia a través de un reconocido staff de profesionales y soporte tecnológico en permanente innovación.

Acompañamos a nuestros pacientes y sus familiares, centrándonos en su bienestar físico y emocional, desde una mirada integral de la Salud.



Central de Turnos: (011) 2150-0000 .  +54 9 11 4166-5537

info@mevaterapia.com.ar . www.mevaterapia.com.ar

Casa Central: Tte. Gral. J. D. Perón 3931/3937 (C1198AAW) . C.A.B.A. Argentina

Centros de Atención: Barrio Norte . Recoleta . Almagro . San Justo .

Avellaneda . Lomas de Zamora . Quilmes

 **HOSPITAL ITALIANO**
de Buenos Aires

Servicio de Radioterapia

Posgrado

Maestrías, Doctorados y Especializaciones

Potenciá tu carrera en salud con
programas de especialización.

La medicina del futuro.



Universidad
Hospital Italiano

posgrado.hospitalitaliano.edu.ar



REVISTA DEL HOSPITAL ITALIANO DE BUENOS AIRES

Editor jefe

Dr. Daniel Matusevich (Hospital Italiano, Argentina)

Editor asociado

Dr. Alfredo Eymann (Universidad Hospital Italiano, Argentina)

Comité Editorial

Lic. Griselda Almada (Hospital Italiano, Argentina), Lic. Ignacio Barbagallo (Hospital Italiano, Argentina), Dr. Luis A. Boccalatte (Hospital Universitario Sagrat Cor, Barcelona, España), Dr. Alberto Bonetto (Profesional independiente, Argentina), Dr. Horacio Castro (Hospital Italiano, Argentina), Dra. María Soledad Dawson (Hospital Italiano, Argentina), Dr. Rodrigo A. Gasque (Clínica Universitaria Reina Fabiola, Argentina), Dr. Carlos Lazzarino (Hospital Municipal de Oncología "Marie Curie", Argentina), Dra. María Florencia Martínez (Hospital Italiano, Argentina), Dra. Vanina Pagotto (Hospital Italiano, Argentina), Dra. Rosa Pace (Hospital Italiano, Argentina), Dra. María Paula Russo (Hospital Italiano, Argentina), Dr. Julián Sánchez Viamonte (Universidad Nacional de Mar del Plata, Argentina), Dr. Fernando J. Vázquez (Hospital Italiano, Argentina), Dr. Carlos Wahren (Hospital Italiano, Argentina), Dr. Joaquín Álvarez Gregori (Universidad de Salamanca, España), Dr. Juan Dartiguelongue (Hospital de Niños Dr. Ricardo Gutiérrez, Argentina), Dra. Isabel del Cura (Universidad Rey Juan Carlos Campus Móstoles, España), Dr. Marcelo García Dieguez (Centro de Estudios en Educación de Profesionales de la Salud, Universidad Nacional del Sur, Argentina), Dr. Fernando Goldenberg (University of Chicago, EE. UU.), Prof. Dr. Jorge Manrique (Hospital Interzonal General de Agudos "Eva Perón", Argentina), Dr. Hernán Rowensztein (Hospital de Pediatría "Prof. Dr. Juan P. Garrahan", Argentina), Dra. Romina Rubin (Fundación Nuevo Hogar Ledorvador, Argentina), Dra. Dalia Szulik (Instituto de Estudios de América Latina y el Caribe, Argentina), Dra. Carla Venturi (Hospital Clinic, España).

Comité asesor

Dr. Diego Faingold (Universidad Hospital Italiano, Argentina), Lic. Virginia Garrote (Universidad Hospital Italiano, Argentina), Dr. Pablo Knoblovits (Hospital Italiano, Argentina), Dr. Luis J. Catoggio (Hospital Italiano, Argentina), Dr. Marcelo Figari (Hospital Italiano, Argentina), Dr. Mario Perman (Hospital Italiano, Argentina), Dr. Enrique Soriano (Hospital Italiano, Argentina)

Coordinadora editorial y página Web

Mariana Rapoport (Universidad Hospital Italiano, Argentina)

Corrección de textos

Prof. María Isabel Siracusa (Buenos Aires, Argentina)

Indización colaborativa para base de datos Lilacs

Bib. Aldana Ljeschak (Universidad Hospital Italiano, Argentina)

Gestión bibliográfica

Lic. Virginia Garrote (Universidad Hospital Italiano, Argentina)

Editor responsable

Publicación de la Universidad Hospital Italiano

Tte. Juan D. Perón 4190 | 1° piso | Escalera J | C1199ABB | C.A.B.A. | Argentina

URL: <http://ojs.hospitalitaliano.org.ar/index.php/revistahi/> | E-mail: revista@hospitalitaliano.org.ar

Tel. 54-11-4959-0200 int. 9293 | Fax. 54-11-4959-0383

Instagram: /revistadelhiba

Facebook: /RevistadelHIBA

LinkedIn: /revista-del-hiba

Sociedad Italiana de Beneficiencia

en Buenos Aires

Consejo Directivo Ejercicio 2025/2026

Presidente: Arq. Aldo Brunetta

Vicepresidente 1°: Sr. Eduardo Tarditi

Vicepresidente 2°: Sr. Dr. Marcelo Marchetti

Consejeros: Lic. Daniel Artana, Roberto Baccanelli,

Roberto Bonatti, Guillermo Jaim Etcheverry,

Atilio Migues, Renato Montefiore, Elio Squillari

Hospital Italiano

Director Médico: Dr. Esteban Langlois

Director de Gestión: Dr. Matías Antonini

Vice-Director Médico de Gestión Operativa: Dr.

Guillermo Arbues

Vice-Director Médico de Planeamiento Estratégico:

Dr. Leonardo Garfi

Vice-Director Médico de Plan de Salud:

Dra. Lucila Hornstein

Vice-Director Médico Hospital Italiano de

San Justo Agustín Rocca: Dr. Santiago Bongiovanni

Directores Honorarios: Dres. Jorge Sívori,

Héctor Marchitelli, Atilio Migues

Consejo de Administración Fundación Instituto

Universitario

Presidente: Arq. Aldo Brunetta

Vicepresidente: Ing. Franco Livini

Vocales: Giorgio Allia di Montereale, Daniel Artana,

Renato Montefiore, Diego Faingold

Universidad Hospital Italiano

Consejo Superior

Rector: Diego Faingold

Vicerrector: Mariano Falconi

Secretaría Académica: María Laura Eder

Secretaría Administrativa y de Gestión Financiera:

Alberto Llebara

Directores de Deptos./Carreras: Nicolas Cacchiarelli,

Roberta Ladenheim, Gisela Schwartzman, María

Rezzónico, Matteo Baccanelli, Virginia Garrote, Karín

Kopitowski, Eduardo Durante, Ana Mosca, Susana

Llesuy, Patricia Chavarría, Marcelo Risk

Perfil de la Revista

La Revista del Hospital Italiano de Buenos Aires (E-ISSN 2314-3312) es una publicación científica de publicación continua, en idioma castellano e inglés, de la Universidad Hospital Italiano (UHI) que se edita desde 1981. Su contenido es revisado por pares bajo la modalidad doble ciego, mientras que su difusión se realiza en línea. Es de acceso abierto bajo la licencia CC BY-NC-SA y el proceso de publicación para los autores es no arancelado. Publica trabajos originales relevantes para las ciencias de la salud y las áreas relacionadas. Se aceptan contribuciones de autores tanto pertenecientes a la Institución como externos a ella, en idioma castellano e inglés.

Scope

The Revista del Hospital Italiano de Buenos Aires (E-ISSN 2314-3312) is a scientific journal published continuously in Spanish and English by the Universidad Hospital Italiano (UHI) and has been published since 1981. Its double-blind peer-reviewed content is available online. It is open access under the license CC BY-NC-SA, and the publishing process for authors is fee-free. It offers original papers relevant to the health sciences and related areas. It accepts contributions by authors from within and outside the Institution, in Spanish and English.

DOI: <https://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3>

Instrucciones para autores, ver: <http://ojs.hospitalitaliano.org.ar/index.php/revistahi/instrucciones>

Indexada en DOAJ (Directory of Open Access Journals), Latindex catálogo 2.0, Núcleo Básico de Revistas Científicas Argentinas, LILACS

(Literatura Latinoamericana y del Caribe en Ciencias de la Salud).

ISSN 2314-3312 (en línea)



INDICE

EDITORIAL

132 El precio de revisar gratis: el trabajo invisible que sostiene la ciencia

Alfredo Eymann, Mariana Rapoport, Daniel Matusevich

La revisión por pares es una tarea esencial, altamente especializada y mayormente no remunerada, que sostiene la calidad científica. Frente a la incorporación de inteligencia artificial, el desafío consiste en definir usos legítimos y transparentes que asistan al revisor sin sustituir su juicio experto ni su responsabilidad.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1372>

134 Otoesclerosis en el adulto mayor: la indicación más allá de la edad cronológica

Graciela L. Tiberti

Este editorial, basado en el artículo de Redivo, Fernández Long y Pérez Raffo sobre cirugía estapediovestibular en pacientes con otoesclerosis de 65 años o más, analiza sus resultados junto con la evidencia disponible y aborda el papel de la edad, los resultados auditivos, la seguridad y la rehabilitación posterior.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1373>

136 Evaluación clínica del bocio y riesgo de malignidad

Marcelo F. Figari

A partir del trabajo de Daza sobre predictores de malignidad en bocio, se analiza el valor de la anamnesis y la semiología inicial. En entornos con recursos limitados, identificar hallazgos clínicos de alarma resulta clave para priorizar el acceso oportuno a estudios ecográficos y citológicos de mayor complejidad.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1377>

138 ARTÍCULO ORIGINAL

Factores clínicos y demográficos asociados a carcinoma diferenciado de tiroides en pacientes con bocio de larga data en Bolivia

Álvaro B. Daza Mier

Una investigación en Bolivia determinó que, en pacientes con bocio de larga duración, los siguientes signos pueden ser indicativos de cáncer de tiroides: 1) un bulto rígido que permanece inmóvil, 2) un aumento rápido en crecimiento y 3) contexto urbano. Contribuyen a la detección temprana de casos, sobre todo en áreas con escasos recursos.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1290>

146 Cirugía estapediovestibular en pacientes con otoesclerosis de 65 años o más: resultados auditivos posoperatorios. Experiencia de 20 años

Micaela Redivo, Federica Fernández Long, Gabriela Pérez Raffo

Este estudio evalúa los resultados auditivos de la cirugía estapediovestibular en pacientes mayores de 65 años, que demuestran alta tasa de éxito, mejoría significativa de la audición y baja necesidad de audífonos, respaldando su indicación como tratamiento eficaz y seguro en adultos mayores.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1326>

153 Síndrome de seno silente: resultados quirúrgicos y propuesta de algoritmo para la reconstrucción orbitaria

Antonella Giorgi, Micaela Redivo, Carlos S. Ruggeri

El síndrome de seno silente (SSS) es una entidad infrecuente caracterizada por atelectasia maxilar progresiva con descenso del piso orbitario, que genera enoftalmos e hipoglobo. Describimos los resultados del tratamiento quirúrgico en pacientes con SSS y proponemos un algoritmo para la indicación de reconstrucción orbitaria.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1342>

159 Sinusitis odontógena: ¿por dónde empezar? La importancia del abordaje multidisciplinario

Candelaria Serrano, Facundo Zamar, Sebastián Aragón

Existe una falta de consenso sobre el manejo de la sinusitis odontógena. La colaboración entre otorrinolaringólogos y cirujanos maxilofaciales es fundamental para alcanzar resultados óptimos. La cirugía combinada representa una estrategia fiable y curativa. Se presenta el algoritmo de acción utilizado en el Hospital Italiano de Buenos Aires.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1318>

REVISIÓN NARRATIVA

167 El papel del *benchmarking* en la gestión sanitaria: más allá de la medición, una estrategia para aprender y mejorar

María F. Grande Ratti, Noelia S. De Masi, José A. Orozco Niño, Sebastián Marciano

Existe una falta de consenso sobre el manejo de la sinusitis odontógena. La colaboración entre otorrinolaringólogos y cirujanos maxilofaciales es fundamental para alcanzar resultados óptimos. La cirugía combinada representa una estrategia fiable y curativa. Se presenta el algoritmo de acción utilizado en el Hospital Italiano de Buenos Aires.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1317>

REVISIÓN BREVE

177 Nuevas terapias para enfermedad hepática por depósito graso/metabólica (MASLD/MASH): resultados clínicos y retos en implementación

Javier J. Flores, Elizabeth P. Aguilera

El trabajo revisa la evidencia reciente sobre nuevas terapias para MASLD/MASH, analiza sus resultados clínicos en eficacia y seguridad, y discute los retos para su implementación en la práctica real. Incluye consideraciones sobre acceso, equidad, adherencia, comorbilidades y el impacto en los sistemas de salud.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1314>

CASO CLÍNICO

183 Aneurisma aórtico micótico por enterobacterias multirresistentes en paciente anciano: desafío diagnóstico

Anabela Games, Sofía Rossetti, Ana Braslavsky, Ignacio Martingano, Tomás Barrera

Presentamos el caso de un paciente anciano con hallazgo de aneurisma aórtico micótico asociado a bacteriemia por enterobacterias multirresistentes. Este informe de caso resalta la importancia de considerar complicaciones vasculares ante infecciones persistentes y el papel clave de métodos imagenológicos avanzados para su diagnóstico.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1325>

188 Lesión hepática grave asociada al consumo de productos fitoterapéuticos no regulados: Informe de caso y revisión de la literatura

María A. De la Rosa, José Sanabria Vanegas, Gabriela Beltrán Gil

El uso de terapias alternativas suele estar impulsado por la idea errónea de que los productos naturales son inherentemente seguros y pueden proporcionar importantes beneficios para la salud sin efectos adversos. Existe una preocupación creciente debido a la asociación entre estos productos y el daño hepático significativo, lo que llevó a la definición de la lesión hepática inducida por hierbas (HILI).

DOI: <https://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1333>

NOTAS SOBRE ESTADÍSTICA E INVESTIGACIÓN

194 En búsqueda de una $p < 0,05$: el problema de las comparaciones múltiples

María E. Knorre, Mariano Bergier, Arturo M. Cagide

¿Qué tienen en común el azar y la medicina? Este artículo reflexiona sobre cómo múltiples comparaciones pueden generar resultados engañosamente significativos. A través de ejemplos claros y analogías, se aborda el uso correcto del valor de p y la importancia de una adecuada interpretación estadística en investigación clínica.

DOI: <http://doi.org/10.51987/rev.hosp.ital.b.aires.v46i3.1301>

199 Desenlaces subrogados en investigación clínica

Mariana A. Burgos, Silvana B. Soli, Habik A. Páez Álvarez, María F. Grande Ratti

Este trabajo explica qué son los resultados subrogados: mediciones intermedias usadas para estimar beneficios clínicos como mortalidad o calidad de vida. Pese a su uso frecuente en investigación, seguimiento clínico y decisiones de cobertura, no siempre reflejan beneficios reales. El manuscrito propone una mirada crítica y prioriza desenlaces relevantes para los pacientes.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1312>

HUMANIDADES

205 Decisiones médicas compartidas: una perspectiva desde la Ciencia de la Complejidad

Alicia Rodríguez López

Las decisiones médicas compartidas nos invitan a superar el modelo paternalista y reconocer la participación activa del paciente. Desde la ciencia de la complejidad, decidir en salud implica integrar evidencia, experiencia, valores y contexto, comprendiendo que cada situación es dinámica e incierta. Esta mirada favorece decisiones más humanas, participativas y adaptativas.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1319>

212 ¿Puede un antidepresivo cambiar una vida? Evidencia, experiencia y relevancia clínica

Nahuel E. Baca

El artículo analiza el debate sobre la eficacia de los antidepresivos y cuestiona la interpretación de los tamaños de efecto promedio como único criterio de relevancia clínica. Propone una perspectiva integradora que articula la evidencia científica con el juicio clínico, el contexto asistencial y la experiencia del paciente.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1320>



INDEX

EDITORIAL

132 The Cost of Reviewing for Free: The Invisible Work that Sustains Science

Alfredo Eymann, Mariana Rapoport, Daniel Matusevich

Peer review is an essential, highly specialized, and largely unpaid task that supports the quality of scientific publishing. With the growing use of artificial intelligence, the challenge is to define legitimate and transparent applications that assist reviewers without replacing their expert judgment or responsibility.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1372>

134 Otosclerosis in Older Adults: Indications Beyond Chronological Age

Graciela L. Tiberti

This editorial, based on the article by Redivo, Fernández Long, and Pérez Raffo on stapes surgery in patients aged 65 years or older with otosclerosis, analyzes its findings in light of the available evidence and discusses the role of age, hearing outcomes, safety, and subsequent auditory rehabilitation.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1373>

136 Clinical Evaluation of Goiter and Risk of Malignancy

Marcelo F. Figari

Based on the study by Daza on predictors of malignancy in goiter, this commentary analyzes the value of the initial clinical history and physical examination. In resource-limited settings, identifying clinical red flags is key to prioritizing timely access to more complex ultrasound and cytological evaluations.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1377>

ORIGINAL ARTICLE

138 Risk Factors for Differentiated Thyroid Carcinoma in Long-Standing Goiter: A Bolivian Clinical-Demographic Study

Álvaro B. Daza Mier

A study in Bolivia found that, in patients with long-standing goiter, the following signs may indicate thyroid cancer: 1) a firm lump that remains immobile, 2) rapid growth, and 3) an urban setting. These factors contribute to the early detection of cases, especially in resource-limited areas.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1290>

146 Postoperative Hearing Outcomes of Stapedio-Vestibular Surgery in Patients with Otosclerosis Aged 65 Years or Older: A 20-Year Experience

Micaela Redivo, Federica Fernández Long, Gabriela Pérez Raffo

This study evaluates the hearing outcomes of stapes surgery in patients over 65 years of age, demonstrating a high success rate, significant improvement in hearing, and a low need for hearing aids, supporting its use as an effective and safe treatment in older adults.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1326>

153 Silent sinus syndrome: surgical outcomes and a proposed algorithm for orbital reconstruction

Antonella Giorgi, Micaela Redivo, Carlos S. Ruggeri

Silent sinus syndrome (SSS) is a rare condition characterized by progressive maxillary atelectasis with a lowering of the orbital floor, leading to enophthalmos and hypoglobus. We describe the results of surgical treatment in patients with SSS and propose an algorithm for determining the indication for orbital reconstruction.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1342>

159 Odontogenic Sinusitis: Where to Start? The Importance of a Multidisciplinary Approach

Candelaria Serrano, Facundo Zamar, Sebastián Aragón

There is a lack of consensus regarding the management of odontogenic sinusitis. Collaboration between otolaryngologists and maxillofacial surgeons is essential to achieving optimal results. Combined surgery represents a reliable and curative strategy. This article presents the treatment algorithm used at the Hospital Italiano in Buenos Aires.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1318>

NARRATIVE REVIEW

167 The Role of Benchmarking in Healthcare Management: Beyond Measurement, a Strategy for Learning and Improvement

María F. Grande Ratti, Noelia S. De Masi, José A. Orozco Niño, Sebastián Marciano

There is a lack of consensus regarding the management of odontogenic sinusitis. Collaboration between otolaryngologists and maxillofacial surgeons is essential to achieving optimal results. Combined surgery represents a reliable and curative strategy. This article presents the treatment algorithm used at the Hospital Italiano in Buenos Aires.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1317>

BRIEF REVIEW

177 New Therapies for Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD/MASH): Clinical Outcomes and Implementation Challenge

Javier J. Flores, Elizabeth P. Aguilera

The article reviews the recent evidence on new therapies for MASLD/MASH, analyzes their clinical outcomes in terms of efficacy and safety, and discusses the challenges of implementing them in real-world practice. It includes considerations regarding access, equity, adherence, comorbidities, and the impact on healthcare systems.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1314>

CLINIC CASE

183 Mycotic Aortic Aneurysm Caused by Multidrug-Resistant Enterobacteriaceae in an Elderly Patient: A Diagnostic Challenge

Anabela Games, Sofía Rossetti, Ana Braslavsky, Ignacio Martingano, Tomás Barrera

We present the case of an elderly patient with a fungal aortic aneurysm associated with bacteremia caused by multidrug-resistant Enterobacteriaceae. This case report highlights the importance of considering vascular complications in cases of persistent infections and the key role of advanced imaging methods in their diagnosis.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1325>

188 Severe Liver Injury Associated with the Use of Unregulated Herbal Products: A Case Report and Literature Review

María A. De la Rosa, José Sanabria Vanegas, Gabriela Beltrán Gil

The use of alternative therapies is often driven by the misconception that natural products are inherently safe and can provide substantial health benefits devoid of adverse effects. There is growing concern as the association between these products and significant liver damage, leading to the definition of Herb-Induced Liver Injury.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1333>

NOTES ON STATISTICS AND RESEARCH

194 In Search of a $p < 0.05$: The Problem of Multiple Comparisons

María E. Knorre, Mariano Bergier, Arturo M. Cagide

What do chance and medicine have in common? This article reflects on how multiple comparisons can generate misleadingly significant results. Through clear examples and analogies, it addresses the proper use of the p-value and the importance of appropriate statistical interpretation in clinical research.

DOI: <http://doi.org/10.51987/rev.hosp.ital.b.aires.v46i3.1301>

199 Surrogate Outcomes in Clinical Research

Mariana A. Burgos, Silvana B. Soli, Habik A. Páez Álvarez, María F. Grande Ratti

This paper explains what surrogate outcomes are: intermediate measures used to estimate clinical benefits such as mortality or quality of life. Despite their frequent use in research, clinical follow-up, and coverage decisions, they do not always reflect actual benefits. The manuscript offers a critical perspective and prioritizes outcomes that are relevant to patients.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1312>

HUMANITIES

205 Shared Medical Decision-Making: A Complexity Science Perspective

Alicia Rodríguez López

Shared medical decision-making invites us to move beyond the paternalistic model and recognize the patient's active participation. From the perspective of complexity science, making decisions in healthcare involves integrating evidence, experience, values, and context, while acknowledging that each situation is dynamic and uncertain. This perspective promotes more human-centered, participatory, and adaptive decisions.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1319>

212 Can an Antidepressant Change a Life? Evidence, Experience, and Clinical Relevance

Nahuel E. Baca

The article examines the debate surrounding the efficacy of antidepressants and questions the interpretation of average effect sizes as the sole criterion of clinical relevance. It proposes an integrative perspective that combines scientific evidence with clinical judgment, the healthcare context, and the patient's experience.

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1320>

El precio de revisar gratis: el trabajo invisible que sostiene la ciencia

Alfredo Eymann[■], Mariana Rapoport[■] y Daniel Matusevich[■]

Revista del Hospital Italiano de Buenos Aires, Universidad Hospital Italiano. Buenos Aires, Argentina

La revisión por pares acompaña a la comunicación científica desde hace siglos. Sus antecedentes pueden rastrearse hasta *Philosophical Transactions*, fundada en 1665, y, en el ámbito médico, hasta *Medical Essays and Observations*, publicada en Edimburgo desde 1731, donde los manuscritos eran distribuidos entre miembros considerados especialmente versados en cada materia antes de decidir su publicación. El sistema que hoy conocemos se desarrolló progresivamente durante los siglos siguientes hasta convertirse en uno de los pilares de la publicación científica¹.

Antes de que un manuscrito llegue a sus lectores, profesionales altamente calificados y con experiencia reconocida dedican tiempo a leerlo críticamente, identificar sus fortalezas y debilidades, evaluar la pertinencia de sus métodos e interpretar la relevancia de sus resultados. Lo hacen, además, con una responsabilidad particular: de sus observaciones puede depender no solo la publicación de un artículo, sino también la calidad con que ese conocimiento se incorpora a la literatura científica.

Se trata de una tarea de alta exigencia intelectual que, paradójicamente, permanece en gran medida invisible. Una buena revisión requiere conocimiento especializado, experiencia en investigación, capacidad crítica y prudencia para distinguir entre un verdadero problema metodológico y una simple diferencia de opinión. También requiere generosidad. El buen revisor no se limita a señalar errores: procura contribuir a que el trabajo sea mejor. En ese sentido, la revisión por pares conserva una dimensión profundamente altruista, basada en la idea de que quienes participan de la comunidad científica tienen también una responsabilidad en la construcción colectiva del conocimiento^{2,3}.

La magnitud de esta contribución es difícil de dimensionar. Se ha estimado que, solo durante 2020, los investigadores dedicaron más de 130 millones de horas a revisar manuscritos científicos, equivalentes a casi 15 000 años de trabajo. El valor económico estimado de ese

tiempo superó los 1500 millones de dólares únicamente para los revisores de los Estados Unidos⁴.

Estas cifras adquieren una dimensión particular cuando se considera la economía contemporánea de la publicación científica. La revisión puede entenderse como parte de la reciprocidad que sostiene a una comunidad académica: hoy evaluamos el trabajo de otros y mañana otros evaluarán el nuestro. Sin embargo, esta explicación resulta menos sencilla cuando se observa el modelo económico de una parte importante de la publicación científica. Los investigadores producen el conocimiento, escriben los manuscritos y los revisores aportan gratuitamente su tiempo y experiencia; al mismo tiempo, según el modelo editorial, los autores o sus instituciones pueden pagar cargos para publicar y, asimismo, los lectores o sus bibliotecas pueden pagar para acceder a los contenidos.

La paradoja resulta particularmente visible en las revistas pertenecientes a grandes grupos editoriales con fines de lucro: una parte sustancial del valor intelectual del proceso es aportada por la propia comunidad científica sin remuneración, mientras que esa misma comunidad puede tener que pagar para publicar el conocimiento que produce o para acceder a él una vez publicado. Esto no invalida el principio de reciprocidad que históricamente ha sostenido la revisión por pares, pero obliga a preguntarse hasta qué punto ese principio es equitativo y resulta suficiente para justificar el modelo actual.

Si la ciencia necesita revisores competentes, también necesita preguntarse cómo valorar su tiempo, acreditar su tarea y evitar que una actividad esencial para garantizar la calidad de las publicaciones dependa indefinidamente de una disponibilidad altruista que parece darse por sentada.

A esta tensión se agrega ahora otra, más reciente: la inteligencia artificial generativa.

Las recomendaciones editoriales llaman a una especial prudencia en su utilización durante la revisión, fundamentalmente por los riesgos vinculados con la confidencialidad de manuscritos inéditos, la propiedad

Autor para correspondencia: alfredo.eymann@hospitalitaliano.org.ar, Eymann A.

Recibido: 1/09/2026 | Aceptado: 9/09/2026 | Publicado: 18/9/2026

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1372>

Cómo citar: Eymann A, Rapoport M, Matusevich D. El precio de revisar gratis: el trabajo invisible que sostiene la ciencia. *Rev Hosp Ital B.Aires.* 2026;46(3):e0001372

intelectual y la posibilidad de que los sistemas generativos produzcan información incorrecta, incompleta o sesgada^{5,6}. Sin embargo, comienza a consolidarse una perspectiva que procura diferenciar la sustitución del juicio del revisor del uso de la inteligencia artificial como herramienta de apoyo.

La Declaración de Heredia, publicada en Costa Rica en 2024, resulta particularmente interesante en este sentido. Reconoce a la inteligencia artificial como una herramienta que puede incorporarse al proceso editorial, pero establece que no debe sustituir las responsabilidades de quienes ejercen los roles de autoría, revisión y edición. Para la revisión por pares propone un uso complementario, bajo principios de transparencia y responsabilidad humana⁶.

La discusión ha avanzado rápidamente. En 2025, editores de JAMA y de revistas de JAMA Network plantearon explícitamente la posibilidad de estrategias híbridas de revisión que incorporen inteligencia artificial manteniendo al ser humano como responsable del proceso. Entre los posibles usos señalaron tareas como resumir información, verificar la presencia de elementos requeridos o identificar problemas de informe, mientras reconocieron que algunas dimensiones como la originalidad, la relevancia clínica o la interpretación contextual continúan requiriendo especialmente el juicio experto⁷. En 2026, Zavaleta-Monestel y cols. retomaron esta propuesta y advirtieron que la incorporación de inteligencia artificial a la revisión no debería orientarse simplemente a acelerar el proceso, sino también a mejorar su calidad, evaluación y la formación de los revisores⁸.

Quizá la pregunta relevante ya no sea simplemente si los revisores deben utilizar inteligencia artificial, sino para qué podrían utilizarla y bajo qué condiciones. No parece equivalente delegar en una herramienta la valoración de la originalidad, la relevancia o la solidez científica de un manuscrito que utilizarla para asistir en tareas acotadas, como organizar observaciones propias, comprobar aspectos formales o mejorar la claridad de un comentario. Las primeras pertenecen al núcleo del juicio experto; las segundas podrían, bajo condiciones apropiadas, constituir herramientas de apoyo⁵⁻⁸.

Esta distinción no elimina los problemas. La confidencialidad de un manuscrito inédito debe estar garantizada, el uso de estas herramientas debe ajustarse a la política de cada revista y su utilización debe ser transparente cuando corresponda. Sobre todo, el revisor debe conservar el control intelectual del proceso: la responsabilidad por una

revisión no puede atribuirse a un algoritmo; pertenece al revisor que la firma^{5,6}.

La inteligencia artificial ya forma parte de numerosas actividades de investigación, escritura y edición científica. La revisión por pares necesita del conocimiento, la experiencia y, especialmente, del juicio de quienes la realizan. Ninguna herramienta debería sustituirlos. Pero preservar la dimensión humana de la revisión no necesariamente exige prescindir de la tecnología. Quizás el desafío no consista en establecer una frontera entre revisiones “con IA” y “sin IA”, sino entre aquello que puede ser asistido y aquello que no debería ser delegado. Tal vez la pregunta que debamos hacernos ya no sea si los revisores utilizarán inteligencia artificial, sino qué usos estamos dispuestos a considerar legítimos, transparentes y compatibles con la responsabilidad que les confiamos.

Contribuciones de autores: concepción (A.E., M.R., D.M.), redacción (A.E.), revisión y aceptación del manuscrito (A.E., M.R., D.M.).

Conflictos de intereses: los autores declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: los autores declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Benos DJ, Bashari E, Chaves JM, et al. The ups and downs of peer review. *Adv Physiol Educ.* 2007;31(2):145-152. <https://doi.org/10.1152/advan.00104.2006>.
2. Tennant JP, Dugan JM, Graziotin D, et al. A multi-disciplinary perspective on emergent and future innovations in peer review. *F1000Res.* 2017;6:1151. <https://doi.org/10.12688/f1000research.12037.3>.
3. Grande Ratti MF, Mollerach S, Perez Manelli R. La revisión por pares en ciencias de la salud. *Rev Hosp Ital B.Aires.* 2025;45(2):e0000351. <https://doi.org/10.51987/rev.hosp.ital.baires.v45i2.351>.
4. Aczel B, Szaszi B, Holcombe AO. A billion-dollar donation: estimating the cost of researchers' time spent on peer review. *Res Integr Peer Rev.* 2021;6(1):14. <https://doi.org/10.1186/s41073-021-00118-2>.
5. International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. Philadelphia, PA: American College of Physicians; updated Jan 2026 [citado 2026 sep 9]. Disponible en: <https://www.icmje.org/recommendations/>.
6. Penabad-Camacho L, Penabad-Camacho MA, Mora-Campos A, et al. Declaración de Heredia: principios sobre el uso de inteligencia artificial en la edición científica. *Rev Electrón Educare.* 2024;28(S):1-10. <https://doi.org/10.15359/ree.28-S.19967>.
7. Perlis RH, Christakis DA, Bressler NM, et al. Artificial intelligence in peer review. *JAMA.* 2025;334(17):1520-1522. <https://doi.org/10.1001/jama.2025.15827>.
8. Zavaleta-Monestel E, Chaverri-Fernández JM, Arguedas-Chacón S. AI in peer review. *JAMA.* 2026;335(6):549-550. <https://doi.org/10.1001/jama.2025.22266>.

Otosclerosis in Older Adults: Indications Beyond Chronological Age

Graciela L. Tiberti 

Área de Otolología, Hospital Británico de Buenos Aires. Buenos Aires, Argentina

La cirugía del estribo continúa ocupando un lugar central en el tratamiento de la otosclerosis cuando existe un componente conductivo significativo y una reserva coclear que permita esperar beneficio auditivo. Las recomendaciones basadas en evidencia de la Brazilian Society of Otolology incluyen, dentro del abordaje terapéutico, la cirugía estapedial, los audífonos y los dispositivos implantables, con selección individual según las características audiológicas y clínicas. En este marco, la edad cronológica ha sido objeto de análisis específico porque en los adultos mayores puede coexistir un componente neurosensorial y puede ser necesaria rehabilitación auditiva aun después de una cirugía técnicamente exitosa^{1,2}.

El trabajo de Redivo, Fernández Long y Pérez Raffo aporta una serie institucional de 20 años sobre 63 pacientes y 72 oídos con otosclerosis y seguimiento posoperatorio mayor de un año. La cohorte fue definida por pacientes que tenían 65 años o más al 1 de enero de 2025: 41 fueron operados antes de los 65 años y 22 a los 65 años o más. El 98,6% presentaba hipoacusia mixta. El cierre del gap aéreo-óseo (ABG) ≤ 10 dB se alcanzó, en promedio, en el 85,4% de los oídos, y la variación de la conducción ósea se mantuvo ≤ 10 dB en el 85%. El seguimiento medio fue de 7,9 años; el 22% requirió audífonos después de la cirugía y se registraron complicaciones en el 9,52% de los pacientes. Estos datos deben interpretarse junto con la composición etaria de la cohorte y con las limitaciones que los propios autores señalan: diseño retrospectivo y descriptivo, selección de pacientes quirúrgicos y ausencia de un grupo comparador más joven¹.

La literatura específica sobre cirugía estapedial en pacientes de edad avanzada es consistente en un punto: los resultados auditivos pueden ser comparables a los obtenidos en pacientes más jóvenes cuidadosamente seleccionados. Vartiainen comparó 46 oídos operados en pacientes mayores de 60 años con 330 oídos de pacientes menores y observó

resultados similares en el ABG, sin mayor frecuencia de hipoacusia neurosensorial posoperatoria en el grupo de mayor edad. Ayache y cols. estudiaron 16 pacientes mayores de 65 años: todos alcanzaron un ABG < 20 dB, el 87,5% logró un ABG < 10 dB al mes y el 81,8% mantuvo ese resultado en la última evaluación. En una serie prospectiva de 3050 estapedectomías, Vincent y cols. informaron cierre del ABG ≤ 10 dB en el 94,5% del subgrupo de pacientes mayores de 65 años y no identificaron un aumento del riesgo de complicaciones atribuible a la edad³⁻⁵.

Los estudios más recientes tampoco demuestran que la edad, por sí sola, determine el resultado quirúrgico. Dhooge y cols., en una serie con seguimiento prolongado, obtuvieron un ABG ≤ 10 dB en el 70,7% de los casos en la evaluación tardía y no hallaron a la edad como predictor independiente del éxito a largo plazo. Zhang y cols. evaluaron 188 oídos y observaron mejoras significativas de la vía aérea y del ABG en todos los grupos etarios; aunque la ganancia de conducción aérea mostró una correlación negativa modesta con la edad, no hubo diferencias significativas entre grupos jóvenes y mayores en los cambios de conducción aérea, conducción ósea o cierre del ABG. Los autores concluyeron que la edad no debería considerarse una contraindicación cuando se cumplen los criterios audiológicos^{6,7}.

El seguimiento prolongado es especialmente relevante en la otosclerosis porque el éxito en la corrección del componente conductivo no elimina la posibilidad de progresión neurosensorial. Ishai y cols., en 357 oídos con una media de 14 años de seguimiento, observaron que el 34% presentó progresión clínicamente significativa de la hipoacusia neurosensorial por encima de la esperada para la edad. Este hallazgo ofrece un contexto importante para la serie del Hospital Italiano: un ABG cerrado y una conducción ósea preservada en el posoperatorio no excluyen la necesidad futura de amplificación^{1,8}.

Autor para correspondencia: ltibertiorl@gmail.com, Tiberti GL.

Recibido: 13/09/2026 | Aceptado: 21/09/2026 | Publicado: 25/09/2026

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1373>

Cómo citar: Tiberti GL. Otosclerosis in Older Adults: Indications Beyond Chronological Age. Rev Hosp Ital B.Aires. 2026;46(3):e0001373

La relación entre cirugía y audífonos, por lo tanto, no debe plantearse necesariamente como una elección excluyente. Wardenga y cols. analizaron 374 cirugías en pacientes con hipoacusia mixta: el 44,9% dejó de necesitar audífono después de la cirugía y, entre quienes continuaron requiriéndolo, un audífono de potencia podía satisfacer los requerimientos audiológicos en las cuatro frecuencias estudiadas en el 81,6% de los casos. En la serie de Fernández Long y cols., 14 pacientes (22%) requirieron audífonos luego de la intervención, con una media de 3,9 años entre cirugía y adaptación. Estos resultados sostienen la utilidad de considerar la cirugía y la rehabilitación auditiva como etapas potencialmente complementarias del tratamiento^{1,9}.

La seguridad también forma parte central de la indicación. En el registro sueco analizado por Pauli y cols., que incluyó 832 pacientes sometidos a estapedectomías, las complicaciones fueron poco frecuentes y el 91% refirió una audición mejor o mucho mejor al año. En el estudio del Hospital Italiano no se informaron secuelas auditivas graves entre las complicaciones descritas; se registraron perforación timpánica residual, granuloma, extrusión protésica, reagudización transitoria del acúfeno y un síndrome de Ramsay Hunt, con mantenimiento de resultados audiológicos favorables^{1,10}.

El aporte principal de esta experiencia institucional es sumar seguimiento prolongado en una población de adultos mayores, un grupo menos representado en muchas series quirúrgicas. Al mismo tiempo, la lectura del estudio exige distinguir entre la edad del paciente durante el seguimiento y la edad al momento de la intervención, dado que el 65,1% de la cohorte fue operado antes de los 65 años. La evidencia publicada que incluye comparaciones por edad y subgrupos específicamente operados después de los 60-65 años coincide en que la edad cronológica aislada no predice un peor cierre del ABG ni un aumento consistente de las complicaciones. La indicación debe sostenerse en la evaluación audiológica, la reserva coclear y la expectativa de rehabilitación posterior, incluyendo audífonos cuando sean necesarios. En ese contexto, la

serie presentada refuerza una conducta respaldada por varias décadas de evidencia: en pacientes seleccionados: alcanzar una edad avanzada no constituye por sí mismo un motivo para excluir la cirugía estapedial^{1,3-7,9,10}.

Conflictos de intereses: la autora declara no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: la autora declara que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Redivo M, Fernández Long F, Pérez Raffo G. Cirugía estapediovestibular en pacientes con otosclerosis de 65 años o más: resultados auditivos posoperatorios. Experiencia de 20 años. *Rev Hosp Ital B.Aires.* 2026;46(3). <https://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1326>
2. Silva VAR, Pauna HF, Lavinsky J, et al. Brazilian Society of Otolaryngology task force - Otosclerosis: evaluation and treatment. *Braz J Otorhinolaryngol.* 2023;89(5):101303. <https://doi.org/10.1016/j.bjorl.2023.101303>
3. Vartiainen E. Surgery in elderly patients with otosclerosis. *Am J Otol.* 1995;16(4):536-8.
4. Ayache D, Corre A, Van Prooyen S, et al. Surgical treatment of otosclerosis in elderly patients. *Otolaryngol Head Neck Surg.* 2003;129(6):674-7. <https://doi.org/10.1016/j.otohns.2003.08.013>
5. Vincent R, Sperling NM, Oates J, et al. Surgical findings and long-term hearing results in 3,050 stapedotomies for primary otosclerosis: a prospective study with the otology-neurotology database. *Otol Neurotol.* 2006;27(8 Suppl 2):S25-S47. <https://doi.org/10.1097/01.mao.0000235311.80066.df>
6. Dhooze I, Desmedt S, Maly T, et al. Long-term hearing results of stapedotomy: analysis of factors affecting outcome. *Eur Arch Otorhinolaryngol.* 2018;275(5):1111-1119. <https://doi.org/10.1007/s00405-018-4899-2>
7. Zhang N, Ma X, Wang D, et al. Impact of age on hearing outcomes following stapes surgery for otosclerosis: a retrospective study. *Acta Otolaryngol.* 2025;145(8):682-686. <https://doi.org/10.1080/00016489.2025.2504040>
8. Ishai R, Halpin CF, Shin JJ, et al. Long-term incidence and degree of sensorineural hearing loss in otosclerosis. *Otol Neurotol.* 2016;37(10):1489-1496. <https://doi.org/10.1097/MAO.0000000000001234>
9. Wardenga N, Diedrich V, Waldmann B, et al. Hearing aid treatment in patients with mixed hearing loss. Part I: expected benefit and limitations after stapes surgery. *Audiol Neurotol.* 2020;25(3):125-132. <https://doi.org/10.1159/000502994>
10. Pauli N, Strömbäck K, Lundman L, et al. Surgical technique in stapedotomy hearing outcome and complications. *Laryngoscope.* 2020;130(3):790-796. <https://doi.org/10.1002/lary.28072>

Evaluación clínica del bocio y riesgo de malignidad

Marcelo F. Figari 

Sector de Cabeza y Cuello, Servicio de Cirugía General, Hospital Italiano. Buenos Aires, Argentina

Ante un paciente con bocio de larga data, uno de los primeros interrogantes diagnósticos es definir qué hallazgos clínicos pueden permitirnos modificar la impresión inicial de benignidad y realizar una investigación más profunda a fin de confirmar la potencial malignidad. Si bien la ecografía, la citología y las pruebas moleculares han mejorado drásticamente nuestras posibilidades diagnósticas, la anamnesis y la semiología inicial mantienen un valor significativo. Esto es aún más relevante en contextos socioeconómicos desfavorables con dificultades de acceso a la medicina de alta complejidad.

El Dr. Daza presenta en este número de la Revista del Hospital Italiano un estudio retrospectivo realizado sobre pacientes de un hospital de tercer nivel de Santa Cruz de la Sierra, en Bolivia. Se trata de una cohorte de 193 pacientes operados por bocio a lo largo de 11 años, de los cuales 76 presentan un carcinoma diferenciado de tiroides. El objetivo del estudio es encontrar asociaciones entre los hallazgos clínicos y de los auxiliares diagnósticos al momento de la presentación y la presencia de malignidad. Así, analiza la consistencia duro-pétrea o adherencia a planos profundos, el crecimiento rápido advertido por el paciente en el último año, el hecho de pertenecer a la población urbana o haber migrado de la rural a la urbana más de 10 años atrás, las categorías ecográficas de sospecha (TI-RADS ≥ 4) y las categorías de Bethesda IV a VI en la citología por punción. En el análisis multivariado comunicado por el autor, la consistencia pétrea o adherencia es la única variable que conserva significación estadística independiente.¹

El trabajo del Dr. Daza tiene el valor de recuperar una dimensión fundamental de la consulta clínica: reconocer cambios en una enfermedad que el paciente refiere como estable desde hace años. Sin embargo, la interpretación de estos hallazgos requiere precisar qué es lo que se observa y en qué momento. En el caso de la consistencia

duro-pétrea y la adherencia, la inclusión de hallazgos intraoperatorios en esta variable limita su interpretación como predictor disponible en la consulta inicial. La consistencia duro-pétrea, la disminución de movilidad y la adherencia constituyen características relacionadas, pero no equivalentes. Ninguna de ellas permite afirmar contundentemente, por sí misma, que exista malignidad.

El autor discute también el valor del crecimiento glandular o nodular como probable indicio de malignidad. Daza lo define mediante la percepción del paciente durante los doce meses previos a la consulta, pero no documenta mediciones objetivas seriadas. Además, esa apreciación tiene un peso subjetivo inevitable, no comparable con el crecimiento documentado por ecografía. Una revisión de Bernet y Chindris (2021) destaca que, en nódulos con citología benigna, las características ecográficas tienen un papel relevante para guiar el seguimiento y la necesidad de nuevas evaluaciones.²

En el bocio multinodular, además, el volumen de la glándula y el tamaño del nódulo dominante pueden desviar la atención de otras lesiones relevantes. Luo y colaboradores encuentran, en una serie quirúrgica de 838 pacientes operados por bocio multinodular, que solo el 37 % de los cánceres se localiza en el nódulo de mayor tamaño.³ Precisamente por eso es aconsejable realizar una evaluación ecográfica que seleccione las lesiones por sus características de sospecha, evitando que el tamaño sea el único criterio para decidir la punción.

La citología aporta información diagnóstica específica. La clasificación de Bethesda distingue categorías citológicas con diferentes riesgos de malignidad, cuya interpretación debe vincularse con la ecografía y el contexto individual. Las guías de la European Thyroid Association sostienen, precisamente, una evaluación integrada de factores clínicos, laboratorio y estratificación ecográfica para definir la indicación de punción.⁴

Autor para correspondencia: marcelo.figari@hospitalitaliano.org.ar, Figari MF

Recibido: 25/09/2026 | Aceptado: 28/09/2026 | Publicado: 29/09/2026

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1377>

Cómo citar: Figari MF. Evaluación clínica del bocio y riesgo de malignidad. Rev Hosp Ital B.Aires. 2026;46(3):e0001377

Respecto a la extensión anatómica del bocio, aun en los bocios cérvico-torácicos y retroesternales, la posibilidad de malignidad es variable. Prete y colaboradores encuentran proporciones similares de cáncer inesperado en bocios retroesternales y cervicales sin sospecha preoperatoria.⁵ Las guías ATA (2025) añaden una perspectiva importante respecto al comportamiento clínico: en la evaluación del cáncer diferenciado de tiroides, el crecimiento rápido, la fijación a estructuras cervicales, la disfagia progresiva o los cambios significativos de la voz pueden sugerir enfermedad localmente invasiva. Ante esa sospecha, recomiendan complementar el examen físico y la ecografía con tomografía o resonancia con contraste.⁶ Los hallazgos mencionados orientan el estudio de la probable extensión tumoral, sin constituir por sí solos una confirmación de invasión. Sin embargo, una vez definida la misma, necesariamente impactará sobre el tipo de tratamiento o la necesidad de derivación, de no contarse con los recursos necesarios.

Estas consideraciones alcanzan a la interpretación del estudio de Daza. Su diseño retrospectivo y la selección de pacientes operados limitan la extrapolación de los resultados a la población general con bocio.

Sin embargo, el aporte del trabajo puede situarse en la necesidad de reconocer y estudiar oportunamente los hallazgos clínicos de alarma. En entornos con recursos limitados, esa identificación temprana puede contribuir a priorizar el acceso a ecografía, citología y evaluación más especializada. En suma, una anamnesis cuidadosa

y un examen físico preciso siguen siendo componentes indispensables de una evaluación que debe integrar la evidencia disponible y las circunstancias de cada paciente.

Conflictos de intereses: el autor declara no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: el autor declara que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Daza AB. Factores clínicos y demográficos asociados a carcinoma diferenciado de tiroides en pacientes con bocio de larga data en Bolivia. *Rev Hosp Ital B.Aires.* 2026;46(3):e0001290. <https://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1290>
2. Bernet VJ, Chindris AM. Update on the evaluation of thyroid nodules. *J Nucl Med.* 2021;62 Suppl 2:13S-19S. <https://doi.org/10.2967/jnumed.120.246025>
3. Luo J, McManus C, Chen H, et al. Are there predictors of malignancy in patients with multinodular goiter? *J Surg Res.* 2012;174(2):207-10. <https://doi.org/10.1016/j.jss.2011.11.1035>
4. Durante C, Hegedüs L, Czarniecka A, et al. 2023 European Thyroid Association Clinical Practice Guidelines for thyroid nodule management. *Eur Thyroid J.* 2023;12(5):e230067. <https://doi.org/10.1530/ETJ-23-0067>
5. Prete FP, De Luca GM, Sgaramella LI, et al. Prevalence and clinical risk factors of thyroid cancer in retrosternal goiter: a retrospective comparative study with cervical multinodular goiter. *J Clin Med.* 2025;14(2):489. <https://doi.org/10.3390/jcm14020489>
6. Ringel MD, Sosa JA, Baloch Z, et al. 2025 American Thyroid Association Management Guidelines for Adult Patients with Differentiated Thyroid Cancer. *Thyroid.* 2025;35(8):841-985. <https://doi.org/10.1177/10507256251363120>

Factores clínicos y demográficos asociados a carcinoma diferenciado de tiroides en pacientes con bocio de larga data en Bolivia

Álvaro B. Daza Mier¹

Sección de Cirugía de Cabeza y Cuello, Hospital San Juan de Dios, Santa Cruz de la Sierra, Bolivia

RESUMEN

Introducción: en áreas endémicas como Bolivia, el cáncer diferenciado de tiroides (CDT) en pacientes con bocio crónico supone un reto diagnóstico. El objetivo de este estudio fue describir los factores clínicos asociados a la presencia de CDT en una serie de pacientes operados por bocio de larga data en un hospital de tercer nivel.

Métodos: estudio observacional, descriptivo y retrospectivo (2014-2024) que abarcó a 193 pacientes operados por bocio en el Hospital San Juan de Dios (Santa Cruz, Bolivia), con confirmación histopatológica de CDT en 76 casos. Se empleó regresión logística en SPSS v23.0® para examinar variables clínicas, demográficas y diagnósticas, determinando razones de probabilidades (OR) con intervalos de confianza del 95 %.

Resultados y discusión: se detectaron cinco factores clínicos relevantes ($p < 0,05$), con una frecuencia de malignidad (CDT) del 42 % dentro de la cohorte operada. La adherencia a planos profundos o la consistencia dura fue el factor más relevante (85,5 % de los casos, OR 6,66, IC 95 %: 3,19-13,89), seguido del crecimiento rápido en el último año (72,4 %, OR 5,89, IC 95 %: 3,54-12,84) y de la procedencia urbana o migración de más de 10 años (72,4 %, OR 2,5, IC 95 %: 1,34-4,62). También se encontró asociación significativa entre los hallazgos citológicos (Bethesda $\geq$ IV, OR 6,69) e imagenológicos (TI-RADS $\geq$ IV, OR 3,5). Estos hallazgos son congruentes con informes internacionales que subrayan la importancia predictiva de las características clínicas en contextos con recursos escasos. La regresión multivariada identificó la adherencia a planos profundos como la única variable independiente ($p=0,024$). La asociación con procedencia urbana podría relacionarse tanto con factores ambientales como con un mayor acceso a estudios diagnósticos en estas zonas; se requieren investigaciones específicas en el contexto boliviano para esclarecer tal relación.

Conclusión: en esta serie de pacientes operados por bocio crónico, la presencia de una masa adherida a planos profundos, la procedencia de área urbana y los valores elevados en los sistemas TI-RADS/Bethesda se asociaron con la presencia de CDT. Estos hallazgos podrían contribuir a mejorar los criterios de derivación quirúrgica en áreas endémicas con recursos limitados. Futuras investigaciones deberían explorar el papel que desempeñan factores geoquímicos locales (como el litio y el calcio) en la carcinogénesis tiroidea, dado que su asociación permanece como hipótesis no confirmada en el contexto boliviano.

Palabras clave: cáncer de tiroides, bocio endémico, factores de riesgo, diagnóstico clínico.

Autor de correspondencia: adaza@ucb.edu.bo, Daza AB.

Recibido: 12/08/2025 | Aceptado: 02/09/2026 | Publicado: 28/09/2026

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1290>

Cómo citar: Daza Mier AB. Factores clínicos y demográficos asociados a carcinoma diferenciado de tiroides en pacientes con bocio de larga data en Bolivia. Rev Hosp Ital B.Aires. 2026;46(3):e0001290

Risk Factors for Differentiated Thyroid Carcinoma in Long-Standing Goiter: A Bolivian Clinical-Demographic Study

ABSTRACT

Introduction: in endemic areas such as Bolivia, differentiated thyroid cancer (DTC) in patients with chronic goiter represents a diagnostic challenge. The objective of this study was to describe the clinical factors associated with the presence of DTC in a series of patients operated on for long-standing goiter at a tertiary-level hospital.

Methods: observational, descriptive, and retrospective study (2014-2024) that included 193 patients operated on for goiter at Hospital San Juan de Dios (Santa Cruz, Bolivia), with histopathological confirmation of DTC in 76 cases. Logistic regression was performed in SPSS v23.0 to examine clinical, demographic, and diagnostic variables, determining odds ratios (OR) with 95 % confidence intervals.

Results and discussion: five relevant clinical factors were detected ($p < 0.05$), with a malignancy frequency (DTC) of 42 % within the operated cohort. Adherence to deep planes or hard consistency was the most relevant factor (85.5 % of cases, OR 6.66, 95 % CI: 3.19-13.89), followed by rapid growth in the last year (72.4 %, OR 6.75, 95 % CI: 3.54-12.84), and urban origin or migration of more than 10 years (72.4 %, OR 2.5, 95 % CI: 1.34-4.62). A significant association was also found between cytological findings (Bethesda $\geq$ IV, OR 6.69) and imaging findings (TI-RADS $\geq$ IV, OR 3.5). These findings are consistent with international reports that emphasize the predictive importance of clinical characteristics in resource-limited settings. Multivariate regression identified adherence to deep planes as the only independent variable ($p = 0.024$). The association with urban origin could be related to both environmental factors and greater access to diagnostic studies in these areas; specific research is needed in the Bolivian context to clarify this relationship.

Conclusion: in this series of patients operated on for chronic goiter, the presence of a mass adherent to deep planes, urban origin, and elevated TI-RADS/Bethesda scores were associated with the presence of DTC. These findings could help improve surgical referral criteria in endemic areas with limited resources. Future research should explore the role of local geochemical factors (such as lithium and calcium) in thyroid carcinogenesis, given that this association remains an unconfirmed hypothesis in the Bolivian context

Keywords: thyroid cancer, endemic goiter, risk factors, clinical diagnosis.

INTRODUCCIÓN

Tradicionalmente y de manera general, el término cáncer diferenciado de tiroides (CDT) se utiliza para los dos tipos de enfermedades oncológicas de la glándula tiroides más frecuentes, ambas originadas en células de estirpe epitelial¹. Existe actualmente debate en la literatura sobre la clasificación unificada como CDT, pues, aunque el carcinoma papilar y folicular comparten esta categorización internacional, presentan diferencias significativas en su comportamiento clínico. Varios estudios recientes demuestran que estos subtipos exhiben patrones evolutivos distintos, tasas de recurrencia variables y pronósticos divergentes, lo que cuestiona la pertinencia de su agrupación bajo un mismo término. Esta controversia tiene implicaciones directas en la estandarización de protocolos terapéuticos y seguimiento¹.

El bocio endémico, caracterizado por el aumento de tamaño de la glándula tiroides, no solo representa un problema estético, sino que se relaciona epidemiológicamente con el cretinismo endémico, la sordomudez y

el retraso mental². Su etiología es multifactorial: además de la deficiencia de yodo, se han descrito otros determinantes como la deficiencia de micronutrientes (p. ej., selenio) y la ingesta de sustancias bociógenas en la dieta³, que pueden interferir en la función tiroidea junto con los factores geoquímicos locales que se describen a continuación.

Bolivia, país con antecedentes históricos de bocio endémico (descrito desde la época precolombina, con prevalencias que alcanzaron hasta el 60 % de la población en algunas regiones⁴), fue declarado libre de trastornos por deficiencia de yodo en 1997 tras un programa nacional de yodación evaluado por la Organización Panamericana de la Salud (OPS), el Fondo de las Naciones Unidas para la Infancia (UNICEF) y otros organismos internacionales⁴. Sin embargo, persisten zonas rurales de altura donde el bocio continúa presente, aunque con baja prevalencia tras los programas de yodación, como lo evidenció un estudio en estudiantes de Camata Marka (La Paz), que encontró una prevalencia de bocio del 1,4 % mediante palpación

semiológica y respaldo ultrasonográfico⁵. Adicionalmente, se ha señalado que el agua de consumo en algunas regiones de Bolivia posee elevadas concentraciones de calcio, lo que podría dificultar la unión del yodo en la función tiroidea, y que la sal de uso común contiene litio (un elemento reconocido como bociógeno), que se ha demostrado que inhibe la organificación del yodo y altera la estructura de la tiroglobulina, produciendo hipotiroidismo y bocio compensador⁶. Sin embargo, hasta donde pudimos identificar, no existe evidencia directa que cuantifique una exposición ambiental relevante al litio a través de la sal de consumo en Bolivia; esta asociación permanece como hipótesis que requeriría estudios geoquímicos y toxicológicos específicos.

Factores genéticos, demográficos y ambientales pueden intervenir de forma sinérgica en la evolución del bocio difuso y nodular; los nódulos que se desarrollan en un bocio suelen originarse también por mutaciones genéticas somáticas que conducen a transformación neoplásica³. Estos elementos deben interpretarse en el contexto epidemiológico particular de Bolivia como país con antecedentes de bocio endémico y claramente en el contexto de desarrollo del CDT en países con similar característica.

El objetivo del presente estudio fue describir los factores clínicos asociados a la presencia de CDT en una serie de pacientes operados por bocio de larga data en un hospital de tercer nivel de Santa Cruz de la Sierra, Bolivia.

MATERIALES Y MÉTODOS

Se realizó un estudio observacional, descriptivo, retrospectivo y transversal, mediante la conformación

dinámica y anual de una cohorte quirúrgica durante 11 años (gestiones 2014-2024).

La población de estudio fue un Censo de pacientes operados por patología tiroidea en el Hospital San Juan de Dios (HSJD). De 193 pacientes intervenidos por bocio de larga evolución sin cirugía previa, 76 tuvieron confirmación anatomopatológica de CDT (papilar o folicular).

Para esta cohorte de pacientes con CDT, los criterios de inclusión fueron: ambos sexos, mayores de 14 años, tiroidectomía por bocio o sospecha de neoplasia, con registros preoperatorios y posoperatorios completos. También indicaciones quirúrgicas como aumento de tamaño, síntomas compresivos, sospecha de malignidad por imagen/citología, o motivos estéticos. Los criterios de exclusión fueron: pacientes con antecedente de cirugía tiroidea previa (reintervenciones), aquellos con anatomía patológica negativa para malignidad y casos de carcinoma anaplásico (CAT, n = 10) o medular (CMT, n = 2), excluidos por no corresponder al objetivo del estudio; estos 12 casos se excluyeron del análisis de factores asociados y se informan únicamente a efectos descriptivos en esta sección, se presenta la figura 1 para mejor comprensión.

Para estandarizar la recolección de datos, las variables clínicas se definieron operativamente de la siguiente manera: adherencia a planos profundos/consistencia dura (pétrea) como hallazgo clínico por palpación cervical preoperatoria o confirmado intraoperatoriamente por el cirujano. No se emplearon criterios ecográficos para definir esta variable. Como crecimiento rápido definimos la percepción subjetiva del paciente de aumento notorio del bocio en los últimos 12 meses, en una glándula con antecedente de bocio de larga evolución previamente

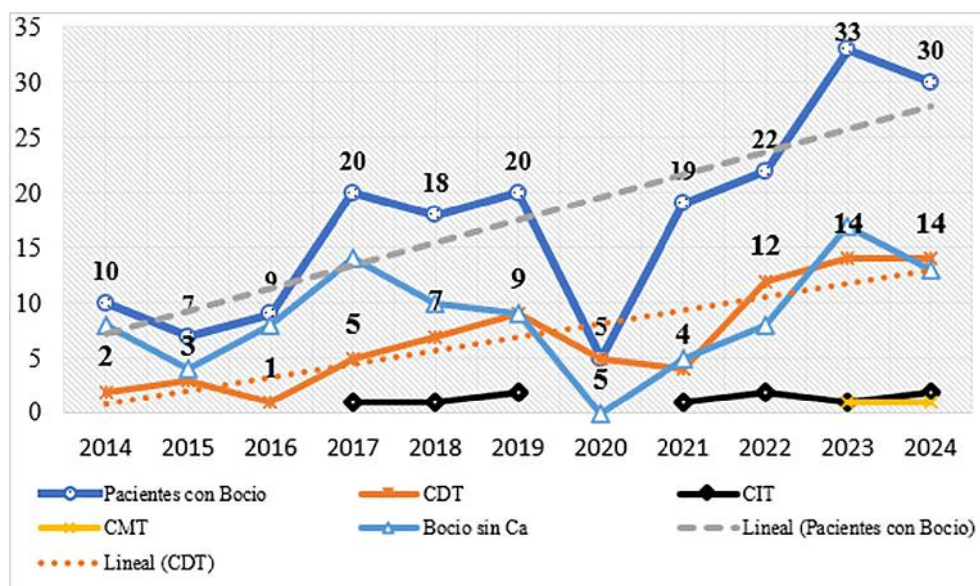


Figura 1. Frecuencia de pacientes atendidos por Ca y bocio simple durante 11 años en el HSJD. El año 2020 solo se operaron pacientes con evidente signo de desarrollo de cáncer (los cinco operados tenían CDT). Es el año de la pandemia COVID-19. Los datos de CMT y CIT solo se presentan con fines descriptivos. CDT: cáncer diferenciado de tiroides. CIT: cáncer indiferenciado de tiroides (anaplásico). CMT: cáncer medular de tiroides. Ca: cáncer. Fuente: figura modificada de Daza A., 2024⁷.

estable; se reconoce como limitación al no basarse en mediciones objetivas seriadas. La variable de procedencia urbana/rural, fue una definición *ad hoc* del estudio; urbano incluyó migración desde zona rural >10 años. El corte de edad se analizó en los puntos de 40 y 60 años. El corte de 40 años se adoptó sobre la base de un estudio previo del mismo autor⁷, en el cual se documentó mayor riesgo de recidiva de CDT a partir de esa edad en la población boliviana, reconociendo que este corte difiere del establecido por la 8.^a edición del sistema de estadificación del American Joint Committee on Cancer (AJCC), que utiliza 55 años como límite pronóstico⁸.

Se diseñó una ficha de cotejo estandarizada desde historias clínicas, complementada por entrevista posoperatoria cuando fue necesario. Los datos luego se codificaron de forma anónima, eliminando los números de historia clínica del análisis final para garantizar la confidencialidad, en concordancia con los principios éticos de la investigación. El peso de la pieza quirúrgica se registró mediante balanza de patología tras la resección.

Se realizó punción-aspiración con aguja fina (PAAF) en los nódulos con hallazgos ecográficos sospechosos de malignidad. Esto implica que la muestra citológica no representa la totalidad de los nódulos operados, sino a aquellos previamente seleccionados por sospecha ecográfica, lo que debe considerarse como una limitación al interpretar la asociación entre categoría Bethesda y CDT, dado el potencial sesgo de selección hacia resultados citológicos de mayor riesgo. Los resultados se clasificaron

según el sistema Bethesda, y los hallazgos ecográficos según la clasificación TI-RADS.

RESULTADOS

Durante un período de 11 años en el Hospital San Juan de Dios se evaluaron 193 pacientes intervenidos quirúrgicamente por bocio de larga evolución y sin tratamiento quirúrgico previo (reintervenciones de cualquier tipo). El estudio anatomopatológico confirmó 76 casos de CDT, lo que representó una frecuencia de malignidad del 42 % dentro de la cohorte quirúrgica operada. Los diez casos de carcinoma anaplásico (CAT) y los dos casos de carcinoma medular (CMT) identificados durante el período fueron excluidos del presente análisis por corresponder a criterios de exclusión, reservándose su descripción detallada para futuras publicaciones.

Con respecto a las características basales de la cohorte, la edad de los pacientes con CDT presentó una media de $47,6 \pm 14,4$ años (mediana 52 años; rango 16-71 años), con predominio del sexo femenino (60/76; 78,9 %) sobre el masculino (16/76; 21,1 %), en relación 3,75:1. El peso de la pieza quirúrgica tuvo una media de $97,9 \pm 79,6$ g (mediana 95,65 g; rango 10,45-400 g). Con respecto a la indicación quirúrgica, 29 casos (38,2 %) se indicaron por punción aspirativa con aguja fina (PAAF) sospechosa sola, 25 casos (32,9 %) por PAAF sospechosa asociada a masa clínicamente aparente, 15 casos (19,7 %) por masa clínicamente aparente sin PAAF sospechosa, y 6 casos (7,9 %) por incidentaloma; en 1 caso (1,3 %) el registro no permitió precisar la indicación (Tabla 1).

Tabla 1. Características clínicas y demográficas de los pacientes con CDT (n = 76)

Variable	n (%) o media $\pm$ DE
Edad, media $\pm$ DE (años)	47,6 $\pm$ 14,4
Edad, mediana (RIC)	52 (38-59,75)
Sexo femenino	60 (78,9%)
Sexo masculino	16 (21,1%)
Peso glandular, media $\pm$ DE (g)	97,9 $\pm$ 79,6
Peso glandular, mediana (RIC) (g)	95,65 (40,75-123,63)
Indicación quirúrgica	
PAAF sospechosa sola	29 (38,2%)
PAAF sospechosa + masa aparente	25 (32,9%)
Masa clínicamente aparente	15 (19,7%)
PAAF sospechosa + incidentaloma	6 (7,9%)
No especificado en registro	1 (1,3%)
Clasificación Bethesda	
I / II / III	1 (1,3%) / 7 (9,2%) / 8 (10,5%)
IV / V / VI	42 (55,3%) / 12 (15,8%) / 6 (7,9%)
Clasificación TI-RADS	
I / II / III	4 (5,3%) / 11 (14,5%) / 17 (22,4%)
IV / V	35 (46,1%) / 9 (11,8%)

N: número de casos (frecuencia). DE: desviación estándar. RIC: rango intercuartílico. Fuente: elaboración propia.

Se analizaron múltiples variables para identificar factores relacionados con la presencia de CDT en pacientes con bocio de largo tiempo de evolución. El análisis estadístico de 18 variables clínicas permitió identificar cinco factores significativamente asociados al desarrollo de CDT en pacientes con antecedentes de bocio crónico o de larga data. La procedencia urbana o migración a zonas pobladas por más de 10 años mostró una clara asociación, presente en el 72,4 % (55/76) de los casos de CDT frente a un 27,6 % (21/76) que provienen netamente del área rural; con un OR de 2,5 (IC 95 %: 1,34-4,62), los pacientes de zonas urbanas tienen más riesgo que los de zona rural. La consistencia dura (pétrea) o adherida a planos profundos del bocio fue otro hallazgo relevante, observado en el 85,5 % (65/76) de los pacientes con CDT, con un OR de 6,66 (IC 95 %: 3,19-13,89), constituyéndose en la variable con mayor riesgo. La percepción subjetiva de crecimiento rápido del bocio en el último año antes de la consulta con el especialista fue referida por el 72,4 % (55/76) de los casos y mostró un OR de 5,89 (IC 95 %: 3,54-12,84); esta variable en nuestra serie ocupó el segundo lugar, mostrando los factores clínicos como aporte importante en el diagnóstico.

En cuanto a los estudios diagnósticos, la clasificación TI-RADS $\geq$ IV en la ecografía tiroidea se asoció con el 57,9 % (44/76) de los casos de CDT, con un OR de 3,5 (IC 95 %: 1,91-6,43); es decir, si nuestros pacientes llegan con un valor de alta sospecha de nódulo maligno por ecografía es un factor importante de riesgo para sospecha diagnóstica de CDT. Similarmente, la citología por PAAF con clasificación Bethesda $\geq$ IV se encontró en el 78,9 % (60/76) de los casos confirmados de CDT, mostrando un OR de 6,69 (IC 95 %: 3,43-13,07). Todos estos resultados fueron estadísticamente significativos ($p < 0,05$) mediante la prueba estadística de Fisher exacto.

En este estudio se encontró asociación entre CDT y los pacientes con bocio de larga evolución que presentan características clínicas de consistencia dura/adherida, crecimiento rápido tomando en cuenta el último año o meses de evolución, procedencia urbana o migración prolongada con más de 10 años en zonas pobladas, junto con hallazgos ecográficos de TI-RADS $\geq$ IV y citológicos por PAAF en gradación de Bethesda $\geq$ IV o altamente sospechosos, factores clínicos presentes en la mayoría de los casos de CDT en esta serie. Dichos factores podrían ser útiles para estratificar el riesgo preoperatorio en esta población.

Los resultados del OR y X^2 se presentan en la tabla 2 junto con el valor de p obtenida por el estadístico de Fisher; los datos se tabularon para cada variable relacionada como factor clínico asociado a desarrollo de CDT en bocios de larga data de evolución.

Por último, se realizó un estudio de regresión logística binaria multivariada en el SPSS V23.0; los resultados muestran, en el estadístico de R cuadrado de Nagelkerke de 0,798 y en los pasos sucesivos de la regresión, la única variable que se considera independiente de las cinco variables estudiadas con significancia estadística de factor de riesgo: es la variable de que la masa sea dura o esté

adherida a planos profundos con valor de $p = 0,024$. Es decir, en los pasos sucesivos de la regresión o *likelihood ratio*, solo queda como variable independiente que la masa sea dura (pétrea) o adherida a planos profundos, estos datos se presentan en la figura 2.

Una limitación de este estudio es la falta de registro sistemático de las características histológicas detalladas de los tumores (subtipo histológico específico, tamaño tumoral, multicentricidad, extensión extraglandular y presencia de metástasis ganglionares o a distancia), lo que impide una caracterización más completa de la cohorte y su comparación con la estadificación histopatológica informada en otras series. Futuros estudios prospectivos en nuestro medio deberían incorporar el registro sistemático de estas variables.

DISCUSIÓN

La adherencia de la masa tiroidea a planos profundos o estructuras musculares se asocia significativamente con la presencia de CDT en el estudio histopatológico final. Nuestros resultados muestran que el 85,5 % de los casos de CDT presentaron esta característica, con un OR 6,66, lo que la convierte en uno de los predictores clínicos más sólidos de malignidad en pacientes con bocio crónico. Este hallazgo concuerda con estudios internacionales recientes⁹, donde Sharma informó que el 82 % de los CDT en bocios multinodulares presentaban adherencia (OR: 5,4; IC 95 %: 2,9-10,1)¹⁰, aunque no existe un consenso universal sobre los criterios ecográficos o físicos para definir "adherencia"¹¹. La consistencia dura o pétrea al examen físico semiológico, como se evidenció en nuestro estudio, emerge como un signo clave para sospechar malignidad, especialmente en contextos con limitado acceso a imágenes avanzadas. Esta observación es respaldada por Burman y Wartofsky, quienes señalan que un hallazgo clínico importante que indica invasión local es la adherencia del bocio o nódulo a los planos profundos¹². Por lo tanto, en entornos clínicos donde la evaluación preoperatoria depende fuertemente del examen físico, la palpación de una masa adherida o de consistencia pétrea debería priorizar la derivación a cirugía y estudio histopatológico.

Dentro de la variable de crecimiento rápido, si la masa tiene crecimiento evidente el último año antes de la consulta, se constituye como un factor clínico asociado a CDT en un bocio que durante años no presentaba crecimiento^{13,14}. Algunos estudios regionales describen el crecimiento del CDT como lento y de larga data¹⁵; según esos criterios no existiría correlación con nuestros hallazgos. Sin embargo, un estudio multicéntrico en Asia Oriental halló que el 80 % de CDT en bocios presentaban crecimiento acelerado vs. 30 % en benignos¹⁶. También en Japón, un estudio determinó un 85 % de CDT en pacientes con bocio multinodular con crecimiento rápido vs. 22 % en benignos ($p < 0,001$), con un OR ajustado de 7,2¹⁷. El crecimiento rápido de una masa en la tiroides está relacionado mayormente con un carcinoma anaplásico (CAT)¹⁸. Si bien estos datos se relacionan mayormente con el CAT, en nuestro estudio pudimos evidenciar que

Tabla 2. Variables analizadas como factores de riesgo para CDT

Variables	CDT		OR	IC al 95%	χ^2	Fisher exacto (p)
	Si	No				
Edad $\geq$ 60 años						
$\geq$ 60 años	19 (25%)	31 (26,5%)	0,9247	0,9247	0,0537	0,8675
< 60 años	57 (75%)	86 (73,5%)				
Edad $\geq$ 40 años						
$\geq$ 40 años	56 (73,7%)	83 (78,3%)	1,147	1,147	0,1722	0,744
< 40 años	20 (26,3%)	34 (21,7%)				
Procedencia						
Urbano $\geq$ 10 años	55 (72,4%)	60 (51,3%)	2,5	2,5	8,5068	0,004
Rural	21 (27,6%)	57 (48,7%)				
Sexo						
Femenino	60 (21,1%)	89 (76,1%)	1,1798	1,1798	0,217	0,7265
Masculino	16 (78,9%)	28 (23,9%)				
Sexo/edad Femenino						
Mujer $\geq$ 60 años	15 (25%)	23 (25,8%)	0,957	0,957	0,0134	1
Mujer < 60 años	45 (75%)	66 (74,2%)				
Sexo/edad Femenino						
Mujer $\geq$ 40 años	47 (78,3%)	64 (71,9%)	1,412	1,412	0,7783	0,445
Mujer < 40 años	13 (21,7%)	25 (28,9%)				
Sexo/edad Masculino						
Varón $\geq$ 60 años	4 (25%)	8 (38%)	0,8333	0,8333	0,0655	1
Varón < 60 años	12 (75%)	20 (62%)				
Sexo/edad Masculino						
Varón $\geq$ 40 años	9 (56,3%)	18 (64,3%)	0,714	0,714	0,2773	0,7494
Varón < 40 años	7 (43,7%)	10 (35,7%)				
Pétrea/adherida						
Si	65 (85,5%)	55 (47%)	6,6612	6,6612	29,065	0,0001
No	11 (14,5%)	62 (53%)				
Peso del bocio						
$\geq$ 100 gr	40 (37,1%)	45 (35%)	1,7778	1,7778	3,7537	0,056
< 100 gr	36 (62,9%)	72 (75%)				
Crecimiento rápido						
Creció último año	55 (72,4%)	36 (27,9%)	5,89	5,89	36,22	0,001
No creció últi. año	21 (27,6%)	81 (71,1%)				
Diabetes						
Si	27 (72,4%)	39 (27,9%)	1,058	1,058	0,1772	0,7558
No	48 (27,6%)	79 (71%)				
Fumar						
Si	25 (72,4%)	32 (27,9%)	1,3	1,3	0,6805	0,4236
No	51 (27,6%)	85 (71%)				
Antecedentes						
Si	30 (44,7%)	61 (48,4%)	0,864	0,864	0,2321	0,6523
No	37 (55,3%)	65 (51,6%)				
TI-RADS (eco)						
$\geq$ IV	44 (57,9%)	33 (23,9%)	3,5	3,5	16,936	0,0001
< IV	32 (42,1%)	84 (76,1%)				
Bethesda PAAF						
$\geq$ IV	60 (78,9%)	15 (23,4%)	6,69	6,69	15,145	0,0001
< IV	16 (21,1%)	49 (76,6%)				
Hipertiroidismo						
Si	17 (37,5%)	37 (36,7%)	0,623	0,623	1,9586	0,1904
No	58 (62,5%)	80 (63,3%)				
Hipotiroidismo						
Si	25 (44,4%)	44 (32,2%)	1,6842	1,6842	1,4398	0,2756
No	41 (55,6%)	73 (67,8%)				

CDT: cáncer diferenciado de tiroides. OR: odds ratio. IC: intervalo de confianza. X2: chi cuadrado. Fuente: elaboración propia.

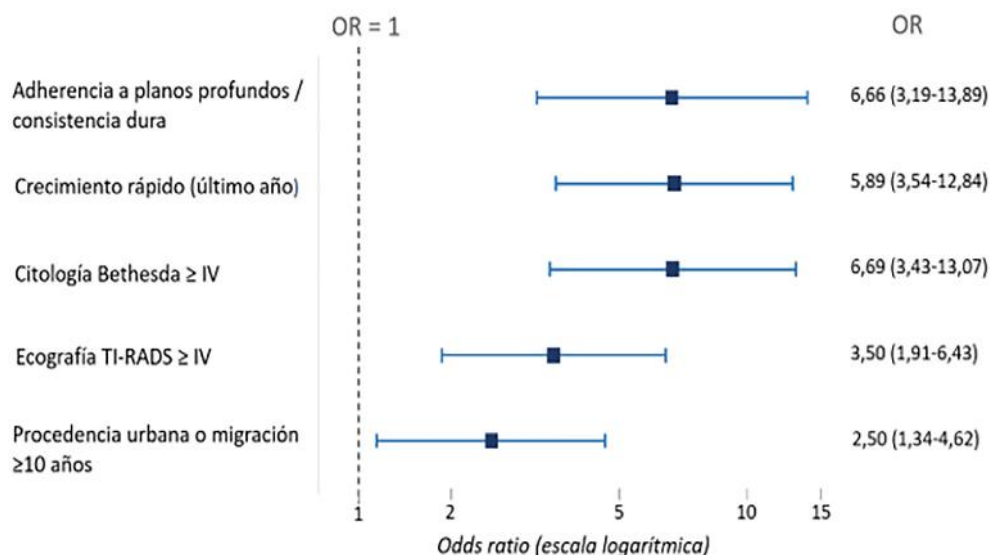


Figura 2. Factores clínicos asociados a CDT (análisis univariado OR e IC 95 %). OR: odds ratio; IC 95 % representado por las líneas horizontales. Todos los valores con $p < 0,05$ (prueba exacta de Fisher). Fuente: elaboración propia.

la percepción de un crecimiento más rápido es un factor asociado a CDT; este signo fue el primer motivo de consulta en el 75 % de los casos positivos para CDT estudiados.

Al analizar la procedencia de los pacientes, encontramos una asociación significativa que podría vincularse a múltiples factores ambientales y de acceso a servicios de salud propios del entorno urbano. Un estudio reciente en Estados Unidos, que analizó 92 794 casos de cáncer de tiroides, documentó disparidades significativas entre pacientes urbanos y rurales a lo largo de todo el continuo de atención¹⁹; en Bolivia, esos factores geográficos también podrían estar presentes en nuestras urbes. Sin embargo, la asociación observada con la procedencia urbana podría reflejar también un mayor acceso a estudios diagnósticos en zonas urbanas, un posible factor de confusión no controlado. En contraposición, en el área rural es sabida la baja calidad de la sal aun siendo yodada; respecto al litio y al calcio en su función bociógena, existe evidencia de que puede causar hipotiroidismo subclínico⁶ y tiroiditis, aunque la evidencia disponible proviene principalmente de contextos de tratamiento farmacológico con litio. En Bolivia solo encontramos la referencia de Cardona², quien señala que el agua consumida contiene altos niveles de calcio.

La ecografía desempeña un papel importante para el diagnóstico y la toma de conducta quirúrgica en relación con el CDT. En el presente estudio existen 6,73 veces más probabilidades de encontrar un CDT si se informa una ecografía con clasificación TI-RADS $\geq$ IV. Según un estudio realizado en Singapur, los nódulos con calificaciones de 4 o 5 mostraron tasas de cáncer del 10,9 % y el 71,4 %, respectivamente, tras las tiroidectomías²⁰, hallazgos relacionados con nuestros resultados.

La clasificación por PAAF de Bethesda también se constituye en un factor de riesgo; en nuestro estudio, los pacientes con un grado $\geq$ IV presentan una probabilidad 5,94 veces mayor de CDT en el estudio histopatológico. Investigaciones sobre cohortes quirúrgicas indican que la probabilidad de malignidad se incrementa con cada categoría Bethesda, siendo particularmente alta en los grados V y VI, con tasa de malignidad mayor del 80 %^{21,22}. En nuestra serie los datos concuerdan con la literatura mundial. Estos hallazgos son coherentes con las recomendaciones de la American Thyroid Association, que enfatizan el uso combinado de la clasificación ecográfica y citológica para estratificar el riesgo de malignidad²³.

El sexo también ha sido descripto como un factor clínico asociado al desarrollo de CDT. Cardona² señala que el sexo femenino tiene mayor probabilidad de desarrollar CDT que el varón, aunque cuando aparece en el sexo masculino es más probable que sea de más difícil erradicación y con mayores recidivas¹⁵. En nuestra serie no encontramos significancia estadística en la relación entre sexo y CDT; la mayoría de los pacientes con CDT fueron mujeres, en una relación de 3,75:1, sin asociación estadísticamente significativa.

Con respecto a la edad, tampoco se presentó como factor de riesgo; se analizaron cortes de 60 y 40 años, sin hallar significancia en ninguno. El CDT puede desarrollarse a cualquier edad, pero se tomaron estos cortes por los riesgos de recidiva descriptos en pacientes mayores de 40 años⁷.

Otro factor estudiado fueron los antecedentes hereditarios; en nuestra serie tampoco se identificó riesgo significativo. Según Orois, la mayor parte de los pacientes con cáncer de tiroides tienen subtipos histológicos no medulares, y alrededor del 9 % tiene historial familiar²⁴.

CONCLUSIONES

En nuestro estudio descubrimos cinco elementos clínicos que se relacionan de forma significativa con la aparición de CDT en pacientes que padecen bocio durante un largo período. La adherencia a planos profundos o la consistencia dura se desarrolló como el predictor más sólido entre estos. En ambientes con recursos escasos, donde el examen físico continúa siendo una herramienta de diagnóstico crucial, este descubrimiento cobra particular importancia. La detección de este signo en el transcurso de la evaluación clínica debería tomarse como una señal de alerta que amerite una remisión quirúrgica inmediata, aun cuando no existan otros estudios adicionales.

El crecimiento “rápido” que los pacientes mencionaron durante el último año anterior a la consulta mostró una asociación significativa. Este dato indica que las modificaciones en el patrón de crecimiento de un bocio preexistente deben interpretarse como una señal de potencial malignidad, sobre todo si suceden en glándulas que han padecido enfermedades durante mucho tiempo.

La procedencia urbana o migración prolongada a zonas urbanas (>10 años) mostró una asociación significativa. Este descubrimiento da pie a líneas de investigación acerca de los hipotéticos factores ambientales urbanos que podrían estar favoreciendo la carcinogénesis en la tiroides en nuestro ambiente, tales como la exposición a radiaciones no ionizantes, la contaminación ambiental o las alteraciones en los hábitos alimentarios.

La importancia de los sistemas de clasificación TI-RADS y Bethesda en Bolivia se ve reforzada por nuestros hallazgos, que demuestran que su uso sistemático podría contribuir a detectar CDT de manera más temprana. Al combinar las clasificaciones con los resultados clínicos (crecimiento y consistencia), se obtiene un algoritmo de diagnóstico que es accesible incluso para centros con recursos tecnológicos escasos.

Los datos que hemos encontrado se enmarcan en las particularidades de Bolivia. Siendo un país con un historial de bocio endémico y, teniendo en cuenta la singularidad geoquímica de sus fuentes de agua y sal (elevada concentración de calcio y litio), es necesario interpretar nuestros hallazgos dentro de este marco específico. El hecho de que haya persistencia de bocio en áreas rurales, a pesar de los programas de yodación, indica que los factores bociogénicos en Bolivia podrían ser múltiples y requerir un estudio más detallado.

Conflictos de intereses: el autor declara no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: el autor declara que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

- Sitges-Serra A. Estado actual: el carcinoma papilar de tiroides después de Mazzaferri. En: Gómez-Ramírez J, editor. Controversias en el manejo del carcinoma papilar de tiroides. Madrid: Arán; 2018. p. 11-2. (Monografías de la AEC).
- Cardona A. Estudio médico psicológico del bocio en una población del área rural. *Gac Med Bolivia*. 1989;13(1):1-8.
- Sernaqué-Mechato R, Pavia-Puga S. Bocio multinodular intratorácico. *Rev Soc Peru Med Interna*. 2024;37(1):43-47. <https://doi.org/10.36393/spmiv371.837>.
- Virtual elimination of iodine deficiency disorders in Bolivia. *Rev Panam Salud Publica*. 1997;2(3):215-219.
- Urquieta Maldonado MA, Urquieta Márquez MA, Gutiérrez Cruz D, et al. Perfil bociógeno en estudiantes de Camata Marka, mediante palpación semiológica y respaldo ultrasonográfico. Un estudio de corte transversal. *Gestión 2014. Rev Med La Paz*. 2015;21(2):32-37.
- Calé J, Lutfi RJ. El litio y su influencia en la función tiroidea. *Rev Argent Endocrinol Metab*. 2009;46(3):37-42.
- Daza A. Cáncer diferenciado de tiroides atención en un hospital de tercer nivel de Bolivia. *Gac Med Bol*. 2024;47(2):40-48. <https://doi.org/10.47993/gmb.v47i2.900>.
- Amin MB, Edge SB, Greene FL, et al., eds. *AJCC cancer staging manual*. 8th ed. Cham: Springer; 2017.
- Orellana-Suárez KD, Alcívar-Quiroz BM, Salazar-Figueroa GV. Factores de riesgo y manifestaciones clínicas en cáncer de tiroides en adultos en Latinoamérica. *MQRInvestigar*. 2023;7(3):868-884. <https://doi.org/10.56048/MQR20225.7.3.2023.868-884>.
- Ali MA, ElMakki Ahmed M. Risk of malignancy in long-standing goiters: a retrospective study at Khartoum Teaching Hospital, Sudan. *Cureus*. 2025;17(6):e86931. <https://doi.org/10.7759/cureus.86931>.
- Pitoia F, Califano I, Vázquez A, et al. Consenso intersocietario sobre tratamiento y seguimiento de pacientes con cáncer diferenciado de tiroides. *Rev Argent Endocrinol Metab*. 2014;51(2):85-118.
- Burman KD, Wartofsky L. Clinical practice. Thyroid nodules. *N Engl J Med*. 2015;373(24):2347-2356. <https://doi.org/10.1056/NEJMc1415786>.
- Kim M, Jeon MJ, Han M, et al. Tumor growth rate does not predict malignancy in surgically resected thyroid nodules classified as Bethesda category III with architectural atypia. *Thyroid*. 2019;29(2):216-221. <https://doi.org/10.1089/thy.2018.0366>.
- Smith JJ, Chen X, Schneider DF, et al. Cancer after thyroidectomy: a multi-institutional experience with 1,523 patients. *J Am Coll Surg*. 2013;216(4):571-577. <https://doi.org/10.1016/j.jamcollsurg.2012.12.022>.
- American Cancer Society. Thyroid cancer risk factors [Internet]. Atlanta: American Cancer Society; 2019 [citado 2025 jul 10]. Available from: <https://www.cancer.org/cancer/types/thyroid-cancer/causes-risks-prevention/risk-factors.html>.
- Lan Z, Yang F, Zhang J, et al. Diagnostic value and challenges of BRAF V600E molecular testing and thyroid fine-needle aspiration cytology: a retrospective study from a tertiary institution in Southern Hunan Province, China. *Acta Cytol*. 2023;67(6):629-638. <https://doi.org/10.1159/000534138>.
- Asanuma K, Kobayashi S, Shingu K, et al. The rate of tumour growth does not distinguish between malignant and benign thyroid nodules. *Eur J Surg*. 2001;167(2):102-105. <https://doi.org/10.1080/110241501750070538>.
- Mota Gamboa JD, García Tamayo J. Carcinoma indiferenciado de la glándula tiroides: estudio histológico e inmunohistoquímico en cinco casos. *Rev Fac Med*. 2002;25(1):115-120.
- Huston-Paterson HH, Mao Y, Tseng CH, et al. Rural-urban disparities in the continuum of thyroid cancer care: analysis of 92,794 cases. *Thyroid*. 2024;34(5):635-645. <https://doi.org/10.1089/thy.2023.0357>.
- Huang EYF, Kao NH, Lin SY, et al. Concordance of the ACR TI-RADS classification with Bethesda scoring and histopathology Risk stratification of thyroid nodules. *JAMA Netw Open*. 2023;6(9):e2331612. <https://doi.org/10.1001/jamanetworkopen.2023.31612>.
- Linhares SM, Handelsman R, Picado O, et al. Fine needle aspiration and the Bethesda system: correlation with histopathology in 1,228 surgical patients. *Surgery*. 2021;170(5):1364-8. DOI: <https://doi.org/10.1016/j.surg.2021.05.016>.
- Apodaca Cruz A, Vázquez Islas G, Torres Jiménez J, et al. Urgencias metabólicas. En: Herrera-Gómez A, Namendys-Silva SA, Meneses-García A, editores. *Manual de oncología: procedimientos médico-quirúrgicos*. 6a ed. México: McGraw-Hill Education; 2017. p. 152-159.
- Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2016;26(1):1-133. <https://doi.org/10.1089/thy.2015.0020>.
- Orois A, Mora M, Halperin I, et al. Carcinoma diferenciado de tiroides familiar: más allá de las formas sindrómicas. *Endocrinol Diabetes Nutr*. 2021;68(4):260-269. <https://doi.org/10.1016/j.endinu.2020.08.002>.

Cirugía estapediovestibular en pacientes con otosclerosis de 65 años o más: resultados auditivos posoperatorios. Experiencia de 20 años

Micaela Redivo¹, Federica Fernández Long² y Gabriela Pérez Raffo³

Servicio de Otorrinolaringología, Hospital Italiano. Buenos Aires, Argentina

RESUMEN

Introducción: la otosclerosis es una osteodistrofia de la cápsula ótica que afecta el oído medio e interno, con una prevalencia aproximada del 1% en la población general, predominio femenino y presentación habitual entre los 30 y 40 años. Se manifiesta principalmente como hipoacusia conductiva o mixta. El tratamiento quirúrgico es de elección en pacientes con adecuada reserva coclear, con tasas de éxito, definidas como cierre del gap aéreo-óseo (ABG) y variación de conducción ósea (BC) ≤ 10 dB, que oscilan entre 67 y 94%. Sin embargo, al tratarse de una enfermedad progresiva, puede persistir o progresar la hipoacusia neurosensorial, requiriendo rehabilitación auditiva posterior.

Objetivo: evaluar los resultados auditivos en pacientes ≥ 65 años sometidos a cirugía estapedial.

Materiales y métodos: estudio observacional, retrospectivo y descriptivo. Se incluyeron 63 pacientes (72 oídos) con diagnóstico de otosclerosis, intervenidos entre enero de 2004 y diciembre de 2024. Se analizaron variables demográficas y parámetros audiológicos preoperatorios y posoperatorios, incluyendo ABG, conducción ósea (BC) y promedio tonal puro (PTA) en frecuencias de 0,5, 1, 2 y 4 kHz.

Resultados: el 58,7% de los pacientes fueron mujeres (relación 1,4:1). El 98,6% presentó hipoacusia mixta. La afección fue bilateral en el 52,3% de los casos. Se logró un cierre de ABG ≤ 10 dB en el 85,4% de los oídos y una variación de BC ≤ 10 dB en el 85%. El PTA mostró una mejoría promedio de -6,25 dB en vía aérea y -6,25 dB en vía ósea. El seguimiento medio fue de 7,9 años. El 22% de los pacientes requirió audífonos en el posoperatorio. Se registraron complicaciones menores en el 9,52% de los casos.

Conclusión: La cirugía estapedial en pacientes mayores de 65 años es un procedimiento eficaz y seguro, con altas tasas de éxito audiológico y baja incidencia de complicaciones, constituyendo una opción terapéutica válida en este grupo etario.

Palabras clave: otosclerosis, estapedotomía, cirugía del estribo, pérdida auditiva funcional, hipoacusia, resultados auditivos.

Postoperative Hearing Outcomes of Stapedio-Vestibular Surgery in Patients with Otosclerosis Aged 65 Years or Older: A 20-Year Experience

ABSTRACT

Introduction: otosclerosis is an osteodystrophy of the otic capsule that affects the middle and inner ear, with an approximate prevalence of 1% in the general population, a female predominance, and usual onset between 30 and 40 years of age. It primarily manifests as conductive or mixed hearing loss.

Autora para correspondencia: federica.fernandezlong@hospitalitaliano.org.ar, Fernández Long F.

Recibido: 1/12/2025 | Aceptado: 10/08/2026 | Publicado: 14/09/2026

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1326>

Cómo citar: Redivo M, Fernández Long F, Pérez Raffo G. Cirugía estapediovestibular en pacientes con otosclerosis de 65 años o más: resultados auditivos posoperatorios. Experiencia de 20 años. *Rev Hosp Ital B. Aires.* 2026;46(3):e0001326

Surgical treatment is the treatment of choice in patients with adequate cochlear reserve, with success rates, defined as an air-bone gap (ABG) closure and bone conduction (BC) change ≤ 10 dB, ranging from 67% to 94%. However, as the disease is progressive, sensorineural hearing loss may persist or progress, requiring subsequent auditory rehabilitation.

Objective: to evaluate hearing outcomes in patients aged ≥ 65 years who underwent stapes surgery.

Materials and methods: observational, retrospective, descriptive study. Sixty-three patients (72 ears) diagnosed with otosclerosis who underwent surgery between January 2004 and December 2024 were included. Demographic variables and pre- and postoperative audiological parameters were analyzed, including ABG, bone conduction (BC), and pure-tone average (PTA) at frequencies of 0.5, 1, 2, and 4 kHz.

Results: of the patients, 58.7% were women (female-to-male ratio, 1.4:1). Mixed hearing loss was present in 98.6% of patients. Bilateral involvement was observed in 52.3% of cases. ABG closure ≤ 10 dB was achieved in 85.4% of ears, and a BC change ≤ 10 dB was achieved in 85%. The PTA showed a mean improvement of -6.25 dB for air conduction and -6.25 dB for bone conduction. Mean follow-up was 7.9 years. Postoperatively, 22% of patients required hearing aids. Minor complications occurred in 9.52% of cases.

Conclusion: stapes surgery in patients aged over 65 years is an effective and safe procedure, with high rates of audiological success and a low incidence of complications, constituting a valid therapeutic option for this age group.

Keywords: stapedotomy, stapes surgery, functional hearing loss, hearing results, hearing loss.

INTRODUCCIÓN

La otosclerosis es una osteodistrofia de la cápsula ótica caracterizada por una fase inicial de reabsorción ósea seguida de neoformación¹. Históricamente, fue descrita por Valsalva en 1735; Toynbee identificó la anquilosis del estribo como causa de sordera en 1860; von Troeltsch introdujo el término “esclerosis” en 1881, Politzer acuñó el término “otosclerosis” en 1894 y Siebenmann propuso “otospongiosis” para la fase activa de la enfermedad en 1912²⁻³.

Los focos suelen originarse en la fisura *ante fenestram* o comprometer la cóclea, determinando formas fenestral y retrofenestral⁴. Su etiopatogenia es multifactorial e incierta, con una prevalencia clínica cercana al 1%, predominio femenino y herencia autosómica dominante de penetrancia incompleta^{1,4,5}. El diagnóstico es clínico-audiológico; la enfermedad se presenta principalmente como hipoacusia conductiva con gap aéreo-óseo significativo; puede asociarse a un componente neurosensorial en estadios avanzados⁶⁻⁸. La clasificación tomográfica de Symons y Fanning permite valorar la extensión de la enfermedad⁹⁻¹⁰.

El tratamiento es individualizado. Si bien existen alternativas médicas para estabilizar la progresión, la cirugía estapedial constituye el tratamiento de elección en pacientes con hipoacusia conductiva significativa y adecuada reserva coclear, mientras que la rehabilitación auditiva se reserva para casos seleccionados^{1,10,11}.

Aunque la cirugía estapedial ha demostrado excelentes resultados audiológicos en la población general, la evidencia sobre sus resultados en pacientes adultos

mayores continúa siendo limitada, especialmente con seguimiento prolongado. En este contexto, el presente estudio evalúa los resultados auditivos posoperatorios en pacientes de 65 años o más sometidos a cirugía estapedial en nuestro centro.

Objetivo principal

Evaluar los resultados auditivos a largo plazo en pacientes con otosclerosis de 65 años o más que fueron sometidos a cirugía estapedial en el Hospital Italiano de Buenos Aires entre enero de 2004 y diciembre de 2024.

Objetivos secundarios

1. Conocer la prevalencia hospitalaria de pacientes ≥ 65 años que hayan sido sometidos a cirugía estapedial en algún momento de su vida, en el período de enero de 2004 a diciembre de 2024.

2. Describir las características demográficas y establecer la forma de presentación clínica mediante audiometría, así como el grado de otosclerosis según la clasificación de Symmons y Fanning en tomografía de hueso temporal.

3. Establecer los resultados audiológicos posoperatorios, en relación con la mejoría del umbral óseo (BC) y la reducción del gap aéreo-óseo (ABG).

4. Conocer el tiempo medio de seguimiento posoperatorio y caracterizar al grupo de pacientes que requirió uso de dispositivos de amplificación tras la cirugía.

MATERIALES Y MÉTODOS

Se realizó un estudio de tipo observacional, retrospectivo, descriptivo.

Se realizó revisión de historia clínica electrónica de pacientes mayores de 65 años, que fueron sometidos a cirugía estapedial por otosclerosis en el Hospital Italiano entre los años enero 2004-diciembre 2024.

Se confeccionó fuente de datos en planilla de Excel para estudio de variables: sexo, edad actual y edad de cirugía, presencia de enfermedad de otosclerosis unilateral o bilateral y tipo de patrón de hipoacusia establecida en la paciente constatada por audiometría preoperatoria (mixta, conductiva, neurosensorial), tipo de cirugía realizada, conservación de cuerda de tímpano. Se evaluó según resultados de audiometría preoperatoria y posoperatoria, individualmente para las frecuencias 500-1000-2000-4000 Hz, el valor del BC (conducción ósea), CA (conducción aérea) y el ABG (diferencia osteo-aérea). Se calculó también el PTA (*Pure Tone Average*) como el promedio de los umbrales auditivos en determinadas frecuencias, usado para resumir el nivel de audición de un paciente. Para ello se utilizó la siguiente fórmula recomendada por la American Academy of Otolaryngology–Head and Neck Surgery (AAO-HNS) para evaluar otosclerosis:

$$PTA = (500\text{Hz} + 1000\text{Hz} + 2000\text{Hz} + 4000\text{Hz}) / 4^{12,13}$$

Se indicó el grado de otosclerosis establecido por clasificación de Symons y Fanning en tomografía *multislice* de huesos temporales. Además, se identificaron secundariamente, las complicaciones posoperatorias, así como el tiempo de seguimiento y el requerimiento del uso de dispositivos de amplificación tras la cirugía.

Los resultados audiológicos posoperatorios se evaluaron mediante dos desenlaces independientes: el cierre del gap aéreo-óseo (ABG) a ≤ 10 dB HL y (2) la preservación

de la conducción ósea (BC), definida como una variación posoperatoria ≤ 10 dB HL respecto de la audiometría preoperatoria para cada frecuencia analizada. Ambos parámetros se analizaron de forma independiente.

El siguiente trabajo cuenta con aprobación del protocolo de Investigación por parte del Comité de Ética y Protocolos de investigación del Hospital Italiano con N.º 7148. También, se encuentra inscripto en la Plataforma de Registro Informatizado de Investigaciones en Salud de Buenos Aires (PRIISA) con N.º 13372.

Población

Pacientes con diagnóstico clínico, audiológico y/o tomográfico de otosclerosis que fueron sometidos a cirugía estapedial en el Hospital Italiano entre enero de 2004 y diciembre de 2024, y que tenían 65 años o más al 1 de enero de 2025.

Criterios de inclusión

Pacientes con diagnóstico de otosclerosis atendidos en el Hospital Italiano, que contaban con procedimientos de cirugía estapedial realizados entre enero de 2004 y diciembre de 2024 en el centro. Además, se consideró que contaran con disponibilidad de estudios de audiometría y logaudiometría en historia clínica electrónica y seguimiento posoperatorio mayor de 1 año.

Criterios de exclusión

Fueron excluidos aquellos pacientes que a la fecha 1 de enero de 2025 fueran menores de 65 años. También, aquellos casos que tuvieran ausencia de estudios complementarios (audiometría y logaudiometría) en la historia clínica y seguimiento posoperatorio insuficiente menor de 1 año (Fig. 1).

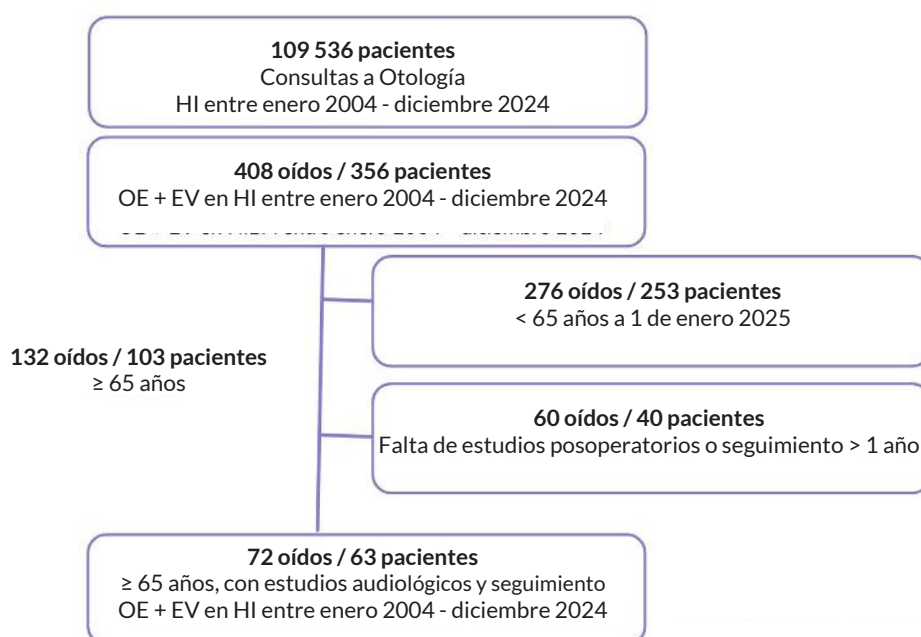


Figura 1. Población en estudio. Criterios de inclusión y exclusión. OE: Otosclerosis. EV: cirugía estapediovestibular. HI: Hospital Italiano (Buenos Aires, Argentina)

RESULTADOS

Se registraron entre enero de 2004 y diciembre de 2024, en el Sector de Otolología del Hospital Italiano, 109 536 consultas. La prevalencia hospitalaria de otoesclerosis en pacientes sometidos a cirugía estapedio-vestibular fue de 0,33% (365 casos/109 536 pacientes atendidos). En tanto que, la proporción de cirugías estapediovestibulares (EV) realizadas a pacientes ≥ 65 años fue de 28,8% (103 casos/365 pacientes con cirugía EV) en el mismo centro y periodo.

En este estudio retrospectivo, descriptivo y observacional, se incluyeron 63 pacientes (72 oídos), mayores de 65 años, operados de cirugía estapedial por otoesclerosis en nuestro centro entre enero 2004-diciembre 2024. Fueron intervenidos 26 hombres y 37 mujeres. En la muestra, las mujeres constituyeron el 58,7% del total, con una razón mujer: hombre de 1,4:1.

De ellos, 41 pacientes (65,1%) fueron operados a edad menor de 65 años, con una media de edad de 58,7 (47-64). En tanto que 22 pacientes (34,9%) fueron operados a edad ≥ 65 años, con un promedio de edad de 68,4 años (65-89).

En 55 (88%) de los casos, se realizó estapedotomía y en 17 (23,6%) estapedectomía. Durante la cirugía, en 31 (58,3%) de los pacientes se conservó la cuerda del tímpano. No se identificaron diferencias entre las prótesis utilizadas.

La selección de candidatos quirúrgicos se realizó sobre la base de los pacientes con sospecha de otoesclerosis por valoración clínica, estudios audiométricos y tomografía computarizada de huesos temporales cuando fue posible. En ellos, 71 oídos (98,6%) presentaron hipoa-cusia mixta y 1 (1,4%) conductiva.

De los 63 pacientes, 33 (52,3%) presentaron otoesclerosis bilateral y 30 (47,6%) unilateral. De estos últimos, 13 fueron intervenidos del oído derecho y 17 del oído izquierdo. De los pacientes que presentaron enfermedad bilateral, solamente 9 fueron intervenidos de ambos oídos, en tanto que el resto fue operado solamente de uno de sus dos oídos.

Se encontró que 46 oídos intervenidos (63,8%) contaban con tomografía computarizada de huesos temporales registrada en la historia clínica. La clasificación de Symons y Fanning se realizó considerando los hallazgos de ambos huesos temporales (ambos oídos) de cada paciente. De estos, 6 pacientes (8,3%) corresponden al grado 0, 34 (47,2%) al grado I, 3 (4,2%) al grado II y 3 (4,2%) al grado III. En 26 oídos intervenidos (36,1%) no se encontraron estudios tomográficos preoperatorios registrados en la historia clínica.

Los resultados audiológicos posoperatorios se analizaron mediante dos parámetros independientes: el cierre del ABG ≤ 10 dB HL y la preservación de la conducción ósea, definida como una variación ≤ 10 dB HL respecto de la audiometría preoperatoria.

La media de valor de gap preoperatorio en la audiometría para la frecuencia 500 Hz fue de 40 [20-70 dB] logrando un cierre del ABG ≤ 10 dB HL en 59,7% (43 cirugías). Para la frecuencia 1000 Hz, la media de gap preoperatorio fue de 35 dB [10-65 dB] y el cierre del ABG se produjo en el 87,5% (63 cirugías). En la frecuencia 2000 Hz el valor del gap precirugía fue de 20 dB [0-50 dB] y el cierre del ABG fue total en el 98,6% (71 cirugías).

En tanto que para la frecuencia 4000 Hz la media preoperatoria fue de 15 dB [0- 40 dB] y el cierre completo se logró en 95, 8% (69 cirugías).

El promedio de ABG preoperatorio para las cuatro frecuencias estudiadas tuvo una media de 26,25 dB y para el posoperatorio de 5 dB. Entretanto, el promedio de cierre de ABG ≤ 10 (dB) para las mismas frecuencias fue de 85,4% (Tablas 1 y 2).

Por otro lado, las vías óseas (VO) preoperatorias en la audiometría se encontraron en promedio en 15 dB [0-40 dB] para la frecuencia 500 Hz, 30 dB [0-50 dB] en la 024 Hz; 40 dB [15-70 dB] en la 2000 Hz y 50 dB [15-80 dB] en la 4000 Hz. Se evidenció BC con variación ≤ 10 dB HL en todas las frecuencias, siendo del 63,80% (46 cirugías) en la frecuencia 500 Hz, 87,5% (63 cirugías) en la 1000 Hz y 94,40% (68 cirugías) en las frecuencias 2000 y 4000 Hz. En resumen, el promedio de cierre de BC ≤ 10 (dB) fue de 85% considerando todas las frecuencias (Fig. 2, Tabla 3).

Se calculó el valor del PTA (*Pure Tone Average*) según la American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS)^{13,14} comparando los resultados preoperatorios y posoperatorios. Así se encontró un valor de PTA para la VO preoperatoria con una media de: 33,75 dB, y para la posoperatoria de 27,5 dB. El Delta valor para los PTA de los pacientes, es decir la diferencia

Tabla 1. Valores de ABG previo a la cirugía y posterior a ella discriminados por frecuencia

Frecuencia (Hz)	ABG preoperatorio	ABG posoperatorio
500 Hz	40 dB [20-70 dB]	10 dB [0-40 dB]
1000 Hz	30 dB [10-65 dB]	10 dB [0-40 dB]
2000 Hz	20 dB [0-50 dB]	5 dB [0-20 dB]
4000 Hz	15 dB [0-40 dB]	0 dB [0-40 dB]
Promedio de ABG	26,25 dB	5 dB

Hz: Hertz. dB: decibeles

Tabla 2. Valoración del promedio de ABG "Air-Bone Gap" (diferencia entre VA: vía aérea y VO: vía ósea) ≤ 10 (dB) discriminada por frecuencia (Hz)

Frecuencia (Hz)	ABG ≤ 10 (dB)
500 Hz	59,7% (43 cirugías)
1000 Hz	87,5% (63 cirugías)
2000 Hz	98,6% (71 cirugías)
4000 Hz	95,8% (69 cirugías)
Promedio de ABG ≤ 10 (dB)	85,40%

Hz: Hertz. dB: decibeles

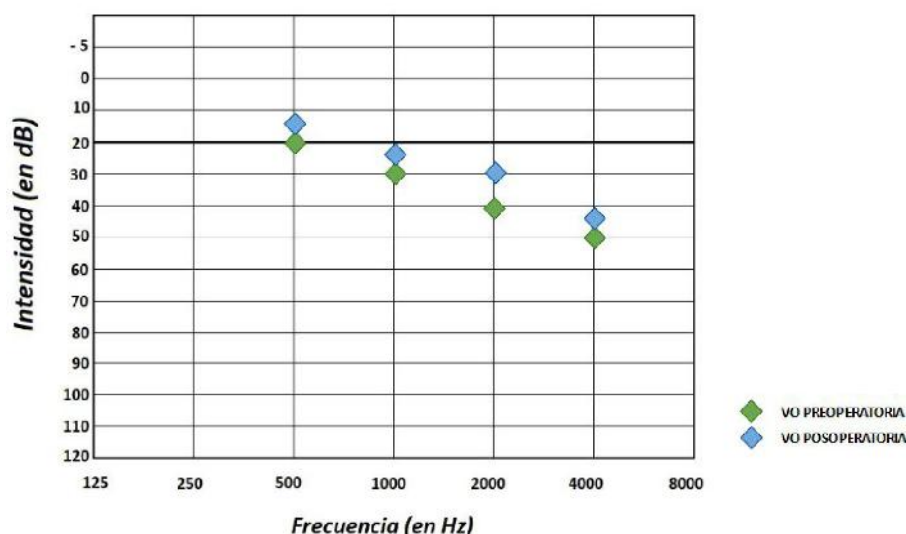


Figura 2. Valoración del promedio ponderado de las VO preoperatorias y posoperatorias. (Hz: Hertz. dB: decibeles).

Tabla 3. Valoración del promedio de BC "Bone Conduction" ≤ 10 (dB HL) discriminada por frecuencia.

Frecuencia (Hz)	BC ≤ 10 (dB)
500 Hz	63,8% (46 cirugías)
1000 Hz	87,5% (63 cirugías)
2000 Hz	94,4% (68 cirugías)
4000 Hz	94,4% (68 cirugías)
Promedio de BC ≤ 10 (dB)	85%

Hz: Hertz. dB: decibeles

del PTA posoperatorio y preoperatorio tuvo una media de: $-6,25$ dB (mejora promedio).

Se valoraron las vías aéreas (VA) preoperatorias discriminadas por frecuencias y se encontró que, la media de PTA preoperatorio para las vías aéreas fue de $62,5$ dB [60-70 dB]. En tanto que, para el posoperatorio, el PTA arrojó una media de vía aérez $36,25$ dB [25-50 dB]. La diferencia entre los valores de PTA posoperatorio y preoperatorio (Δ PTA) fue de $-6,25$ dB para las vías óseas, lo que representa una mejoría considerable con cierre de ABG.

El seguimiento medio de los pacientes posterior a la cirugía fue de 7,9 años. Posteriormente a ella, 14 pacientes (22%) requirieron audífonos a una edad media de 71,7 años. La media entre la cirugía y la utilización de los audífonos fue de 3,9 años. Todos ellos presentaron luego de la intervención BC ≤ 10 dB y ABG ≤ 10 dB. El PTA de VO en el oído contralateral a la cirugía fue de $32,5$ dB. El 64,8% (9 pacientes) fueron intervenidos a edad ≥ 65 años. Se encontró que el 93% (13 pacientes) requirieron audífonos en ambos oídos, en tanto que solo uno de ellos se equipó de forma unilateral en el oído operado. No obstante, se debe considerar que 5 pacientes requirieron audífonos en el oído contralateral sin otosclerosis, por presentar patrón audiométrico de presbiacusia.

Además, 2 pacientes con otosclerosis bilateral se encontraron equipados de ambos oídos antes de la cirugía, y solo 1 paciente se equipó antes de la cirugía del oído contralateral por presentar presbiacusia sin signos de osteodistrofia en el hueso temporal.

Se encontró en el análisis de este subgrupo que 2 de los 3 pacientes de la serie que presentaron grado III de osteodistrofia evidenciado por clasificación de Symons y Fanning requirieron audífonos. No obstante, no se evidenció una correlación lineal entre los grados imagenológicos más graves de otosclerosis, el rendimiento auditivo posoperatorio o la necesidad de utilización de OA.

Se evidenció un 9,52% (6 pacientes) de complicaciones posoperatorias. Se presentó un caso de perforación timpánica residual que requirió timpanoplastia posterior, así como un caso de granuloma que se controló por consultorio. Otro caso presentó extrusión de prótesis, que se intervino colocando una prótesis de mayor longitud (4,75 mm). Además, 2 pacientes experimentaron reagudización de su acúfeno previo, que resolvió bajo tratamiento médico y psicoterapia. Otro paciente presentó síndrome de Ramsay Hunt al mes de la intervención, sin secuelas auditivas.

DISCUSIÓN

La otosclerosis presenta una prevalencia clínica estimada de 0,4% en poblaciones de ascendencia europea, mientras que los estudios histopatológicos evidencian hasta un 2,5% de compromiso en huesos temporales¹⁵. La proporción de pacientes sometidos a cirugía estapedial varía entre 35-54% en estudios poblacionales, mientras que en series quirúrgicas los mayores de 65 años representan entre 3-16% de los casos¹⁶. En nuestra serie, este grupo correspondió al 28,9% de las cirugías realizadas.

La enfermedad presenta predominio femenino, con una relación aproximada de 2:1 según series previas^{7,15}. En nuestra muestra, las mujeres representaron el 58,7%, con una relación mujer: hombre de 1,4:1. La presentación

bilateral es frecuente¹⁷; en esta serie fue del 58,3% y la unilateral del 41,6%.

En estudios previos, la hipoacusia preoperatoria puede presentarse como conductiva o mixta, con variaciones según la población analizada^{17,18}. En nuestra cohorte, el 98,6% presentó hipoacusia mixta y el 1,4% conductiva.

La clasificación tomográfica de Symons y Fanning presenta una alta concordancia interobservador e intraobservador ($\kappa \approx 0,92$), siendo útil para valorar la extensión de la enfermedad antes de la cirugía²⁰. En nuestra serie, el 63,8% contó con tomografía de huesos temporales; sin embargo, no se observó correlación entre grados tomográficos más avanzados, resultados auditivos posoperatorios y necesidad de audífonos.

Sin embargo, a pesar de presentar una alta tasa de mejoría auditiva, los parámetros audiológicos evaluados en este estudio, es decir, el cierre de la ABG ≤ 10 dB HL, así como una BC sin cambios, varían entre diferentes estudios en la literatura. Considerando estos parámetros, en el estudio realizado por Salvador, el 67,2% de los pacientes presentó una cirugía EV exitosa¹⁶. Estos resultados son coherentes con otros autores que informan una tasa de cierre del ABG dentro de 10 dB en un rango del 62-94%^{21,22}.

Vincent y cols. evaluaron 93 casos de cirugía EV y encontraron que el 59% de los pacientes lograron una mejora del ABG posoperatorio ≤ 10 dB, y el 57% una ganancia de BC ≥ 20 dB. Además, identificaron como factor pronóstico positivo un ABG más grande²³. En el presente estudio, se evidenció un cierre del ABG ≤ 10 dB de 85,4%, para un promedio preoperatorio de este de 26,3 dB. En tanto, la variación de BC ≤ 10 dB HL se produjo en el 85% de los casos en las mismas frecuencias.

Carhart describió la disminución de la conducción ósea a 2000 Hz (muesca de Carhart) como un hallazgo característico de la otosclerosis, atribuible al aumento de la impedancia en la ventana oval. Tras la cirugía estapedial, la mejoría de la conducción ósea, especialmente de la muesca de Carhart, se considera un indicador de buen rendimiento auditivo posoperatorio²⁴. En nuestra casuística, encontramos que el promedio ponderado de las BC preoperatorias para la frecuencia 2000 Hz fue de 40 dB y 30 dB en el posoperatorio, evidenciando una variación del BC ≤ 10 dB en 94,4% (68 cirugías), y un cierre en la muesca de Carhart en concordancia con la bibliografía.

Recientes estudios indican mejoras significativas en los umbrales de BC y el cierre de ABG en todos los grupos de edad ($p < 0,05$). Los umbrales de BC mejoraron significativamente en pacientes < 50 años. La regresión lineal reveló una modesta correlación negativa entre la edad y la mejora de BC ($R^2 = 0,18$, $p = 0,006$). A pesar de las ganancias menores en pacientes de mayor edad, no se encontraron diferencias significativas entre los grupos más jóvenes y mayores en los cambios en el ABG y BC ($p > 0,05$). Los resultados mostraron que la cirugía EV es recomendable incluso en personas mayores con discapacidad auditiva mixta, siempre que el ABG se limite a 20-30 dB²⁵. En esta revisión, se encontró que, para pacientes operados mayores de 65 años, el Delta valor para el PTA por VO, fue de -6,25 dB, mejorando el promedio. Esto indica una

mejora considerable en la audición comparable, según la bibliografía, con pacientes más jóvenes.

En estudios clínicos retrospectivos, la cirugía EV en pacientes mayores de 65 años, ha demostrado ser tan segura y eficaz como en pacientes más jóvenes. En un análisis de 42 pacientes (> 60 años) 36 comparados con 275 pacientes menores de 60, los resultados audiológicos y la incidencia de complicaciones fueron equivalentes entre ambos grupos²⁶. En otra serie con 16 pacientes mayores de 65, entre 473 cirugías primarias, todos lograron reducción del ABG a < 20 dB. El 87% cerró a < 10 dB tras un mes y el 82% lo mantuvo durante un seguimiento medio de 20 meses²⁷. Estos datos respaldan que la cirugía estapedial en adultos mayores es una opción válida, con resultados clínicos comparables en eficacia y seguridad.

En relación con el tratamiento, en casos seleccionados, los audífonos por sí solos podrían servir como una opción de tratamiento no invasivo para la otosclerosis. Sin embargo, a menudo no ofrece suficientes beneficios en términos de discriminación del habla. Por lo tanto, el tratamiento combinado de la cirugía y el audífono para pacientes con umbrales de VO posoperatorios > 30 dB HL ha sido reconocido, como el tratamiento más adecuado para pacientes con pérdida auditiva mixta²⁸.

En el estudio realizado por Wardenga²⁷, se observó que en 44,9% de los pacientes no necesitaban un audífono después de la cirugía. En tanto que un audífono de potencia cumpliría criterios audiológicos en las 4 frecuencias en el 81,6% de los casos que necesitaban un audífono luego de la cirugía²⁸. En estos pacientes, cuando fue necesario se pudo indicar un audífono de menor potencia, y esto fue considerado exitoso²⁵. En comparación con nuestra casuística, el 22% (14 pacientes) debió utilizar audífonos posteriores a la cirugía en una media de 3,9 años. Todos ellos presentaron luego de la intervención un BC y ABG ≤ 10 .

En un estudio realizado por Ishai y cols., se indica que la pérdida auditiva neurosensorial promedio a largo plazo debida a la otosclerosis fue estadísticamente mayor que la observada solo por la edad en cada frecuencia. Aproximadamente un tercio de los pacientes con otosclerosis mostraron una progresión clínicamente significativa del componente neurosensorial de la pérdida auditiva, con un BC promedio superior a los cambios esperados relacionados con la edad, que oscilaron entre 10 y 15 dB para las frecuencias evaluadas en este subgrupo⁸. Los pacientes requirieron audífonos con un valor de PTA de VO promedio en el oído contralateral a la cirugía de 32,5 dB.

En series recientes, las complicaciones graves son infrecuentes, con hipoacusia neurosensorial profunda y grave de 0,2% y 3,2%, respectivamente, y una tasa de revisión del 4,5%. El vértigo posoperatorio ocurre en el 18% (persistente en 6%), mientras que *tinnitus* (acúfeno) y alteración del gusto se presentan en 3,6% y 4,7%; complicaciones menos frecuentes, como paresia facial o fistula perilinfática, rondan el 1%²⁹.

En nuestra serie, la tasa de complicaciones fue del 9,52% (6 pacientes), incluyendo perforación timpánica residual, granuloma, extrusión protésica con reintervención, reagudización del acúfeno y un síndrome de Ramsay Hunt. No obstante, todos mantuvieron resultados audiológicos favorables.

Limitaciones

Este estudio presenta algunas limitaciones. Su diseño retrospectivo y descriptivo puede introducir sesgos inherentes a la revisión de historias clínicas y limitar el control sobre variables potencialmente confusoras. Asimismo, la inclusión exclusiva de pacientes seleccionados para cirugía puede generar sesgo de selección. Finalmente, la ausencia de un grupo comparador de pacientes más jóvenes impide establecer comparaciones directas entre grupos etarios, por lo que los resultados deben interpretarse en el contexto de una serie descriptiva.

CONCLUSIÓN

La cirugía estapedial en pacientes mayores de 65 años se asoció con una mejoría de los resultados audiológicos, evidenciada por el cierre del gap aéreo-óseo y la mejoría de la muesca de Carhart, incluyendo la frecuencia de 2000 Hz. Asimismo, la necesidad de audífonos en el seguimiento fue baja, ya que solo el 22% de los pacientes requirió amplificación posterior a la cirugía. Estos hallazgos respaldan a la cirugía estapedial como una opción terapéutica para pacientes mayores de 65 años con indicación quirúrgica.

Contribuciones de las autoras: Conceptualización, Metodología, Investigación, Visualización, Administración de proyecto (FFL, MR, GPR). Redacción-Borrador original (FFL, MR). Redacción-Revisión y edición (FFL, MR, GPR). Supervisión (GPR).

Conflictos de intereses: las autoras declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo

Financiamiento: las autoras declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

- Silva VAR, Pauna HF, Lavinsky J, et al. Brazilian Society of Otolaryngology task force - Otosclerosis: evaluation and treatment. *Braz J Otorhinolaryngol*. 2023;89(5):101303. <https://doi.org/10.1016/j.bjorl.2023.101303>.
- Rajan D, Cureoglu S, Adams ME, et al. Otosclerosis and the evolution of stapes surgery: a historical and otopathological study. *Laryngoscope Invest Otolaryngol*. 2024;9(6):e70045. <https://doi.org/10.1002/lio.270045>.
- Mudry A, Adam Politzer (1835-1920) and the description of otosclerosis. *Otol Neurotol*. 2006;27(2):276-281. <https://doi.org/10.1097/O1.mao.0000187050.04286.80>.
- Casas JS, Rodríguez DD, Miranda GG, et al. Otoesclerosis: revisión de aspectos etiopatogénicos, clínico-demográficos e imagenológicos. *Rev Chil Radiol*. 2016;22(3):108-113. <https://doi.org/10.1016/j.rchira.2016.08.002>.
- Ricci G, Gambacorta V, Lapenna R, et al. The effect of female hormone in otosclerosis. A comparative study and speculation about their effect on the ossicular chain based on the clinical results. *Eur Arch Otorhinolaryngol*. 2022;279(10):4831-4838. <https://doi.org/10.1007/s00405-022-07295-w>.
- Bittermann AJN, Wegner I, Noordman BJ, et al. An introduction of genetics in otosclerosis: a systematic review. *Otolaryngol Head Neck Surg*. 2014;150(1):34-39. <https://doi.org/10.1177/0194599813509951>.
- Chen W, Meyer NC, McKenna MJ, et al. Single-nucleotide polymorphisms in the COL1A1 regulatory regions are associated with otosclerosis. *Clin Genet*. 2007;71(5):406-414. <https://doi.org/10.1111/j.1399-0004.2007.00794.x>.
- Crompton M, Cadge BA, Ziff JL, et al. The epidemiology of otosclerosis in a British cohort. *Otol Neurotol*. 2019;40(1):22-30. <https://doi.org/10.1097/MAO.0000000000002047>.
- Ishai R, Halpin CF, Shin JJ, et al. Long-term incidence and degree of sensorineural hearing loss in otosclerosis. *Otol Neurotol*. 2016;37(10):1489-1496. <https://doi.org/10.1097/MAO.0000000000001234>.
- Lee TC, Aviv RI, Chen JM, et al. CT grading of otosclerosis. *AJNR Am J Neuroradiol*. 2009;30(7):1435-1443. <https://doi.org/10.3174/ajnr.A1558>.
- Gogoulou PP, Sideris G, Nikolopoulos T, et al. Conservative otosclerosis treatment with sodium fluoride and other modern formulations: a systematic review. *Cureus*. 2023;15(2):e34850. <https://doi.org/10.7759/cureus.34850>.
- Handke V, Agha-Mir-Salim P, James P, et al. Audiological outcome after stapes surgery in relation to prosthesis type. *Eur Arch Otorhinolaryngol*. 2023;280(7):3171-3176. <https://doi.org/10.1007/s00405-023-07822-3>.
- Committee on Hearing and Equilibrium. Committee on Hearing and Equilibrium guidelines for the evaluation of results of treatment of conductive hearing loss. *Otolaryngol Head Neck Surg*. 1995;113(3):186-187. [https://doi.org/10.1016/S0194-5998\(95\)70103-6](https://doi.org/10.1016/S0194-5998(95)70103-6).
- Paulsen AJ, Fischer ME, Pinto A, et al. Incidence of hearing impairment and changes in pure-tone average across generations. *JAMA Otolaryngol Head Neck Surg*. 2021;147(2):151-158. <https://doi.org/10.1001/jamaoto.2020.4352>.
- Declau F, van Spaendonck M, Timmermans JP, et al. Prevalence of histologic otosclerosis: an unbiased temporal bone study in Caucasians. *Adv Otorhinolaryngol*. 2007;65:6-16. <https://doi.org/10.1159/000098663>.
- Marinelli JP, Totten DJ, Chauhan KK, et al. The rise and fall of otosclerosis: a population-based study of disease incidence spanning 70 years. *Otol Neurotol*. 2020;41(9):e1082-e1090. <https://doi.org/10.1097/MAO.0000000000002763>.
- Gredilla Molinero J, Mancheño Losa M, Santamaría Guinea N, et al. Actualización en el diagnóstico radiológico de la otosclerosis. *Radiología*. 2016;58(4):246-256. <https://doi.org/10.1016/j.rx.2016.04.008>.
- Romdhane N, Chiboub D, Nessib A, et al. Functional outcomes of otosclerosis surgery. *Braz J Otorhinolaryngol*. 2025;91(6):101679. <https://doi.org/10.1016/j.bjorl.2025.101679>.
- Alzhrani F, Mokhatrish MM, Al-Momani MO, et al. Effectiveness of stapedotomy in improving hearing sensitivity for 53 otosclerotic patients: retrospective review. *Ann Saudi Med*. 2017;37(1):49-55. <https://doi.org/10.5144/0256-4947.2017.49>.
- Quesnel AM, Ishai R, Meehan T, et al. Computed tomography density as a bio-marker for histologic grade of otosclerosis: a human temporal bone pathology study. *Otol Neurotol*. 2022;43(6):e605-e612. <https://doi.org/10.1097/MAO.0000000000003535>.
- Salvador P, Costa R, Silva F, et al. Primary stapedotomy: influence of prosthesis diameter on hearing outcome. *Acta Otorrinolaringol Esp*. 2021;72(4):238-245. <https://doi.org/10.1016/j.otorri.2020.06.004>.
- Dhooge I, Desmedt S, Maly T, et al. Long-term hearing results of stapedotomy: analysis of factors affecting outcome. *Eur Arch Otorhinolaryngol*. 2018;275(5):1111-1119. <https://doi.org/10.1007/s00405-018-4899-2>.
- Vincent R, Sperling NM, Oates J, et al. Surgical findings and long-term hearing results in 3,050 stapedotomies for primary otosclerosis: a prospective study with the otology-neurotology database. *Otol Neurotol*. 2006;27(8 Suppl 2):S25-S47. <https://doi.org/10.1097/O1.mao.0000235311.80066.df>.
- Zhang N, Ma X, Wang D, et al. Impact of age on hearing outcomes following stapes surgery for otosclerosis: a retrospective study. *Acta Otolaryngol*. 2025;145(8):682-686. <https://doi.org/10.1080/00016489.2025.2504040>.
- Vartiainen E. Surgery in elderly patients with otosclerosis. *Am J Otol*. 1995;16(4):536-538.
- Ayache D, Corre A, Van Prooyen S, et al. Surgical treatment of otosclerosis in elderly patients. *Otolaryngol Head Neck Surg*. 2003;129(6):674-647. <https://doi.org/10.1016/j.jot.2003.08.013>.
- Wardenga N, Diedrich V, Waldmann B, et al. Hearing aid treatment in patients with mixed hearing loss. Part I: expected benefit and limitations after stapes surgery. *Audiol Neurotol*. 2020;25(3):125-132. <https://doi.org/10.1159/000502994>.
- Pauli N, Strömbäck K, Lundman L, et al. Surgical technique in stapedotomy: hearing outcome and complications. *Laryngoscope*. 2020;130(3):790-796. <https://doi.org/10.1002/lary.28072>.

Síndrome de seno silente: resultados quirúrgicos y propuesta de algoritmo para la reconstrucción orbitaria

Antonella Giorgi¹, Micaela Redivo² y Carlos S. Ruggeri³

Servicio de Otorrinolaringología, Hospital Italiano. Buenos Aires, Argentina.

RESUMEN

Introducción: el síndrome de seno silente (SSS) es una entidad infrecuente caracterizada por atelectasia maxilar progresiva asociada al descenso del piso orbitario, que provoca enoftalmos e hipoglobos en ausencia de síntomas rinosinuales significativos. Aunque el tratamiento de elección consiste en la ventilación del seno maxilar mediante cirugía endoscópica, la indicación de reconstrucción simultánea del piso orbitario continúa siendo motivo de controversia.

Objetivo: describir los resultados del tratamiento quirúrgico del síndrome de seno silente y proponer un algoritmo terapéutico para la indicación de reconstrucción del piso orbitario.

Métodos: se realizó un estudio descriptivo retrospectivo que incluyó pacientes con diagnóstico de SSS tratados quirúrgicamente entre marzo de 2018 y octubre de 2025.

Resultados: se trataron cinco pacientes; todos presentaron enoftalmos y uno asoció diplopía. En todos los casos se realizó cirugía endoscópica endonasal con antrostomía maxilar media para restablecer la ventilación sinusal. En dos pacientes con enoftalmos grave y diplopía asociada se efectuó reconstrucción simultánea del piso orbitario. Durante el seguimiento, todas las antrostomías permanecieron permeables. Se observó mejoría del enoftalmos en todos los pacientes y resolución de la diplopía en el caso afectado.

Conclusiones: nuestros hallazgos sugieren que la cirugía endoscópica maxilar puede ser suficiente en casos leves de SSS, mientras que los pacientes con deformidad orbitaria grave o síntomas visuales podrían beneficiarse de la reconstrucción simultánea del piso orbitario.

Palabras clave: síndrome de seno silente, atelectasia maxilar crónica, enoftalmos, órbita, cirugía endoscópica nasal.

Silent sinus syndrome: surgical outcomes and a proposed algorithm for orbital reconstruction

ABSTRACT

Introduction: silent sinus syndrome (SSS) is an uncommon entity characterized by progressive maxillary atelectasis associated with inferior displacement of the orbital floor, resulting in enophthalmos and hypoglobus in the absence of significant sinonasal symptoms. Although endoscopic maxillary sinus ventilation is considered the treatment of choice, the indication for simultaneous orbital floor reconstruction remains controversial.

Objective: to describe the surgical outcomes of patients with silent sinus syndrome and to propose a therapeutic algorithm for orbital floor reconstruction.

Autor para correspondencia: carlos.ruggeri@hospitalitaliano.org.ar, Ruggeri SC.

Recibido: 16/05/2026 | Aceptado: 18/08/2026 | Publicado: 17/09/2026

DOL: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1342>

Cómo citar: Giorgi A, Redivo M, Ruggeri CS. Síndrome de seno silente: resultados quirúrgicos y propuesta de algoritmo para la reconstrucción orbitaria. Rev Hosp Ital B.Aires. 2026;46(3):e0001342

Methods: a retrospective descriptive study was conducted including patients diagnosed with SSS who underwent surgical treatment between March 2018 and October 2025.

Results: five patients were treated; all presented with enophthalmos and one also had diplopia. All patients underwent endoscopic endonasal surgery with middle meatal antrostomy to restore maxillary sinus ventilation. Simultaneous orbital floor reconstruction was performed in two patients with severe enophthalmos and associated diplopia. During follow-up, all maxillary antrostomies remained patent. Improvement of enophthalmos was observed in all patients, and resolution of diplopia was achieved in the affected patient.

Conclusions: our findings suggest that endoscopic maxillary sinus surgery alone may be sufficient in mild cases of SSS, whereas patients with severe orbital deformity or visual symptoms may benefit from simultaneous orbital floor reconstruction.

Keywords: silent sinus syndrome, chronic maxillary atelectasis, enophthalmos, orbit, endoscopic sinus surgery.

INTRODUCCIÓN

El síndrome de seno silente (SSS) es una patología infrecuente que afecta principalmente al seno maxilar.

Se caracteriza por un colapso centripeto de las paredes óseas del maxilar, incluso de su pared superior que conforma el piso de la órbita.

Produce enoftalmos e hipoglobo y puede provocar alteraciones estéticas y en la visión, sin la presencia de síntomas rinosinuales¹.

La ventilación del seno maxilar mediante una antrostomía media endonasal es el principal tratamiento, pero la indicación de cuándo reconstruir el piso de la órbita es controvertida.

Describimos tres casos clínicos de síndrome de seno silente que fueron tratados mediante cirugía endoscópica maxilar y dos con un abordaje combinado, endonasal y externo con reconstrucción del piso de la órbita.

Diseñamos un algoritmo para establecer cuándo es necesario –además del drenaje sinusal– reconstruir el piso de la órbita.

OBJETIVOS

Describir los resultados del tratamiento quirúrgico del síndrome del seno silente, y establecer un algoritmo para indicar la reconstrucción del piso de la órbita.

Diseño descriptivo y retrospectivo

MÉTODO

Revisión de las historias clínicas electrónicas de todos los pacientes tratados por SSS en el Hospital Italiano de Buenos Aires entre marzo de 2018 y octubre de 2025.

Se incluyeron los pacientes que tuvieron asimetría facial, enoftalmos y, en la tomografía, presentaron colapso de la pared ósea superior del seno (piso de órbita).

Se recolectaron en una tabla Excel® los siguientes datos: edad y sexo; lateralidad; signos y síntomas oftalmológicos; si se reconstruyó el piso de órbita; mejoría del enoftalmos/hipoglobo, de los síntomas visuales, y el tiempo de seguimiento.

Todos los pacientes fueron evaluados mediante endoscopia nasal y tomografía computarizada de macizo facial y órbita sin contraste (TC). A quienes refirieron síntomas visuales o tuvieron enoftalmos grave se les realizó una evaluación oftalmológica.

El grado de enoftalmos se determinó de forma retrospectiva mediante TC de senos paranasales.

Se identificó el corte axial que pasara por el ecuador del globo ocular de ambos lados. Se trazó una línea de referencia entre ambos arcos cigomáticos (línea intercigomática), utilizada como plano de referencia óseo fijo no afectado por el proceso patológico orbitario. Desde esta línea se midió, de forma perpendicular, la distancia hasta el punto más anterior de la córnea en cada globo ocular. El enoftalmos se calculó como la diferencia entre ambas mediciones (lado afectado vs. lado no afectado), expresada en milímetros.

Se consideró enoftalmos radiológicamente significativo una diferencia interocular ≥ 2 mm. Los casos se clasificaron según gravedad en: leve (2 mm), moderado (3-4 mm) y grave (≥ 5 mm)².

El tratamiento fue quirúrgico, bajo anestesia general y consistió en realizar en todos los pacientes una cirugía endonasal del seno maxilar con endoscopios de 0 grado. Se hizo una uncinectomía retrógrada o se identificó directamente la zona del *ostium* cuando la unciforme no pudo identificarse y se realizó una antrostomía maxilar media amplia. Se aspiró moco espeso del seno.

Cuando hubo necesidad de reconstruir el piso de la órbita, se hizo una incisión cutánea subtarsal y se disecaron las fibras del músculo orbicular hasta localizar el reborde orbitario inferior. Se incidió el periostio y, con asistencia del endoscopio, se disecó la periórbita del piso óseo cuando existió, evitando la extrusión de la grasa orbitaria.

Se colocó a continuación una micromalla de titanio, que se fijó con tornillos al reborde orbitario para realizar la reconstrucción del piso de la órbita.

Se controló por vía endonasal a través de la antrostomía media que la colocación de la micromalla fuera correcta.

Se suturó la incisión subtarsal con nailon 4 cero.

Los controles se realizaron en consultorio mediante endoscopia nasal y por TC al año.

Mediante endoscopia nasal se constató la permeabilidad de la antrostomía maxilar.

El resultado fue determinado evaluando la resolución de los síntomas visuales, la mejoría del enoftalmos (reducción de la diferencia en la distancia entre la línea intercigomática y córnea entre el ojo sano y el afectado, evaluación basada en comparación fotográfica y apreciación estética subjetiva del paciente: sin mejoría, parcial: < 50%, marcada > 50% y < 100% y resolución) y mediante correlación tomográfica posoperatoria, de acuerdo con la ventilación adecuada del seno maxilar.

El protocolo del presente estudio fue aprobado por el Comité de ética institucional del Hospital Italiano de Buenos Aires (número de protocolo 19546) y fue conducido de acuerdo con los lineamientos asentados por la declaración modificada de Helsinki.

RESULTADOS

Cinco pacientes fueron tratados quirúrgicamente por padecer SSS, tres fueron mujeres y dos hombres. La edad promedio fue 44,8, el más joven tuvo 9 y el mayor 66 años.

Ninguno tuvo síntomas rinosinuales y solo uno tuvo diplopía. En la tomografía computarizada se observó en todos los pacientes un colapso de las paredes óseas y una ocupación por secreciones del seno maxilar.

En dos pacientes se reconstruyó el piso de la órbita con una micromalla de titanio porquén en una mujer, el enoftalmos e hipoglobo impactó estéticamente y, en otro, además porque tuvo diplopía en supravversión.

En este último al realizar la orbitotomía inferior se detectó la falta total de hueso en la pared inferior de la órbita.

Ningún paciente presentó complicaciones ni fue necesaria una reintervención. Todas las antrostomías maxilares fueron permeables y todos mejoraron el enoftalmos y uno además la diplopía.

El seguimiento promedio fue 2 años (Tabla 1 y Fig. 1).

DISCUSIÓN

El síndrome de seno silente (SSS) representa la forma grave de la atelectasia maxilar crónica, caracterizada por colapso progresivo del seno maxilar asociado a enoftalmos e hipoglobo³.

Su fisiopatología no está completamente establecida, aunque la obstrucción del *ostium* maxilar con hipovenilación, reabsorción gaseosa y generación de presión negativa constituye la hipótesis más aceptada^{4,5}.

Rose y cols.⁶ compararon la fisiopatología del SSS con el mecanismo de producción de la retracción timpánica y la producción de una otitis serosa por oclusión de la trompa de Eustaquio.

Si bien este mecanismo explicaría el colapso sinusal, la baja incidencia del SSS en comparación con la sinusitis crónica sugiere la participación de factores adicionales. En nuestra serie, los antecedentes de cirugía nasal en pacientes con mayor compromiso orbitario podrían haber contribuido a la alteración del complejo ostiomeatal, en línea con informes previos que vinculan procedimientos nasales y traumatismos con el desarrollo de esta entidad⁷⁻⁹.

Dos pacientes de nuestra serie en los que realizamos una reconstrucción del piso de órbita tuvieron antecedentes de cirugías previas (septoplastia más turbinoplastia y rinoplastia). Uno de ellos, 12 meses antes de la consulta. Refirió que previamente a la septoplastia y turbinoplastia inferior no tuvo asimetría facial ni diplopía.

El diagnóstico se basa en la tomografía computarizada, que permite evaluar el colapso sinusal y la posición del globo ocular¹⁰, así como descartar diagnósticos diferenciales como: la hipoplasia sinusal (es estable y congénita), hemiatrofia facial, lipodistrofia progresiva, síndrome de Horner y los mucocelos¹¹ (el seno maxilar tiene contenido mucoso que produce una reabsorción ósea del piso orbitario, pero por la presión que ejerce la acumulación de secreciones, el ojo se desplaza hacia arriba).

El tratamiento consiste en la restauración de la ventilación del seno maxilar mediante cirugía endoscópica¹².

Sin embargo, la indicación de reconstrucción del piso orbitario continúa siendo motivo de controversia.

Tabla 1. Pacientes tratados por síndrome del seno silente

N	Sexo/Edad	Localización	Signos/síntomas oftalmológicos	Reconstrucción piso órbita	Mejoría enoftalmos, síntomas visuales	Tiempo seguimiento (años)
1	M/9años	izquierdo	leve enoftalmos	no	sí (marcada)	3
2	M/49años	derecho	leve enoftalmos	no	sí (marcada)	1
3	H/66años	derecho	leve enoftalmos	no	sí (resolución)	3
4	M/51años	derecho	enoftalmos grave	sí	sí (resolución)	2
5	H/49 años	derecho	enoftalmos grave/diplopía supravversión	sí	sí (resolución)	1

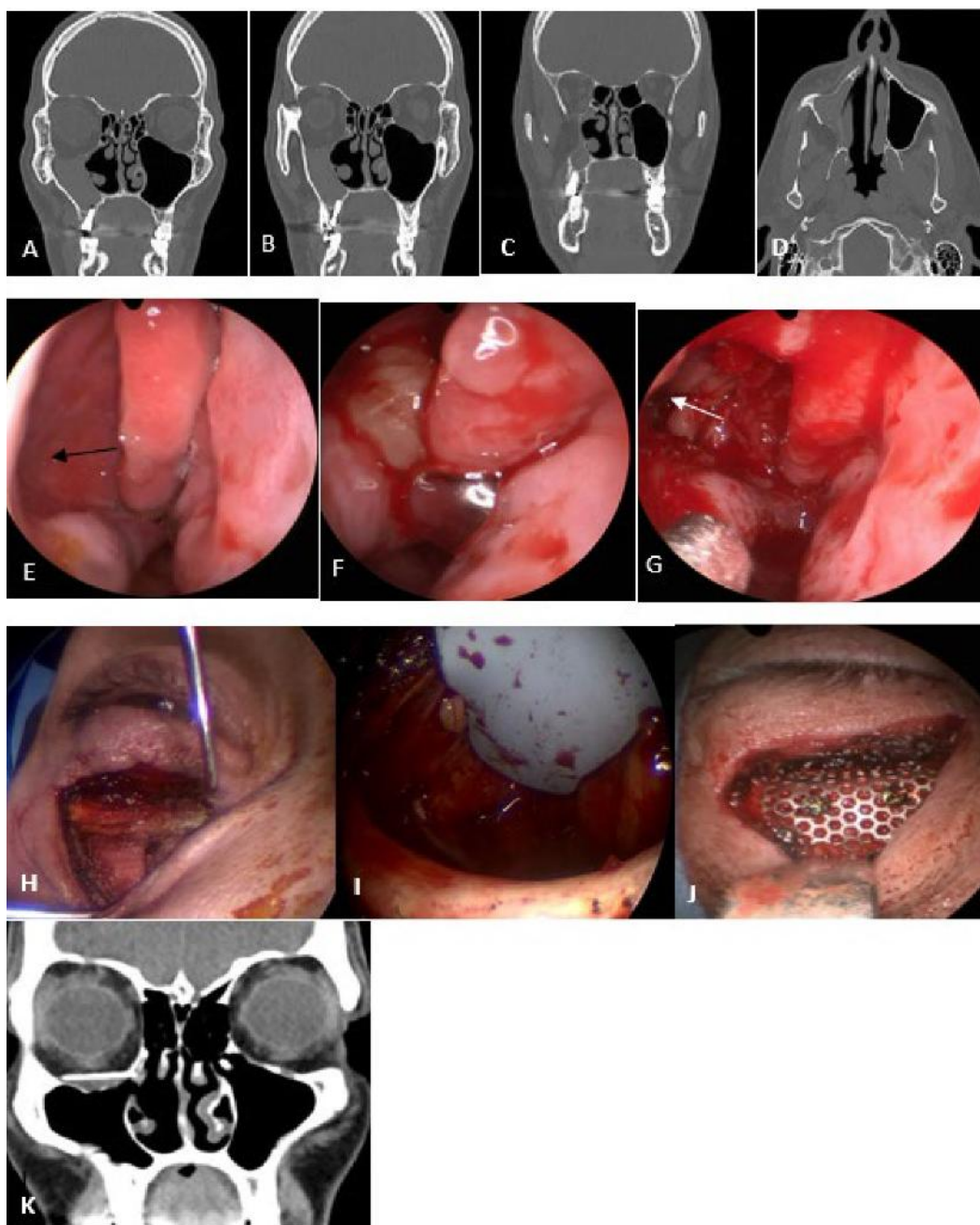


Figura 1. A, B, C: tomografía computarizada de macizo facial sin contraste; cortes coronales: se observa seno maxilar derecho ocupado por secreciones hiperdensas con colapso de la pared medial y marcado hipoglobo por erosión del piso de la órbita. D: TC corte axial: ocupación del seno por secreciones y colapso de su pared anterior, medial y posterior. Cirugía endoscópica maxilar - E: visión de la pared medial del seno atelectásica (flecha negra), F: antrostomía y aspiración de moco, G: antrostomía maxilar media (flecha blanca). H: incisión subtarsal y exposición del reborde orbitario, I: incisión de la periórbita y disección asistida con endoscopio para exponer el piso orbitario, J: colocación de micromalla de titanio, K: se observa, en la TC posoperatoria, la antrostomía maxilar media permeable, el seno maxilar ventilado y la malla de titanio en posición adecuada en el piso de la órbita derecha.

Mientras algunos autores sostienen que el drenaje aislado puede revertir el enoftalmos, otros recomiendan un abordaje diferido o combinado. La uncinectomía retrógrada es de elección para evitar lesionar el ojo que se

encuentra descendido y vulnerable en el seno maxilar. En tres de los pacientes la apófisis unciforme no fue visible debido a la grave atelectasia de la pared medial del maxilar. En estos casos se realizó la antrostomía medial

ingresando en el seno por encima del cornete inferior a la altura de la cabeza del cornete medio.

Algunos estudios informaron que el drenaje y la ventilación del seno maxilar colapsado son suficientes para revertir el enoftalmos, sin reconstruir el piso orbitario, y mencionaron que también pueden detener la reabsorción ósea^{13,14}.

Otros sugieren ventilar el seno maxilar y esperar 6 meses para realizar la reconstrucción del piso orbitario si fuese necesaria, a fin de evitar colocar el implante en un sitio con infección¹⁵.

En varios estudios prefieren realizar un abordaje combinado en un solo tiempo quirúrgico con el objetivo de una rehabilitación rápida^{4,16-18}.

En un estudio retrospectivo compararon los resultados del drenaje y la ventilación endonasal del maxilar con el drenaje y la reconstrucción del piso de la órbita en el mismo tiempo quirúrgico.

Los pacientes tratados mediante ventilación maxilar y reconstrucción de la órbita tuvieron una mejoría más significativa del enoftalmos e hipoglobo que los tratados solamente por drenaje maxilar endonasal¹⁹.

En nuestra experiencia, los pacientes con mayor grado de enoftalmos e hipoglobo se beneficiaron de la reconstrucción orbitaria en un mismo tiempo quirúrgico, sin observar complicaciones infecciosas. Estos hallazgos coinciden con los estudios que informan mejores resultados estéticos con el abordaje combinado.

Consideramos que la reconstrucción orbitaria debe reservarse para pacientes con deformidad significativa o compromiso visual, mientras que –en casos leves– el drenaje endoscópico podría ser suficiente.

Las limitaciones de este estudio incluyen el reducido tamaño muestral y su diseño retrospectivo (Fig. 2).

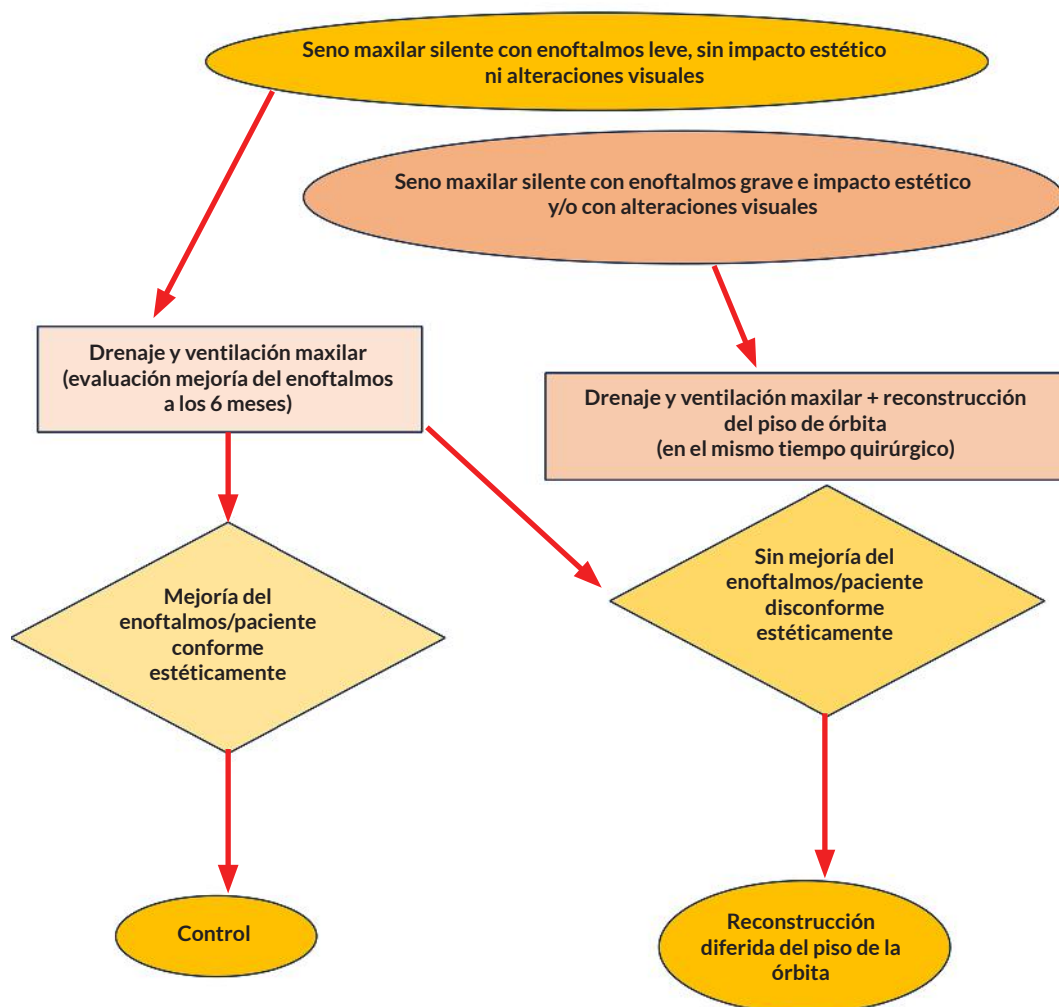


Figura 2. Diagrama de flujo: Algoritmo de tratamiento quirúrgico del seno maxilar silente maxilar.

CONCLUSIONES

El drenaje y ventilación del seno maxilar realizado por vía endonasal con endoscopios es de elección para tratar el síndrome de seno silente. Todos los pacientes tratados mejoraron el enoftalmos.

En pacientes con enoftalmos grave que impacta estéticamente en la calidad de vida del paciente y/o con síntomas visuales, es mejor reconstruir el piso de la órbita en el mismo tiempo quirúrgico.

Nuestros resultados sugieren que la gravedad clínica inicial debería guiar la indicación de reconstrucción orbitaria inmediata.

Contribuciones de los autores: Conceptualización, curación de datos, análisis formal, investigación, metodología, recursos, **software**, supervisión, validación, visualización, redacción-borrador original y redacción-revisión y edición (CSR, MR, ADG). Administración del proyecto (CSR).

Conflictos de intereses: los autores declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: los autores declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

- de Dorlodot C, Collet S, Rombaux P, et al. Chronic maxillary atelectasis and silent sinus syndrome: two faces of the same clinical entity. *Eur Arch Otorhinolaryngol*. 2017;274(9):3367-3373. <https://doi.org/10.1007/s00405-017-4622-8>.
- Rullan-Oliver B, del Toro-Diez E, Portela-Arraiza JC. Natural progression of bilateral maxillary silent sinus syndrome: a metachronous case report. *SAGE Open Med Case Rep*. 2020;8:2050313X20902339. <https://doi.org/10.1177/2050313X20902339>.
- Brandt MG, Wright ED. The silent sinus syndrome is a form of chronic maxillary atelectasis: a systematic review of all reported cases. *Am J Rhinol*. 2008;22(1):68-73. <https://doi.org/10.2500/ajr.2008.22.3118>.
- Soparkar CN, Patrinely JR, Cuaycong MJ, et al. The silent sinus syndrome. A cause of spontaneous enophthalmos. *Ophthalmology*. 1994;101(4):772-778. [https://doi.org/10.1016/S0161-6420\(94\)31267-X](https://doi.org/10.1016/S0161-6420(94)31267-X).
- Cobb AR, Murthy R, Cousin GC, et al. Silent sinus syndrome. *Br J Oral Maxillofac Surg*. 2012;50(6):e81-e85. <https://doi.org/10.1016/j.bjoms.2011.10.001>.
- Rose GE, Lund VJ. Clinical features and treatment of late enophthalmos after orbital decompression: a condition suggesting cause for idiopathic "imploding antrum" (silent sinus) syndrome. *Ophthalmology*. 2003;110(4):819-826. [https://doi.org/10.1016/S0161-6420\(02\)01994-2](https://doi.org/10.1016/S0161-6420(02)01994-2).
- Strabbing EM, Engin O, Telleman MAJ, et al. Post-traumatic and iatrogenic silent sinus syndrome: a case series. *Oral Maxillofac Surg*. 2025;29(1):106. <https://doi.org/10.1007/s10006-025-01391-x>.
- Shoji MK, Al-Sharif E, Dib El Jalbout N, et al. Orbital causes of ostiomeatal complex occlusion. *Ophthalmic Plast Reconstr Surg*. 2026;42(1):28-34. <https://doi.org/10.1097/IOP.0000000000002965>.
- Eloy JA, Jacobson AS, Elahi E, et al. Enophthalmos as a complication of rhinoplasty. *Laryngoscope*. 2006;116(6):1035-1038. <https://doi.org/10.1097/01.mlg.0000217254.22699.09>.
- Bossolesi P, Autelitano L, Brusati R, et al. The silent sinus syndrome: diagnosis and surgical treatment. *Rhinology*. 2008;46(4):308-316.
- Tousidonis M, Alvarez-Mokthari S, Khayat S, et al. Contemporary treatment of silent sinus syndrome: a case report and literature review. *Cureus*. 2024;16(4):e57577. <https://doi.org/10.7759/cureus.57577>.
- Amin D, Chitguppi C, Xu V, et al. Volumetric analysis of the sinus and orbit in silent sinus syndrome after endoscopic sinus surgery. *Otolaryngol Head Neck Surg*. 2023;169(1):151-156. <https://doi.org/10.1002/ohn.259>.
- Rosso C, Saibene AM, Felisati G, et al. Silent sinus syndrome: systematic review and proposal of definition, diagnosis and management. *Acta Otorhinolaryngol Ital*. 2022;42(4):305-316. <https://doi.org/10.14639/0392-100X-N1598>.
- Sivasubramaniam R, Sacks R, Thornton M. Silent sinus syndrome: dynamic changes in the position of the orbital floor after restoration of normal sinus pressure. *J Laryngol Otol*. 2011;125(12):1239-1243. <https://doi.org/10.1017/S0022215111001952>.
- Babar-Craig H, Kayhanian H, De Silva DJ, et al. Spontaneous silent sinus syndrome (imploding antrum syndrome): case series of 16 patients. *Rhinology*. 2011;49(3):315-317. <https://doi.org/10.4193/Rhino10.103>.
- Sesenna E, Oretti G, Anghinoni ML, et al. Simultaneous management of the enophthalmos and sinus pathology in silent sinus syndrome: a report of three cases. *J Craniomaxillofac Surg*. 2010;38(6):469-472. <https://doi.org/10.1016/j.jcms.2009.12.003>.
- Arnon R, Gluck O, Winter H, et al. Combined single-step procedure for correction of silent sinus syndrome. *Case Rep Ophthalmol*. 2019;10(1):95-100. <https://doi.org/10.1159/000498964>.
- Clarós P, Sobolewska AZ, Cardesa A, et al. Silent sinus syndrome: combined sinus surgery and orbital reconstruction - report of 15 cases. *Acta Otolaryngol*. 2019;139(1):64-69. <https://doi.org/10.1080/00016489.2018.1542161>.
- Shoji MK, Legala AR, Al-Sharif E, et al. Outcomes of sinus surgery and orbital reconstruction in silent sinus syndrome: a retrospective cohort study. *J Craniofac Surg*. 2025;36(8):e1247-e1251. <https://doi.org/10.1097/SCS.00000000000011897>.

Sinusitis odontógena: ¿por dónde empezar? La importancia del abordaje multidisciplinario

Candelaria Serrano^{1, }, Facundo Zamar^{2, } y Sebastián Aragón^{1, }

1. Servicio de Otorrinolaringología, Hospital Italiano. Buenos Aires, Argentina

2. Servicio Cirugía General, Hospital Italiano. Buenos Aires, Argentina

RESUMEN

Introducción: la sinusitis odontógena puede ser secundaria a una patología infecciosa dental adyacente o a lesiones iatrogénicas derivadas de procedimientos dentales. Representa hasta el 75% de los casos de opacificación unilateral de seno maxilar por tomografía computarizada. Existe una falta de consenso claro con respecto al manejo óptimo y el momento adecuado para las intervenciones. El objetivo es proponer un enfoque centrado en la colaboración activa entre los Servicios de Otorrinolaringología y Maxilofacial y presentar el algoritmo de acción utilizado.

Materiales y métodos: estudio descriptivo retrospectivo de pacientes con sinusitis odontógena que recibieron abordaje multidisciplinario y cirugía combinada.

Resultados: veintiún (21) pacientes fueron sometidos a cirugía combinada para tratamiento odontológico y nasal. A la altura sinusal se realizó cirugía endoscópica nasal con antrostomía maxilar según ocupación evidenciada por tomografía computarizada. A nivel dentario, las piezas dentarias afectadas fueron removidas, al igual que los quistes odontogénicos. La comunicación y las fístulas oroantrales fueron cerradas con colgajos en el mismo tiempo quirúrgico. Todos los pacientes resolvieron sin complicaciones.

Discusión: si el origen es una infección dental tratable, el abordaje inicial debe incluir tratamiento odontológico específico en conjunto con manejo médico. Existen indicaciones de cirugía endoscópica temprana como impacto en la calidad de vida o riesgo de complicaciones. Es notable que cuando se abordan tanto la patología dental como la sinusitis, puede esperarse una resolución del cuadro en el 90%, con una mejoría sintomática sinusal temprana. En nuestra experiencia, preferimos realizar el tratamiento combinado cuando no es factible tratar la patología dental de forma conservadora y cuando es necesario extraer la pieza dental o el implante responsable.

Conclusión: consideramos que la colaboración estrecha entre especialistas es fundamental para alcanzar resultados óptimos. Cuando está indicada la intervención quirúrgica en ambos campos, la cirugía combinada simultánea representa una estrategia directa, fiable y curativa, con baja morbilidad, baja incidencia de complicaciones y recuperación rápida.

Palabras clave: sinusitis, infección focal dental, otorrinolaringología, cirugía endoscópica nasal, cirugía maxilofacial.

Autora para correspondencia: candelaria.serrano@hospitalitaliano.org.ar, Serrano C.

Recibido: 23/2/2026 | Aceptado: 10/08/2026 | Publicado: 18/09/2026

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1318>

Cómo citar: Candelaria S, Zamar F, Aragón S. Sinusitis odontógena: ¿por dónde empezar? La importancia del abordaje multidisciplinario. Rev Hosp Ital B.Aires. 2026;46(3):e0001318

<https://ojs.hospitalitaliano.org.ar/>

 BY-NC-SA 4.0

ISSN (en línea) 2314-3312

Odontogenic Sinusitis: Where to Start? The Importance of a Multidisciplinary Approach

ABSTRACT

Introduction: odontogenic sinusitis can be secondary to an adjacent infectious dental pathology or iatrogenic lesions resulting from dental procedures. It accounts for up to 75 % of cases of unilateral maxillary sinus opacification seen in computed tomography. There is a lack of clear consensus regarding the optimal management and the appropriate timing for interventions. The goal is to propose an approach centered on active collaboration between the Otolaryngology and Maxillofacial services and present the action algorithm used.

Materials and methods: retrospective descriptive study of patients with odontogenic sinusitis who received a multidisciplinary approach and combined surgery.

Results: 21 patients underwent combined surgery for dental and nasal treatment. All patients underwent combined surgery. Nasal surgery was performed endoscopically with maxillary antrostomy based on the opacification seen in the computed tomography. In the dental aspect, the affected teeth and odontogenic cysts were removed. Oral-antral communications and fistulas were closed with flaps during the same surgical procedure. All patients resolved without complications.

Discussion: if the origin is a treatable dental infection, the initial approach should include specific dental treatment alongside medical management. There are indications for early endoscopic surgery, such as impact on quality of life or risk of complications. Notably, when both the dental pathology and sinusitis are addressed, a resolution rate of 90 % can be expected, with early symptomatic improvement of the sinus. In our experience, we prefer to perform combined treatment when it is not feasible to treat the dental pathology conservatively and when the responsible dental piece or implant needs to be extracted.

Conclusion: we believe that close collaboration between specialists is essential to achieve optimal outcomes. When surgical intervention is indicated in both fields, simultaneous combined surgery represents a direct, reliable, and curative strategy, with low morbidity, a low incidence of complications, and rapid recovery.

Keywords: sinusitis, dental infection, otolaryngology, endoscopic nasal surgery, maxillofacial surgery.

INTRODUCCIÓN

La sinusitis odontógena puede ser secundaria a una patología infecciosa dental adyacente o a lesiones iatrogénicas derivadas de procedimientos dentales. Suele representar entre el 25 y el 40% de todos los casos de sinusitis maxilar crónica y representa entre el 45 y el 75% de los casos de opacificación unilateral del seno maxilar en tomografía computarizada (TC)¹. Su diagnóstico puede ser difícil de establecer debido a la inespecificidad de los síntomas nasosinusales y a la escasa sintomatología dental.

Las posibles patologías dentales incluyen periodontitis apical (enfermedad endodóntica que surge de la necrosis pulpar), periodontitis marginal, comunicación o fístula oroantral (COA o FOA), o cuerpos extraños relacionados con tratamientos dentales dentro del seno maxilar.

Los otorrinolaringólogos deben confirmar la sinusitis mediante endoscopia nasal evaluando la presencia de pus en el meato medio, edema o pólipos, y mediante tomografía computarizada de macizo craneofacial. Los profesionales dentales deben confirmar la patología dental mediante exámenes e imágenes apropiadas como la radiografía panorámica².

Actualmente, existe una falta de consenso claro con respecto al manejo óptimo y el momento adecuado para

la intervención quirúrgica en casos de sinusitis odontógena. La literatura muestra una variabilidad significativa en cuanto a las indicaciones, secuencia y tiempos de las intervenciones dentales y otorrinolaringológicas, lo que ha llevado a una práctica clínica heterogénea y, en ocasiones, subóptima. En la práctica clínica contemporánea, es fundamental identificar, comprender, abordar de manera adecuada la sinusitis odontogénica para lograr resultados óptimos en los pacientes.

El objetivo de este trabajo es proponer un enfoque centrado en la colaboración activa entre los Servicios de Otorrinolaringología y Maxilofacial, con el fin de garantizar una evaluación integral, una planificación terapéutica adecuada y una intervención oportuna. A continuación se presenta el algoritmo de acción desarrollado en un hospital privado de Buenos Aires, el cual estandariza el abordaje diagnóstico-terapéutico con el objetivo de optimizar los resultados clínicos y reducir el riesgo de recurrencia.

OBJETIVO

Evaluar los resultados del tratamiento de sinusitis odontógena con un abordaje centrado en la colaboración activa entre los Servicios de Otorrinolaringología y Maxilofacial.

MATERIALES Y MÉTODOS

Estudio descriptivo retrospectivo de pacientes con sinusitis odontógena que recibieron abordaje multidisciplinario y cirugía combinada. Se obtuvieron variables demográficas (sexo y edad); la afectación sinusal (según escala tomográfica de Lund-Mackay), la causa odontológica, el tratamiento realizado (procedimiento odontológico y cirugía endoscópica nasal). Para la descripción de los pacientes se utilizó estadística descriptiva. Las variables categóricas se informan como frecuencias relativas y/o porcentajes; mientras que las variables cuantitativas se expresan como media y desvío estándar, o mediana y rango intercuartílico (RIC), de acuerdo con la normalidad. Se utilizó STATA Versión 16®.

El protocolo del presente estudio fue aprobado por el Comité de Ética institucional (CEPI#7717 PRIISA#18150) y fue conducido de acuerdo con los lineamientos asentados por la declaración modificada de Helsinki.

RESULTADOS

Veintiún (21) pacientes fueron sometidos a cirugía combinada para tratamiento odontológico y nasal. Once (11) eran mujeres y 10 hombres con una edad promedio de 55 años. En el 71% (15 pacientes), la causa odontológica fue infección a la altura de la pieza dentaria con proceso perirradicular. De estos, 3 pacientes presentaban comunicación oroantral evidente por tomografía. De los pacientes restantes, 3 presentaron quistes odontogénicos y 3 presentaron sinusopatía posterior a extracción dentaria con fístula oroantral posterior a extracción. El seno más frecuentemente afectado fue el izquierdo, con 12 pacientes con sinusitis maxilar derecha.

Todos los pacientes fueron sometidos a cirugía combinada. A nivel sinusal se realizó cirugía endoscópica nasal (CENS) con antrostomía maxilar según ocupación evidenciada por tomografía computarizada. A nivel dentario, las piezas dentarias afectadas fueron removidas, al igual que los quistes odontogénicos. La comunicación o defecto posterior a la extracción intraquirúrgica fue cerrado con colgajo vestibular. En el caso de las fistulas oroantrales crónicas fueron cerradas con colgajos de doble capa (marginal + bola de Bichat o marginal + vestibular). Todas las cirugías se realizaron de forma ambulatoria y por el mismo equipo quirúrgico. Todos los pacientes presentaron resolución sintomática sin complicaciones (Tabla 1).

DISCUSIÓN

De acuerdo con su origen, la fuente dental puede deberse a: origen endodóntico, que incluye necrosis pulpar, inflamación periapical, fractura radicular y otras infecciones de origen endodóntico; origen periodontal, que comprende defectos periodontales asociados con reabsorción ósea alveolar grave; lesiones endoperiodontales, que corresponden a dientes afectados por una infección combinada de origen endodóntico y periodontal; y otros orígenes, que incluyen la presencia de una fístula oroantral o de cuerpos extraños desplazados al interior del seno maxilar durante una extracción dental u otras cirugías orales (Figs. 1, 2 y 3). La proximidad de las raíces

de los dientes al piso del seno maxilar crea una relación anatómica delicada. La afección de estos puede alterar la membrana del seno, proporcionando una vía para que las bacterias, material extraño y mediadores inflamatorios invadan la cavidad del seno. Esta invasión desencadena inflamación de la mucosa, deterioro de la limpieza mucociliar y, si no se trata, una eventual propagación pansinusal. La colaboración con los colegas odontólogos es esencial; la radiografía panorámica y la tomografía computarizada pueden revelar sutiles patologías periapicales responsables de la sinusitis³.

La sinusitis odontógena tiende a pasar inadvertida o a ser mal diagnosticada. Algunos casos incluso se han encontrado asintomáticos en cuanto a síntomas dentales y sinusales. Turfe y cols. analizaron sinusopatías unilaterales donde el 44% informaba rinorrea fétida y el 87% presentaba material purulento en la región media del meato nasal⁴. Es obligatorio que los otorrinolaringólogos deriven a los pacientes que presenten opacificación unilateral del seno maxilar a especialistas dentales para una evaluación de posibles problemas dentales, incluso en los casos en los que no se detecten estas anomalías por tomografía³.

Además, la relevancia del tratamiento odontológico suele ser subestimada, y con frecuencia se recurre a la cirugía endoscópica nasal (CENS) cuando no se observa mejoría clínica con la antibioticoterapia. No obstante, es fundamental considerar que, si el origen de la sinusitis es una infección dental tratable (como caries profunda, absceso periapical o necrosis pulpar), el abordaje inicial debe incluir tratamiento odontológico específico en conjunto con manejo médico (antibióticos, medidas de higiene oral y tratamiento local). Diversos estudios han demostrado que una proporción significativa de estos casos puede resolverse sin necesidad de intervención quirúrgica sinusal. De los 33 pacientes incluidos en un estudio⁵, 22 (67%) fueron curados mediante tratamiento farmacológico y odontológico mientras que 11 pacientes (33%) necesitaron someterse a CENS tras el fracaso del tratamiento dental y médico para resolver la sinusitis.

Una de las causas más comunes es la infección endodóntica, que se origina por la necrosis pulpar con extensión a través de los conductos radiculares hacia los tejidos de soporte circundantes, lo que provoca periodontitis apical. Esta suele ser manejada más comúnmente a través de la terapia de conducto (endodoncia) resolviendo tanto la periodontitis como la sinusitis. La endodoncia permite conservar la dentición natural; sin embargo, la anatomía del sistema de conductos radiculares es compleja y requiere técnicas microquirúrgicas especializadas e instrumentación adecuada para optimizar el éxito técnico. Posteriormente, las restauraciones coronales deben garantizar un sellado eficaz para evitar filtraciones y recontaminación del espacio endodóntico. Sin embargo, existe mayor evidencia sobre el éxito de la extracción dental en el tratamiento de la sinusitis odontogénica ya que la enfermedad pulpar y periapical puede ser erradicada por completo mediante la remoción de la pieza dental⁶.

Se recomienda considerar la cirugía de forma más temprana en pacientes que presentan síntomas sinusales

Tabla 1. Características demográficas, diagnóstico y tratamiento de la serie

N.º paciente	Sexo	Edad	Patología sinusal	Patología odontológica	Tratamiento
1	F	55	Sinusopatía maxilo-etmoidal izquierda	Pieza 26 + quiste apical que compromete piso sinusal.	CENS + exodoncia, cierre de comunicación y relleno óseo para preservación alveolar
2	M	30	Sinusopatía maxilar izquierda	Pieza 26	CENS + exodoncia de 26 + cierre de comunicación buco-sinusal
3	F	35	Sinusopatía maxilar izquierda	Pieza 26	CENS + exodoncia de pieza dental 26 + colgajo de avance vestibular para cierre de esta
4	F	62	Sinusopatía maxilar derecha	Raíces de molar retenido	CENS + exodoncia de pieza dental 17 + colgajo vestibular
5	F	81	Sinusopatía maxilar derecha	Pieza 16	CENS + exodoncia de 16 + colgajo vestibular
6	M	78	Sinusopatía maxilar derecha	Comunicación bucosinusal posexodoncia de pieza 18.	CENS + colgajo bola + vestibular
7	M	38	Sinusopatía maxiloetmoidal izquierda	Pieza 26	CENS + exodoncia de pieza dental 26 + colgajo vestibular
8	F	48	Sinusopatía maxilar izquierda	Pieza 25	CENS + exodoncia de pieza dental 25 + colgajo vestibular
9	F	44	Sinusopatía maxilar izquierda	Pieza 27 + fístula oroantral	CENS + exodoncia de pieza dental 27 + colgajo vestibular
10	F	25	Sinusopatía maxilar derecha + quiste maxilar	3.º molar superior derecho retenido	CENS + exodoncia 18,19
11	M	30	Sinusopatía maxilar izquierda	2.º molar izquierdo en mal estado	CENS + exodoncia de 27 y 28 + colgajo vestibular
12	M	67	Sinusopatía maxilar izquierda	Dos focos odontogénicos en la zona e imagen compatible con pequeña fístula oroantral	CENS + exodoncia de piezas dentales 26 y 27 + colgajo vestibular
13	F	59	Sinusopatía maxilar izquierda	Resto radicular de molar superior izquierdo	CENS + exodoncia de piezas dentales 24 y 27 + colgajo vestibular
14	F	65	Sinusopatía maxilar izquierda	Fístula oroantral posterior a exodoncia	CENS + cierre de fístula con colgajo marginal y vestibular
15	M	76	Sinusopatía maxilar derecha	Pieza dental 17 con proceso infeccioso perirradicular	CENS + exodoncia de pieza 17 y cierre vestibular
16	M	58	Sinusopatía maxilar derecha	Queratoquiste	CENS + resección y cierre vestibular
17	M	84	Sinusopatía maxilar derecha	Pieza 18 con quiste periapical	CENS + exodoncia de 18 + colgajo vestibular
18	F	72	Sinusopatía maxilar izquierda	Pieza 26	CENS + exodoncia de 26 + colgajo vestibular
19	M	59	Sinusopatía maxilar derecha	Puente sobre piezas dentarias 17 y 15, con movilidad asociada, falta de continuidad ósea ápico-periapical	CENS + exodoncia 17 y 15
20	F	61	Sinusopatía maxilar derecha + fístula	Fístula oroantral posterior a exodoncia	CENS + cierre marginal + vestibular
21	M	45	Sinusopatía maxilar izquierda	Piezas 27 y 28	CENS + exodoncia de 27 y 28 + cierre vestibular



Figura 1. Tomografía macizo craneofacial sin contraste. Corte coronal: lesión lítica periapical que afecta al segundo molar superior derecho y sinusopatía maxilar derecha.

intensos con impacto significativo en la calidad de vida, extensión de la enfermedad más allá de los senos o riesgo de desarrollar complicaciones¹. Cuando la inflamación de la sinusitis maxilar se disemina, la infección puede extenderse desde el seno maxilar hacia los tejidos blandos faciales, periorbitarios, intraorbitarios e incluso hacia el seno cavernoso o la región intracraneal, lo que puede provocar celulitis facial, celulitis orbitaria, absceso orbitario y tromboflebitis del seno cavernoso o infección intracraneal³.

El hábito tabáquico y un puntaje elevado en la escala de Lund-Mackay podrían considerarse factores predictivos de una pobre respuesta a la antibioticoterapia, por lo que en estos pacientes podría estar indicada la CENS en una etapa más temprana del tratamiento⁷.

Particularmente en los casos de sinusitis relacionada con implantes dentales, se ha observado una mayor tasa de éxito con la CENS como tratamiento primario, en comparación con la extracción del implante. La extracción de implantes sanos osteointegrados no sería necesaria, a menos que exista periimplantitis o el implante presente movilidad¹. En los casos en que sea posible, la CENS debe realizarse con el objetivo de preservar el implante dental.



Figura 2. Tomografía macizo craneofacial sin contraste. Corte coronal: lesión quística perirradicular en relación con el segundo molar superior con comunicación oroantral y sinusopatía maxilar izquierda.

Es notable que, cuando se abordan tanto la patología dental como la sinusitis, puede esperarse una resolución del cuadro en el 90-100% de los casos con una mejoría sintomatológica sinusal temprana¹. Craig y cols. demostraron, en un estudio prospectivo con 37 pacientes, que aquellos que fueron sometidos a cirugía endoscópica nasosinusal presentaron una resolución significativamente más rápida de los síntomas de sinusitis y de los hallazgos endoscópicos, en un plazo de 7 a 12 días, en comparación con los pacientes tratados inicialmente con endodoncia o extracción dental, cuyo tiempo de resolución fue de 35 a 56 días.

No cabe duda de que tanto el tratamiento rinológico como el dental son necesarios, pero la secuencia y el momento ideal aún son controvertidos. Consideramos que la variante ideal para el tratamiento es realizar una cirugía combinada, es decir, la cirugía endoscópica funcional junto con el abordaje intraoral en un mismo tiempo quirúrgico. Varios estudios informan que una tasa de éxito mediante un abordaje combinado con CENS fue del 97%²⁻⁸.

En nuestra experiencia, preferimos realizar el tratamiento combinado cuando no es factible tratar la patología

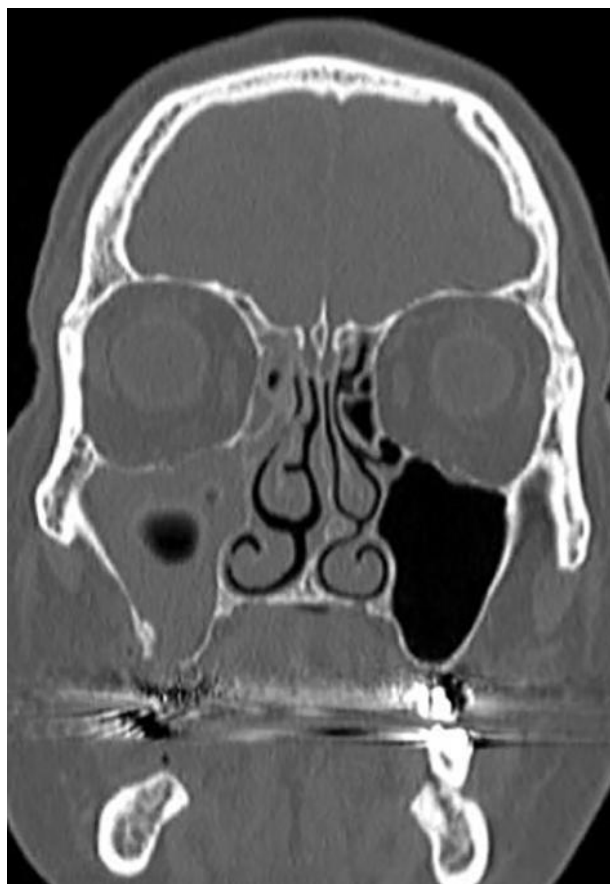


Figura 3. Tomografía macizo craneofacial sin contraste. Corte coronal: comunicación oroantral posterior a exodoncia de pieza dentaria 18 con sinusopatía maxilar derecha.

dental de forma conservadora y cuando es necesario extraer la pieza dental o el implante responsable, con el consiguiente riesgo de generar una comunicación bucosinusal.

Cuando un odontólogo realiza la extracción de la pieza dental o implante en presencia de sinusitis activa, debe considerar la posibilidad de una comunicación bucosinusal y, cuando esta no es adecuadamente tratada, puede desarrollar una fístula oroantral. En estos casos, realizar un cierre inmediato con colgajo sin evaluar previamente el estado del complejo osteomeatal (COM) puede aumentar significativamente el riesgo de fallo en la cicatrización y de fístula. Es posible que la incidencia de fístula sea mayor cuando las extracciones se realizan en el contexto de una sinusitis aguda con presencia de pus drenando a través de la comunicación. Por lo tanto, podría argumentarse que realizar la CENS como tratamiento primario contribuiría a reducir el riesgo de comunicación y fístula oroantral al eliminar previamente la infección del seno maxilar. Como ha sido demostrado por otros autores, el cierre de la comunicación tiene una

probabilidad significativamente mayor de éxito cuando se restablece la permeabilidad del COM⁸. Otro motivo por el cual sería prudente plantear las intervenciones en un mismo tiempo quirúrgico (CENS, extracción dental y cierre de comunicación).

El cierre de la comunicación oroantral suele realizarse mediante un procedimiento quirúrgico, según el tamaño y puede lograrse utilizando diferentes métodos⁹. Suturar la encía circundante podría ser un medio eficaz para el cierre en caso de una fístula que mida menos de 5 mm, mientras que se recomienda un procedimiento con colgajo si la abertura del trayecto fistuloso es mayor de 5 mm. Los procedimientos de colgajo local más ampliamente utilizados incluyen el colgajo bucal, el de grasa bucal (bola de Bichat) o el rotacional palatino¹⁰.

El colgajo bucal consiste en el diseño de un colgajo mucoperiostico trapezoidal de base amplia y su colocación y sutura sobre el defecto. Es sencillo de ejecutar y la literatura ha informado una alta tasa de supervivencia asociada a esta técnica. El colgajo de bola de Bichat se realiza mediante una incisión vestibular seguida de una disección roma y una tracción suave de la grasa hacia el sitio receptor. Sus ventajas incluyen la posibilidad de cierre en doble capa y una cobertura completa por la mucosa oral. El rico suministro sanguíneo permite que la grasa expuesta se epitelice por sí sola. Finalmente, el colgajo palatino presenta un mayor grosor, lo que lo hace menos susceptible de desgarrarse; sin embargo, su manipulación representa un desafío debido a su rigidez y suele asociarse con un mayor dolor posoperatorio¹⁰ (Fig. 4).

Acorde con lo descrito, se extiende el siguiente algoritmo de manejo (Fig. 5).

Consideramos que la colaboración estrecha entre especialistas es fundamental ya que la integración de conocimientos especializados en ambas áreas es necesaria para alcanzar resultados óptimos. Cuando está indicada la intervención quirúrgica en ambos campos, la cirugía combinada simultánea representa una estrategia directa, fiable y curativa, con baja morbilidad, baja incidencia de complicaciones y recuperación generalmente rápida. Este enfoque evita la prolongación de múltiples tratamientos conservadores fallidos y facilita reconstrucciones dentales sin complicaciones.

Se pueden esperar altas tasas de éxito cuando se tratan tanto las patologías dentales como las sinusales; sin embargo, los tratamientos instaurados y el orden en que se realicen los procedimientos pueden variar según el tipo de patología dental y la carga sintomática nasosinusal.

CONCLUSIONES

Nuestro estudio resalta la importancia de un abordaje multidisciplinario entre otorrinolaringólogos y especialistas odontológicos en el manejo de la sinusitis odontogénica, enfatizando la alta tasa de éxito, fiabilidad y predictibilidad de la cirugía combinada en un solo tiempo.

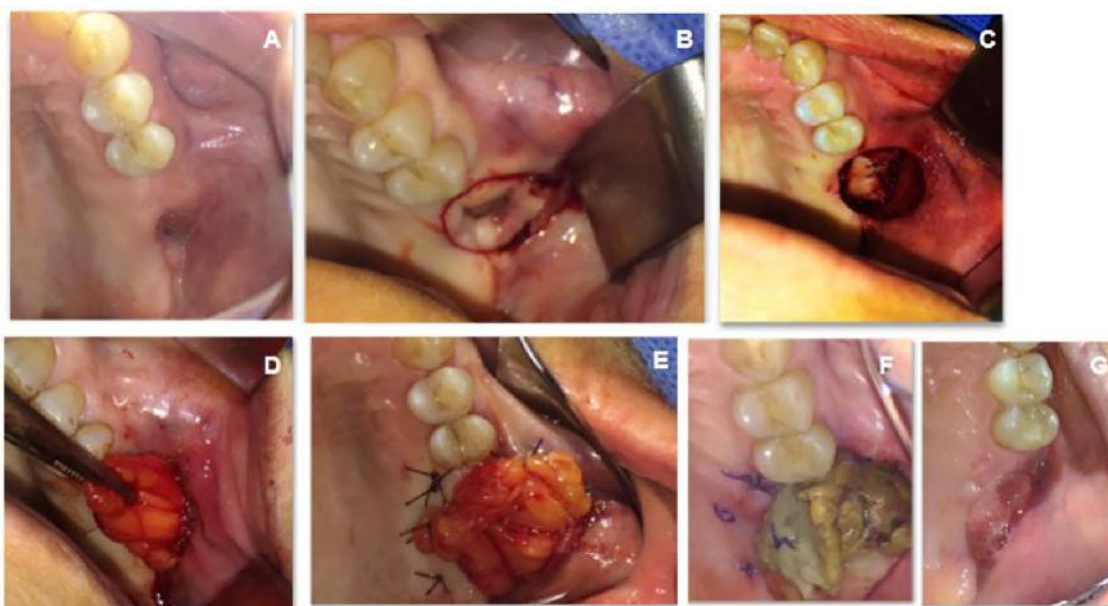


Figura 4. Cavidad oral. A: fistula oroantral crónica. B: tallado de colgajo marginal. C: cierre de defecto con colgajo marginal. D: disección de bola de Bichat. E: cierre en segunda capa con bola de Bichat. F: epitización. G: control al mes, mucosa vestibular completamente epitizada.

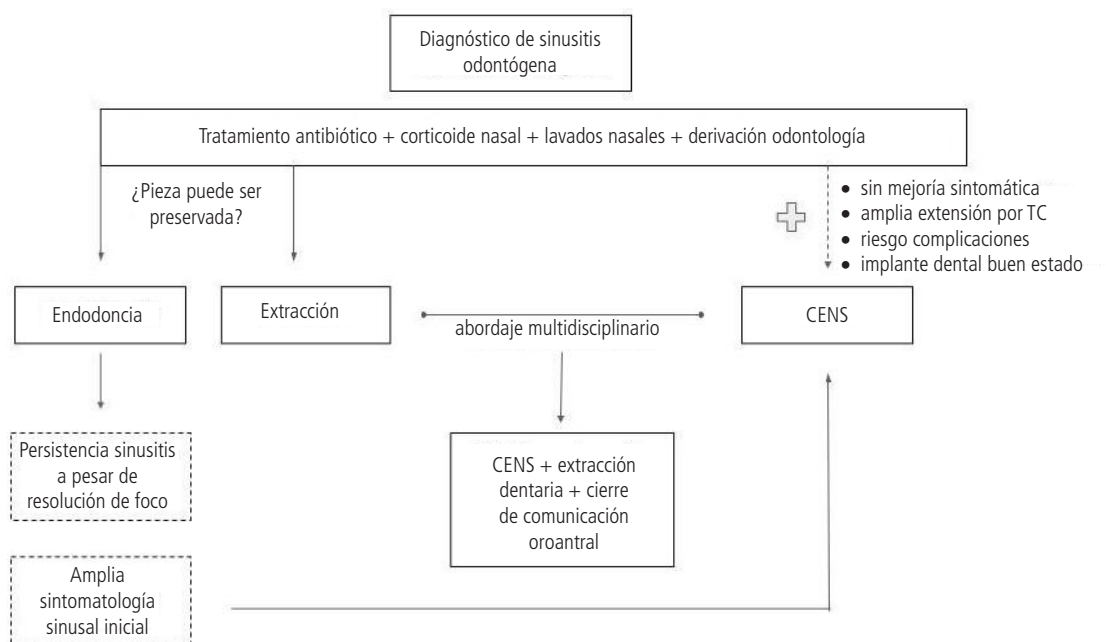


Figura 5. Algoritmo de manejo: abordaje multidisciplinario.

Contribución de los autores: Conceptualización, Investigación (CS, ZF, AS). Curación de datos, Análisis formal, Metodología, Administración del proyecto (CS).

Conflicto de intereses: los autores declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: los autores declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Craig JR. Odontogenic sinusitis: a state-of-the-art review. World J Otorhinolaryngol Head Neck Surg. 2022;8(1):8-15.<https://doi.org/10.1002/wjo2.9>.

2. De Corso E, Rigante M, Mele DA, et al. Real-life experience in the management of sinonasal complications of dental disease or treatments. *J Pers Med*. 2022;12(12):2078.<https://doi.org/10.3390/jpm12122078>.
3. Lin J, Wang C, Wang X, et al. Expert consensus on odontogenic maxillary sinusitis multi-disciplinary treatment. *Int J Oral Sci*. 2024;16(1):11. <https://doi.org/10.1038/s41368-024-00278-z>.
4. Harvey R. Dental infection and sinusitis...more common than we think. Darlinghurst: University of New South Wales. Sydney ENT Clinic; 2025 Nov 7 [citado 2026 feb 23]. Disponible en:<https://richardharvey.sydneyclinic.com/2025/11/07/dental-infection-and-sinusitis-more-common-than-we-think/>.
5. Yoo BJ, Jung SM, Lee HN, et al. Treatment strategy for odontogenic sinusitis. *Am J Rhinol Allergy*. 2021;35(2):206-212.<https://doi.org/10.1177/1945892420946969>.
6. Pannkuk TF, Craig JR, Tušas P, et al. Management of endodontic disease for odontogenic sinusitis. *Otolaryngol Clin North Am*. 2024;57(6):1119-1138.<https://doi.org/10.1016/j.otc.2024.07.002>.
7. Mattos JL, Ferguson BJ, Lee S. Predictive factors in patients undergoing endoscopic sinus surgery for odontogenic sinusitis. *Int Forum Allergy Rhinol*. 2016;6(7):697-700.<https://doi.org/10.1002/alr.21736>.
8. Kocum P, Šedý J, Traboulsi J, et al. One-stage combined ENT and dental surgical treatment of odontogenic sinusitis: a prospective study. *Eur Arch Otorhinolaryngol*. 2024;281(3):1347-1356.<https://doi.org/10.1007/s00405-023-08332-y>.
9. Visscher SH, van Minnen B, Bos RRM. Closure of oroantral communications: a review of the literature. *J Oral Maxillofac Surg*. 2010;68(6):1384-1391.<https://doi.org/10.1016/j.joms.2009.07.044>.
10. Kwon MS, Lee BS, Choi BJ, et al. Closure of oroantral fistula: a review of local flap techniques. *J Korean Assoc Oral Maxillofac Surg*. 2020;46(1):58-65.<https://doi.org/10.5125/jkaoms.2020.46.1.58>.

El papel del *benchmarking* en la gestión sanitaria: más allá de la medición, una estrategia para aprender y mejorar

María F. Grande Ratti¹, Noelia S. De Masi², José A. Orozco Niño¹ y Sebastián Marciano³

1. Área de Investigación en Medicina Interna, Servicio de Clínica Médica, Hospital Italiano. Buenos Aires, Argentina

2. Departamento Experiencia del Cliente, Hospital Italiano. Buenos Aires, Argentina

3. Secretaría de Investigación Clínica, Universidad Hospital Italiano. Buenos Aires, Argentina

RESUMEN

Introducción: el *benchmarking* se ha consolidado como una herramienta estratégica clave en la gestión sanitaria, alineada con el principio de que aquello que no se mide no puede mejorarse. Aunque es un anglicismo, su uso está plenamente aceptado en la literatura científico-sanitaria. Entendido como un proceso sistemático de medición, comparación y aprendizaje, permite evaluar el desempeño organizacional a partir de indicadores válidos y comparables. En un contexto de creciente demanda de eficiencia, seguridad y resultados centrados en el paciente, el *benchmarking* facilita la mejora continua de la calidad asistencial y el posicionamiento institucional mediante la transparencia y el aprendizaje organizacional.

Estado del arte: la evidencia muestra que el *benchmarking* en salud se apoya en el uso de indicadores de estructura, proceso y resultado, según el modelo de Donabedian. Diversas revisiones sistemáticas han demostrado una asociación positiva entre la implementación de *benchmarking* y mejoras en la calidad asistencial, tanto en procesos como en resultados clínicos. No obstante, su impacto suele potenciarse cuando se combina con intervenciones complementarias como auditorías, retroalimentación, planes de mejora o incentivos, lo que dificulta aislar su efecto específico. Asimismo, persisten desafíos metodológicos relevantes, como la comparabilidad entre instituciones, la necesidad de ajustes por riesgo, la calidad y disponibilidad de los datos, y la sostenibilidad de las mejoras a largo plazo. En este marco, adquiere creciente relevancia la integración de indicadores de experiencia del paciente y datos cualitativos, que amplían la evaluación tradicional del desempeño hacia una perspectiva más integral y centrada en la persona.

Discusión: no debe entenderse como una simple comparación ni como un sustituto de los programas de calidad, sino como un componente esencial de estos. Su implementación efectiva requiere sistemas de información robustos, indicadores bien definidos, equipos interdisciplinarios, liderazgo institucional y una cultura organizacional orientada al aprendizaje y la mejora continua. En conclusión, la ciencia de la implementación resulta clave para cerrar la brecha entre la evidencia y la práctica, y la evaluación comparativa basada en datos de alta calidad se ha vuelto una práctica prácticamente obligatoria en los sistemas de salud modernos.

Palabras clave: benchmarking, gestión de la información en salud, calidad de la atención en salud, indicadores de gestión, mejoramiento de la calidad, satisfacción del paciente.

Autor para correspondencia: maria.grande@hospitalitaliano.org.ar, Grande Ratti MF.

Recibido: 3/02/2026 | Aceptado: 15/09/2026 | Publicado: 28/09/2026

DOL: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1317>

Cómo citar: Grande Ratti MF, De Masi NS, Orozco Niño JA, Marciano S. El papel del *benchmarking* en la gestión sanitaria: más allá de la medición, una estrategia para aprender y mejorar. *Rev Hosp Ital B. Aires.* 2026;46(3):e0001317

The Role of Benchmarking in Healthcare Management: Beyond Measurement, a Strategy for Learning and Improvement

ABSTRACT

Introduction: benchmarking has become a key strategic tool in healthcare management, aligned with the principle that what is not measured cannot be improved. Although the term is an anglicism, its use is fully accepted in the scientific and healthcare literature. Understood as a systematic process of measurement, comparison, and learning, it allows for the evaluation of organizational performance based on valid and comparable indicators. In a context of increasing demand for efficiency, safety, and patient-centered outcomes, benchmarking facilitates the continuous improvement of healthcare quality and institutional positioning through transparency and organizational learning.

State of the art: evidence shows that benchmarking in healthcare relies on the use of structure, process, and outcome indicators, according to the Donabedian model. Several systematic reviews have demonstrated a positive association between the implementation of benchmarking and improvements in healthcare quality, both in processes and clinical outcomes. However, its impact is usually enhanced when combined with complementary interventions such as audits, feedback, improvement plans, or incentives, which makes it difficult to isolate its specific effect. Furthermore, significant methodological challenges persist, such as comparability between institutions, the need for risk adjustments, data quality and availability, and the long-term sustainability of improvements. Within this framework, the integration of patient experience indicators and qualitative data is gaining increasing importance, broadening the traditional performance evaluation toward a more holistic and person-centered perspective.

Discussion: this should not be understood as a simple comparison or a substitute for quality programs, but rather as an essential component of them. Its effective implementation requires robust information systems, well-defined indicators, interdisciplinary teams, institutional leadership, and an organizational culture oriented toward learning and continuous improvement. In conclusion, implementation science is key to bridging the gap between evidence and practice, and benchmarking based on high-quality data has become virtually mandatory in modern healthcare systems.

Keywords: benchmarking, health information management, quality of health care, management indicators, quality improvement, patient satisfaction.

INTRODUCCIÓN

La presente revisión narrativa tiene como objetivo integrar y analizar los principales componentes conceptuales, metodológicos y organizacionales que condicionan la utilidad del *benchmarking* en la gestión sanitaria. Si bien la medición y comparación sistemática del desempeño constituyen herramientas fundamentales para identificar brechas de calidad y orientar procesos de mejora, la utilidad no depende únicamente de la disponibilidad de indicadores, sino también de su pertinencia, calidad y comparabilidad, así como de la adecuada interpretación de las diferencias observadas en función del contexto en el que se producen.

Desde tal perspectiva, este artículo aborda los fundamentos conceptuales del *benchmarking* y sus principales modalidades de aplicación en salud, examinando el papel de los indicadores clave de desempeño (KPIs por *Key Performance Indicators*), los desafíos asociados a la medición y comparabilidad del desempeño, así como la incorporación de dimensiones vinculadas con la experiencia del paciente.

Asimismo, se analizan las condiciones organizacionales necesarias para que los resultados de la comparación puedan trascender la identificación de diferencias y traducirse en aprendizaje, adopción de mejores prácticas y acciones sostenibles de mejora.

La medición de indicadores permite cuantificar determinados aspectos del desempeño; la monitorización del desempeño implica el seguimiento sistemático de estos indicadores a lo largo del tiempo; los tableros de gestión (*dashboards*) constituyen herramientas para visualizar y comunicar información de desempeño; y los programas de mejora de la calidad comprenden intervenciones destinadas a modificar los procesos y resultados. En contraste, se entiende el *benchmarking* como un proceso sistemático de comparación del desempeño con un referente, orientado a identificar brechas, comprender diferencias y favorecer el aprendizaje y la mejora.

La contribución consiste en articular estas dimensiones desde la gestión sanitaria, destacando que no debe entenderse únicamente como una herramienta

de medición, sino como un mecanismo de aprendizaje organizacional y de mejora continua, que permite identificar referentes, comprender las diferencias observadas y analizar los procesos que podrían explicarlas. El valor no reside en identificar quién obtiene mejores resultados, sino en comprender las diferencias, interpretar los factores que las explican y determinar qué prácticas pueden ser adaptadas al contexto propio. Esta revisión busca aportar un marco conceptual con orientación más pragmática para comprender bajo qué condiciones el *benchmarking* puede constituirse en una herramienta efectiva para mejorar el desempeño y favorecer una atención sanitaria más equitativa, transparente y orientada a la mejora continua.

ESTADO DEL ARTE

“BENCHMARKING” EN SALUD: DEFINICIÓN Y FUNDAMENTOS

La frase “Lo que no se puede medir, no se puede mejorar” suele atribuirse a Peter Drucker, por parafraseos conceptuales similares en sus obras, que hacen referencia a la importancia de medir para gestionar. Sin embargo, fue publicada por primera vez por Ridgway en 1956¹. En salud, el *benchmarking* es un proceso que se basa precisamente en medir, comparar y aprender.

Ante todo, reconocemos que se trata de un anglicismo cuya traducción al español podría ser “evaluación comparativa” o “comparación con estándares de referencia”, aunque no resulte plenamente equivalente ni capte con precisión el uso consolidado del término en la literatura internacional de gestión y evaluación en salud. Dado que el *benchmarking* es un concepto técnico ampliamente aceptado en el ecosistema científico-sanitario, en el presente artículo mantendremos la denominación original en inglés.

En el ámbito sanitario puede entenderse como un proceso sistemático de medición, comparación e interpretación del desempeño respecto de referentes definidos, orientado al aprendizaje y la mejora. Se trata de un campo en expansión y se ha incorporado como una herramienta estratégica de gestión orientada a la mejora continua de la calidad asistencial. Etimológicamente, proviene del concepto de “punto de referencia” (*benchmark*) y su uso se extendió desde el ámbito empresarial, adquiriendo una formulación sistemática como herramienta de gestión a partir de la experiencia de Xerox durante la década de 1970, y posteriormente documentada por Camp en 1989².

En un entorno cada vez más exigente en términos de eficiencia, seguridad y resultados centrados en el paciente³, el *benchmarking* ofrece un marco estructurado para la evaluación y el aprendizaje de una organización. Además de impulsar mejoras internas, permite a las instituciones posicionarse al mostrar resultados transparentes y sostenidos en la calidad, la seguridad y la experiencia.

Recientemente, Byrne y Haut nos recordaron que la atención de los pacientes debe cumplir con un estándar, independientemente de la ubicación geográfica donde se produzca la enfermedad/lesión, o del centro de atención, o del equipo profesional tratante⁴. En otras palabras, siempre se debería garantizar el tratamiento adecuado, en el lugar y el momento apropiados para lograr el mejor

resultado posible⁵. Desafortunadamente, la persistencia de variaciones en el desempeño asistencial pone de manifiesto la necesidad de contar con métodos que permitan medir, comparar e interpretar sistemáticamente esas diferencias. El *benchmarking* fomenta el uso de indicadores válidos, comparables y ajustados, agrupados por factores de riesgo, complejidad o patologías en salud. No es una simple comparación, sino que requiere análisis crítico, interpretación contextual y capacidad de adaptación local. Su uso también abarca la gestión (p. ej., rediseño de procesos y espacios, comunicación, experiencia del paciente), ya que facilita la mejora continua y la toma de decisiones basada en evidencia, orienta la priorización de oportunidades de mejora y promueve el aprendizaje organizacional a partir de experiencias propias y de otras instituciones, contribuyendo así a fortalecer la experiencia de pacientes, acompañantes y equipos de trabajo.

Su eficacia aumenta cuando los equipos participan activamente en la definición e interpretación de los indicadores (instrumentos de medición), y cuando existe liderazgo institucional para traducir los hallazgos en acciones de mejora⁶. Esta metodología se sirve de indicadores correspondientes a las tres dimensiones de calidad (estructura, procesos y resultados) para evaluar y comparar sistemáticamente el desempeño entre organizaciones, servicios o procesos, establecer referentes e identificar prácticas potencialmente transferibles cuya adaptación al contexto local pueda ser evaluada⁷.

TIPOS DE BENCHMARKING

Ya en el año 1994, Benson informaba cuatro tipos diferentes⁸ (Tabla 1).

- Interno, que compara las funciones dentro de una organización.
- Competitivo, ya que compara productos o servicios directamente, que operan en el mismo mercado.
- Funcional, que compara un aspecto o una función específica, como nóminas o compras (p. ej., gestión de turnos), con organizaciones que no necesariamente pertenecen al mismo sector, pero que realizan ese proceso de manera destacada.
- Genérico, permite una comparación más amplia ya que busca identificar y adaptar principios de gestión, enfoques estratégicos o metodologías transversales aplicables a cualquier organización. Se trata de mirar cómo otras empresas hacen las cosas de manera eficiente y pensar en cómo adaptar esos métodos a nuestro contexto. El pensamiento Lean, inspirado en el sistema de producción Toyota, busca maximizar el valor para el usuario mediante la identificación y eliminación de actividades que no agregan valor, la reducción del desperdicio y la mejora continua de los procesos⁹.

INDICADORES Y KPIs: SELECCIÓN, DEFINICIÓN Y ESTANDARIZACIÓN

Según el modelo de Donabedian, los indicadores pueden clasificarse en *estructura*, *procesos* y *resultados*¹⁰. La tabla 2 muestra ejemplos en la atención clínica de personas con diabetes.

Tabla 1. Tipos de *benchmarking* según Benson con ejemplos clínicos y de gestión

Tipo	Ejemplo clínico	Ejemplo de gestión
Interno	Comparación entre distintos servicios de urgencias de una misma institución (Hospital Italiano)	Comparación de la eficiencia en la asignación de turnos entre diferentes unidades administrativas del mismo establecimiento
Competitivo	Tasa de reconsultas a las 72 horas en la Central de Emergencia de Adultos, sede central vs. San Justo Comparación de mortalidad en lista de espera o tiempo de resolución quirúrgica entre hospitales de una misma red	Demora (en días) para la asignación de turno ambulatorio programado entre Ginecología/Obstetricia vs. Hematología
Funcional	Comparación de la gestión del flujo del paciente crítico con centros de emergencias de otro sistema que han optimizado triaje y derivación	Comparación del costo por paciente atendido entre centros de salud de la misma región Comparación de la gestión de turnos o logística de insumos con organizaciones externas
Genérico	Reducir tiempos de espera en consultas externas o mejorar el recorrido del paciente	Un hospital podría inspirarse en la logística de supermercados para organizar la entrega de medicamentos

Fuente: elaboración propia

Tabla 2. Clasificación de indicadores según el modelo de Donabedian

Categoría	Definición	Ejemplo
Estructura	Recursos físicos, humanos y organizacionales disponibles para la atención de salud	Porcentaje de centros ambulatorios que cuentan con al menos un profesional certificado en educación diabetológica Proporción de centros con disponibilidad de HbA1c en laboratorio
Procesos	Actividades y procedimientos mediante los cuales se brinda la atención	Porcentaje de pacientes con HbA1c medida al menos dos veces en los últimos 12 meses Porcentaje de pacientes con evaluación de fondo de ojo anual
Resultados	Cambios en el estado de salud atribuibles a la atención recibida	Proporción de pacientes con HbA1c < 7% (o según metas individualizadas) Porcentaje con LDL < 100 mg/dL (o < 70 mg/dL en alto riesgo)

Fuente: elaboración propia

Su aplicación en salud está relacionada con un trabajo de equipo, que incluya al médico especialista en tema/patología/población que se está evaluando, así como otros miembros con conocimientos técnicos (p. ej., informática, epidemiología, bioestadística, análisis de datos y otras disciplinas afines que permitan la adecuada reutilización de la información derivada de la historia clínica electrónica) y de comunicación –para la medición cuantitativa y su difusión–, y experiencia en gestión de proyectos¹¹.

Los objetivos de las métricas habitualmente más utilizadas en salud (productividad, calidad, tiempos y costos) son fundamentalmente los siguientes:

- Identificar oportunidades de mejora en procesos clínicos (p. ej., tiempo a antibióticos, mortalidad), administrativos o de gestión (p. ej., readmisiones).
- Adoptar mejores prácticas que hayan demostrado ser eficaces en contextos similares (p. ej., resultados en procedimientos complejos como trasplante hepático).
- Reducir la variabilidad injustificada entre unidades, servicios o instituciones.
- Mejorar la toma de decisiones mediante datos comparativos y evidencia.
- Estimular una cultura de calidad centrada en el aprendizaje organizacional (p. ej., establecer planes de mejora tras el análisis de desviaciones).

Dado que los procesos de resultados son complejos y costosos de medir, muchos sistemas sanitarios han recurrido a restringir el uso de mediciones indirectas y más sencillas, como las de estructura y de proceso¹².

Los indicadores clave de desempeño (KPIs por *Key Performance Indicators*) y su construcción deben considerarse entre los pasos principales para estructurar un sistema de monitorización continua que permita evaluar desempeño, identificar brechas y guiar acciones de mejora¹³. Para el diseño de los paneles de gestión hospitalaria, la propia participación de los usuarios principales en la selección y la priorización contribuye a su éxito¹⁴.

Una revisión sistemática de 2024 identificó y clasificó 64 indicadores clave (financieros, gerenciales y clínicos) utilizados para medir el desempeño en la cadena de abastecimiento hospitalaria¹⁵. El estudio remarcó que, para los hospitales, identificar y generar estos indicadores representa un enorme desafío, debido a la frecuente falta de fuentes de datos completas y comparables. El problema mayor radica en que la validez y la fiabilidad dependen de la disponibilidad de datos de alta calidad, en concreto, de los elementos relevantes del conjunto mínimo de datos¹⁶.

Un estudio italiano de 2013 revisó la evidencia y los datos hospitalarios para analizar la relación entre el volumen de actividad (cantidad de casos atendidos) y los resultados sanitarios (especialmente mortalidad intrahospitalaria a 30 días)¹⁷. No obstante, los propios autores advierten que no siempre fue posible fijar umbrales universales y que las diferencias entre condiciones o patologías clínicas obligan a un enfoque contextualizado¹⁷.

Desde la perspectiva de Experiencia del Paciente, el *benchmarking* incorpora una dimensión adicional de interés poblacional al incluir indicadores de percepción

y de participación de pacientes/familias. La integración sistemática de esta información cualitativa en los *dashboards* no solo permite comparar resultados clínicos y de proceso, sino también comprender cómo dichos resultados son vividos y significados por los propios usuarios. En consecuencia, es fundamental integrar datos cualitativos (p. ej., voz del paciente, encuestas de experiencia y relatos narrativos), dentro de los indicadores de resultado¹⁸. Estos insumos aportan dimensiones que no pueden ser captadas exclusivamente por mediciones cuantitativas y permiten una comprensión más profunda¹⁹. No obstante, su interpretación exige una adecuada contextualización según las características de la población atendida, el tipo de servicio y las particularidades de cada proceso. Este enfoque amplía la interpretación tradicional del desempeño institucional, al aportar información contextual, subjetiva y relacional, que resulta esencial para evaluar la calidad del cuidado de manera integral y centrada en la persona.

Las iniciativas orientadas a estandarizar resultados relevantes para las personas, como ICHOM (por International Consortium for Health Outcomes Measurement)²⁰, desplazan el foco desde la comparación exclusiva de actividad o desempeño operativo hacia la comparación de resultados que importan a quienes reciben la atención, ayudando a operacionalizar qué resultados deberían medirse. En este sentido, han desarrollado conjuntos estandarizados de resultados para determinadas patologías (p. ej., fibrilación auricular, enfermedad renal crónica, insuficiencia cardíaca)^{21–23}.

De manera similar, PROMs (por *Patient-Reported Outcome Measures*) y PREMs (por *Patient-Reported Experience Measures*) son fuentes de información para construir medidas de desempeño que permiten incorporar la perspectiva de las personas²⁴. Mientras que las primeras evalúan qué resultado tiene la enfermedad o la atención sobre la salud de la persona, desde su propia perspectiva (p. ej., calidad de vida, capacidad funcional)²⁵, las segundas abordan cómo fue la experiencia¹⁸. Entonces, PROMs/PREMs proporcionan mecanismos para capturar directamente la perspectiva de las personas. Según Jamieson Gilmore y cols.²⁶, los PREMs se utilizan principalmente “a nivel micro” para mejorar la calidad de la atención, mientras que su aplicación “a nivel meso” se vincula con la mejora del desempeño y la financiación basada en calidad, y su uso “a nivel macro” aún continúa siendo limitado.

CALIDAD Y GOBERNANZA DE LOS DATOS

La evidencia destaca la importancia de los datos para tomar decisiones informadas. Por ende, la recolección y la gestión de datos constituyen el pilar. Medir, analizar e interpretar datos permite no solo conocer el estado actual del sistema de salud, sino también anticipar necesidades, corregir desviaciones y promover mejoras continuas.

Sin embargo, la disponibilidad de grandes volúmenes de información no garantiza, por sí misma, que las comparaciones sean válidas. Los datos utilizados para *benchmarking* deben ser suficientemente completos, exactos, consistentes, oportunos, pertinentes y comparables, y contar con definiciones operativas y reglas de cálculo

claramente establecidas (p. ej., criterios homogéneos de inclusión y exclusión, períodos de observación equivalentes y procesos suficientemente estandarizados como para que las diferencias observadas puedan atribuirse al desempeño y no a diferencias en la forma de registrar o brindar la atención). Una revisión sistemática reciente identificó múltiples dimensiones de calidad de los datos en los sistemas de información sanitaria, entre ellas exactitud, completitud, consistencia, oportunidad, fiabilidad, accesibilidad y seguridad, destacando la heterogeneidad existente en su definición y evaluación²⁷.

En este contexto, la gobernanza de datos complementa la evaluación de su calidad. No se limita a aspectos técnicos, sino que comprende definir qué información se recolectará, con qué propósito, mediante qué fuentes, utilizando qué definiciones y bajo qué procedimientos de validación. También requiere establecer mecanismos para detectar errores, datos faltantes, cambios en los procesos de registro y modificaciones en los sistemas de información²⁸.

Entonces, la medición requiere un proceso sistemático y estandarizado que garantice la validez (si los datos realmente miden lo que se supone que deberían, ¿estamos midiendo lo correcto?), fiabilidad (si los datos son consistentes y estables cuando se miden repetidas veces, ¿da el mismo resultado cada vez que se mide?) y comparabilidad de los datos (si los datos se pueden co-tejar con otros de diferentes lugares, tiempos o grupos, ¿podemos comparar estos datos con otros y que dicha

comparación tenga sentido?), para lo cual existen 4 pasos fundamentales:

- Definición clara de los indicadores: fuente de datos, método de cálculo, frecuencia de medición y umbrales esperados. Por ejemplo, mortalidad en lista de espera para trasplante hepático en adultos, estimada considerando el trasplante como riesgo competitivo, informada anualmente y comparada con los datos del mismo año de la United Network for Organ Sharing (UNOS). En una aproximación más realista, no basta con hacer un Kaplan-Meier simple, porque eso podría sobreestimar o subestimar el riesgo de muerte al no considerar que algunos se trasplantan antes de morir.

- Sistemas de información integrados: historias clínicas electrónicas, bases de datos administrativas, registros específicos.

- Uso de tecnologías digitales: herramientas de visualización que permitan procesar grandes volúmenes de información en tiempo real, por ejemplo *dashboards*⁵.

- Capacitación del personal: todos los actores implicados deben comprender la importancia de registrar los datos adecuadamente y utilizarlos éticamente²⁹.

El “Marco 5-P” del proceso de selección de indicadores (Fig. 1) consta de cinco dominios³⁰: Propósito (objetivo mediante un proceso operativo o alcance mediante un plan estratégico), Política (gobernanza o comité directivo), Preparación (búsqueda bibliográfica y opinión del usuario final), Procedimiento (método reconocido de construcción de consenso) y Probar (pruebas piloto de calidad, entrenamientos).

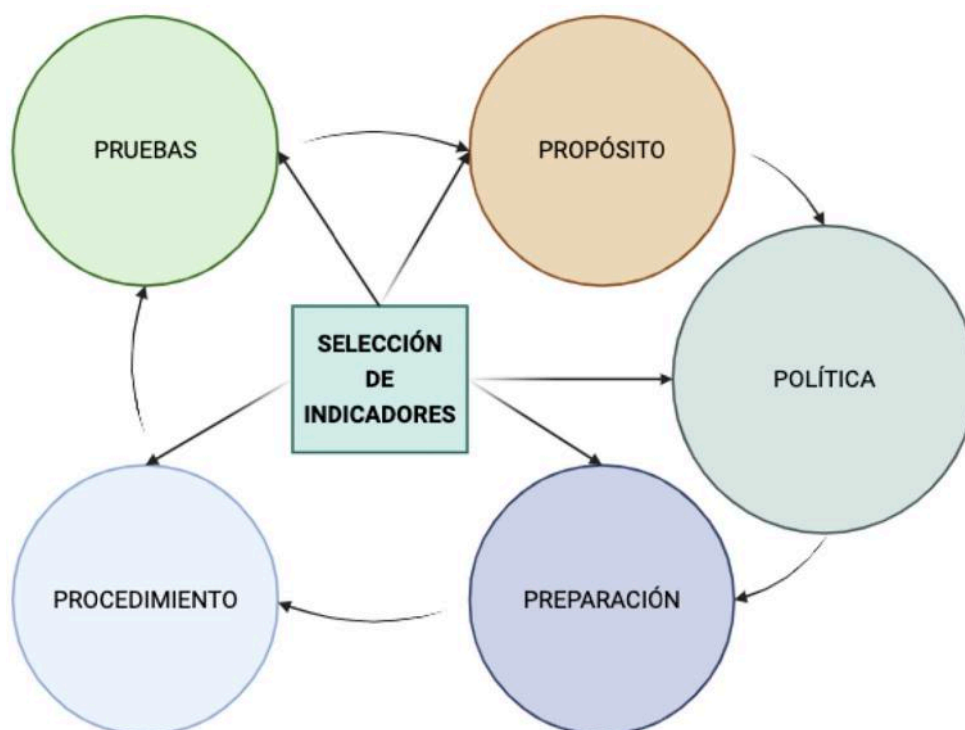


Figura 1. Los cinco dominios ('5-P') del proceso de selección de indicadores.

Fuente: elaboración propia, creada con BioRender.com®.

¿QUÉ SIGNIFICA COMPARAR EN SALUD?

Implica mucho más que disponer de indicadores y establecer diferencias entre dos o más organizaciones, servicios o unidades asistenciales. La comparación permite identificar variaciones en el desempeño y formular preguntas acerca de su magnitud, consistencia y posible relevancia. Sin embargo, una diferencia observada no necesariamente representa una diferencia de calidad. Puede reflejar diferencias en las características de las poblaciones atendidas, complejidad de los casos, disponibilidad de recursos, estructura organizacional, modelos de atención, sistemas de información o contexto socioeconómico. Estas diferencias adquieren especial relevancia en América Latina, donde coexisten sistemas de salud fragmentados y heterogéneos, con variaciones importantes en cobertura, acceso, disponibilidad de recursos y capacidad de los sistemas de información entre países y, en algunos casos, dentro de un mismo país. Por ello, interpretar una comparación requiere considerar tanto las características de las unidades comparadas como las condiciones en las que se generan los resultados.

Este aspecto resulta particularmente relevante en salud, donde la heterogeneidad entre instituciones y poblaciones es considerable. La comparación de tasas de mortalidad, reingreso, complicaciones, tiempos de espera, utilización de recursos o experiencia de los pacientes puede ser informativa, pero su interpretación requiere determinar si las diferencias observadas son esperables, atribuibles al riesgo basal de las poblaciones, consecuencia de diferencias en los procesos de atención o potencialmente evitables. En consecuencia, la identificación de una brecha constituye el punto de partida de la indagación y no necesariamente una conclusión sobre el desempeño. Una comparación orientada al aprendizaje busca avanzar hacia preguntas de mayor profundidad: ¿por qué existen estas diferencias?, ¿qué procesos, prácticas o condiciones podrían explicarlas?, ¿qué podemos aprender de quienes obtienen mejores resultados? y ¿qué elementos podrían adaptarse a nuestro propio contexto? Es en este punto donde la comparación adquiere sentido como componente del *benchmarking*.

En el contexto de la evaluación comparativa es importante reconocer que los indicadores pueden no ser directamente comparables entre instituciones debido a diferencias en sus definiciones operativas, las características de las poblaciones atendidas, los procesos asistenciales y el contexto organizacional, que Byrne y Haut describieron haciendo uso de la expresión *comparing apples to oranges*, que se conoce coloquialmente en nuestro contexto como “peras con manzanas”⁴. Particularmente en América Latina y la Argentina, la heterogeneidad entre instituciones, modelos de atención, disponibilidad de recursos y sistemas de información puede limitar la comparabilidad directa. Incluso los modelos de ajuste de riesgo empleados para generar informes de *benchmarking* podrían no capturar plenamente las diferencias en las características de los pacientes atendidos entre hospitales, lo que puede conducir a comparaciones potencialmente injustas.

La interpretación de las comparaciones en salud requiere integrar consideraciones epidemiológicas, estadísticas y económicas, dado que las poblaciones y los contextos asistenciales rara vez son idénticos. Por ejemplo, si una guardia hospitalaria presentara una demora promedio de dos horas y el *benchmark* establece una hora, la interpretación de esta diferencia requiere considerar el “contexto” y la metodología utilizada para el cálculo del indicador. En este mismo sentido, Van Vliet y cols. compararon tres hospitales oftalmológicos del Reino Unido, Estados Unidos y Países Bajos³¹. A pesar de tratarse de importantes centros internacionales de atención terciaria y formación en Oftalmología, los autores se vieron obligados a informar factores ambientales que influyen en los procedimientos de cataratas, viéndose obligados a proporcionar detalles sobre la ubicación, el tipo, la condición de hospital universitario y el volumen operativo de las organizaciones.

De forma similar, Byrne y col. citaron un ejemplo quirúrgico en hospitales de alta complejidad, donde algunos reciben derivaciones de casos de trauma graves (p. ej., heridas de bala), que incrementan su riesgo basal de mortalidad⁴. En tales casos, se sugiere excluir a esta subpoblación para permitir comparaciones más equitativas y justas frente a hospitales que no reciben esas derivaciones.

COMPARABILIDAD E INTERPRETACIÓN DE LAS DIFERENCIAS (Tabla 3)

Ya hacia el año 2006, Arah y cols. hablaban fundamentalmente de un marco metodológico para la medición y comparación del desempeño de sistemas, particularmente de indicadores de calidad de la atención sanitaria entre países pertenecientes a la Organización para la Cooperación y el Desarrollo Económicos (OCDE, es decir, Organization for Economic Cooperation and Development)³². El punto de partida de los autores es muy interesante: antes de seleccionar indicadores hay que preguntarse: ¿El desempeño de qué estamos midiendo, y con qué propósito? Distinguen fundamentalmente entre dos tipos: un marco de desempeño en salud, que adopta una perspectiva más amplia, social o de salud pública, sobre la determinación de la salud; mientras que un marco de desempeño en atención adopta una perspectiva más restringida, principalmente clínica o técnica, sobre la atención médica en relación con las necesidades de salud. Los autores argumentan que el *benchmarking* no debería comenzar simplemente por elegir indicadores y comparar números en un microsistema, sino que debería ir más allá de la medición y favorecer el aprendizaje de procesos.

Kruk y cols. en 2018³³ plantearon que históricamente se ha puesto mucho énfasis en cobertura, acceso y utilización, mientras que la calidad de la atención ha recibido relativamente menos atención. Esto genera una situación paradójica: aumentar la cobertura de servicios de baja calidad puede producir poco beneficio e incluso daño. En consecuencia, propusieron pensar la calidad como algo que debe estar presente en todo el sistema sanitario, y no solamente como características de una prestación indivi-

Tabla 3. De la definición de resultados a la comparación del desempeño: conceptos y herramientas clave

Concepto	¿Para qué sirve?
Benchmarking	Permite comparar el desempeño y aprender de las diferencias
VBHC	Ayuda a definir qué resultados importan
ICHOM	Aporta un ejemplo de estandarización sobre qué y cómo medir (para que los resultados sean comparables)
PROMs/PREMs	Contribuyen a medir resultados y experiencia desde la perspectiva del paciente
VBHC: <i>Value-Based Healthcare</i>	
ICHOM: International Consortium for Health Outcomes Measurement	
PROMs: <i>Patient-Reported Outcome Measures</i> . PREMs: <i>Patient-Reported Experience Measures</i>	

dual. Por ende, su aporte es más amplio: proporciona una justificación conceptual para medir y comparar calidad, identificar brechas y utilizar esa información para orientar la mejora. El artículo utiliza datos comparativos y muestra grandes variaciones en indicadores de calidad de atención, incluso entre países con niveles similares de desarrollo o recursos. La comparación permite hacer visible una cuestión fundamental: que existan diferencias de desempeño entre organizaciones abre la posibilidad de preguntarse qué están haciendo diferente aquellos que obtienen mejores resultados, haciendo hincapié en cómo transformar la comparación en aprendizaje organizacional.

Los indicadores permiten detectar esta brecha; el *benchmarking* busca comprenderla y aprender de quienes obtienen mejores resultados. El marco conceptual de atención basada en valor (VBHC, por *Value-Based Healthcare*)³⁴ ayuda a responder preguntas complementarias fundamentales: ¿Qué entendemos por valor? o ¿qué resultados importan? o ¿qué significa realmente “mejor desempeño” en salud?

DEL BENCHMARKING A LA MEJORA CONTINUA

En una revisión sistemática del año 2022 se incluyeron 17 trabajos que evidenciaron una asociación positiva entre la implementación de *benchmarking* y las mejoras en la calidad asistencial: 10 estudios mostraron mejoras en procesos, 13 en resultados y 7 en ambos tipos de indicadores³⁵. Willmington y cols. informaron que la mayoría de los estudios (12 de 17) combinaron el *benchmarking* con al menos una intervención complementaria (p. ej., reuniones entre participantes clave, auditorías y retroalimentación, planes de mejora y/o incentivos financieros), y en 10 estudios se usaron múltiples estrategias de forma simultánea³⁵. Esto último dificulta discernir cuál es el efecto atribuible estrictamente al *benchmarking* y subraya que su eficacia parece reforzarse cuando se acompaña de intervenciones adicionales. Otra limitación mencionada es que muchos estudios tienen períodos de seguimiento cortos (p. ej., entre seis meses y dos años), lo que genera dudas sobre la sostenibilidad del efecto observado, pues las mejoras iniciales podrían

deberse al efecto posimplementación inmediato del proyecto más que a un cambio organizacional y cultural perdurable a largo plazo³⁵. En este sentido, el *benchmarking* debería considerarse parte de los programas de gestión de calidad, y no un sustituto³⁶.

DISCUSIÓN

Los cambios organizacionales que podrían mejorar la salud de los pacientes suelen ser difíciles de implementar en la práctica, incluso cuando están respaldados por evidencia sólida; se requieren equipos interdisciplinarios y una cultura colaborativa³⁷. La “ciencia de la implementación” es el estudio de métodos y estrategias para promover la adopción e integración de tratamientos clínicos, prácticas, intervenciones organizacionales y/o de gestión comprobados en la práctica clínica habitual y, por lo tanto, mejorar la salud³⁸. Es aquella que cierra el salto “de la teoría a la práctica”, estudiando cómo las intervenciones o las prácticas basadas en la evidencia se difunden en el mundo real y comprendiendo las barreras que las limitan³⁸. En particular, existen herramientas que pueden utilizarse antes, durante o después del proceso de comparación para transformar los hallazgos del *benchmarking* en acciones de mejora como: a) Lean, que ayuda a identificar desperdicios, variabilidad y actividades que no agregan valor³⁹; b) Six Sigma, que aporta herramientas para reducir variabilidad y defectos, y mejorar la capacidad de los procesos⁴⁰ y c) ciclos PDSA (de *plan-do-study-act*), un enfoque complementario que proporciona un mecanismo iterativo para probar y ajustar cambios⁴¹.

Es el paso final clave para llevar la atención de las “mejores prácticas” y potencialmente transferibles a un gran número de pacientes⁴². Sin embargo, la gran cantidad y complejidad de los datos del mundo real también exigen el desarrollo de técnicas de procesamiento y análisis de datos más apropiadas, sofisticadas e innovadoras, manteniendo el rigor científico de los hallazgos de investigación y prestando atención a la ética de los datos para aprovechar al máximo su potencial⁴³.

Muchas organizaciones aún no entienden que poner al paciente en el centro es el propósito principal de la

atención sanitaria (p. ej., escepticismo o miedo a alguna aplicación punitiva), lo que implica integrarlo activamente en las mejoras del sistema, algo que sigue generando resistencia entre algunos líderes y profesionales³. Un proceso continuo de evaluación y comunicación puede contribuir a fomentar una cultura de transparencia, confianza y aprendizaje.

El objetivo final de medir la calidad de la atención es proporcionar nuevos conocimientos sobre cómo mejorar el rendimiento: qué cambios se pueden realizar en las estructuras o procesos existentes para mejorar los resultados de los pacientes. La evaluación comparativa del desempeño se ha incorporado ampliamente a las estrategias de evaluación y mejora de la calidad de la atención sanitaria⁴. Sin embargo, su utilidad no depende únicamente de la disponibilidad de indicadores comparables, sino también de la calidad de los datos, la validez de las medidas y la consideración de las diferencias en el riesgo, el contexto de las poblaciones evaluadas, gobernanza clara, transparencia y una cultura de mejora no punitiva.

CONCLUSIONES

El *benchmarking* puede constituir una herramienta de alto valor para la gestión sanitaria al transformar la medición del desempeño en información orientada a la mejora. Su utilidad radica en identificar brechas, comprender las diferencias observadas y favorecer el aprendizaje y la adopción de prácticas basadas en la evidencia, considerando el contexto clínico, institucional y poblacional.

Su capacidad para generar valor depende, sin embargo, de comparaciones metodológicamente válidas, sustentadas en datos de calidad, indicadores comparables y ajustes adecuados. Integrado a estrategias de mejora de la calidad y a organizaciones que favorezcan la transparencia, la confianza y el aprendizaje, el *benchmarking* puede contribuir a traducir los hallazgos en acciones de mejora. Por el contrario, datos incompletos, indicadores no comparables o ajustes insuficientes pueden conducir a interpretaciones erróneas o comparaciones injustas. La incorporación de la experiencia del paciente puede contribuir a orientar estas evaluaciones hacia resultados relevantes para las personas.

Contribuciones de los autores: Conceptualización, Administración del proyecto (MFGR, SM). Curación de datos, Análisis formal, Investigación, Metodología (MFGR, JAON, SM). Recursos, Supervisión (MFGR). Validación, Visualización: MFGR, SM. Redacción–borrador original, Redacción–revisión y edición (MFGR, NSDM, JAON, SM).

Conflictos de intereses: los autores declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: los autores declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Ridgway VF. Dysfunctional consequences of performance measurements. *Adm Sci Q*. 1956;1(2):240-247. <https://doi.org/10.2307/2390989>.
2. Camp RC. Benchmarking: the search for industry best practices that lead to superior performance. New York: CRC Press; 2024. <https://doi.org/10.4324/9781003578871>.
3. Edgman-Levitan S, Schoenbaum SC. Patient-centered care: achieving higher quality by designing care through the patient's eyes. *Isr J Health Policy Res*. 2021;10(1):21. <https://doi.org/10.1186/s13584-021-00459-9>.
4. Byrne JP, Haut ER. Facts and fallacy of benchmark performance indicators. *Adv Surg*. 2022;56(1):89-109. <https://doi.org/10.1016/j.yasu.2022.03.002>.
5. Frutos EL, Iannicelli A, Luna D, et al. Implementation of a real-time dashboard for the emergency department. *Stud Health Technol Inform*. 2025;329:184-188. <https://doi.org/10.3233/SHTI250826>.
6. Nilsen P, Seing I, Ericsson C, et al. Characteristics of successful changes in health care organizations: an interview study with physicians, registered nurses and assistant nurses. *BMC Health Serv Res*. 2020;20:147. <https://doi.org/10.1186/s12913-020-4999-8>.
7. Grande Ratti MF, Luna DR. Indicadores en salud: uso en gestión para mejorar la toma de decisiones. *Rev Hosp Ital B Aires*. 2021;41(2):82-83. <https://doi.org/10.51987/revhospitalbaires.v41i2.92>.
8. Benson HR. An introduction to benchmarking in healthcare. *Radiol Manage*. 1994;16(4):35-39.
9. Grande Ratti MF, De Masi NS, Fedele GV, et al. Implementación del método 5S en salud: de la industria automotriz a los traslados intrahospitalarios. *Evid Actual Pract Ambul*. 2026;29(2):e007223. <https://doi.org/10.51987/evidencia.v29i3.7223>.
10. Breyer JZ, Giacomazzi J, Kuhmmer R, et al. Hospital quality indicators: a systematic review. *Int J Health Care Qual Assur*. 2019;32(2):474-487. <https://doi.org/10.1108/IJHCQA-04-2018-0091>.
11. Walsh KE, Razzaghi H, Hartley DM, et al. Testing the use of data drawn from the electronic health record to compare quality. *Pediatr Qual Saf*. 2021;6(4):e432. <https://doi.org/10.1097/pq.0000000000000432>.
12. Salamessy BH, Portrait FRM, van der Hijden E, et al. On the correlation between outcome indicators and the structure and process indicators used to proxy them in public health care reporting. *Eur J Health Econ*. 2021;22(8):1239-1251. <https://doi.org/10.1007/s10198-021-01333-w>.
13. Cossio-Gil Y, Pérez-Sádaba FJ, Ribera J, et al. Identifying potential predictable indicators for the management of tertiary hospitals. *Int J Health Plann Manage*. 2024;39(2):278-292. <https://doi.org/10.1002/hpm.3710>.
14. Nabovati E, Farrahi R, Sadeqi Jabali M, et al. Identifying and prioritizing the key performance indicators for hospital management dashboard at a national level: viewpoint of hospital managers. *Health Informatics J*. 2023;29(4):14604582231221139. <https://doi.org/10.1177/14604582231221139>.
15. Fallahnezhad M, Langarizadeh M, Vahabzadeh A. Key performance indicators of hospital supply chain: a systematic review. *BMC Health Serv Res*. 2024;24:1610. <https://doi.org/10.1186/s12913-024-11954-5>.
16. Gannon B, Jones C, Wakai A, O'Sullivan R. The cost of data collection for performance monitoring in hospitals: protocol for a systematic review. *Syst Rev*. 2014;3:65. <https://doi.org/10.1186/2046-4053-3-65>.
17. Amato L, Colais P, Davoli M, et al. Volumi di attività ed esiti delle cure: prove scientifiche dalla letteratura e dalle valutazioni empiriche in Italia. *Epidemiol Prev*. 2013;37(2-3 Suppl 2):1-100.
18. Grande Ratti MF, De Masi N, Garfi L, et al. A case report of a mixed-methods assessment of patient experiences to inform quality improvement in an emergency department in Argentina. *Int J Health Plann Manage*. 2025;40(3):620-629. <https://doi.org/10.1002/hpm.3906>.
19. Medina-Orozco I. La narrativa médica como herramienta educativa. *Rev Med Inst Mex Seguro Soc*. 2025;63(6):e6840. <https://doi.org/10.5281/zenodo.17342650>.
20. Kelley TA. International Consortium for Health Outcomes Measurement (ICHOM). *Trials*. 2015;16(Suppl 3):O4. <https://doi.org/10.1186/1745-6215-16-S3-O4>.

21. Seligman WH, Das-Gupta Z, Jobi-Odeneye AO, et al. Development of an international standard set of outcome measures for patients with atrial fibrillation: a report of the International Consortium for Health Outcomes Measurement (ICHOM) atrial fibrillation working group. *Eur Heart J*. 2020;41(10):1132-1140. <https://doi.org/10.1093/eurheartj/ehz871>.
22. Verberne WR, Das-Gupta Z, Allegretti AS, et al. Development of an international standard set of value-based outcome measures for patients with chronic kidney disease: a report of the International Consortium for Health Outcomes Measurement (ICHOM) CKD Working Group. *Am J Kidney Dis*. 2019;73(3):372-384. <https://doi.org/10.1053/j.ajkd.2018.10.007>.
23. Burns DJP, Arora J, Okunade O, et al. International Consortium for Health Outcomes Measurement (ICHOM): standardized patient-centered outcomes measurement set for heart failure patients. *JACC Heart Fail*. 2020;8(3):212-222. <https://doi.org/10.1016/j.jchf.2019.09.007>.
24. Bull C, Teede H, Watson D, et al. Selecting and implementing patient-reported outcome and experience measures to assess health system performance. *JAMA Health Forum*. 2022;3(4):e220326. <https://doi.org/10.1001/jamahealthforum.2022.0326>.
25. Gil Folgar ML, Grande Ratti MF, Perman G, et al. Resultados reportados por pacientes muy mayores tras implante valvular aórtico percutáneo. *Arch Peru Cardiol Cir Cardiovasc*. 2026;7(2):99-107. <https://doi.org/10.47487/apcyccv.7i2.634>.
26. Jamieson Gilmore K, Corazza I, Coletta L, et al. The uses of Patient Reported Experience Measures in health systems: a systematic narrative review. *Health Policy*. 2023;128:1-10. <https://doi.org/10.1016/j.healthpol.2022.07.008>.
27. Ghalavand H, Shirshahi S, Rahimi A, et al. Common data quality elements for health information systems: a systematic review. *BMC Med Inform Decis Mak*. 2024;24:243. <https://doi.org/10.1186/s12911-024-02644-7>.
28. Hooshafza S, Mc Quaid L, Stephens G, et al. Development of a framework to assess the quality of data sources in healthcare settings. *J Am Med Inform Assoc*. 2022;29(5):944-952. <https://doi.org/10.1093/jamia/ocac017>.
29. Abbassi F, Barkun J. Artificial intelligence in surgical outcomes reporting: the next best thing, or just artificially intelligent "garbage in, garbage out"? *Ann Surg*. 2025;282(6):897-899. <https://doi.org/10.1097/SLA.0000000000006789>.
30. Heenan MA, Randall GE, Evans JM. Selecting performance indicators and targets in health care: an international scoping review and standardized process framework. *Risk Manag Healthc Policy*. 2022;15:747-764. <https://doi.org/10.2147/RMHP.S357561>.
31. Van Vliet EJ, Bredenhoff E, Sermeus W, et al. Exploring the relation between process design and efficiency in high-volume cataract pathways from a lean thinking perspective. *Int J Qual Health Care*. 2011;23(1):83-93. <https://doi.org/10.1093/intqhc/mzq071>.
32. Arah OA, Westert GP, Hurst J, et al. A conceptual framework for the OECD Health Care Quality Indicators Project. *Int J Qual Health Care*. 2006;18(Suppl 1):5-13. <https://doi.org/10.1093/intqhc/mzl024>.
33. Kruk ME, Gage AD, Arsenault C, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. *Lancet Glob Health*. 2018;6(11):e1196-e1252. [https://doi.org/10.1016/S2214-109X\(18\)30386-3](https://doi.org/10.1016/S2214-109X(18)30386-3). Errata en: *Lancet Glob Health*. 2018;6(11):e1162. [https://doi.org/10.1016/S2214-109X\(18\)30456-X](https://doi.org/10.1016/S2214-109X(18)30456-X); *Lancet Glob Health*. 2021;9(8):e1067. [https://doi.org/10.1016/S2214-109X\(21\)00250-3](https://doi.org/10.1016/S2214-109X(21)00250-3).
34. Teisberg E, Wallace S, O'Hara S. Defining and implementing value-based health care: a strategic framework. *Acad Med*. 2020;95(5):682-685. <https://doi.org/10.1097/ACM.00000000000003122>.
35. Willmington C, Belardi P, Murante AM, et al. The contribution of benchmarking to quality improvement in healthcare: a systematic literature review. *BMC Health Serv Res*. 2022;22:139. <https://doi.org/10.1186/s12913-022-07467-8>.
36. Reponen E, Rundall TG, Shortell SM, et al. Benchmarking outcomes on multiple contextual levels in lean healthcare: a systematic review, development of a conceptual framework, and a research agenda. *BMC Health Serv Res*. 2021;21:161. <https://doi.org/10.1186/s12913-021-06160-6>.
37. Pronovost PJ, Berenholtz SM, Needham DM. Translating evidence into practice: a model for large scale knowledge translation. *BMJ*. 2008;337:a1714. <https://doi.org/10.1136/bmj.a1714>.
38. Neuman HB, Kaji AH, Haut ER. Practical guide to implementation science. *JAMA Surg*. 2020;155(5):434-435. <https://doi.org/10.1001/jamasurg.2019.5149>.
39. Cohen RI. Lean methodology in health care. *Chest*. 2018;154(6):1448-1454. <https://doi.org/10.1016/j.chest.2018.06.005>.
40. Niñerola A, Sánchez-Rebull MV, Hernández-Lara AB. Quality improvement in healthcare: Six Sigma systematic review. *Health Policy*. 2020;124(4):438-445. <https://doi.org/10.1016/j.healthpol.2020.01.002>.
41. Leis JA, Shojania KG. A primer on PDSA: executing plan-do-study-act cycles in practice, not just in name. *BMJ Qual Saf*. 2017;26(7):572-577. <https://doi.org/10.1136/bmjqs-2016-006245>.
42. Dang A. Real-world evidence: a primer. *Pharmaceut Med*. 2023;37(1):25-36. <https://doi.org/10.1007/s40290-022-00456-6>.
43. Liu F, Panagiotakos D. Real-world data: a brief review of the methods, applications, challenges and opportunities. *BMC Med Res Methodol*. 2022;22:287. <https://doi.org/10.1186/s12874-022-01768-6>.

Nuevas terapias para enfermedad hepática por depósito graso/metabólica (MASLD/MASH): resultados clínicos y retos en implementación

Javier J. Flores[■] y Elizabeth P. Aguilera[■]

Universidad Católica Santiago de Guayaquil. Guayaquil, Ecuador.

RESUMEN

Introducción: la enfermedad hepática por depósito graso asociada a disfunción metabólica (MASLD) y su forma inflamatoria progresiva, la esteatohepatitis asociada (MASH), representan un desafío creciente por su alta prevalencia y su potencial evolución hacia fibrosis avanzada y cirrosis. En los últimos años, el abordaje terapéutico ha trascendido las medidas higiénico-dietéticas, impulsando el desarrollo de alternativas farmacológicas orientadas a modificar la historia natural de la enfermedad.

Estado del arte: la investigación reciente se ha centrado en fármacos que actúan sobre vías metabólicas, inflamatorias y antifibróticas implicadas en la fisiopatología de MASLD/MASH. Los ensayos clínicos de fases II y III han mostrado resultados prometedores, destacando los agonistas del receptor del péptido similar al glucagón tipo 1 (GLP-1) y los agonistas selectivos del receptor beta de hormona tiroidea, los cuales han evidenciado mejoría en la resolución de la esteatohepatitis y disminución del grado de fibrosis hepática.

Discusión: a pesar de los avances observados, la eficacia terapéutica no es uniforme entre los pacientes, lo que sugiere heterogeneidad fisiopatológica y necesidad de estrategias de medicina personalizada. Asimismo, persisten interrogantes con respecto a la seguridad a largo plazo, la adherencia terapéutica y la relación costo-beneficio en distintos contextos sanitarios, aspectos clave para su incorporación en la práctica clínica habitual.

Conclusión: las terapias farmacológicas emergentes para MASLD y MASH constituyen un avance relevante en el manejo de estas entidades, al ofrecer beneficios clínicos e histológicos significativos. Sin embargo, la optimización de su implementación requiere mayor evidencia sobre seguridad prolongada, selección adecuada de pacientes y estrategias que faciliten su integración efectiva en los sistemas de salud.

Palabras clave: MASLD, MASH, terapias emergentes, fibrosis hepática, agonistas FXR, GLP-1.

Autor para correspondencia: javier.flores@cu.ucsg.edu.ec, Flores JJ.

Recibido: 10/01/2026 | Aceptado: 2/09/2026 | Publicado: 28/09/2026

DOL: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1314>

Cómo citar: Flores JJ, Aguilera EP. Nuevas terapias para enfermedad hepática por depósito graso/metabólica (MASLD/MASH): resultados clínicos y retos en implementación. *Rev Hosp Ital B.Aires.* 2026;46(3):e0001314

New Therapies for Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD/MASH): Clinical Outcomes and Implementation Challenges

ABSTRACT

Introduction: metabolic dysfunction-associated fatty liver disease (MASLD) and its progressive inflammatory form, associated steatohepatitis (MASH), represent a growing challenge due to their high prevalence and potential progression to advanced fibrosis and cirrhosis. In recent years, the therapeutic approach has gone beyond lifestyle and dietary measures, driving the development of pharmacological alternatives aimed at modifying the natural history of the disease.

State of the art: recent research has focused on drugs that act on metabolic, inflammatory, and antifibrotic pathways involved in the pathophysiology of MASLD/MASH. Phase II and III clinical trials have shown promising results, notably glucagon-like peptide-1 (GLP-1) receptor agonists and selective thyroid hormone beta receptor agonists, which have shown improvement in the resolution of steatohepatitis and a decrease in the degree of liver fibrosis.

Discussion: despite the advances observed, therapeutic efficacy is not uniform among patients, suggesting pathophysiological heterogeneity and the need for personalized medicine strategies. Questions also remain regarding long-term safety, therapeutic adherence, and cost-benefit ratios in different healthcare settings, which are key aspects for their incorporation into routine clinical practice.

Conclusion: emerging drug therapies for MASLD and MASH represent a significant advance in the management of these conditions, offering significant clinical and histological benefits. However, optimizing their implementation requires further evidence on long-term safety, appropriate patient selection, and strategies to facilitate their effective integration into healthcare systems.

Keywords: MASLD, MASH, emerging therapies, liver fibrosis, FXR agonists, GLP-1.

INTRODUCCIÓN

La MASLD se caracteriza por la acumulación de grasa hepática asociada a condiciones metabólicas como obesidad, dislipidemia y resistencia a la insulina. La búsqueda de terapias efectivas ha sido un reto debido a la compleja fisiopatología multifactorial de la enfermedad. Las intervenciones actuales incluyen modificaciones del estilo de vida, control metabólico y el uso emergente de terapias farmacológicas específicas¹.

En la actualidad, el tratamiento de la enfermedad hepática por depósito graso/metabólica (MASLD/MASH) está experimentando una rápida evolución con la aparición de nuevas terapias dirigidas a mecanismos metabólicos, inflamatorios y fibrogénicos. El pilar sigue siendo la intervención sobre el estilo de vida, pero la baja adherencia y eficacia sostenida han impulsado el desarrollo de diversos fármacos con evidencia clínica robusta que se resumen en la tabla 1^{2,3}.

La medicina de laboratorio ha impulsado la integración de biomarcadores séricos directos que evalúan el recambio activo de la matriz extracelular (MEC). Estos marcadores miden de forma cuantitativa la síntesis y degradación de la matriz por parte de las células estrelladas hepáticas activadas⁴.

Panel Directo (ELF Score): algoritmo que combina el ácido hialurónico (HA), el propéptido aminoterminal del procolágeno tipo III (PIIINP) y el inhibidor tisular de metalo-

proteinasas-1 (TIMP-1). El ELF Score permite una evaluación directa del remodelado tisular, independiente de la citólisis; valores < 7,7 permiten descartar fibrosis hepática avanzada (ALF) con un elevado valor predictivo negativo, mientras que valores ≥ 9,8 identifican a pacientes con MASLD en alto riesgo de descompensación y eventos vasculares⁴.

Algoritmo Mixto Metabólico (ADAPT Score): diseñado específicamente para responder a la fisiopatología del MASLD, el modelo ADAPT combina la edad, la presencia de diabetes tipo 2 (T2D) y el recuento de plaquetas con el PRO-C3 (marcador neoepitópico de la formación activa de colágeno tipo III). Al incorporar el estado glucémico y el nivel de fibrogénesis en tiempo real, el ADAPT Score alcanza una precisión diagnóstica superior [AUROC (área bajo la curva de características operativas del receptor) > 0,86 en la detección de fibrosis avanzada en MASLD frente a las pruebas indirectas convencionales⁵.

En cuanto a resultados clínicos, la progresión de MASLD a MASH ocurre en aproximadamente 15-40% de los pacientes, y la fibrosis avanza más rápidamente en MASH que en la esteatosis simple. La presencia de fibrosis significativa (≥ F2) es el principal predictor de complicaciones hepáticas y mortalidad, y la regresión de fibrosis se asocia con mejores pronósticos^{1,6}.

Los retos en la implementación incluyen la baja tasa de diagnóstico temprano, especialmente en poblaciones

Tabla 1. Lista de fármacos investigados para el tratamiento de la enfermedad hepática esteatósica asociada a disfunción metabólica

Fármacos	Mecanismo de acción	Relación fisiopatológica
Resmetirom	Agonista selectivo del receptor de la hormona tiroidea (THR)	Estimula la β -oxidación hepática y la degradación del colesterol, disminuye la lipogénesis <i>de novo</i> , combate el “hipotiroidismo hepático” y reduce directamente la esteatosis, la inflamación y la fibrosis
Ácido obeticoólico	Agonista del receptor farnesoide X (FXR)	Reduce la lipogénesis <i>de novo</i> , aumenta la oxidación de ácidos grasos, mejora la sensibilidad a la insulina y reduce la inflamación/fibrosis hepática
Tropifexor	Agonista de FXR	Comparte el mismo eje metabólico que el ácido obeticoólico, inhibiendo la acumulación lipídica y atenuando la cascada proinflamatoria y profibrótica de las células estrelladas hepáticas
Pioglitazona	Agonista del receptor activado por el proliferador de peroxisomas (PPAR γ)	Sensibiliza a la insulina en tejido adiposo y músculo, redistribuye la grasa fuera del hígado y reduce la lipotoxicidad y el estrés oxidativo
Saroglitazar	Agonista de PPAR α/γ	PPAR α disminuye triglicéridos y acelera la β -oxidación; PPAR δ mejora la resistencia a la insulina, atacando simultáneamente la dislipidemia y la esteatosis
Elafibranor	Agonista de PPAR α/δ	PPAR α reduce lípidos e inflamación; PPAR δ mejora la homeostasis de glucosa y la esteatosis, atenuando la activación de las células estrelladas
Semaglutida	Agonista del péptido similar al glucagón (GLP-1)	Promueve pérdida de peso significativa, mejora la secreción de insulina, disminuye el vaciamiento gástrico y reduce la lipotoxicidad sistémica e inflamación
Liraglutida	Agonista de GLP-1	Similar a semaglutida: mejora el control glucémico, reduce el peso corporal y disminuye la acumulación lipídica hepática de forma indirecta y directa
Tirzepatida	Agonista de GLP-1 y GIP	Potencia el efecto de pérdida de peso y control glucémico por sinergia de receptores, reduciendo drásticamente el contenido de grasa hepática y la inflamación
Pegbelfermin	Análogo del factor de crecimiento de fibroblastos (FGF21)	Regula el metabolismo lipídico y glucídico, aumenta la adiponectina, mejora la sensibilidad a la insulina y reduce la esteatosis y la fibrosis
Efruxifermina	Análogo de FGF21	Normaliza el metabolismo hepático, disminuye el estrés del retículo endoplasmático, reduce la inflamación y revierte la fibrosis hepática
Aldafermin	Análogo de FGF19	Reprime la síntesis de ácidos biliares (vía CYP7A1, citocromo P450 familia 7 subfamilia A miembro 1), disminuye la lipogénesis <i>de novo</i> e inhibe la respuesta inflamatoria y fibrocítica
Aramchol	Inhibidor de la síntesis <i>de novo</i> de lípidos	Modula el metabolismo de ácidos grasos en el hepatocito, reduciendo el almacenamiento de triglicéridos y la inflamación asociada
Cenicriviroc	Agonista de los receptores de quimiocinas CCR2 y CCR5	Bloquea el reclutamiento de monocitos/macrófagos proinflamatorios al hígado, interrumpiendo la cascada inflamatoria y profibrótica
Selonsertib	Inhibidor de la quinasa reguladora de la señal de apoptosis (ASK1)	Inhibe la vía de señalización p38/JNK activada por estrés oxidativo, previniendo la apoptosis del hepatocito y la fibrosis asociada

GIP: péptido insulínico dependiente de la glucosa

de alto riesgo como personas con diabetes tipo 2, obesidad y factores cardiometabólicos, para quienes el cribado con índices como FIB-4 (índice de fibrosis 4) es recomendado pero subutilizado^{2,7}.

Las nuevas terapias para la enfermedad hepática por depósito graso/metabólica (MASLD/MASH), como los agonistas de GLP-1 (semaglutida, tirzepatida) y el sofora β del receptor de la hormona tiroidea (resmetirom, aprobado por la FDA [Administración de Alimentos y Medicamentos] de Estados Unidos para MASH no cirrótico con fibrosis F2-F3), se comparan con las opciones tradicionales (modificación de estilo de vida, manejo de comorbilidades metabólicas y, en algunos casos, vitamina E o pioglitazona) en varios aspectos de eficacia clínica^{8,9}.

En términos de reducción de esteatosis y mejoría de parámetros metabólicos, los agonistas de GLP-1 han demostrado superioridad sobre las intervenciones convencionales, especialmente en la reducción de peso y mejoría de la resistencia a la insulina, factores clave en la fisiopatología de MASLD/MASH¹⁰. Los estudios clínicos han mostrado que semaglutida y tirzepatida logran una mayor resolución de la esteatohepatitis y regresión de fibrosis en comparación con placebo y con las opciones farmacológicas previas, aunque aún no están aprobados específicamente para MASH fuera de Estados Unidos^{11,12}. Sin embargo, la eficacia clínica de estas nuevas terapias debe contextualizarse: aunque muestran superioridad en *endpoints* histológicos y metabólicos, su impacto en desenlaces duros (como progresión a cirrosis, eventos hepáticos mayores o mortalidad) aún está en evaluación a largo plazo, ya que, los datos sobre mortalidad y progresión a cirrosis provienen de las extensiones a 54 meses de los ensayos Fase III (p. ej., MAESTRO-NASH), cuyos resultados se esperan hacia 2026-2027. Además, las opciones tradicionales siguen siendo fundamentales, especialmente en estadios iniciales y en pacientes con contraindicación para nuevos fármacos^{13,14}.

ESTADO DEL ARTE

Terapias emergentes (Tabla 2)

1. Agonistas del receptor FXR

Los agonistas del receptor de farnesoide X (FXR), como el ácido obeticoico y el cilofexor, modulan la homeostasis de los ácidos biliares, reduciendo el estrés hepatocitario y la fibrosis en MASH¹⁵. Aunque las fases III muestran mejoría en la inflamación y la fibrosis, su trayectoria regulatoria se ha visto limitada por efectos adversos clave, principalmente la elevación de los niveles de colesterol de lipoproteínas de baja densidad (LDL-C) que típicamente requiere manejo concomitante con estatinas y la aparición de prurito dosis-dependiente^{16,17}.

2. Agonistas GLP-1

Los agonistas del receptor GLP-1, como semaglutida, han mostrado efectos robustos en la reducción de la grasa hepática y en la pérdida ponderal, lo que contribuye indirectamente a la mejoría de la resistencia a la insulina y la inflamación hepática¹⁸.

En el ensayo fase III ESSENCE, semaglutida en dosis de 2,4 mg semanales durante 72 semanas logró la resolución de MASH sin empeoramiento de fibrosis en el 62,9% de los pacientes, en comparación con 34,3% en placebo, y mejoría de fibrosis en el 36,8% frente a 22,4% en placebo. Sin embargo, el impacto sobre la regresión de la fibrosis es más modesto y no se ha demostrado eficacia en pacientes con cirrosis¹⁹.

3. Terapias antifibróticas

Moléculas dirigidas a mecanismos de fibrosis, como pegbelfermin e inhibidores de integrinas, están en evaluación. Los resultados preliminares indican reducción de marcadores de fibrosis y mejorías histológicas. Las terapias antifibróticas dirigidas a mecanismos fibrogénicos incluyen análogos de FGF21 como pegbelfermin y pegozafermin, así como inhibidores de integrinas²⁰.

Pegozafermin, en dosis altas, logró regresión de fibrosis en hasta el 20% de los pacientes y resolución de MASH en el 24% después de seis meses de tratamiento. Los inhibidores de integrinas están en fases tempranas de investigación, con datos iniciales que sugieren reducción de marcadores de activación de células estrelladas hepáticas, pero sin resultados robustos en *endpoints* histológicos²¹.

Desafíos en la implementación

La implementación de nuevas terapias para MASLD/MASH enfrenta múltiples desafíos clínicos, regulatorios y logísticos; la heterogeneidad fisiopatológica de MASLD/MASH dificulta la selección de tratamientos óptimos para cada paciente. La enfermedad involucra vías metabólicas, inflamatorias y fibrogénicas complejas, lo que limita la eficacia de terapias dirigidas a un solo mecanismo^{22,23}.

La implementación clínica de estas terapias enfrenta barreras adicionales:

- Acceso y costo: los nuevos medicamentos suelen tener precios elevados y acceso limitado, lo que puede aumentar las desigualdades en salud.
- Polifarmacia y comorbilidades: especialmente en adultos mayores, la presencia de múltiples comorbilidades y el uso concomitante de varios fármacos aumentan el riesgo de interacciones y efectos adversos, requiriendo una revisión cuidadosa de la medicación y ajuste de objetivos terapéuticos.
- Falta de datos en poblaciones especiales: aunque los ensayos clínicos recientes han incluido adultos mayores, aún se necesita más evidencia en poblaciones pediátricas, pacientes con cirrosis y aquellos con múltiples comorbilidades.
- Adherencia y sostenibilidad: la eficacia de las intervenciones de estilo de vida sigue siendo limitada en la práctica real, y la adherencia a tratamientos farmacológicos a largo plazo es un reto.

CONCLUSIONES

La evidencia derivada de ensayos clínicos fases II y III indica que múltiples clases terapéuticas incluyendo agonistas de receptores de incretinas, moduladores de

Tabla 2. Terapias farmacológicas en MASLD/MASH : tabla integrada

Clase de fármaco	Mecanismo de acción	Fibrosis hepática	Esteatosis hepática	Efecto en peso	Efectos cardiorrenales	Resolución de MASH	Limitaciones principales
Agonistas PPAR (p. ej., lanifibranor, pioglitazona)	Mejoran el metabolismo lipídico y la sensibilidad a la insulina vía activación PPAR	Mejoría moderada	Reducción moderada-significativa	Aumento o neutro	Beneficios metabólicos; riesgo IC con pioglitazona	Moderada	Aumento de peso, edemas, seguridad a largo plazo
Agonistas GLP-1 (sema-glutida, tirzepatida)	Aumentan secreción de insulina, reducen apetito, efecto antiinflamatorio	Mejoría significativa en marcadores e histología	Fuerte reducción de grasa hepática	Pérdida de peso significativa	Beneficio cardiovascular y potencial renal	Alta en ensayos clínicos	Costo, efectos GI, datos a largo plazo limitados
Inhibidores SGLT-2 (Co-transportador sodio-glucosa tipo 2) (empagliflozina, dapagliflozina)	Bloquean reabsorción renal de glucosa	Mejoría modesta	Reducción leve	Pérdida de peso moderada	Protección cardiorrenal bien establecida	Evidencia limitada	Evidencia histológica limitada
Inhibidores Agonistas de FXR (p. ej., ácido obetónico)	Reduce la lipogénesis hepática y la producción de ácidos biliares activando el receptor farnesoide X	Mejoría significativa de la fibrosis (≥ 1 estadio en ensayos)	Reducción modesta de la esteatosis	Peso neutro	Posible empeoramiento del perfil lipídico ($\uparrow$ LDL)	No demostrada	El prurito y la dislipidemia se encuentran entre los efectos secundarios; problemas de seguridad impiden la aprobación

receptores nucleares (FXR, PPAR), agentes que impactan el metabolismo lipídico y glucídico, y compuestos con actividad antifibrótica logran reducciones clínicamente relevantes en la actividad necroinflamatoria hepática y, en subgrupos seleccionados, inducen regresión de la fibrosis.

La marcada heterogeneidad biológica y clínica de MASLD/MASH limita la extrapolación uniforme de los resultados de los ensayos, subrayando la necesidad de estrategias de medicina de precisión basadas en fenotipos metabólicos, perfiles genéticos y biomarcadores no invasivos validados. Adicionalmente, persisten interrogantes en relación con la durabilidad de la respuesta terapéutica, la seguridad a largo plazo y el impacto real sobre desenlaces

con mal pronóstico, tales como descompensación hepática, carcinoma hepatocelular y mortalidad global.

Contribución de los autores: Conceptualización, Análisis formal, Investigación, Administración del proyecto, *Software*, Visualización, Redacción-borrador original: EPA. Curación de datos, Captación de fondos, Metodología, Recursos, Supervisión, Validación, Redacción-revisión y edición: JJF.

Conflictos de intereses: los autores declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: los autores declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Yadav P, Quadri K, Kadian R, et al. New approaches to the treatment of metabolic dysfunction-associated steatotic liver with natural products. *ILIVER*. 2024;3(4):100131. <https://doi.org/10.1016/j.iliver.2024.100131>.
2. Branković M, Dukić M, Gmizić T, et al. New therapeutic approaches for the treatment of patients with metabolic dysfunction-associated steatotic liver disease (MASLD) and increased cardiovascular risk. *Diagnostics (Basel)*. 2024;14(2):229. <https://doi.org/10.3390/diagnostics14020229>.
3. Patel P. Treatment landscape of metabolic-dysfunction-associated steatotic liver disease. *J Clin Med*. 2025;14(17):6060. <https://doi.org/10.3390/jcm14176060>.
4. Vidal-Cevallos P, Sorroza-Martínez AP, Chávez-Tapia NC, et al. The relationship between pathogenesis and possible treatments for the MASLD-cirrhosis spectrum. *Int J Mol Sci*. 2024;25(8):4397. <https://doi.org/10.3390/ijms25084397>.
5. Krintus M, Panteghini M. Detection of advanced liver fibrosis using blood noninvasive tests: a laboratory medicine perspective. *Clin Chem*. 2026;72(3):337-346. <https://doi.org/10.1093/clinchem/hvaf193>.
6. Li Y, Sun W, Liu H, et al. Tirzepatide, a dual GIP/GLP-1 receptor agonist, alleviates metabolic dysfunction-associated steatotic liver disease by reducing the expression of CD36 and OBP2A. *Genes Dis*. 2025;12(6):101761. <https://doi.org/10.1016/j.gendis.2025.101761>.
7. Shakeel L, Shaukat A, Akilimali A. Resmetirom: a breakthrough in the treatment of metabolic dysfunction-associated steatotic liver disease (MASLD). *Health Sci Rep*. 2025;8(6):e70920. <https://doi.org/10.1002/hsr2.70920>.
8. Barb D, Kalavalapalli S, Godinez Leiva E, et al. Pan-PPAR agonist lanifibranor improves insulin resistance and hepatic steatosis in patients with T2D and MASLD. *J Hepatol*. 2025;82(6):979-991. <https://doi.org/10.1016/j.jhep.2024.12.045>.
9. Choi R, Vemuri J, Poloju A, et al. Current and emerging treatments for metabolic associated steatotic liver disease and diabetes: a narrative review. *Endocrines*. 2025;6(2):27. <https://doi.org/10.3390/endocrines6020027>.
10. Vilarrasa N, Pellitero S. GLP-1-based therapies for obesity: impact on comorbidities or obesity-related diseases. *Med Clin (Barc)*. 2025;165(6):107184. <https://doi.org/10.1016/j.medcli.2025.107184>.
11. Kamrul-Hasan ABM, Boroza S, Jena S, et al. Safety and efficacy of efruxifermin in metabolic dysfunction-associated steatohepatitis: a systematic review. *World J Gastrointest Pharmacol Ther*. 2025;16(3):1-15. <https://doi.org/10.4292/wjgpt.v16.i3.110709>.
12. Cusi K, Abdelmalek MF, Apovian CM, et al. Metabolic dysfunction-associated steatotic liver disease (MASLD) in people with diabetes: the need for screening and early intervention. A consensus report of the American Diabetes Association. *Diabetes Care*. 2025;48(7):1057-1082. <https://doi.org/10.2337/dci24-0094>.
13. Romero M. New molecules for treating metabolic-associated dysfunction steatotic liver disease (MASLD). *Ann Hepatol*. 2025;30(2):102157. <https://doi.org/10.1016/j.aohp.2025.102157>.
14. Drygalski K. Pharmacological treatment of MASLD: contemporary treatment and future perspectives. *Int J Mol Sci*. 2025;26(13):6518. <https://doi.org/10.3390/ijms26136518>.
15. De la Hoz Gil L, Martínez-Echevarría SV, Lozano Estevan MDC. Enfermedad hepática esteatótica asociada a disfunción metabólica (MASLD) y fibrosis hepática. Propuesta de score APRI+FIB4+NFS. *Nutr Hosp*. 2025;42(3):1-6. <https://doi.org/10.20960/nh.05252>.
16. Srikanth P, Shaik KM, Patibandla V, et al. Significance of FXR agonists in MASLD treatment: a deep dive into lipid alteration by analytical techniques. *Explor Endocr Metab Dis*. 2025;2:1-24. <https://doi.org/10.37349/eemd.2025.101425>.
17. Zhou D, Fan J. Drug treatment for metabolic dysfunction-associated steatotic liver disease: progress and direction. *Chin Med J (Engl)*. 2024;137(22):1-10. <https://doi.org/10.1097/CM9.0000000000003355>.
18. Havranek B, Loh R, Torre B, et al. Glucagon-like peptide-1 receptor agonists improve metabolic dysfunction-associated steatotic liver disease outcomes. *Sci Rep*. 2025;15(1):1-11. <https://doi.org/10.1038/s41598-025-89408-z>.
19. Suki M, Amer J, Milgrom Y, et al. Semaglutide in MASLD patients: improved survival and liver outcomes. *Pharmaceuticals (Basel)*. 2025;18(7):1075. <https://doi.org/10.3390/ph18071075>.
20. Schwabe RF, Tacke F, Sugimoto A, et al. Antifibrotic therapies for metabolic dysfunction-associated steatotic liver disease. *JHEP Rep*. 2025;7(8):1-29. <https://doi.org/10.1016/j.jhepr.2025.101421>.
21. Harrison SA, Rolph T, Knott M, et al. FGF21 agonists: an emerging therapeutic for metabolic dysfunction-associated steatohepatitis and beyond. *J Hepatol*. 2024;81(3):562-576. <https://doi.org/10.1016/j.jhep.2024.05.011>.
22. Chiang JYL. New drug therapies for metabolic dysfunction-associated steatohepatitis. *Liver Res*. 2025;9(2):94-103. <https://doi.org/10.1016/j.livres.2025.01.001>.
23. Stefan N, Yki-Järvinen H, Neuschwander-Tetri BA. Metabolic dysfunction-associated steatotic liver disease: heterogeneous pathomechanisms and effectiveness of metabolism-based treatment. *Lancet Diabetes Endocrinol*. 2025;13(2):1-15. [https://doi.org/10.1016/S2213-8587\(24\)00318-8](https://doi.org/10.1016/S2213-8587(24)00318-8).

Aneurisma aórtico micótico por enterobacterias multirresistentes en paciente anciano: desafío diagnóstico

Anabela Games^{1,✉}, Sofía Rossetti^{1,✉}, Ana Braslavsky^{2,✉}, Ignacio Martingano^{1,✉} y Tomás Barrera^{1,✉}

1. Servicio de Clínica Médica, Hospital Italiano. Buenos Aires, Argentina.

2. Servicio de Clínica Pediátrica, Hospital Italiano. Buenos Aires, Argentina.

RESUMEN

Se presenta el caso de un paciente de 86 años con antecedente de vía urinaria instrumentada que consultó por fiebre y debilidad generalizada. Se documentó bacteriemia, inicialmente interpretada como foco urinario por evidencia de litiasis y urocultivo positivo con enterobacterias multirresistentes, y se inició tratamiento antibiótico dirigido. Durante la evolución presentó persistencia de fiebre y marcadores inflamatorios elevados pese a tratamiento terapéutico adecuado. Por tomografía computarizada se evidenció un aneurisma sacular de la aorta abdominal, confirmado como aneurisma micótico mediante tomografía con emisión de positrones. Dada la edad avanzada y la fragilidad del paciente, se optó por un manejo conservador con antibioticoterapia prolongada. El paciente presentó una evolución clínica favorable.

Palabras clave: aneurisma de la aorta, aneurisma infectado, Enterobacteriaceae, anciano, bacteriemia, fragilidad.

Mycotic Aortic Aneurysm Caused by Multidrug-Resistant Enterobacteriaceae in an Elderly Patient: A Diagnostic Challenge

ABSTRACT

We report the case of an 86-year-old patient with a complex urological history who presented with fever and generalized weakness. Bloodstream infection was documented and initially attributed to a urinary source, based on the presence of urolithiasis and a positive urine culture for multidrug-resistant Enterobacteriaceae. Targeted antibiotic therapy was therefore initiated.

However, the patient showed persistent fever and sustained elevation of inflammatory markers despite appropriate antimicrobial treatment. Computed tomography subsequently revealed a sacular aneurysm of the abdominal aorta, which was further confirmed as a mycotic aneurysm by positron emission tomography. Given the patient's advanced age and frailty, a conservative management strategy with prolonged antibiotic therapy was selected. The patient showed a favorable clinical outcome.

Keywords: aortic aneurysm, aneurysm infected, Enterobacteriaceae, aged, bacteremia, frailty.

Autora para correspondencia: anabel.games@hospitalitaliano.org.ar, Games AM.

Recibido: 18/03/2026 | Aceptado: 08/07/2026 | Publicado: 28/09/2026

DOL: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1325>

Cómo citar: Games AM, Rosetti S, Braslavsky A, Martingano I, Barrera T. Aneurisma aórtico micótico por enterobacterias multirresistentes en paciente anciano: desafío diagnóstico. *Rev Hosp Ital B.Aires.* 2026;46(3):e0001325

INTRODUCCIÓN

Los aneurismas micóticos aórticos corresponden a dilataciones infecciosas de la pared de la aorta y constituyen una entidad infrecuente, que representa aproximadamente entre 0,7 y 4,5% de todos los aneurismas aórticos. A pesar de su baja incidencia se asocia con elevada morbilidad debido a la dificultad diagnóstica y al riesgo de complicaciones graves como sepsis, rotura aneurismática y compromiso de estructuras adyacentes¹.

El término “micótico” se originó por la apariencia macroscópica de estas lesiones y no necesariamente implica etiología fúngica. En la mayoría de los casos, el origen es bacteriano. Los microorganismos más frecuentemente involucrados son especies de *Salmonella*, *Staphylococcus aureus* y especies de *Streptococcus*, mientras que las enterobacterias representan una causa menos habitual².

El diagnóstico se basa en la integración de datos clínicos, microbiológicos e imagenológicos. Clínicamente, suele manifestarse con fiebre de origen desconocido (presente en aproximadamente el 70% de los casos) y dolor según la localización del aneurisma^{3,4}. Los hallazgos de laboratorio habitualmente incluyen elevación de reactantes de fase aguda. La tomografía computarizada (TC) contrastada constituye el estudio inicial de elección para evaluar la morfología vascular y alteraciones en la pared arterial, mientras que la tomografía por emisión de positrones asociada a tomografía computarizada (PET-TC) incrementa la certeza diagnóstica al detectar inflamación activa de la pared vascular⁴⁻⁶.

El tratamiento estándar consiste en la resección quirúrgica con antibioticoterapia dirigida. Las técnicas endovasculares son una alternativa en casos seleccionados, mientras que el manejo médico exclusivo se reserva para casos de elevado riesgo quirúrgico o en los cuales la intervención quirúrgica resulta inviable^{7,8}.

Presentamos este caso por la combinación de etiología por enterobacterias multirresistentes, confirmación diagnóstica mediante PET-TC y manejo exclusivamente médico en un paciente anciano frágil.

CASO CLÍNICO

Varón de 86 años que consultó por debilidad generalizada de cuatro días de evolución asociada a escalofríos en las últimas 48 horas.

Como antecedentes presentaba múltiples cirugías urológicas por estenosis uretral. Seis meses antes de la consulta había sido sometido a reconstrucción uretral con colgajo, complicada con fistula uretral distal y requerimiento de talla vesical.

Al ingreso se encontraba lúcido, en buen estado general y febril (38 °C), con signos vitales estables. El laboratorio mostró leucocitosis de 10 900/mm³ con neutrofilia, creatinina 1,33 mg/dL (previa 0,9), urea 69 mg/dL y proteína C reactiva de 230 mg/L. La gasometría arterial evidenció acidosis metabólica con lactato de 6,2 mmol/L. El sedimento urinario presentó leucocituria, hematuria y abundantes bacterias. Se realizó ecografía renal donde se descartó la presencia de litiasis, absceso y dilatación pielocalicial. Ante la presencia de reactantes de fase aguda

elevados con fallo renal no presente en estudios previos y sedimento urinario patológico, se interpretó el cuadro como infección urinaria alta. Se obtuvieron hemocultivos y urocultivo e inició tratamiento empírico con ceftriaxona. Posteriormente, los cultivos aislaron *Escherichia coli* productora de betalactamasas de espectro extendido (BLEE) y *Klebsiella pneumoniae* resistente a ampicilina, motivo por el cual se ajustó el tratamiento a imipenem durante ocho días.

Durante la evolución presentó lumbalgia persistente. La TC abdominopélvica evidenció una litiasis uretral distal derecha de 6 mm con dilatación de la vía excretora y cambios inflamatorios perirrenales. Se reinterpretó el cuadro como una infección urinaria complicada con litiasis en tránsito, la cual fue resuelta por litotricia uretral endoscópica con colocación de catéter doble J. Se retiró una semana después tras la resolución del cuadro obstructivo, con mejoría inicial del dolor, curva térmica y reactantes de fase aguda.

Posteriormente, el paciente presentó recurrencia de fiebre y escalofríos; el laboratorio mostró nuevamente elevación de reactantes de fase aguda y leucocitosis. Una nueva TC abdominopélvica con contraste evidenció dilatación pielocalicial derecha con estriación de planos grasos, pero sin causal obstructivo, así como la presencia de un aneurisma sacular de la aorta abdominal, hallazgo que no se encontraba presente en estudios previos (Fig. 1).

Los hemocultivos resultaron nuevamente positivos para *Escherichia coli*, BLEE y el urocultivo para *Citrobacter koserii*, por lo que se reinició tratamiento con meropenem.

Ante la persistencia de bacteriemia, se ampliaron estudios para evaluar focos secundarios. El ecocardiograma transtorácico mostró función ventricular conservada y ausencia de vegetaciones, mientras que el fondo de ojo no evidenció lesiones compatibles con embolias sépticas; se descartó endocarditis infecciosa como diagnóstico probable.

Tras una semana de antibioticoterapia dirigida, el paciente persistía febril y con marcadores inflamatorios elevados. La TC previa había evidenciado un aneurisma sacular de reciente aparición en la aorta abdominal. Ante estos hallazgos se solicitó una PET-TC para evaluar un posible foco vascular infeccioso. El estudio evidenció una intensa captación de fluorodesoxiglucosa (FDG) en la pared aneurismática, compatible con aneurisma micótico de la aorta abdominal, con aumento del tamaño respecto del estudio tomográfico previo (Figs. 2 y 3).

Dado el rápido crecimiento del aneurisma y el elevado riesgo de complicación, el paciente fue trasladado a la unidad de cuidados intensivos para monitorización hemodinámica.

El caso fue evaluado en forma multidisciplinaria. Debido a la fragilidad clínica moderada del paciente, objetivada mediante la *Clinical Frailty Scale*⁹ y a la complejidad técnica del procedimiento, se desestimó el tratamiento quirúrgico o endovascular. Se decidió continuar con manejo médico exclusivo con antibioticoterapia prolongada.

Durante la internación presentó evolución favorable, con resolución de la fiebre y descenso progresivo



Figura 1. TC abdominopélvica con contraste intravenoso, corte coronal en fase arterial, muestra un aneurisma sacular de $13 \times 12,2 \times 14,7$ mm en la aorta abdominal, superior a la bifurcación de las arterias renales, marcado con flecha roja. Se asocia a opacificaciones de la grasa retroperitoneal adyacentes.

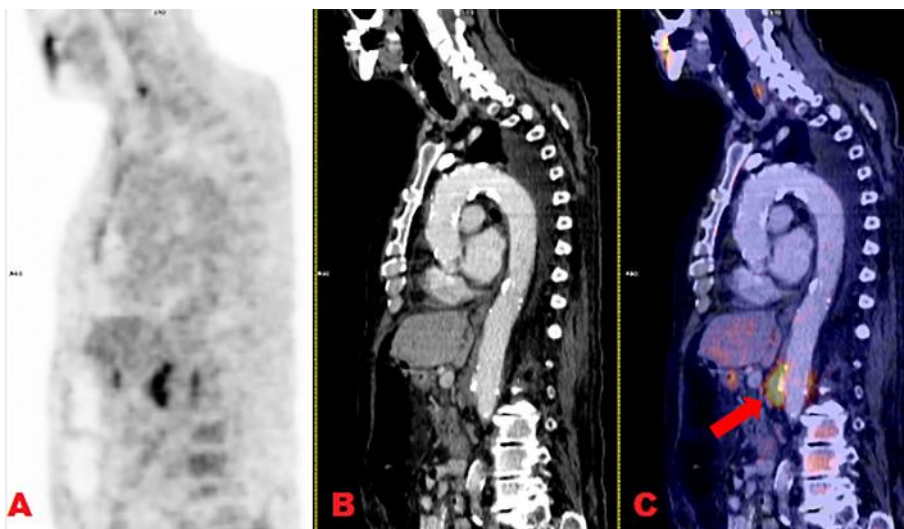


Figura 2. PET-TC con fluorodesoxiglucosa, corte sagital. En la imagen A correspondiente a la proyección PET se identifica una zona hipermetabólica en relación con la aorta abdominal. En la imagen de fusión C, se evidencia un aneurisma micótico de la aorta abdominal de aproximadamente $30 \times 21 \times 41,5$ mm, previa al nacimiento de las arterias renales.

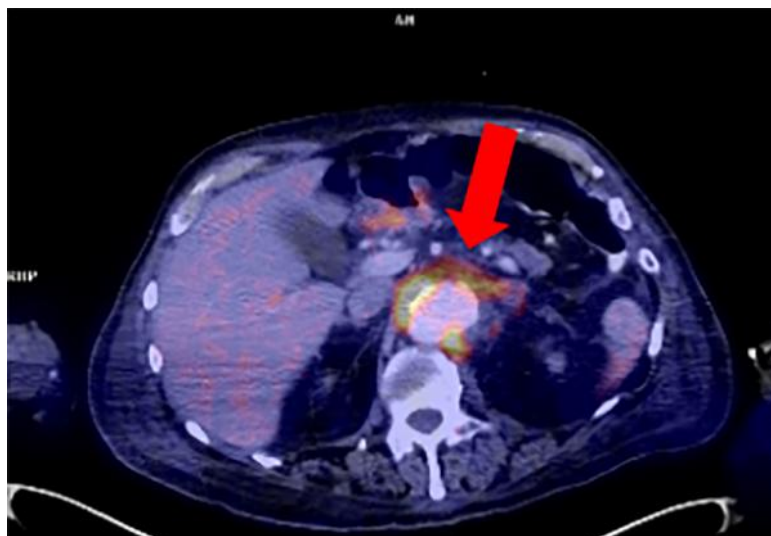


Figura 3. PET-TC corte transversal, imagen de fusión, que evidencia aneurisma micótico de la aorta abdominal. Se acompaña de un tejido con densidad de partes blandas hipermetabólico, con elevada captación de fluorodesoxiglucosa, que rodea la aorta con predominio izquierdo y alcanza un valor de captación estandarizado (*Standardized up take value, SUV*) de 10,8 (valor normal para partes blandas 0,5-1,5).

de parámetros inflamatorios. Fue dado de alta con la indicación de completar al menos seis semanas de antibioticoterapia intravenosa en domicilio, con seguimiento clínico e imagenológico. Posteriormente falleció por causas no relacionadas con el aneurisma micótico, motivo por el cual no se pudo completar el seguimiento posalta.

ASPECTOS ÉTICOS

Se obtuvo la aprobación del Comité de Ética institucional (CEPI 7277, PRIISA 14414).

DISCUSIÓN

Aunque la mayoría de los aneurismas micóticos se asocian a especies de *Salmonella* o *Staphylococcus aureus*, algunos informes han descripto enterobacterias como agentes etiológicos, particularmente en el contexto de infecciones urinarias o bacteriemias de origen abdominal. Sin embargo, los casos asociados a microorganismos multirresistentes continúan siendo escasos en la literatura^{2,7}.

Si bien no existen criterios diagnósticos universalmente aceptados para el aneurisma micótico, en el caso referido el diagnóstico se estableció en concordancia con lo propuesto por las principales revisiones y la declaración científica de la American Heart Association^{2,3}, mediante la integración de hallazgos clínicos, microbiológicos e imagenológicos. La bacteriemia persistente por *Escherichia coli*, BLEE, el síndrome inflamatorio a pesar de antibioticoterapia adecuada, el hallazgo de un aneurisma sacular de rápido crecimiento y cambios inflamatorios periaórticos en la TC, junto con la intensa captación de FDG en la PET-TC, fueron altamente sugestivos de origen infeccioso^{4,6}. Asimismo, la ausencia de vegetaciones en el ecocardiograma y de embolias sépticas en el fondo de ojo, junto con la falta de captación de FDG en la PET-TC,

hicieron poco probable la endocarditis infecciosa como diagnóstico, mientras que la coexistencia de bacteriemia documentada y los hallazgos metabólicos compatibles en la PET-TC apoyaron el diagnóstico de aneurisma micótico por sobre el inflamatorio no infeccioso^{4,6}.

En el tratamiento de los aneurismas micóticos, la cirugía abierta con resección del segmento afectado y reconstrucción vascular se considera el estándar terapéutico.

Las técnicas endovasculares han emergido como alternativa en pacientes seleccionados, con menor mortalidad perioperatoria, aunque con riesgo de complicaciones infecciosas tardías. En el caso presentado, la edad avanzada y la fragilidad del paciente motivaron la decisión de adoptar un enfoque conservador con antibioticoterapia prolongada. Aunque el tratamiento médico exclusivo se asocia en general con peor pronóstico, algunos informes describen resultados favorables cuando se seleccionan cuidadosamente los pacientes y se realiza un seguimiento estrecho^{7,8}.

Este caso destaca la importancia de mantener un alto índice de sospecha de infección endovascular en pacientes con fiebre de origen desconocido y bacteriemia persistente, especialmente cuando no existe respuesta clínica a pesar de la antibioticoterapia dirigida, así como el valor de la PET-TC para orientar el diagnóstico. Asimismo, destaca que, en pacientes seleccionados, previamente frágiles con elevado riesgo quirúrgico, el manejo médico exclusivo puede constituir una terapéutica razonable.

Contribución de los autores: Validación, Visualización, Redacción-revisión y edición (AMG, TB, IM, SR, AB). Contribuciones de los autores: Análisis formal, Administración del proyecto, Recursos (AMG, TB, IM, SR). Supervisión (TB, IM, SR, AB). Investigación, Metodología (AMG, AB). Curación de datos (AMG, TB). Conceptualización, Software, Redacción-borrador original (AMG).

Conflictos de intereses: los autores declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: los autores declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Oderich GS, Panneton JM, Bower TC, et al. Infected aortic aneurysms: aggressive presentation, complicated early outcome, but durable results. *J Vasc Surg.* 2001;34(5):900-908. <https://doi.org/10.1067/mva.2001.118084>.
2. Sörelus K, Budtz-Lilly J, Mani K, Wanhainen A. Systematic review of the management of mycotic aortic aneurysms. *Eur J Vasc Endovasc Surg.* 2019;58(3):426-435. <https://doi.org/10.1016/j.ejvs.2019.05.004>.
3. Wilson WR, Bower TC, Creager MA, et al. Vascular graft infections, mycotic aneurysms, and endovascular infections: a scientific statement from the American Heart Association. *Circulation.* 2016;134(20):e412-e460. <https://doi.org/10.1161/CIR.0000000000000457>.
4. Zhang N, Xiong W, Li Y, et al. Imaging features of mycotic aortic aneurysms. *Quant Imaging Med Surg.* 2021;11(6):2861-2878. <https://doi.org/10.21037/qims-20-941>.
5. Goldstein SA, Evangelista A, Abbata S, et al. Multimodality imaging of diseases of the thoracic aorta in adults: from the American Society of Echocardiography and the European Association of Cardiovascular Imaging: endorsed by the Society of Cardiovascular Computed Tomography and Society for Cardiovascular Magnetic Resonance. *J Am Soc Echocardiogr.* 2015;28(2):119-182. <https://doi.org/10.1016/j.echo.2014.11.015>.
6. Husmann L, Huellner MW, Ledergerber B, et al. Diagnostic accuracy of PET/CT and contrast enhanced CT in patients with suspected infected aortic aneurysms. *Eur J Vasc Endovasc Surg.* 2020;59(6):972-981. <https://doi.org/10.1016/j.ejvs.2020.01.032>.
7. Lin CH, Hsu RB. Primary infected aortic aneurysm: clinical presentation, pathogen, and outcome. *Acta Cardiol Sin.* 2014;30(6):514-521. <https://doi.org/10.6515/ACS20140630A>.
8. Kim YW. Infected aneurysm: current management. *Ann Vasc Dis.* 2010;3(1):7-15. <https://doi.org/10.3400/avd.AVDctia09003>.
9. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ.* 2005;173(5):489-95. <https://doi.org/10.1503/cmaj.050051>.

Lesión hepática grave asociada al consumo de productos fitoterapéuticos no regulados: Informe de caso y revisión de la literatura

María A. De la Rosa^{1, }, José Sanabria Vanegas^{1, } y Gabriela Beltrán Gil^{2, }

1. Departamento de Farmacología Clínica, Universidad de La Sabana, Cundinamarca, Colombia

2. Profesional independiente. Cundinamarca, Colombia

RESUMEN

El uso de productos fitoterapéuticos y suplementos herbales se ha incrementado, en parte por la percepción de que lo “natural” carece de riesgos. Sin embargo, su consumo puede ocasionar lesión hepática inducida por hierbas (HILI), particularmente cuando se utilizan sin supervisión o circulan como productos falsificados. Se presenta el caso de un hombre de 85 años que consultó por dolor en hipocondrio derecho e ictericia, con lesión hepática de patrón hepatocelular (R = 11,3). Durante la anamnesis se documentó el consumo, durante seis meses, de seis productos fitoterapéuticos, entre ellos Redumex, cuyo registro no pudo verificarse, y Vitacerebrina Francesa, lote VTMA513, identificada por el INVIMA como un producto fraudulento con registro sanitario falso. La causalidad se clasificó como posible mediante RUCAM (5 puntos), mientras que la gravedad correspondió a DILIN grado 4. Tras la suspensión de los productos, se observó mejoría clínica y una disminución progresiva de las aminotransferasas y la fosfatasa alcalina, lo que permitió el alta hospitalaria con indicación de evitar la reexposición. Este caso destaca la necesidad de investigar sistemáticamente el consumo de productos herbales y fortalecer la farmacovigilancia y el control regulatorio de productos falsificados.

Palabras clave: enfermedad hepática inducida por sustancias y drogas, preparaciones de plantas, suplementos dietéticos, farmacovigilancia, medicamentos falsificados.

Severe Liver Injury Associated with the Use of Unregulated Herbal Products: A Case Report and Literature Review

Abstract

The use of phytotherapeutic products and herbal supplements has increased, partly because many people believe “natural” products are harmless. However, their consumption may cause herb-induced liver injury (HILI), particularly when they are used without medical supervision or marketed as falsified products. We report the case of an 85-year-old man who presented with right upper quadrant pain and jaundice and was found to have a hepatocellular pattern of liver injury (R-value 11.3). Further history revealed the use of six phytotherapeutic products for six months, including Redumex, whose registration could not be verified, and Vitacerebrina Francesa, lot VTMA513, which INVIMA identified as a fraudulently marketed product bearing a false sanitary registration. Causality was classified as possible using RUCAM (score: 5), whereas severity was classified as DILIN grade 4. Following withdrawal of the

Autor para correspondencia: mariaacoros@unisabana.edu.co, de la Rosa MA.

Recibido: 3/06/2026 | Aceptado: 14/09/2026 | Publicado: 29/09/2026

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1333>

Cómo citar: Aconcha de la Rosa M, Sanabria Vanegas JS, Beltrán Gil G. Lesión hepática grave asociada al consumo de productos fitoterapéuticos no regulados: Informe de caso y revisión de la literatura. *Rev Hosp Ital B.Aires.* 2026;46(3):e0001333

products, the patient showed clinical improvement and a progressive decline in aminotransferase and alkaline phosphatase levels, allowing hospital discharge with instructions to avoid re-exposure. This case highlights the importance of systematically investigating herbal product use and strengthening pharmacovigilance and regulatory control of falsified products.

Keywords: chemical and drug-induced liver injury, plant preparations, dietary supplements, pharmacovigilance, counterfeit drugs.

INTRODUCCIÓN

El uso de productos fitoterapéuticos y suplementos dietarios a base de hierbas es cada vez más frecuente. La Organización Mundial de la Salud estima que aproximadamente el 80% de la población mundial utiliza medicina tradicional, incluidos los preparados a base de hierbas.¹ Estos productos, comercializados en forma de raíces, semillas, hojas, infusiones, polvos, aceites, cremas, cápsulas o inyectables, son ampliamente utilizados debido a que pueden adquirirse sin receta, tienen bajo costo y se perciben como naturales y seguros.²

Estos productos pueden causar una lesión hepática clínicamente manifiesta inducida por productos herbales (*herb-induced liver injury*, HILI), que se distingue de la lesión hepática inducida por fármacos convencionales (*drug-induced liver injury*, DILI). En la Red de Lesión Hepática Inducida por Fármacos de los Estados Unidos (US Drug-Induced Liver Injury Network, DILIN), representan aproximadamente el 20% de los casos.³ Es probable que la frecuencia de HILI esté subestimada, ya que los pacientes pueden no informar su consumo y los casos pueden no reconocerse o notificarse.²⁻⁴ LiverTox proporciona información actualizada sobre el diagnóstico, los patrones de lesión, los desenlaces y los mecanismos.⁵ Las estimaciones regionales varían según los patrones de consumo, las poblaciones estudiadas, las definiciones utilizadas, los métodos de detección, la derivación de pacientes y los sistemas de notificación.^{3,6}

La evaluación de la causalidad resulta difícil porque los productos pueden contener múltiples ingredientes, presentar variaciones entre lotes y tener un etiquetado inexacto. Ante la ausencia de un biomarcador confirmatorio, HILI continúa siendo un diagnóstico de exclusión basado en los antecedentes de exposición, el patrón bioquímico, la evaluación de causas alternativas y la evolución tras la suspensión del producto; el manejo requiere su discontinuación y el seguimiento clínico y bioquímico.^{3,4}

Aunque HILI es una entidad bien establecida, este caso resulta relevante porque involucra múltiples productos de composición incierta en Colombia, entre ellos Vitacebrina Francesa, lote VTMA513, que el INVIMA identificó como fraudulento y con un registro sanitario falso.⁷ Describimos un caso de lesión hepática hepatocelular grave posiblemente asociada con el uso concomitante de productos fitoterapéuticos y analizamos sus implicancias causales, clínicas, regulatorias y de farmacovigilancia. Este caso pone de manifiesto que el origen natural y la

disponibilidad comercial no garantizan la seguridad, la calidad ni la autenticidad de estos productos, y respalda la evaluación sistemática de la exposición, la verificación regulatoria y la notificación oportuna.

PRESENTACIÓN DEL CASO

Un paciente de 85 años, con antecedentes de colecistectomía 30 años antes, hipertensión arterial e hipotiroidismo, consultó en el servicio de emergencias. Su medicación habitual incluía levotiroxina 50 µg/día, enalapril 5 mg/día, que recibía de manera continua desde hacía aproximadamente 30 años, y omeprazol 20 mg/día, que tomaba desde hacía un año. Consultó por un cuadro de un día de evolución caracterizado por dolor abdominal cólico en el hipocondrio derecho, irradiado a la espalda y asociado con ictericia. En el examen físico, el paciente presentaba dolor agudo e ictericia generalizada. El abdomen era doloroso a la palpación en el hipocondrio derecho y el epigastrio; el resto del examen físico no mostró hallazgos relevantes. Al ingreso, los parámetros hemodinámicos se encontraban dentro de los rangos normales, sin evidencia de hipotensión, shock ni hipoperfusión tisular. El paciente refirió consumo ocasional de alcohol, limitado a situaciones sociales.

Los estudios de laboratorio iniciales mostraron elevación de las aminotransferasas, la fosfatasa alcalina y la bilirrubina (Tabla 1). Utilizando los límites superiores de normalidad específicos del laboratorio para ALT (40 U/L) y fosfatasa alcalina (115 U/L), el valor R fue de 11,3, lo que indicó un patrón hepatocelular de lesión hepática.⁴ Según la escala de gravedad de la Drug-Induced Liver Injury Network, el episodio se clasificó como grado 4 (grave), ya que la elevación de las enzimas hepáticas se acompañó de una bilirrubina total de 15,3 mg/dL y un INR de 1,91, cumpliendo así los criterios de bilirrubina total $\geq 2,5$ mg/dL e INR $\geq 1,5$.⁸ Una ecografía hepatobiliar no mostró dilatación de la vía biliar ni alteraciones del parénquima. A pesar del antecedente remoto de colecistectomía, estos hallazgos hicieron menos probable una causa obstructiva biliar. Los estudios diagnósticos resultaron negativos para VIH, hepatitis A, B y C, citomegalovirus y virus de Epstein-Barr. No se realizaron pruebas para hepatitis E ni estudios serológicos para enfermedad hepática autoinmune, incluidos anticuerpos antinucleares, antimusculo liso y antimitocondriales.

Al volver a entrevistar al paciente y a su familia, un familiar confirmó que el paciente había estado consumiendo durante seis meses productos fitoterapéuticos para

Tabla 1. La tabla presenta los datos del paciente al ingreso y la evolución de los parámetros de laboratorio durante la internación; — indica que el parámetro no se midió en esa fecha. AST, aspartato aminotransferasa; ALT, alanina aminotransferasa; ALP, fosfatasa alcalina; PT, tiempo de protrombina; PTT, tiempo parcial de tromboplastina; INR, razón internacional normalizada. Los límites superiores de normalidad específicos del laboratorio fueron 40 U/L para ALT y 115 U/L para ALP. Los resultados de ALT y ALP al ingreso se obtuvieron en la misma fecha. El valor R se calculó de la siguiente manera: $R = (1076/40)/(274/115) = 11,3$, compatible con un patrón hepatocelular de lesión hepática.⁴

Parámetro	15/06/2024	20/06/2024	21/06/2024	22/06/2024	23/06/2024	24/06/2024	25/06/2024
AST (U/L)	892	712,9	620,8	570,4	560,4	336,9	259,2
ALT (U/L)	1076	777,2	609,6	561,7	562,7	424,6	369,7
Bilirrubina total (mg/dL)	15,30	17,60	19,30	—	19,16	16,87	16,73
Bilirrubina directa (mg/dL)	13,60	15,72	16,72	—	16,61	15,14	15,31
Bilirrubina indirecta (mg/dL)	1,84	2,03	2,68	—	2,54	1,72	1,42
ALP (U/L)	274	—	—	—	204	185	194
PT	20,5	21,7	22,2	21,1	19,1	19,0	—
PTT	32,1	31,2	31,8	31,8	29,0	29,3	—
INR	1,91	2,04	2,09	1,98	1,77	1,76	—

el malestar general y la astenia, junto con suplementos dietarios a base de hierbas. Teniendo en cuenta la exposición documentada, la relación temporal y los estudios diagnósticos disponibles, se consideró HILI como un posible diagnóstico. Al ingreso, suspendimos todos los productos fitoterapéuticos y el omeprazol, y mantuvimos el enalapril. Tras una mejoría inicial de las pruebas bioquímicas hepáticas, se reintrodujo el omeprazol el sexto día de internación, sin que posteriormente se observara un empeoramiento de los parámetros bioquímicos. Los niveles de aminotransferasas y fosfatasa alcalina disminuyeron progresivamente durante la internación.

La investigación posterior reveló que el paciente había estado consumiendo seis productos fitoterapéuticos diferentes. Dos de ellos presentaban números de registro sanitario que no coincidían con registros válidos en la base de datos del INVIMA; uno de estos productos era Redumex. El paciente también consumía Vitacerebrina Francesa, lote VTMA513, que la Alerta N.º 020-2021 del INVIMA identificó específicamente como un producto comercializado de manera fraudulenta y con un registro sanitario falso.⁷ Según la información declarada en las etiquetas de los productos, los ingredientes botánicos incluían té verde, moringa, ajo y anamú (*Petiveria alliacea*). Sin embargo, dado que algunos de los productos presentaban registros sanitarios falsificados y no realizamos análisis químicos, no pudimos verificar de manera independiente la presencia ni las concentraciones de los ingredientes declarados.

Sobre la base de los hallazgos disponibles, la asociación entre los productos fitoterapéuticos y la lesión hepática se clasificó como posible (puntaje RUCAM: 5), y no como confirmada.⁹ Dada la mejoría del patrón de lesión hepatocelular y el adecuado control de los síntomas, el paciente recibió el alta con indicaciones estrictas de suspender todos los productos fitoterapéuticos. La Tabla 1 presenta los parámetros de laboratorio

medidos durante la internación, mientras que las figuras 1 y 2 muestran la evolución temporal de los niveles de aminotransferasas, fosfatasa alcalina y bilirrubina. No se dispuso de los resultados de las pruebas bioquímicas hepáticas posteriores al alta para incluirlos en este informe; por lo tanto, no pudimos documentar la normalización bioquímica completa.

DISCUSIÓN

La principal contribución de este caso radica en su contexto local y de farmacovigilancia. En lugar de utilizar un único preparado bien caracterizado, el paciente consumía seis productos fitoterapéuticos comerciales, algunos de ellos con registros sanitarios que no pudieron verificarse. El INVIMA identificó a Vitacerebrina Francesa, lote VTMA513, como un producto fraudulento con un registro sanitario falso.⁷ Estos productos no ofrecen garantías fiables en cuanto a su composición real, las concentraciones de sus ingredientes, las condiciones de fabricación, su seguridad o su eficacia.^{4,7} Aunque las etiquetas mencionaban varios productos botánicos, no fue posible verificar de manera independiente su presencia ni sus concentraciones. Por consiguiente, la exposición simultánea y la incertidumbre respecto de la composición impidieron atribuir la lesión a un único preparado o ingrediente. La alerta regulatoria aumentó la relevancia de esta exposición para la salud pública, pero no permitió establecer una relación causal.

Las etiquetas declaraban té verde, moringa, ajo y anamú (*Petiveria alliacea*), aunque la evidencia sobre hepatotoxicidad difiere considerablemente entre ellos. El extracto de té verde presenta la asociación mejor documentada con lesión hepática clínicamente manifiesta, de patrón predominantemente hepatocelular, que en ocasiones puede ser grave.¹⁰ LiverTox clasifica a la moringa como una causa posible, aunque muy infrecuente, sobre la base principalmente de un caso con

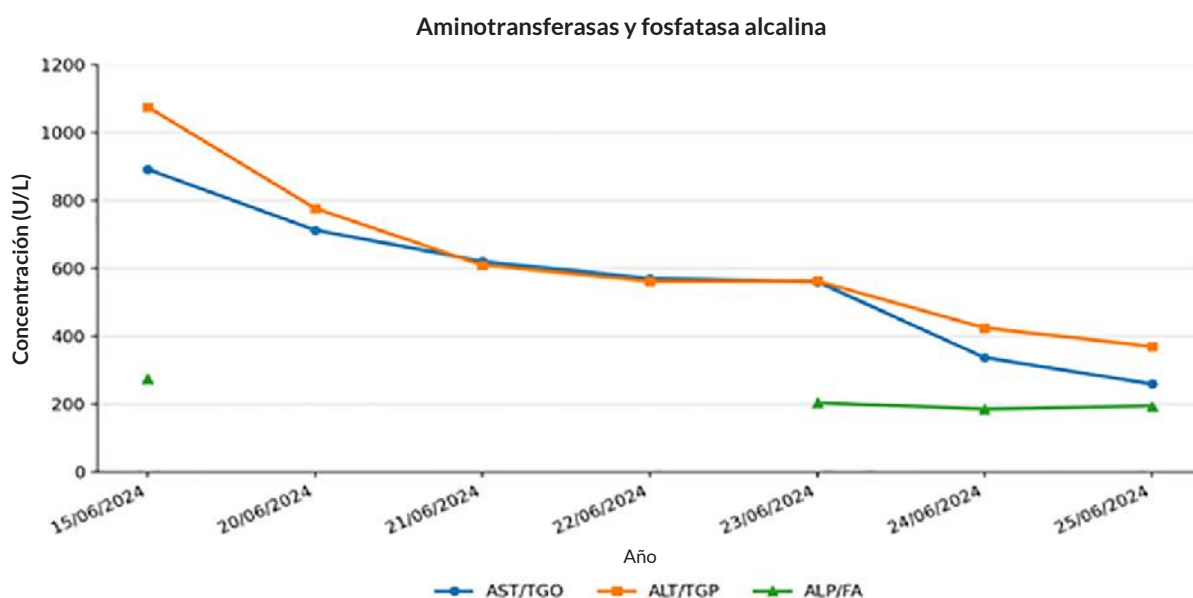


Figura 1. La figura muestra la evolución de las pruebas de función hepática durante la internación del paciente, incluidas AST (aspartato aminotransferasa), ALT (alanina aminotransferasa) y ALP (fosfatasa alcalina).

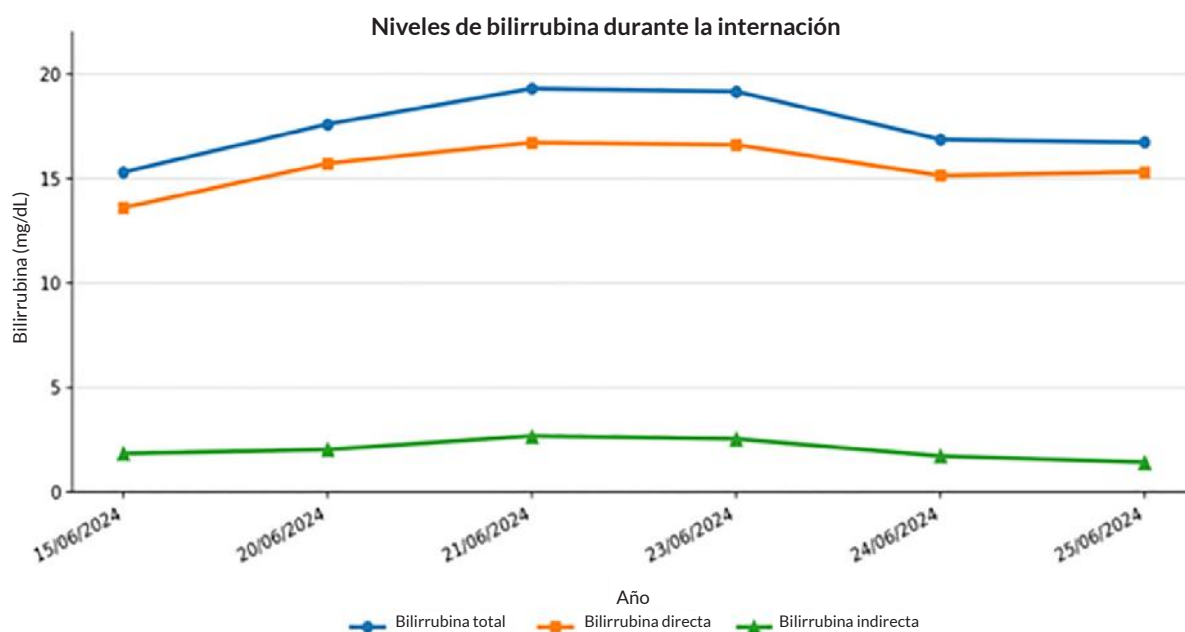


Figura 2. La figura muestra la evolución de los niveles de bilirrubina del paciente durante la internación, incluidos la bilirrubina total (TB), la bilirrubina directa (DB) y la bilirrubina indirecta (IB).

recurrencia tras la reexposición.¹¹ La evidencia sobre el ajo se limita a un informe aislado en un receptor de trasplante hepático,¹² mientras que la evidencia sobre el anamú es predominantemente preclínica y no permite establecer HILI en seres humanos.¹³ Por lo

tanto, ninguno de estos componentes puede considerarse individualmente responsable de la lesión en este paciente, particularmente debido a la exposición a seis productos y a la composición incierta de aquellos con registros falsificados.

El método actualizado de evaluación de causalidad de Roussel Uclaf (*Roussel Uclaf Causality Assessment Method, RUCAM*) arrojó un puntaje de 5, correspondiente a una causalidad “posible”.⁹ La exposición documentada, una relación temporal compatible, un patrón hepatocelular y la mejoría progresiva tras la suspensión de los productos respaldaron el diagnóstico de HILI. Sin embargo, la mejoría no confirmó la causalidad, ya que los seis productos se suspendieron simultáneamente. La medicación concomitante y la exclusión incompleta de etiologías alternativas redujeron aún más la certeza e impidieron atribuir la lesión a un preparado específico. Por lo tanto, los hallazgos respaldan un diagnóstico posible, pero no definitivo, de HILI.

La gravedad se evaluó independientemente del grado de certeza causal. El episodio cumplió los criterios de grado 4 de DILIN porque la elevación de las enzimas hepáticas se acompañó de una bilirrubina total de 15,3 mg/dL y un INR de 1,91.⁸ Por lo tanto, “grave” describe la magnitud clínica del episodio y no una atribución definitiva a los productos fitoterapéuticos. Aunque el patrón hepatocelular presentó un valor R de 11,3, la ALP inicial fue 2,38 veces el límite superior de normalidad. Por consiguiente, el caso no cumplió con la definición bioquímica tradicional de la ley de Hy, que requiere una ALP inferior a dos veces el límite superior de normalidad, y la ley de Hy no se utilizó para determinar la gravedad ni la causalidad.⁴

La evaluación de causalidad tuvo en cuenta el enalapril y el omeprazol. No se dispuso de información sobre la indicación del tratamiento crónico con omeprazol. Aunque ambos fármacos se han asociado en raras ocasiones con lesión hepática idiosincrásica,^{14,15} el paciente había recibido enalapril en una dosis estable durante aproximadamente 30 años. El tratamiento se mantuvo mientras las pruebas hepáticas mejoraban, lo que hizo poco probable que fuera la causa. El paciente había recibido omeprazol durante un año; se suspendió al ingreso y se reintrodujo el sexto día de internación, después de la mejoría inicial, sin que posteriormente se observara un empeoramiento. Dado que la reintroducción se produjo antes de la normalización bioquímica completa, no se consideró una reexposición formalmente negativa según RUCAM.⁹ No obstante, la recuperación sostenida hizo que el omeprazol se considerara una causa menos probable que la exposición combinada a los productos fitoterapéuticos.

Persistieron importantes limitaciones diagnósticas. No se realizaron pruebas para hepatitis E ni estudios serológicos para enfermedad autoinmune, incluidos anticuerpos antinucleares, antímúsculo liso y antimitocondriales; por lo tanto, estas etiologías no pudieron excluirse por completo. Por otra parte, los parámetros hemodinámicos normales y la ausencia de hipotensión o shock hicieron poco probable una lesión isquémica. El consumo ocasional de alcohol, limitado a situaciones sociales, no sugirió una lesión hepática asociada con el alcohol. Estas limitaciones respaldan la cautelosa clasificación de causalidad como “posible” según RUCAM.^{4,9}

Este caso también tiene implicancias para el sistema de salud. Los profesionales de la salud deberían preguntar sistemáticamente y sin emitir juicios sobre el consumo de productos herbales y suplementos, y documentar su nombre comercial, fabricante, registro sanitario, lote, duración del consumo y lugar de compra.⁴ La verificación de estos datos en los registros oficiales puede permitir la identificación de productos fraudulentos, mientras que la notificación oportuna a los programas institucionales de farmacovigilancia y al INVIMA puede contribuir a la detección de señales, la comunicación de riesgos y la adopción coordinada de medidas regulatorias.⁷ La educación de la comunidad debería cuestionar la creencia de que los productos naturales son intrínsecamente seguros y fomentar que los pacientes informen su consumo a los profesionales de la salud.

CONCLUSIÓN

La creciente popularidad de los productos herbales, sumada a la automedicación, pone de manifiesto la necesidad de una regulación rigurosa y de una evaluación de la calidad y la seguridad de estos productos. Más allá del episodio clínico individual, este caso ilustra un riesgo para la salud pública potencialmente prevenible, asociado con la automedicación y la circulación de productos fitoterapéuticos falsificados. Fortalecer la vigilancia regulatoria, promover la notificación oportuna en los sistemas de farmacovigilancia y educar a los pacientes y a los profesionales de la salud son medidas esenciales para mejorar la detección y reducir la aparición de episodios similares.

Contribuciones de los autores: conceptualización, Curación de datos, Análisis formal, Investigación, Metodología, Administración del proyecto, Recursos, Software, Supervisión, Validación, Visualización, Redacción-borrador original: MAR, JSV. Redacción-revisión y edición: MAR, JSV, GBG.

Conflictos de intereses: los autores declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: los autores declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. World Health Organization. WHO establishes the Global Centre for Traditional Medicine in India [Internet]. Geneva: WHO; 2022 Mar 25 [citado 2026 ago 30]. Disponible en: <https://www.who.int/news/item/25-03-2022-who-establishes-the-global-centre-for-traditional-medicine-in-india>.
2. Girón Peláez D, Marulanda Fernández H, Otero Regino W. Conceptos emergentes: injuria hepática inducida por productos herbales (HILI). Revisión de tema. Rev Colomb Gastroenterol. 2019;34(1):61-68. <https://doi.org/10.22516/25007440.355>.
3. Navarro VJ, Khan I, Björnsson E, et al. Liver injury from herbal and dietary supplements. Hepatology. 2017;65(1):363-373. <https://doi.org/10.1002/hep.28813>.
4. Fontana RJ, Liou I, Reuben A, et al. AASLD practice guidance on drug, herbal, and dietary supplement-induced liver injury. Hepatology. 2023;77(3):1036-1065. <https://doi.org/10.1002/hep.32689>.
5. National Institute of Diabetes and Digestive and Kidney Diseases.

- LiverTox: clinical and research information on drug-induced liver injury [Internet]. Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases; 2012- [citado 2026 ago 30]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK547852/>.
6. Mengual-Moreno E, Lizarzábal-García M, Hernández-Rincón I, et al. Hepatotoxicidad asociada a hierbas y productos nutricionales de origen botánico. *Invest Clin*. 2015;56(3):320-335.
 7. Instituto Nacional de Vigilancia de Medicamentos y Alimentos. Alerta sanitaria No. 020-2021: Vitacerebrina francesa lote VTMA513 [Internet]. Bogotá: INVIMA; 2021 Feb 9 [citado 2026 ago 30]. Disponible en: <https://www.invima.gov.co/biblioteca/alerta-sanitaria-vitacerebrina-francesa-vtma513>.
 8. National Institute of Diabetes and Digestive and Kidney Diseases. Severity grading in drug induced liver injury. En: LiverTox: clinical and research information on drug-induced liver injury [Internet]. Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases; 2012 ; actualizado 2019 may 4 [citado 2026 ago 30]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK548241/>.
 9. Danan G, Teschke R. RUCAM in drug and herb induced liver injury: the update. *Int J Mol Sci*. 2015;17(1):14. <https://doi.org/10.3390/ijms17010014>.
 10. Hoofnagle JH, Bonkovsky HL, Phillips EJ, et al. HLA-B*35:01 and green tea-induced liver injury. *Hepatology*. 2021;73(6):2484-2493. <https://doi.org/10.1002/hep.31538>.
 11. National Institute of Diabetes and Digestive and Kidney Diseases. Moringa. En: LiverTox: clinical and research information on drug-induced liver injury [Internet]. Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases; 2012; updated 2024 Jul 15 [citado 2026 ago 30]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK605172/>.
 12. Shaikh SA, Tischer S, Choi EK, et al. Good for the lung but bad for the liver? Garlic-induced hepatotoxicity following liver transplantation. *J Clin Pharm Ther*. 2017;42(5):646-648. <https://doi.org/10.1111/jcpt.12558>.
 13. Conceição BCD, Silva TAD, Pantoja LVPDS, et al. Amazonian plants: a global bibliometric approach to *Petiveria alliacea* L. Pharmacological and toxicological properties. *Plants (Basel)*. 2023;12(18):3343. <https://doi.org/10.3390/plants12183343>.
 14. National Institute of Diabetes and Digestive and Kidney Diseases. Enalapril. En: LiverTox: clinical and research information on drug-induced liver injury [Internet]. Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases; 2012; updated 2018 Feb 11 [citado 2026 ago 30]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK547936/>.
 15. National Institute of Diabetes and Digestive and Kidney Diseases. Omeprazole. En: LiverTox: clinical and research information on drug-induced liver injury [Internet]. Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases; 2012; updated 2019 Apr 15 [citado 2026 ago 30]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK548771/>.

En búsqueda de una $p < 0,05$: el problema de las comparaciones múltiples

María E. Knorre , Mariano Bergier  y Arturo M. Cagide 

Servicio de Cardiología, Hospital Italiano. Buenos Aires, Argentina

RESUMEN

El artículo analiza el papel del azar y la estadística en la investigación médica a partir del concepto de significación estadística y el valor de p . Si bien la aleatorización en los ensayos clínicos permite controlar el azar, el riesgo de obtener resultados falsamente positivos persiste cuando se realizan múltiples comparaciones sin los ajustes estadísticos correspondientes.

Esta problemática se ilustra mediante una analogía con una apuesta ganadora en el casino, en la que se presenta un resultado exitoso sin considerar los intentos fallidos previos, situación comparable al análisis *post hoc* sin corrección. Se describen distintas estrategias para reducir este sesgo, como la definición *a priori* de hipótesis y puntos finales, la aplicación de correcciones estadísticas por comparaciones múltiples y la validación de los hallazgos en cohortes independientes. Finalmente, se destaca que la inteligencia artificial tampoco está exenta de este problema si no se aplican criterios estadísticos adecuados.

Palabras clave: análisis estadístico, significación estadística, análisis de varianza, investigación clínica, inteligencia artificial.

In Search of a $p < 0.05$: The Problem of Multiple Comparisons

ABSTRACT

This article examines the role of chance and statistics in medical research, focusing on statistical significance and the p -value. Although randomization in clinical trials helps control randomness, the risk of false positive findings persists when multiple comparisons are performed without appropriate statistical adjustment.

This issue is illustrated through an analogy involving a winning casino bet, in which a successful outcome is reported without disclosing previous failed attempts, a situation comparable to uncorrected *post hoc* analyses. Strategies to mitigate this bias are discussed, including the *a priori* definition of hypotheses and endpoints, the application of statistical corrections for multiple testing, and the validation of findings in independent cohorts. Finally, the article emphasizes that artificial intelligence is also susceptible to this problem when data are not properly managed.

Keywords: statistics as topic, statistical significance, analysis of variance, clinical trials as topic, artificial intelligence.

Autora para correspondencia: eugenia.knorre@hospitalitaliano.org.ar, Knorre ME.

Recibido: 1/12/2025 | Aceptado: 28/04/2025 | Publicado: 9/09/2026

DOI: <http://doi.org/10.51987/rev.hosp.ital.b.aires.v46i3.1301>

Cómo citar: Knorre ME, Bergier M, Cagide AM. En búsqueda de una $p < 0,05$: el problema de las comparaciones múltiples. Rev Hosp Ital B.Aires. 2026;46(3):e0001301

INTRODUCCIÓN

¿Existen puntos de contacto entre el azar y la medicina? En principio la respuesta es negativa.

En medicina, las decisiones tienen un fundamento lógico basado en evidencias y en la experiencia del profesional. El azar solo interviene cuando, en un ensayo de investigación experimental en forma aleatoria, se asigna el tratamiento o el control que recibirán los grupos en comparación. La base conceptual es que el azar asigna por igual aquellas variables que pudieran afectar la evolución, de modo que la única diferencia es el tratamiento, objetivo del estudio. Bajo esta premisa, una $p < 0,05$ indica que la posibilidad de que la conclusión resulte falsamente positiva es de solo 5 en 100.

Vamos ahora a contraponer las consideraciones anteriores con la siguiente experiencia de la vida real.

En un grupo de amigos de vacaciones, uno de ellos comenta que la noche anterior concurrió al casino donde obtuvo una ganancia suculenta apostando un pleno al número 18. Sus amigos sorprendidos preguntan detalles. El afortunado dice, "ocurre que..."

La anécdota, perfectamente creíble, queda condicionada por una sutileza inicialmente no comentada.

DESARROLLO

Sobre bolilleros y probabilidades (conceptos estadísticos)

Supongamos que en un bolillero se dispone de 1000 bolillas, cada una rotulada con una determinada diferencia, signo positivo o negativo; y cuya *distribución de*

frecuencias corresponde a una curva de tipo campana, de modo que el valor medio de esas diferencias es igual a cero (Fig. 1). El sector sombreado representa el 5% de las diferencias extremas (50 bolillas) 2,5% con valores negativos (25 bolillas) y 2,5% con positivos (25 bolillas).

La posibilidad de obtener una bolilla con una diferencia, negativa o positiva de las existentes en el sector negro, es de solo el 5%. Podemos incluir la bolilla en el bolillero y efectuar una nueva extracción repitiendo el experimento 100 veces, al cabo de los cuales habríamos obtenido 5 bolillas con la diferencia de las contenidas en el sector sombreado de la curva.

Del bolillero a la estadística inferencial aplicada

Dejemos el azar y vayamos a la investigación clínica¹⁻⁴.

Supongamos que un estudio observacional pretende estimar el valor de un indicador diagnóstico, pronóstico o el resultado de un tratamiento, o, un ensayo con diseño aleatorizado, pretende estimar el efecto de una determinada intervención. En ambos casos se están comparando dos muestras diferenciadas con la variable objeto del estudio (*punto final*) para demostrar si existe o no una "diferencia estadísticamente significativa".

Partiendo del criterio de que no hay diferencias (*hipótesis nula*), las probables diferencias halladas en cuanto al punto final entre ambas muestras tendrían una distribución (*distribución de frecuencias*) como la ilustrada en la figura 2 (similar a la Fig. 1) con un valor medio (*media*) de cero.

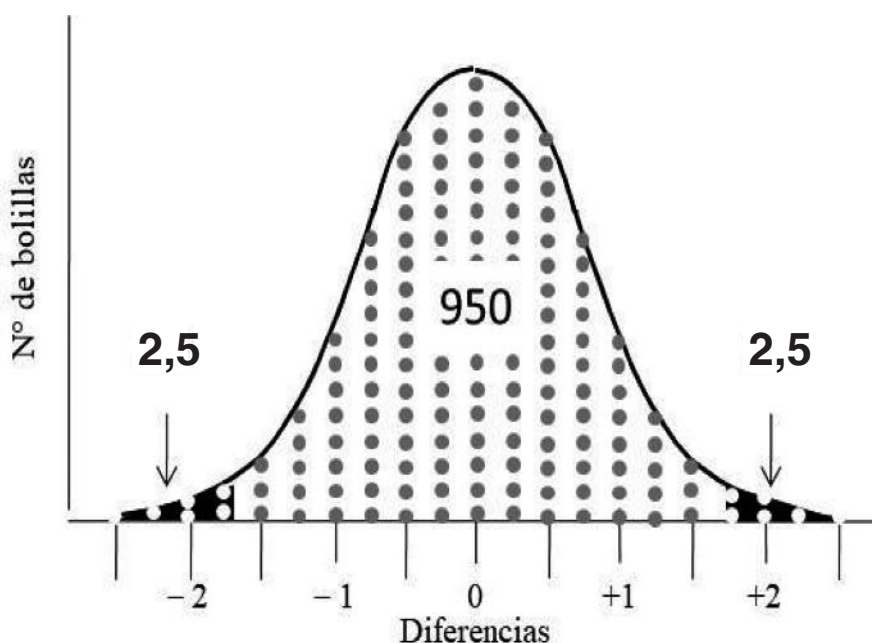


Figura 1. Esta figura describe la probabilidad de hallar en un bolillero una bolilla rotulada con una determinada diferencia (positiva o negativa): si en un bolillero se incluyen 1000 bolillas, 950 con diferencias en el rango de $-1,5$ a $+1,5$ y 50 con diferencias extremas, 25 con $-1,5$ y 25 con $+1,5$, la probabilidad de extraer una de estas bolillas es 5%. Dicha probabilidad se incrementa progresivamente si se repite la prueba resultado que luego de 100 intentos se habrán obtenido 5 bolillas con esas diferencias extremas (o si se prefiere, la relación puede expresarse de 20 a 1). Fuente: elaboración propia.

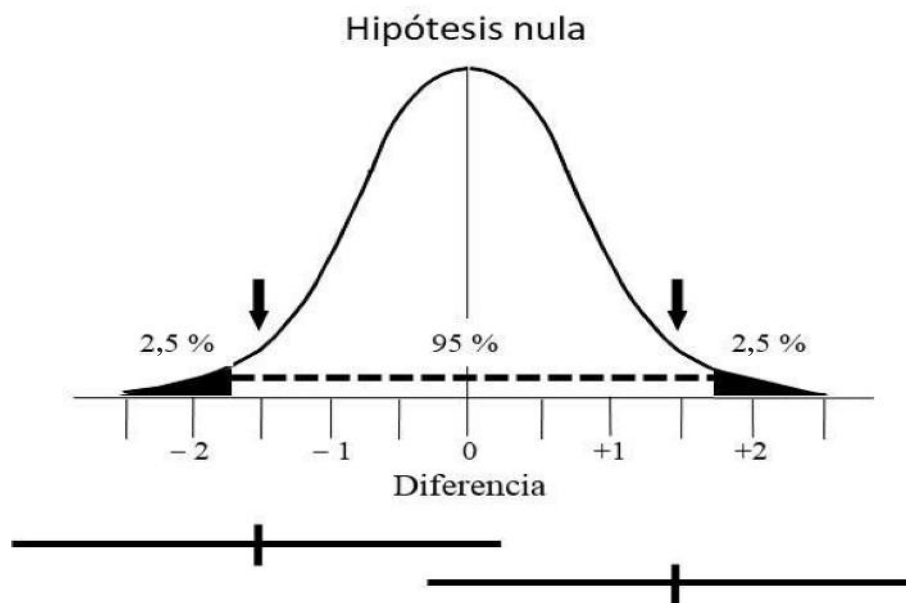


Figura 2. Curva de distribución de frecuencia (curva de Gauss) de la hipótesis nula. Ambas flechas indican el resultado de un ensayo en el que la diferencia hallada entre las muestras (positiva o negativa) no alcanzó una diferencia significativa, ya que se ubica en el sector de la curva correspondiente al 5% para ambas colas. (La línea de puntos delimita el 95% del área de dicha curva). Otra forma de expresar el resultado se representa en la parte inferior donde ambas líneas continuas grafican los intervalos de confianza construidos a partir de las diferencias halladas, indicando en este caso, por no alcanzar el cero, que el estudio fue negativo. Fuente: elaboración propia.

Si el resultado del ensayo demuestra que la diferencia hallada está entre $-1,5$ y $+1,5$ (ambas colas), como indican las flechas de la figura 2, el estudio fue negativo, ya que no se “alcanzó” la significación al “ubicarse” esta en el área de $p \geq 95\%$, (no “llega” a $< 2,5\%$).

En investigación clínica habitualmente se construye un intervalo de confianza (IC) del 95% con la diferencia hallada, $-/+ 1,5$, el cual –al superar el cero de la curva de hipótesis nula– nuevamente indica que la diferencia no es significativa (Fig. 2).

En el ejemplo anterior se empleó la diferencia de porcentajes (DR) entre, por ejemplo, intervención y control. Esto podría expresarse también como un cociente entre ambos, es decir, como *riesgo relativo* (RR). Para el caso, la media de la curva será 1 y la hipótesis nula se rechaza cuando el IC alcanza o sobrepasa la unidad.

De este modo, se *acepta la hipótesis nula* y se concluye que el punto final (criterio diagnóstico/ pronóstico o intervención) no difiere entre las muestras en comparación.

Ahora bien, si se repitiera el estudio 100 veces, en ≥ 95 la diferencia hallada se ubicaría dentro del IC –es decir no significativo– pero en < 5 por fuera de esos límites, por lo que el resultado podría considerarse falsamente como positivo.

Comparaciones múltiples: fundamentos estadísticos

Resulta claro –según lo anteriormente expuesto– que la posibilidad de hallar una diferencia significativa aumenta si se repiten las comparaciones. El valor de p hallado luego de múltiples comparaciones suele denominarse p

nominal, significando una determinación numérica no ubicada en el contexto de la probabilidad real. Existen diferentes procedimientos estadísticos para calcular la p *corregida* al número de comparaciones efectuadas.

Uno de ellos es corregir el valor de p en función del número de comparaciones efectuadas, utilizando procedimientos de ajuste por comparaciones múltiples.

En primer lugar se aplicó la corrección de Bonferroni, en el cual el nivel de significación corregido (α_c) se obtiene dividiendo el valor nominal de α por el número de comparaciones (κ), según la expresión: $\alpha_c = \alpha / \kappa$.

Asimismo, se consideró el cálculo del nivel de significación requerido en función del número de comparaciones, estimado mediante la expresión: $\alpha_c = 1 - (1 - \alpha)^\kappa$, donde α representa el valor nominal de p y κ el número de comparaciones realizadas.

Los decimales llevan coma, no punto.

La tabla 3 muestra, para una p nominal de 0,05, los valores reales de p y los corregidos resultantes al aplicar la fórmula detallada que deberían emplearse en una investigación si se efectuara más de una comparación.

Que de cada 100 comparaciones 5 puedan ser falsamente positivas parece ser una situación poco probable en investigación clínica. Pero si esa proporción se define como 20 a 1, resulta más aceptable. Ahora bien, regresando a la tabla 1, obsérvese que –con solo dos o tres comparaciones– se modifica notablemente la p requerida para alcanzar una diferencia significativa.

Tabla 1. Valores de p nominales y valores ajustados según el número de comparaciones, calculados mediante la corrección de Bonferroni y el método de ajuste del nivel de significación.

α	κ	α_c	α / κ
0,05	1	0,05	0,05
0,05	2	0,10	0,025
0,05	3	0,14	0,017
0,05	4	0,19	0,013
0,05	5	0,22	0,010

Nota: α : valor nominal de p ; α_c : p no corregida; κ : número de comparaciones, α/κ valor de p corregido.

Fuente: elaboración propia.

Comparaciones múltiples en investigación clínica

En investigación clínica, las comparaciones múltiples pueden surgir en diferentes escenarios. En los ensayos clínicos, el investigador suele definir una hipótesis general y un objetivo primario, ya sea para evaluar el efecto de una intervención o el valor pronóstico de un determinado factor. Sin embargo, cuando la comparación inicial no muestra una asociación estadísticamente significativa, puede surgir la tentación de explorar objetivos alternativos o modificar el punto final, repitiendo los análisis hasta obtener un resultado favorable^{1,5}.

Aun sin modificar el punto final primario, las comparaciones múltiples pueden aparecer cuando dicho punto se evalúa en distintos momentos del seguimiento, por ejemplo a los 3, 6 o 12 meses, seleccionando posteriormente aquel que muestra un resultado estadísticamente significativo⁸.

Otra fuente frecuente de comparaciones múltiples es el análisis de subgrupos poblacionales. Ante un resultado global negativo, la muestra original puede fraccionarse para explorar asociaciones en distintos subgrupos, conservando únicamente aquellos con resultados favorables⁷.

Asimismo, es posible aplicar diferentes pruebas estadísticas sobre los mismos datos e informar solo aquella que arroja un valor de p significativo. Estas estrategias, aisladas o combinadas, incrementan sustancialmente la probabilidad de obtener asociaciones espurias⁸.

Estrategias de corrección en investigación clínica

La principal estrategia para otorgar credibilidad a los resultados de una investigación clínica consiste en definir de manera precisa, *a priori*, la hipótesis que se va a evaluar, el objetivo del estudio y el punto final primario, así como la prueba estadística por emplear y el tamaño muestral requerido¹. Los hallazgos adicionales deben presentarse como puntos finales secundarios,

aclarando explícitamente que responden a un análisis *post hoc* y que el valor de p no ha sido corregido por comparaciones múltiples^{5,6}.

En el caso de los análisis secundarios, es posible ajustar el valor de p en función del número de comparaciones realizadas mediante pruebas estadísticas apropiadas, como la corrección de Bonferroni u otros métodos de ajuste^{4,6}.

Otra alternativa consiste en establecer un análisis jerárquico de los puntos finales secundarios. En este enfoque, una vez alcanzada la significación estadística para el punto final primario, se define un orden jerárquico preespecificado para los puntos secundarios. Cuando se obtiene un valor de p no significativo, los análisis subsiguientes deben interpretarse como valores nominales, aun cuando sean inferiores a 0,05¹.

Finalmente, en estudios observacionales, un hallazgo obtenido luego de múltiples comparaciones puede considerarse exploratorio y ser presentado como resultado de una cohorte de derivación, requiriendo su validación posterior en una cohorte independiente^{2,4}.

Ampliando el escenario

La problemática de las comparaciones múltiples es abordada con diferentes criterios según los diferentes escenarios de investigación.

En los *ensayos de intervención, controlados y aleatorizados*, con un punto final primario predefinido, lo habitual para los puntos finales secundarios es aplicar diferentes criterios de ajuste o simplemente indicar que el resultado no está corregido por múltiples comparaciones.

Los *estudios observacionales derivados de grandes bases de datos* que tienen finalidad pronóstica o de intervención (Fase IV) suelen ser más o menos rigurosos, ya que en ocasiones carecen de las correcciones necesarias que ajusten los múltiples análisis, lo que es habitual en este tipo de investigaciones.

Finalmente en *ensayos unicéntricos* son habituales los análisis *post hoc* no explicitados, que exploran –dentro de la formulación general de una hipótesis– múltiples objetivos o puntos finales, sin el correspondiente ajuste por las correcciones realizadas.

De la inteligencia artificial a las comparaciones múltiples

¿Corrige la inteligencia artificial (IA) la problemática de las asociaciones probablemente ficticias consecuencia de comparaciones múltiples? La respuesta al interrogante planteado es no.

Al contrario, la capacidad de la inteligencia artificial de explorar miles de asociaciones a partir de millones de datos puede amplificar el riesgo de identificar relaciones espurias si no se aplican controles estadísticos adecuados⁹.

CONCLUSIÓN

Regresando a la anécdota con que se inició esta comunicación,

“...ocurre que había estado apostando toda la noche, fallando repetidamente en cada una de ellas, hasta que gané con el 18” (no sabemos si nuestro personaje recuperó todo su dinero, pero –en todo caso– esto no cambia la historia).

Ahí precisamente reside el problema: no contar toda la historia, y presentar el resultado como el objetivo original y no como el alcanzado luego de varios intentos que alteran indefectiblemente el valor de p .

Una conclusión sencilla es afirmar que no es igual decir –encontré lo que buscaba– que, –miren qué encontré–. Una sentencia tan simple como esta define el objetivo central de esta presentación.

En este caso, nada mejor que recordar: –si se torturan suficientemente los datos, usted obtendrá lo que quiera oír y, bajo estas circunstancias, el perpetrador puede examinar los datos hasta encontrar una asociación «significativa» entre variables, para luego elaborar una hipótesis que biológicamente se ajuste a esa asociación–). (*“If you torture your data long enough they will tell you whatever you want to hear”; ...the perpetrator simply pores over the data until a “significant” association is found between variables and then devises a biologically plausible hypothesis to fit the association*^{8,10}).

O tal vez, que “una p no sustituye al cerebro” (*A p value is no substitute for a brain*. Anonymous), como ha sido señalado en discusiones contemporáneas sobre el uso e interpretación del valor de p ¹¹. Aunque una mejor consideración al respecto es la del físico Richard Feynman cuando concluye que: “es mucho más interesante vivir con cierta incertidumbre que vivir con respuestas que pueden estar equivocadas” (*“It is much more interesting to live with uncertainty than to live with answers that might be wrong”*)¹².

En medicina esta última aseveración requiere una pequeña corrección: aun con incertidumbre, siempre debemos tomar decisiones ya sea por acción u omisión. El problema es ser conscientes del escenario estrecho al que nos impone aquella condición.

Contribuciones de los autores: Conceptualización, Supervisión, Validación (AMC). Curación de datos, Análisis formal, Captación de fondos, Investigación, Metodología, Administración del proyecto, Recursos, *Software* (MEK, MB, AMC).

Conflictos de intereses: los autores declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: los autores declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Pocock SJ, Stone GW. The primary outcome fails - what next? *N Engl J Med*. 2016;375(9):861-870. <https://doi.org/10.1056/NEJMr1510064>.
2. Sugitani T, Bretz F, Maurer W. A simple and flexible graphical approach for adaptive group-sequential clinical trials. *J Biopharm Stat*. 2016;26(2):202-216. <https://doi.org/10.1080/10543406.2014.972509>.
3. Freemantle N, Calvert MJ. Interpreting composite outcomes in trials. *BMJ*. 2010;341:c3529. <https://doi.org/10.1136/bmj.c3529>.
4. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Stat Methodol*. 1995;57(1):289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
5. Bauer P. Multiple testing in clinical trials. *Stat Med*. 1991;10(6):871-889; discussion 889-890. <https://doi.org/10.1002/sim.4780100609>.
6. Perneger TV. What's wrong with Bonferroni adjustments. *BMJ*. 1998;316(7139):1236-1238. <https://doi.org/10.1136/bmj.316.7139.1236>.
7. Pocock SJ, Assmann SE, Enos LE, et al. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Statist. Med*. 2002;21:2917-2930. <https://doi.org/10.1002/sim.12968>.
8. Bender R, Lange S. Adjusting for multiple testing--when and how? *J Clin Epidemiol*. 2001;54(4):343-349. [https://doi.org/10.1016/s0895-4356\(00\)00314-0](https://doi.org/10.1016/s0895-4356(00)00314-0).
9. Westphal M, Zapf A, Brannath W. A multiple testing framework for diagnostic accuracy studies with co-primary endpoints. *Stat Med*. 2022;41(10):1848-1866. doi:10.1002/sim.9308.
10. Mills JL. Data torturing. *N Engl J Med*. 1993;329(16):1196-1199. <https://doi.org/10.1056/NEJM199310143291613>.
11. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond “ $p < 0.05$ ”. *Am Stat*. 2019;73(Suppl 1):1-19. <https://doi.org/10.1080/00031305.2019.1583913>.
12. Feynman RP. The pleasure of finding things out. Cambridge (MA): Perseus Books; 1999. 270 p.

Desenlaces subrogados en investigación clínica

Mariana A. Burgos¹, Silvana B. Soli², Habik A. Páez Álvarez² y María F. Grande Ratti¹

¹Área de Investigación en Medicina Interna, Servicio de Clínica Médica, Hospital Italiano. Buenos Aires, Argentina

²Secretaría de Investigación, Universidad Hospital Italiano. Buenos Aires, Argentina

RESUMEN

Un desenlace subrogado es un biomarcador o desenlace intermedio usado como sustituto de un desenlace clínico relevante, como la supervivencia o la calidad de vida, propuesto para predecir los beneficios o daños de una intervención. Su utilización en investigación clínica puede reducir la duración y el tamaño de los ensayos clínicos, acelerar la evaluación regulatoria y las decisiones de cobertura de nuevos tratamientos. Sin embargo, no todo biomarcador ni todo desenlace intermedio constituye un subrogado validado. Su uso es objeto de debate, ya que en muchos casos, no han sido validados adecuadamente como predictores fiables de beneficios clínicos significativos para los pacientes. Esta nota estadística ofrece una revisión crítica del uso de desenlaces subrogados, analiza sus principales limitaciones conceptuales y metodológicas, y propone un enfoque que priorice desenlaces clínicamente relevantes y centrados en el paciente.

Palabras clave: desenlaces subrogados, investigación clínica, biomarcadores, desenlaces clínicos, toma de decisiones en salud.

Surrogate Outcomes in Clinical Research

ABSTRACT

A surrogate endpoint is a biomarker or intermediate endpoint used as a substitute for a clinically relevant endpoint, such as survival or quality of life, and proposed to predict the benefits or harms of an intervention. Its use in clinical research may reduce the duration and sample size of clinical trials and expedite regulatory assessment and coverage decisions for new treatments. However, not every biomarker or intermediate endpoint constitutes a validated surrogate endpoint. Their use remains controversial because, in many cases, they have not been adequately validated as reliable predictors of clinically meaningful benefits for patients. This statistical note provides a critical review of the use of surrogate endpoints, examines their main conceptual and methodological limitations, and proposes an approach that prioritizes clinically relevant, patient-centered endpoints.

Keywords: surrogate outcomes, clinical research, biomarkers, clinical endpoints, health decision-making.

Autora para correspondencia: mariana.burgos@hospitalitaliano.org.ar, Burgos MA.

Recibido: 30/12/2025 | Aceptado: 14/08/2026 | Publicado: 25/09/2026

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1312>

Cómo citar: Burgos MA, Soli SB, Páez Álvarez HA, Grande Ratti MF. Desenlaces subrogados en investigación clínica. Rev HospItal B. Aires. 2026;46(3):e0001312

INTRODUCCIÓN

A fines de la década de 2000, el bevacizumab (un anticuerpo monoclonal) se introdujo para el tratamiento del cáncer de mama metastásico, debido a que algunos estudios mostraban una mejora significativa en la supervivencia libre de progresión, utilizada como desenlace subrogado, lo que motivó la aprobación acelerada por la Administración de Alimentos y Medicamentos de los Estados Unidos (Food and Drug Administration [FDA] en inglés) en 2008. Sin embargo, varios estudios posteriores confirmaron que este beneficio no se acompañaba de una mejora en la supervivencia global y se asociaba a toxicidad relevante, lo que llevó a retirar la indicación en 2011¹.

En el año 2006, el torcetrapib (un fármaco experimental desarrollado para el tratamiento de la enfermedad cardiovascular) surgió como herramienta para modificar el perfil lipídico, con aumentos marcados del HDL y reducciones significativas del LDL. Sin embargo, el riesgo de eventos cardiovasculares y mortalidad aumentó de forma significativa. El estudio fue suspendido de manera prematura y torcetrapib nunca fue aprobado, convirtiéndose en un ejemplo claro de cómo un cambio favorable en un desenlace subrogado no garantiza beneficio clínico real².

Estos casos evidencian que los desenlaces subrogados no siempre se corresponden con un beneficio clínico para el paciente. La literatura ofrece múltiples ejemplos que ilustran tal discrepancia (Tabla 1). En este contexto, esta nota estadística tiene como objetivo analizar el papel de dichos desenlaces, revisando su definición, usos y limitaciones en la práctica cotidiana.

CONCEPTO Y DEFINICIÓN

La FDA y el Instituto Nacional de Salud de los Estados Unidos (National Institutes of Health en inglés) proponen las siguientes definiciones conceptuales^{3,4}:

- **Biomarcador:** característica que se mide objetivamente y se evalúa como indicador de procesos biológicos, procesos patológicos o respuestas farmacológicas a una intervención terapéutica.
- **Desenlace intermedio:** variable que puede utilizarse como desenlace en un estudio, pero que no necesariamente ha sido validada para sustituir un desenlace clínicamente relevante.
- **Desenlace subrogado:** biomarcador o desenlace intermedio cuya capacidad para sustituir un desenlace clínicamente relevante y predecir el beneficio o daño de una intervención ha sido demostrada mediante un proceso formal de validación⁵.
- **Desenlace clínico:** característica o variable que refleja cómo se siente, funciona o sobrevive un paciente.

Según Lassere, los biomarcadores y los desenlaces clínicos son variables conceptualmente distintas que pueden ubicarse a lo largo de un continuo que va desde aquellas centradas en la enfermedad hasta aquellas centradas en el paciente⁶.

Un biomarcador es una variable que no tiene significado clínico inmediato y cuyo valor interpretativo depende de la comprensión de los mecanismos biológicos

subyacentes y de la evidencia clínica y epidemiológica acumulada. En consecuencia, carece de validez intrínseca y requiere un proceso formal de validación para ser utilizado como sustituto de un desenlace clínico relevante.

Por el contrario, las variables centradas en el paciente tienen relevancia clínica directa, constituyen resultados significativos por sí mismas, no requieren validación adicional y representan el estándar para la toma de decisiones clínicas y para su utilización como desenlace principal en ensayos terapéuticos. Un desenlace intermedio es una variable situada en la cadena causal de la enfermedad, pero que no ha sido validada formalmente como sustituto clínico.

Un desenlace subrogado se define como un biomarcador o desenlace intermedio que ha sido sometido a un proceso de validación rigurosa que respalda su uso como sustituto de un desenlace clínico relevante⁶ (Figura 1).

Sin un proceso de validación robusta, los desenlaces subrogados pueden derivar en lo que Montori y Yudkin denominaron “idolatría del subrogado”: indicadores que adquieren entidad propia, orientan recomendaciones en guías de práctica clínica y decisiones regulatorias, pero que no necesariamente se traducen en beneficios clínicamente significativos para los pacientes⁷.

USOS Y APLICACIONES EN LA PRÁCTICA

La medición de desenlaces clínicos relevantes para los pacientes, como la calidad de vida o mortalidad, requiere seguimientos prolongados y tamaños muestrales grandes, y por consiguiente una inversión considerable de tiempo y recursos. Los desenlaces subrogados permiten evaluar intervenciones de manera más eficiente y son especialmente útiles en enfermedades poco frecuentes, pediátricas o durante el embarazo, donde la baja incidencia dificulta reunir suficientes casos; en el desarrollo de terapias innovadoras, que requieren evidencia temprana de eficacia biológica; o en fases exploratorias de investigación^{3,8}. Entre los ejemplos mejor establecidos, la reducción de la presión arterial ha demostrado predecir una disminución de eventos cardiovasculares, particularmente del accidente cerebrovascular, en múltiples ensayos y con diferentes clases de antihipertensivos. La reducción del colesterol LDL se asoció con una disminución de eventos cardiovasculares mayores con estatinas. Estos casos muestran que los desenlaces subrogados pueden ofrecer estimaciones fiables del beneficio clínico cuando cuentan con plausibilidad causal y evidencia consistente procedente de múltiples ensayos.

El problema surge cuando su uso deja de ser la excepción y se convierte en la práctica dominante.

LIMITACIONES Y CONTROVERSIAS

El uso indiscriminado de desenlaces subrogados plantea implicancias metodológicas y éticas. La evidencia que los respalda puede ser indirecta: una asociación fisiopatológica entre un biomarcador o desenlace intermedio y el desenlace final puede ser plausible, pero no necesariamente predictiva⁶. Esto reduce la calidad global de la evidencia disponible según el sistema GRADE (del

Tabla 1. Ejemplos documentados en la historia de la investigación clínica sobre el uso de desenlaces subrogados

Condición clínica	Desenlace subrogado	Desenlace clínico	Racional teórico	Estudios de referencia	Evidencia observada
Diabetes mellitus	HbA1c	Reducción de complicaciones microvasculares/macrovasculares	Mayor disminución de la HbA1c se traduciría en disminución de las complicaciones	Estudio ACCORD15:10 251 personas DM2 de alto RCV. Intervención: Terapia intensiva (HbA1c <6%) vs. estándar (HbA1c 7,0-7,9%). Resultados : -HbA1c: 6,4% (Rama intensiva) vs. 7,5% (Rama estándar)-↓ Eventos CV (HR 0,90; IC 95%: 0,78-1,04;p=0,16)-↑ Mortalidad total en grupo intensivo.(HR 1,22; IC,95%: 1,011,46; p=0,04) ¹⁵	El uso de terapia intensiva para alcanzar niveles normales de HbA1c durante 3,5 años aumentó la mortalidad y no redujo significativamente los eventos cardiovasculares mayores. Estos hallazgos identifican un daño previamente no reconocido de la reducción intensiva de la glucosa en pacientes de alto riesgo con DM2
Dislipidemia	HDL LDL	Reducción de mortalidad Reducción de eventos cardiovasculares Reducción de mortalidad	Los niveles de colesterol de lipoproteínas de alta densidad (HDL) son un predictor inverso fuerte e independiente de enfermedad cardiovascular	Estudio Illuminate2: 15 067 pacientes con alto riesgo cardiovascular. Intervención: atorvastatina más torcetrapib vs. atorvastatina sola. Resultados: aumento del 72,1% en el HDL y una disminución del 24,9% en el LDL a los 12 meses (P<0,001). La rama con torcetrapib evidenció un riesgo aumentado de eventos cardiovasculares (HR 1,25; IC 95%: 1,09-1,44; P=0,001) y un riesgo aumentado de muerte por cualquier causa (HR 1,58; IC 95%: 1,14-2,19; P=0,006). El estudio se interrumpió prematuramente debido a los resultados	El uso agregado de torcetrapib generó un aumento importante del colesterol HDL y descenso marcado de los valores de LDL, generando, a su vez, un mayor riesgo de mortalidad y morbilidad por un mecanismo desconocido
Cáncer de mama metastásico HER2 negativo	Supervivencia libre de progresión (SLP)	Sobrevivencia global (SG)	Bevacizumab (Estudio E2100, AVADO, RIBBON-1), demostró mejoría de SLP (subrogado) pero sin beneficio en SG (desenlace primario)	Estudio E21001: 722 pacientes con cáncer de mama metastásico. Terapia combinada (paclitaxel + bevacizumab) vs. monoterapia (paclitaxel). La SLP mediana en el grupo de terapia combinada 11,8 meses vs. paclitaxel solo 5,9 meses (HR de 0,60, IC 95%: 0,47-0,77; P<0,001). La SG media fue similar en el grupo de terapia combinada y en el grupo de paclitaxel solo 26,7 meses y 25,2 meses (HR 0,88; P=0,16)	El uso de terapia combinada de paclitaxel y bevacizumab aumentó la supervivencia libre de progresión y evidenció un aumento de la tasa de supervivencia a 1 año, pero no demostró superioridad en la supervivencia global. Y evidenció mayores efectos tóxicos graves (Grados 3 y 4) en el grupo de terapia combinada

HbA1c: hemoglobina glucosilada; DM2: diabetes mellitus tipo 2; RCV: riesgo cardiovascular. Fuente: elaboración propia



Figura 1. Biomarcadores, desenlaces intermedios y desenlaces subrogados validados⁶.

inglés *Grading of Recommendations, Assessment, Development and Evaluation*), disminuyendo la certeza de la evidencia por indirectitud. Además, una mejoría estadísticamente significativa del subrogado puede no traducirse en beneficio clínico, generar falsos positivos o incluso acompañarse de mayor mortalidad o eventos adversos⁹.

La validez de un desenlace subrogado no es una propiedad universal del marcador, sino que debe establecerse para una enfermedad, un tipo de intervención y un desenlace clínico específico. Por lo tanto, la evidencia obtenida en un contexto determinado no puede extrapolarse automáticamente a otros escenarios clínicos. La FDA señala que, por ejemplo en oncología, la supervivencia libre de progresión va a depender del tipo y estadio de la enfermedad, el contexto de uso, la magnitud del efecto, los tratamientos disponibles y la relación entre beneficios y riesgos. Su capacidad para sustituir la supervivencia global debe evaluarse en cada escenario clínico específico¹⁰.

CONSIDERACIONES GENERALES SOBRE LA VALIDACIÓN

Se han propuesto diversos métodos para determinar si un desenlace subrogado permite predecir un desenlace centrado en el paciente. Su validación requiere integrar plausibilidad biológica, evidencia epidemiológica, validez y fiabilidad de la medición, y evidencia estadística sobre su capacidad para predecir el efecto clínico de una intervención. Inicialmente, los criterios de Prentice establecen que un subrogado válido debe asociarse con el desenlace clínico y capturar completamente el efecto del tratamiento sobre este. Debido a la dificultad para demostrar esta condición, los modelos metanalíticos desarrollados por Buyse y cols. evalúan la asociación entre ambos desenlaces en el plano individual y entre sus respectivos efectos terapéuticos a nivel de ensayo.

Desde una perspectiva causal, los métodos estadísticos se agrupan en enfoques de efectos causales, que analizan la relación entre la intervención, el subrogado y el desenlace clínico, y de asociación causal, que evalúan si el efecto terapéutico sobre el subrogado predice el efecto clínico. Entre estos últimos, los modelos metanalíticos examinan las asociaciones en el ámbito individual y de ensayo. El metanálisis con datos individuales es el enfoque más completo, aunque su elevada demanda de datos y recursos lleva a utilizar también datos agregados. Otra medida es el efecto umbral del subrogado, que estima el cambio mínimo necesario para predecir un efecto clínico, pero depende de la cantidad, calidad y heterogeneidad de los ensayos disponibles¹¹.

En la práctica, Ciani y cols. proponen organizar esta evaluación en tres etapas “jerárquicas”: establecer el nivel de evidencia disponible, valorar la fuerza y consistencia de dichas asociaciones y estimar el efecto clínico esperado a partir del efecto observado sobre el subrogado.

Dado que no existen métodos ni umbrales universalmente aceptados, la validación debe basarse en la evaluación progresiva del conjunto de la evidencia y limitarse al contexto clínico estudiado⁹.

BUENAS PRÁCTICAS EN EL USO DE SUBROGADOS

Las buenas prácticas en el uso de desenlaces subrogados se apoyan en varios elementos interrelacionados que conforman un marco integrado de calidad metodológica:

- Una validación rigurosa asegura que el marcador intermedio tenga una relación sólida con el desenlace clínico¹¹.
- La transparencia en el informe, incluyendo la adherencia a CONSORT-Surrogate, permite reconocer explícitamente su naturaleza, comunicar su grado de validación y discutir sus limitaciones¹².



Figura 2. Lista de verificación para interpretar estudios que utilizan desenlaces subrogados.¹¹

- Los resultados referidos por pacientes (PROMs), capturan dimensiones centrales de calidad de vida y experiencia real de enfermedad¹³.

- Las iniciativas de *Core Outcome Sets* (COS) contribuyen a definir desenlaces mínimos esenciales por condición, asegurando comparabilidad y relevancia clínica¹⁴.

Finalmente, un enfoque centrado en el paciente orienta la selección y priorización de los desenlaces hacia lo que realmente importa para la toma de decisiones en salud. En conjunto, estos elementos operan de manera

sinérgica para promover un uso más reflexivo, transparente y clínicamente significativo de los subrogados.

CONCLUSIÓN

Los desenlaces subrogados representan una herramienta útil en investigación clínica y pueden acelerar el desarrollo de nuevo conocimiento. Sin embargo, su aplicación sin una validación rigurosa implica el riesgo de conducir a la aprobación de intervenciones poco eficaces, costosas o con mayor incidencia de eventos adversos, sin

traducirse en mejoras reales en la supervivencia ni en la calidad de vida de los pacientes. Un enfoque más crítico y transparente en el uso de desenlaces subrogados contribuirá a mejorar la calidad de la evidencia y la toma de decisiones en salud.

Contribuciones de los autores: Conceptualización, Administración del proyecto, Supervisión, Redacción – revisión y edición: MAB, MFGR. Captación de fondos: MFGR. Investigación, Visualización, Redacción – borrador original: MAB, SBS, HAPA. Metodología, Validación: MAB, SBS, HAPA, MFGR.

Conflictos de intereses: las autoras declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: las autoras declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Miller K, Wang M, Gralow J, et al. Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer. *N Engl J Med*. 2007;357(26):2666-2676. DOI: <https://doi.org/10.1056/NEJMoa072113>
2. Barter PJ, Caulfield M, Eriksson M, et al. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med*. 2007;357(21):2109-2122. DOI: <https://doi.org/10.1056/NEJMoa0706628>
3. U.S. Food and Drug Administration. FDA Facts: Biomarkers and Surrogate endpoints [Internet]. Silver Spring (MD): Food and Drug Administration; 2019 [citado 2025 nov 23]. Disponible en: <https://www.fda.gov/about-fda/innovation-fda/fda-facts-biomarkers-and-surrogate-endpoints>
4. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource [Internet]. Silver Spring (MD): Food and Drug Administration; Bethesda (MD): National Institutes of Health; 2016- [citado 2025 nov 23]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK326791/>
5. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69(3):89-95. DOI: <https://doi.org/10.1067/mcp.2001.113989>
6. Lassere MN, Johnson KR, Boers M, et al. Definitions and validation criteria for biomarkers and surrogate endpoints: development and testing of a quantitative hierarchical levels of evidence schema. *J Rheumatol*. 2007;34(3):607-615.
7. Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. *BMJ*. 2011;343:d7995. DOI: <https://doi.org/10.1136/bmj.d7995>
8. Schuster DP. The opportunities and challenges of developing imaging biomarkers to study lung function and disease. *Am J Respir Crit Care Med*. 2007;176(3):224-230. DOI: <https://doi.org/10.1164/rccm.200703-462PP>
9. Ciani O, Buyse M, Drummond M, et al. Time to Review the Role of Surrogate End Points in Health Policy: State of the Art and the Way Forward. *ValueHealth*. 2017;20(3):487-495. DOI: <https://doi.org/10.1016/j.jval.2016.10.01110>
10. U.S. Food and Drug Administration. Guidance Documents [Internet]. Silver Spring (MD): Food and Drug Administration; [citado 2025 nov 23]. Disponible en: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents>
11. Christensen R, Ciani O, Manyara AM, et al. Surrogate endpoints: a key concept in clinical epidemiology. *J Clin Epidemiol*. 2024;167:111242. DOI: <https://doi.org/10.1016/j.jclinepi.2023.111242>
12. Manyara AM, Davies P, Stewart D, et al. Reporting of surrogate endpoints in randomised controlled trial reports (CONSORT-Surrogate): extension checklist with explanation and elaboration. *BMJ*. 2024;386:e078524. DOI: <https://doi.org/10.1136/bmj-2023-078524>
13. Churrua K, Pomare C, Ellis LA, et al. Patient-reported outcome measures (PROMs): a review of generic and condition-specific measures and a discussion of trends and issues. *Health Expect*. 2021;24(4):1015-1024. DOI: <https://doi.org/10.1111/hex.13254>
14. Williamson PR, Altman DG, Bagley H, et al. The COMET Handbook: version 1.0. *Trials*. 2017;18(Suppl 3):280. DOI: <https://doi.org/10.1186/s13063-017-1978-4>
15. Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med*. 2008;358(24):2545-2559. DOI: <https://doi.org/10.1056/NEJMoa0802743>

Decisiones médicas compartidas: una perspectiva desde la Ciencia de la Complejidad

Alicia Rodríguez López[■]

Profesional independiente. España

RESUMEN

La toma de decisiones médicas compartidas (DMC) representa una evolución del modelo paternalista tradicional hacia un enfoque más participativo, donde el paciente deja de ser receptor pasivo y se convierte en agente activo de su salud. Esta transformación integra no solo la evidencia científica, sino también la individualidad de quien consulta: valores, contexto emocional, social, económico y cultural.

La Ciencia de la Complejidad, por su parte, estudia cómo surgen patrones y comportamientos impredecibles a partir de interacciones dinámicas y no lineales entre varios componentes de un sistema. Se propone que las DMC son intervenciones complejas en el contexto de problemas clínicos. Decidir desde esta perspectiva reconoce que las condiciones de salud no son lineales ni simples: involucran múltiples variables, relaciones dinámicas y resultados inciertos. Por lo tanto, la decisión debe surgir de un proceso integrador en el que interactúan conocimientos técnicos, experiencia clínica, juicio experto, intuición y valores personales del paciente. Las DMC no son reducibles a la suma de conocimientos aislados, sino que emergen de la interacción sinérgica entre los participantes.

El enfoque basado en la Complejidad fortalece el proceso de DMC mediante herramientas como modelos mentales compartidos, simulaciones, retroalimentación y análisis multidisciplinario. Estas estrategias permiten visualizar escenarios, anticipar consecuencias y adaptar conductas según la evolución del paciente.

En última instancia, considerar las DMC como intervenciones complejas implica tomar un curso de acción de manera conjunta, a través de un proceso dinámico, humano y profundamente contextual.

Palabras clave: decisiones médicas compartidas, complejidad, cuidados centrados en los pacientes, sistemas complejos.

Shared Medical Decision-Making: A Complexity Science Perspective

ABSTRACT

Shared medical decision-making (SMD) represents an evolution from the traditional paternalistic model toward a more participatory approach, where the patient becomes an active agent in managing his/her own health, rather than a passive recipient. This transformation integrates not only scientific evidence but also the individuality of the patient: values, emotional, social, economic, and cultural context. Complexity, in turn, studies how patterns and unpredictable behaviors emerge from dynamic, non-linear interactions among multiple components of a system.

Autora para correspondencia: ali_n_rodri@yahoo.com, Rodríguez López A.

Recibido: 4/03/2026 | Aceptado: 21/07/2026 | Publicado: 10/09/2026

DOL: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1319>

Cómo citar: Rodríguez López A. Decisiones médicas compartidas: una perspectiva desde la ciencia de la complejidad. *Rev Hosp Ital B.Aires.* 2026;46(3):e0001319

SDM is proposed as a complex intervention in the context of medical conditions. Deciding from this perspective recognizes that health problems are neither linear nor simple; they involve multiple variables, dynamic relationships, and uncertain outcomes. The choice, therefore, must emerge from an integrative process in which technical knowledge, clinical experience, expert judgment, intuition, and the patient's personal values interact. SMD is not reducible to the mere sum of isolated knowledge, but to the synergistic interaction between participants.

The complexity-based approach strengthens the SMD process by incorporating tools such as shared mental models, simulations, feedback, and multidisciplinary analysis. These tools enable the visualization of scenarios, anticipation of consequences, and adaptation of strategies according to the patient's evolving condition.

Ultimately, considering SMD as complex interventions means deciding a course of action jointly, through a dynamic, human, and deeply contextual process.

Keywords: shared medical decisions, complexity, patient-centered care, complex systems.

INTRODUCCIÓN

El presente trabajo tiene como objetivo integrar el marco contemporáneo de la toma de Decisiones Médicas Compartidas (DMC) con los aportes de la Ciencia de la Complejidad. Se propone que las DMC no solo constituyen una técnica comunicacional o deliberativa, sino también una intervención clínica compleja, particularmente relevante en contextos de incertidumbre, pluralidad de valores y distintos cursos de acción razonables.

Decisiones Médicas Compartidas: intervenciones complejas

Las DMC se definen como un proceso estructurado de deliberación conjunta en el cual profesionales de la salud y pacientes colaboran activamente para tomar decisiones clínicas aplicables y prácticas, integrando la mejor evidencia científica disponible sobre las opciones existentes, sus potenciales beneficios y perjuicios, la experiencia clínica del profesional y los valores, preferencias, circunstancias y objetivos vitales de quienes consultan¹.

Por su parte, la Ciencia de la Complejidad estudia cómo surgen patrones y comportamientos impredecibles a partir de interacciones dinámicas y no lineales entre varios componentes de un sistema². Desde este punto de vista, las DMC pueden entenderse como procesos adaptativos que no se limitan a la aplicación de evidencia, de los cuales la decisión clínica emerge de la interacción de profesionales, pacientes y contexto, bajo condiciones de incertidumbre, diferentes valores y resultados no completamente predecibles.

Las DMC se desarrollan dentro de un marco complejo del que forman parte diferentes protagonistas. No se trata únicamente del paciente y el profesional de salud, sino que también intervienen familiares, cuidadores, otros miembros del equipo sanitario, así como las normas culturales y sociales que rodean a la persona³. Esta multiplicidad de participantes se acompaña de una diversidad de valores y objetivos, que muchas veces pueden ser divergentes, incluso contradictorios, cambiantes en

el tiempo y difíciles de jerarquizar, características que definen, precisamente, la Complejidad⁴.

De por sí, las consultas médicas son ya sistemas complejos; y, si se comprenden como tal, la integración de las DMC introduce nuevas dimensiones clínicas, relacionales y organizacionales (Tabla 1). La toma de decisiones en este contexto se ve condicionada por la incertidumbre estructural y la evidencia incompleta, derivadas de datos que pueden ser ambiguos, inconcluyentes o insuficientes, lo que limita la capacidad de predecir resultados con certeza y obliga a los actores a gestionar probabilidades y escenarios inciertos⁵.

Existe una “no linealidad” entre la decisión y los resultados, ya que pequeñas variaciones en la forma de comunicar, en el momento de decidir o en el contexto emocional pueden generar efectos desproporcionados sobre la adherencia del paciente, su satisfacción y los resultados clínicos⁶, lo que en Complejidad se conoce como “Efecto Mariposa”⁷. Las DMC no existen previamente a la interacción médico y paciente, sino que surgen del encuentro clínico a través del diálogo y la deliberación conjunta, y no pueden entenderse como la suma de preferencias individuales⁸. Son un proceso dinámico: las decisiones se revisan, las preferencias cambian y el contexto evoluciona, por lo que deben adaptarse continuamente a las circunstancias emergentes⁸.

Finalmente, algunos factores tales como la asimetría de poder y las influencias contextuales, incluyendo desigualdades sociales, diferencias en el conocimiento y características institucionales, afectan de manera significativa el proceso decisorio, condicionando los resultados de formas que a menudo escapan al control directo de los participantes⁴.

Contexto histórico

Tradicionalmente, las consultas médicas se han desarrollado en un marco de desigualdad, con un modelo médico-paciente predominantemente paternalista, donde los pacientes han tenido un papel más bien pasivo⁹.

Tabla 1. Las consultas médicas son sistemas complejos^{2, 5, 6, 10}

Características de los sistemas complejos	Consultas médicas	Complejidad
Interconexión Los elementos del sistema están interrelacionados	Interacción de médico, paciente, familia, situación clínica, entorno social y sistema de salud	El cambio en un factor (p. ej., diagnóstico) afecta decisiones, emociones y tratamientos
Emergencia Surgen propiedades nuevas que no existían en las partes aisladas	La relación médico-paciente genera confianza, adherencia o resistencia al tratamiento	La calidad del resultado no depende solo del conocimiento médico, sino de la interacción
No linealidad Pequeñas causas pueden producir grandes efectos	Un comentario del médico puede cambiar completamente la actitud del paciente hacia su tratamiento	No existe proporcionalidad directa entre acción y resultado
Retroalimentación Cada resultado influye nuevamente en el sistema	La evolución de síntomas modifica decisiones terapéuticas; la respuesta al tratamiento ajusta el plan	El proceso es dinámico y se reajusta constantemente
Adaptabilidad Hay ajustes continuos a los cambios del entorno	El médico adapta el tratamiento según edad, cultura, recursos y respuesta clínica	No hay un único protocolo rígido aplicable a todos los casos
Autoorganización El orden surge; el sistema se autorregula	El equipo de salud coordina acciones según necesidades del paciente	La organización emerge de la interacción
Dinamismo El sistema cambia con el tiempo	El estado del paciente evoluciona; el diagnóstico puede modificarse	La consulta no es un evento estático, sino un proceso continuo

En el curso del tiempo y paulatinamente, hay una transición lenta de la pasividad al protagonismo, de modo tal que la singularidad de cada consultante es cada vez más relevante. Los modelos tales como la comunicación empática, la entrevista motivacional y la educación terapéutica han representado mejorías al modelo paternalista, aún insuficientes en términos de participación. Han conseguido mejorar la calidad relacional, el cambio conductual o la transmisión del conocimiento, pero no son procesos orientados a tomar una decisión colaborativa.

Las diferencias fundamentales entre estos modelos y las DMC en la Complejidad son evidentes en múltiples niveles. El papel del paciente, por ejemplo, es activo y cocreador de las decisiones⁸. El enfoque deja de ser lineal y unidimensional para ser sistémico y multidimensional. Hay un cambio fundamental en el modo de manejar la incertidumbre: de los protocolos rígidos hacia la flexibilidad, el ajuste dinámico según resultados y la inclusión de la dinámica familiar. La comunicación deja de ser unidireccional para ser multidireccional, incluyendo

a facilitadores de ser necesario⁹. La identificación de problemas, antes centrada en síntomas y diagnósticos, ahora incluye factores sociales, psicológicos y dinámicas interconectadas¹⁰.

Las DMC, entonces, constituyen intervenciones complejas que emergen como procesos deliberativos; toman en cuenta la situación personal de los consultantes, su estructura psíquica, emocional, cultural, social y económica. Esta comprensión, junto con principios éticos de autodeterminación, reconoce el derecho del paciente de participar activamente en sus decisiones de salud¹. Quien consulta, por su parte, debe asumir un papel activo, expresando su voluntad de participar y asegurándose de que su opinión sea valorada¹¹.

Decidir desde la Complejidad: complicado versus complejo

Los problemas de salud son intrínsecamente complejos y ocurren dentro de entornos igualmente complejos. Profesionales y pacientes no enfrentan problemas aislados,

sino sistemas dinámicos y diversos que interactúan¹². Por ejemplo, la falta de respuesta a un tratamiento puede deberse a diversos factores: adherencia, sustitución de medicación, resistencia biológica o desconfianza del paciente (entre muchos otros). Frente a esto, decidir implica la capacidad de tomar acciones efectivas sobre situaciones que requieren cambio, mejora o transformación.

Lejos de ser un juego de palabras, la diferenciación entre complicado y complejo es crucial para la comprensión de las DMC como intervenciones complejas¹³. Los problemas complicados se caracterizan por ser sistemas lineales y predecibles, con relaciones causa-efecto claras. Su resolución puede lograrse mediante pasos, fórmulas o algoritmos, existiendo una solución “correcta” que produce resultados reproducibles, completamente definidos por las condiciones iniciales, donde no hay lugar para el azar. Las interacciones de los elementos son limitadas o simples y la previsibilidad es alta: si se siguen los pasos adecuados, se obtiene el resultado esperado.

En contraste, los problemas complejos ocurren en sistemas dinámicos y no lineales, con interacciones emergentes e interdependientes¹³. No existe una única solución; se requiere colaboración, aprendizaje continuo y participación de distintos actores, considerando la sensibilidad del sistema a pequeños cambios. Los resultados son emergentes y únicos para cada contexto y momento, por lo que la previsibilidad es baja. Las estrategias de resolución deben ser integradas y continuas, considerando cursos simultáneos o secuenciales, experimentación, adaptación, iteración y deliberación conjunta².

¿Qué aporta la perspectiva de la Complejidad a las DMC y al proceso de decisión global?

Es sabido que las DMC no son aplicables a todas las situaciones clínicas. Los escenarios más favorables para introducir las DMC son aquellos donde existen varias opciones razonables (terapéuticas, diagnósticas o de manejo), sin una claramente superior¹⁴. También aquellas donde las investigaciones no muestran un resultado único o concluyente; es decir, la evidencia es insuficiente o ambigua. Otros escenarios son los casos en que las decisiones afectan significativamente aspectos físicos, emocionales o sociales (tal como ocurre en tratamientos oncológicos con marcados efectos secundarios); aquellas que implican riesgos o costos considerables; cuando los valores, creencias o contexto de quien consulta son decisivos para la elección; y en situaciones éticamente sensibles¹⁴.

En los casos en que las DMC son aplicables, la literatura nos muestra que tanto médicos como pacientes se benefician de ellas¹⁵. Quienes consultan tienen una mayor autonomía y control al participar activamente, comprenden sus opciones y pueden expresar sus valores y preferencias. La deliberación conjunta permite que el paciente entienda riesgos, beneficios y alternativas, y las decisiones son más personalizadas. Al sentirse escuchado y comprendido, el paciente se compromete más con el plan terapéutico. El diálogo y la información clara disminuyen la sensación de pérdida de control, reduciendo la ansiedad e incertidumbre¹⁶.

Por su parte, los médicos están más alineados con los valores del paciente, y lo que este considera más importante. La participación activa de quien consulta disminuye la no adherencia o rechazo a los tratamientos. La relación clínica mejora: hay más confianza, respeto mutuo y mejor comunicación. Crece la satisfacción personal del equipo de salud, y hay un desarrollo de las habilidades en comunicación, manejo de incertidumbre y pensamiento sistémico¹⁷.

La comprensión desde la Complejidad reconoce que médicos, pacientes y entorno forman un sistema dinámico que se adapta y evoluciona. Incorporar esta visión implica cambios importantes en la educación para la salud, la formación profesional y los sistemas de salud. En los profesionales sanitarios, se deberá profundizar el desarrollo de competencias complejas, comunicación avanzada, manejo de incertidumbre y ética relacional¹⁸, y el reconocimiento de que las decisiones complejas no pueden descomponerse en pasos lineales, ya que no existe información completa o suficiente para elegir la opción “óptima”. Los resultados, por lo tanto, no son predecibles o estables; los protocolos son útiles hasta cierto punto; se integran la evidencia, intuición y contexto, y se pone en valor el juicio clínico. La interacción médico-paciente se vuelve central y parte de la dinámica de las decisiones¹⁹.

El sistema de salud, por su parte, se verá obligado a diseñar procesos más flexibles y guías que permitan adaptación. Obligatoriamente se trabajará de forma coordinada con equipos interprofesionales, y se deberá incorporar el concepto de Complejidad en políticas y planificación. Las métricas no solo serán de cumplimiento, sino también de satisfacción, calidad de decisión y aprendizaje²⁰.

Herramientas para manejar la Complejidad en DMC

En la práctica clínica y para abordar la riqueza de variables de las DMC, es necesario utilizar herramientas específicas que faciliten la comunicación, el entendimiento mutuo y la anticipación de consecuencias²¹. La tabla 2 sintetiza las principales herramientas empleadas para influir eficazmente en situaciones clínicas complejas.

Estas herramientas no solo permiten un diálogo más eficaz entre el equipo médico y el paciente, sino que también promueven una comprensión profunda de las dimensiones involucradas en las DMC.

La construcción de modelos mentales compartidos facilita que ambas partes visualicen conjuntamente la situación concreta y sus posibles alternativas. Un ejemplo de esto puede verse en un paciente que toma irregularmente su medicación antihipertensiva, por no sentirse enfermo. Explorar el modelo mental del paciente involucra preguntas clave tales como: “¿Qué cree usted que es la hipertensión arterial?”, “¿Qué le preocupa del tratamiento?”, “¿Qué espera lograr?”. De acuerdo con las respuestas obtenidas, el equipo de salud conocerá cuál es la idea del paciente sobre su condición. Siguiendo el ejemplo, respuestas tales como “La tensión arterial es solo un número”, “Si me siento bien, no veo la necesidad de tomar pastillas” o “No quiero depender de

Tabla 2. Herramientas para influir sobre situaciones complejas en DMC

Herramienta	Descripción	Ejemplo de aplicación
Modelos mentales compartidos ³¹	Construcción conjunta de representaciones cognitivas del problema para mejorar el entendimiento mutuo	Médicos y pacientes discuten expectativas y objetivos
Mapas de procesos y sistemas ³²	Visualización de relaciones y dinámicas entre factores para identificar puntos de intervención	Diagramas que muestran interacción de síntomas y tratamientos
Facilitación y mediación ¹	Uso de facilitadores para guiar la comunicación y resolver conflictos en la consulta	Psicólogos que ayudan a clarificar prioridades
Simulación y escenarios ³³	Ejercicios que permiten anticipar consecuencias de diferentes decisiones	<i>Roleplay</i> de decisiones sobre tratamientos
Retroalimentación continua ²⁹	Ajustes periódicos basados en la evaluación constante de resultados y experiencias	Revisiones regulares del plan terapéutico
Comunicación abierta y empática ¹	Creación de un clima de confianza para compartir información y emociones de manera transparente	Entrevistas donde se valoran inquietudes del paciente
Capacitación y educación ³⁴	Formación para mejorar habilidades en la toma de decisiones compartidas y manejo de la complejidad	Talleres para médicos y pacientes sobre comunicación
Análisis multidisciplinario ³⁵	Integración de perspectivas diversas para abordar el problema desde diferentes disciplinas	Equipos de salud con médicos, psicólogos, trabajadores sociales

medicamentos”, permiten que el profesional de la salud integre la información médica pertinente a la condición de salud motivo de la deliberación, transmitiendo, en este caso, que la hipertensión arterial es una enfermedad silenciosa, que aumenta el riesgo de accidente cerebrovascular e infarto de miocardio, y el objetivo del tratamiento es prevenir daños futuros.

Al finalizar la construcción del mapa mental, el paciente puede llegar a la comprensión de que quiere evitar complicaciones, pero sin demasiadas pastillas; se puede elaborar conjuntamente un plan de cambio de estilo de vida y medicación con ajuste progresivo que satisfaga a los actores involucrados. Este ejercicio ha conseguido

que los participantes entiendan el problema de la misma manera, compartan objetivos, se alineen expectativas, se decida en forma conjunta y que haya compromiso con el tratamiento. La retroalimentación continua permitirá ajustar las decisiones en función de la evolución de la enfermedad y las nuevas evidencias.

Por su parte, los mapas de procesos y sistemas son más amplios: por ejemplo, una persona intentando perder peso puede visualizar los factores interrelacionados que impactan en la adherencia a la dieta, tales como el estrés laboral, las oscilaciones emocionales, la frustración, la depresión, el insomnio, el ejercicio físico. Cada punto del mapa señala una posible intervención destinada a obtener

los resultados esperados, y la modificación de uno de los factores incidirá en los demás.

Asimismo, la facilitación y mediación pueden ser decisivas para resolver conflictos de valores o preferencias y mayormente involucran a una tercera persona que ayuda a profesionales de la salud y pacientes a dilucidar aspectos importantes de dilemas diagnósticos y –principalmente– incertidumbres terapéuticas.

En conjunto, estas estrategias enriquecen la toma de decisiones, haciendo posible una práctica clínica más humana, contextualizada y eficaz²².

Es decir, para diseñar estrategias que propicien las DMC con la perspectiva de la Complejidad, tanto profesionales de la salud como pacientes deben indagar en las reglas y patrones de la interacción, disponer de teorías, principios y pautas que permitan construir un conocimiento adaptable a cada situación. Un problema externo debe observarse, reflejarse e interpretarse dentro de un marco de referencia, el cual delimitará lo que se siente, interpreta y comprende.

Diversos marcos conceptuales, cada uno con su lenguaje y formas particulares de caracterizar la situación, facilitan la interpretación del sistema desde varios ángulos (multidisciplinario, transdisciplinario), para obtener un entendimiento más profundo¹. En las DMC, el marco predominante proviene de las ciencias médicas, pero también se incluyen marcos éticos, sociales, humanísticos, educativos y comunicacionales. Cada consulta es única y puede necesitar un marco distinto (o varios combinados) y generar emergentes diferentes²³.

Aunque cada herramienta tiene su impacto causal propio, la sofisticación del sistema impide prever sus resultados con precisión. Respecto de los fenómenos emergentes, aunque no es posible identificar todas sus causas, es factible trabajar con una o varias de estas herramientas para acercarse a los resultantes deseados²⁴.

Integración de DMC y Complejidad

Si bien las DMC y la perspectiva de la Complejidad ofrecen un marco teórico valioso, no deben idealizarse como procesos lineales o perfectos. Su verdadero valor radica en su aporte práctico: guiar la interacción médico-paciente de manera más adaptativa, centrada en los valores del paciente y sensible a la incertidumbre clínica, promoviendo decisiones más contextualizadas, flexibles y sostenibles en la práctica diaria²⁵.

Un ejemplo de aplicación práctica sería el caso de una mujer de 78 años con comorbilidades crónicas, quien ha sido diagnosticada con cáncer de mama de crecimiento lento, en estadio avanzado. La paciente ha recibido ya hormonoterapia asociada a inhibidores CDK4/6, sin haber conseguido evitar la progresión.

La literatura médica indica que la quimioterapia, en estos casos, puede prolongar la sobrevida de manera modesta, aunque con efectos adversos significativos: fatiga intensa, riesgo de infecciones, posible deterioro funcional, hospitalizaciones, mielosupresión, insuficiencia cardíaca²⁶. Si la progresión es lenta y sin compromiso visceral crítico, la decisión terapéutica no es automática. Es necesario

balancear calidad de vida, riesgos y preferencias de la paciente.

La paciente puede preferir mantener calidad de vida y autonomía, y tiempo de calidad con su familia, planteando un claro dilema entre el potencial tratamiento y los valores personales. Este conflicto se intensificaría si hay presión familiar por parte de los hijos, insistentes en un tratamiento agresivo, frente a lo cual la paciente puede percibir que su autonomía se ve limitada, y sentirse abrumada por la autoridad que percibe de médicos y familiares. El equipo médico, por su parte, se enfrenta a información científica que, por una parte, refiere beneficios de la quimioterapia en pacientes de edad avanzada²⁷, y, por otra, alerta sobre el deterioro funcional y la inevitable toxicidad²⁸.

La explicación clara y comprensible de las opciones terapéuticas con sus riesgos y beneficios, el impacto posible sobre la calidad de vida y la consideración de los valores y preferencias de la paciente permitirán que el conjunto decida y se negocien las expectativas familiares al conocer cuáles efectos secundarios la paciente considera aceptables y qué elementos de autonomía son los más importantes para ella. Tal vez esto derive en el diseño de un plan que primariamente respeta sus deseos y preferencias, y los elementos terapéuticos sean más adaptables, con la posibilidad de modificar la estrategia según resultados emergentes.

CONCLUSIONES

En suma, el enfoque de DMC desde la Complejidad busca desarrollar la capacidad de comprender el comportamiento y sentido profundo de las interacciones. Mientras que los elementos simples y complicados pueden abordarse mediante procesos tradicionales de decisión, los componentes complejos requieren estrategias adaptativas, iterativas y colaborativas. Muchos procesos de decisión racional pueden fallar si no consideran la multidimensionalidad del problema⁸.

Hasta ahora, la Complejidad en medicina ha sido considerada de manera superficial⁹. El término se ha utilizado principalmente con un sentido descriptivo, más que como una herramienta conceptual capaz de orientar el diseño, la implementación y la evaluación de las intervenciones en salud. En consecuencia, muchas de estas intervenciones continúan basándose en modelos lineales que no logran captar plenamente la naturaleza intrínsecamente compleja de los problemas y situaciones que se presentan en la práctica sanitaria²⁹.

Se propone reconocer a las DMC como intervenciones complejas³⁰, a pesar de los desafíos que esto implica en implementación, estandarización, medición y reproducibilidad. Se necesitan estudios que permitan evaluar la generalización de este enfoque a distintos contextos clínicos y modelos poblacionales, fortaleciendo la aplicación práctica de la Ciencia de la Complejidad en las decisiones médicas compartidas.

Conflictos de intereses: la autora declara no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: la autora declara que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med.* 2012;27(10):1361-1367. <https://doi.org/10.1007/s11606-012-2077-6>.
2. Plsek PE, Greenhalgh T. Complexity science: the challenge of complexity in health care. *BMJ* 2001;323(7313):625-628. <https://doi.org/10.1136/bmj.323.7313.625>.
3. Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. *N Engl J Med.* 2012;366(9):780-781. <https://doi.org/10.1056/NEJMp1109283>.
4. Légaré F, Witteman HO. Shared decision making: examining key elements and barriers to adoption into routine clinical practice. *Health Aff (Millwood).* 2013;32(2):276-284. <https://doi.org/10.1377/hlthaff.2012.1078>.
5. Han PK, Klein WM, Arora NK. Varieties of uncertainty in health care: a conceptual taxonomy. *Med Decis Making.* 2011;31:6828-838. <https://doi.org/10.1177/0272989x11393976>.
6. Montori VM, Brito JP, Murad MH. The optimal practice of evidence-based medicine: incorporating patient preferences in practice guidelines. *JAMA.* 2013;310(23):2503-2504. <https://doi.org/10.1001/jama.2013.281422>.
7. Shen B-W, Pielke RA Sr, Zeng X. The 50th anniversary of the metaphorical butterfly effect since Lorenz (1972): multistability, multiscale predictability, and sensitivity in numerical models. *Atmosphere.* 2023;14(8):1279. <https://doi.org/10.3390/atmos14081279>.
8. Montori VM, Gafni A, Charles C. A shared treatment decision-making approach between patients with chronic conditions and their clinicians: the case of diabetes. *Health Expect.* 2006;9(1):25-36. <https://doi.org/10.1111/j.1369-7625.2006.00359.x>.
9. Levine DM, Landon BE, Linder JA. Trends in patient-perceived shared decision making among adults in the United States, 2002-2014. *Ann Fam Med.* 2017;15(6):552-556. <https://doi.org/10.1370/afm.2132>.
10. Perron ME, Hudon C, Roux-Levy PH, et al. Shared decision-making with patients with complex care needs: a scoping review. *BMC Prim Care.* 2024;25(1):390. <https://doi.org/10.1186/s12875-024-02633-9>.
11. Elwyn G, Dehlendorf C, Epstein RM, et al. Shared decision making and motivational interviewing: achieving patient-centered care across the spectrum of health care problems. *Ann Fam Med.* 2014;12(3):270-275. <https://doi.org/10.1370/afm.1615>. Errata en: *Ann Fam Med.* 2014;12(4):301.
12. Sturmberg JP, Martin CM, Katerndahl DA. Systems and complexity thinking in the general practice literature: an integrative, historical narrative review. *Ann Fam Med.* 2014;12(1):66-74. <https://doi.org/10.1370/afm.1593>.
13. Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. *BMJ.* 2008;336(7656):1281-1283. <https://doi.org/10.1136/bmj.39569.510521.AD>.
14. Rabi DM, Kunneman M, Montori VM. When guidelines recommend shared decision-making. *JAMA.* 2020;323(14):1345-1346. <https://doi.org/10.1001/jama.2020.1525>.
15. Stigglebout AM, Van der Weijden T, De Wit MP, et al. Shared decision making: really putting patients at the centre of healthcare. *BMJ.* 2012;344:e256. <https://doi.org/10.1136/bmj.e256>.
16. Royal College of Surgeons of Edinburgh. Shared decision making: an essential step in optimal patient care. Edinburgh: Royal College of Surgeons of Edinburgh blog. 2023 Sep [citado 2026 feb 12]. Disponible en: <https://www.rcsed.ac.uk/news-resources/rcsed-blog/2023/september/shared-decision-making-an-essential-step-in-optimal-patient-care>.
17. Elwyn G, Leavitt HB, Vesling M, et al. 'A return on investment'. How shared decision-making benefits doctors: a key informant interview study. *BMJ Open.* 2025;15(12):e107732. <https://doi.org/10.1136/bmjopen-2025-107732>.
18. Lifshitz Guinzberg A, Pomposo García-Cohen A. Las ciencias de la complejidad y la educación médica. *Investig Educ Médica.* 2017;6(24):267-271. <https://doi.org/10.1016/j.riem.2017.06.001>.
19. Karthikeyan G. Evidence-based medicine and clinical judgment: an imaginary divide. *J Am Coll Cardiol.* 2007;49(9):1012; author reply 1012-3. <https://doi.org/10.1016/j.jacc.2006.12.011>.
20. Maree P, Hughes R, Radford J, et al. Integrating patient complexity into health policy: a conceptual framework. *Aust Health Rev.* 2021;45(2):199-206. <https://doi.org/10.1071/AH19290>.
21. Waddell A, Lennox A, Spassova G, et al. Barriers and facilitators to shared decision-making in hospitals from policy to practice: a systematic review. *Implement Sci.* 2021;16(1):74. <https://doi.org/10.1186/s13012-021-01142-y>.
22. Fisher KA, Tan ASL, Matlock DD, et al. Keeping the patient in the center: Common challenges in the practice of shared decision making. *Patient Educ Couns.* 2018;101(12):2195-2201. <https://doi.org/10.1016/j.pec.2018.08.007>.
23. Haesebaert J, Adekpedjou R, Croteau J, et al. Shared decision-making experienced by Canadians facing health care decisions: a Web-based survey. *CMAJ Open.* 2019;7(2):E210-E216. <https://doi.org/10.9778/cmajo.20180202>.
24. Montori VM, Ruissen MM, Hargraves IG, et al. Shared decision-making as a method of care. *BMJ Evid Based Med.* 2023;28(4):213-217. <https://doi.org/10.1136/bmjebm-2022-112068>.
25. Légaré F, Adekpedjou R, Stacey D, et al. Interventions for increasing the use of shared decision making by healthcare professionals. *Cochrane Database Syst Rev.* 2018;7(7):CD006732. <https://doi.org/10.1002/14651858.CD006732.pub4>.
26. Pu S, Xie P, Chen H, et al. Evaluation of outcome of chemotherapy for breast cancer patients older than 70 years: a SEER-based study. *Front Oncol.* 2023;13:992573. <https://doi.org/10.3389/fonc.2023.992573>.
27. Muss HB, Woolf S, Berry D, et al. Adjuvant chemotherapy in older and younger women with lymph node-positive breast cancer. *JAMA.* 2005;293(9):1073-1081. <https://doi.org/10.1001/jama.293.9.1073>.
28. Repetto L. Greater risks of chemotherapy toxicity in elderly patients with cancer. *J Support Oncol.* 2003;1(4 Suppl 2):18-24.
29. Greenhalgh T, Papoutsis C. Studying complexity in health services research: desperately seeking an overdue paradigm shift. *BMC Med.* 2018;16(1):95. <https://doi.org/10.1186/s12916-018-1089-4>.
30. Butterworth JE, Mattick K, Richards SH. It is time to recognise shared decision-making as a complex intervention. *Patient Educ Couns.* 2026;144:109447. <https://doi.org/10.1016/j.pec.2025.109447>.
31. Samost-Williams A, Thomas EJ, Lounsbury O, et al. Bringing team science to the ambulatory diagnostic process: how do patients and clinicians develop shared mental models? *Diagnosis (Berl).* 2024;12(1):25-34. <https://doi.org/10.1515/dx-2024-0115>.
32. Khalil H, Lakhani A. Using systems thinking methodologies to address health care complexities and evidence implementation. *JBIM Evid Implement.* 2021;20(1):3-9. <https://doi.org/10.1097/XEB.0000000000000303>.
33. Marshall DA, Burgos-Liz L, IJerman MJ, et al. Applying dynamic simulation modeling methods in health care delivery research-the SIMULATE checklist: report of the ISPOR simulation modeling emerging good practices task force. *Value Health.* 2015;18(1):5-16. <https://doi.org/10.1016/j.jval.2014.12.001>.
34. Légaré F, Moumjid-Ferdjaoui N, Drolet R, et al. Core competencies for shared decision making training programs: insights from an international, interdisciplinary working group. *J Contin Educ Health Prof.* 2013;33(4):267-273. <https://doi.org/10.1002/chp.21197>.
35. Reeves S, Pelone F, Harrison R, et al. Interprofessional collaboration to improve professional practice and healthcare outcomes. *Cochrane Database Syst Rev.* 2017;6(6):CD000072. <https://doi.org/10.1002/14651858.CD000072.pub3>.

¿Puede un antidepresivo cambiar una vida? Evidencia, experiencia y relevancia clínica

Nahuel E. Baca[■]

Servicio de Salud Mental, Hospital Italiano. Buenos Aires, Argentina

RESUMEN

La eficacia de los antidepresivos ha sido objeto de un intenso debate. Gran parte de la crítica basada en la evidencia se ha centrado en la interpretación de tamaños de efecto promedio obtenidos en ensayos clínicos y metanálisis. En este contexto, muchos resultados frecuentemente descriptos como “clínicamente no significativos” han alimentado lecturas escépticas o incluso nihilistas sobre su utilidad terapéutica. Sin embargo, esta forma de leer la evidencia resulta insuficiente para dar cuenta de la complejidad del tratamiento real y de los procesos que sostienen una recuperación más allá del efecto farmacológico aislado.

Este artículo propone una reflexión desde una perspectiva integradora en psiquiatría, orientada a incluir los datos bioestadísticos con el contexto clínico y la experiencia que se construye en el encuentro entre profesional y paciente, poniendo en cuestión la identificación automática entre significancia estadística y relevancia clínica.

Se argumenta que reducciones modestas en escalas de síntomas pueden adquirir un valor decisivo en contextos de sufrimiento intenso, donde pequeños cambios permiten recuperar funciones básicas, sostener la vida cotidiana o habilitar otros procesos terapéuticos. Finalmente, se propone una lectura integrada que evite tanto la medicalización indiscriminada como el rechazo simplista de intervenciones con efectos promedio modestos, reivindicando el juicio clínico, el vínculo terapéutico y una psiquiatría basada en valores.

Palabras clave: antidepresivos, relevancia clínica, juicio clínico, psiquiatría basada en valores, relación terapéutica.

Can an Antidepressant Change a Life? Evidence, Experience, and Clinical Relevance

ABSTRACT

The efficacy of antidepressants has been the subject of an intense debate. Much of the evidence-based criticism has focused on the interpretation of average effect sizes derived from clinical trials and meta-analyses. In this context, results frequently described as “clinically non-significant” have fueled skeptical, and at times even nihilistic, views regarding their therapeutic usefulness. However, this way of reading the evidence proves insufficient to account for the complexity of real-world treatment and the processes that sustain recovery beyond the isolated pharmacological effect.

Autor para correspondencia: nahuel.baca@hospitalitaliano.org.ar, Baca NE.

Recibido: 9/03/2026 | Aceptado: 16/06/2026 | Publicado: 8/09/2026

DOI: <http://doi.org/10.51987/Rev.Hosp.Ital.B.Aires.v46i3.1320>

Cómo citar: Baca NE. ¿Puede un antidepresivo cambiar una vida? Evidencia, experiencia y relevancia clínica. *Rev Hosp Ital B.Aires.* 2026;46(3):e0001320

This article proposes a reflection from an integrative psychiatric perspective, aimed at incorporating biostatistical data within the clinical context and the experience that emerges from the encounter between clinician and patient, while questioning the automatic identification between statistical significance and clinical relevance.

It is argued that modest reductions in symptom scales may acquire a decisive value in contexts of intense suffering, where small changes allow the recovery of basic functioning, the sustaining of daily life, or the facilitation of other therapeutic processes. Finally, an integrated reading is proposed that avoids both indiscriminate medicalization and the simplistic rejection of interventions with modest average effects, reaffirming clinical judgment, the therapeutic relationship, and a values-based psychiatry.

Keywords: antidepressants, clinical relevance, clinical judgment, values-based psychiatry, therapeutic relationship.

INTRODUCCIÓN

Este artículo propone una reflexión crítica sobre la relevancia clínica de los antidepresivos a la luz del debate contemporáneo en torno a su eficacia. Se trata de un ejercicio de integración entre evidencia cuantitativa, experiencia clínica y los supuestos desde los que el clínico se ubica frente al sufrimiento del paciente. El eje central es interrogar la identificación automática entre significancia estadística y relevancia clínica, especialmente cuando los tamaños de efecto promedio son interpretados como “pequeños” o “clínicamente no significativos”.

El debate contemporáneo sobre la eficacia de los antidepresivos puede ordenarse en tres posiciones. La primera, representada por el metanálisis en red de Cipriani y cols. publicado en *The Lancet* en 2018, sostiene –a partir de 522 ensayos clínicos y más de 116 000 pacientes– que la mayoría de los antidepresivos estudiados resultan significativamente más eficaces que el placebo, con diferencias de eficacia entre las distintas moléculas¹. Esta posición es defendida con énfasis por Goodwin y Nutt, quienes sostienen que el número necesario para tratar (NNT) de los inhibidores selectivos de la recaptación de serotonina (ISRS) es comparable al de muchos fármacos utilizados para condiciones médicas y que la crítica a los antidepresivos refleja sesgos del debate más que limitaciones reales de la evidencia². Hieronymus y cols. refuerzan esta línea argumentando que el umbral de relevancia clínica propuesto por los críticos es arbitrariamente alto e implica exigencias que ningún tratamiento médico cumpliría³.

La segunda posición, asociada principalmente a Kirsch^{4,5}, Moncrieff^{6,7} y Hengartner^{8,9}, sostiene que esa diferencia, aun siendo estadísticamente real, es clínicamente trivial y está sobreestimada por sesgos en el proceso de investigación. Esta línea encuentra soporte adicional en el reanálisis de los datos originales de Cipriani realizado por Munkholm y cols., que cuestiona la calidad de la evidencia bajo criterios GRADE¹⁰.

La tercera, pragmática-contextual, propuesta por Vöhringer y Ghaemi, distingue por gravedad y contexto clínico, sosteniendo que los antidepresivos son claramente eficaces en depresiones moderadas y graves pero probablemente inútiles en las leves¹¹. La presente

reflexión se inscribe en esa línea, pero busca llevarla un paso más allá.

Conviene precisar tres términos centrales en la literatura sobre eficacia antidepresiva, frecuentemente tratados como sinónimos. Siguiendo el consenso establecido por Frank y cols. en 1991 y revisado por la ACNP Task Force en 2006^{12,13}: respuesta designa una reducción $\geq 50\%$ en la gravedad sintomática con respecto al basal; remisión, un estado virtualmente asintomático (*Hamilton Depression Rating Scale* de 17 ítems, HAM-D-17, con puntaje ≤ 7) sostenido por al menos dos a tres semanas; y recuperación, una remisión que se mantiene por al menos cuatro a seis meses, lo suficiente para considerar finalizado el episodio. Esta distinción toma relevancia porque la crítica contemporánea a los antidepresivos suele plantearse en términos de respuesta en ensayos cortos. Cuando ampliamos el horizonte temporal y observamos remisión sostenida o recuperación, los números se vuelven aún más modestos. La pregunta clínica relevante apunta entonces a qué función le asignamos al fármaco dentro del proceso terapéutico. Esta diferencia se vincula con la distinción clásica entre eficacia y efectividad. La eficacia es lo que un fármaco produce en las condiciones controladas de un ensayo clínico aleatorizado (RCT), donde la selección de pacientes por criterios estrictos, la exclusión de comorbilidades complejas, el seguimiento corto y la estandarización del manejo clínico al mínimo buscan aislar el efecto de la molécula. La efectividad es lo que ese mismo fármaco produce en la práctica clínica real, donde la biografía del paciente, las comorbilidades que el ensayo habría excluido, la adherencia variable y el encuadre terapéutico operan como parte del fenómeno que se intenta tratar.

La mayor parte de la evidencia que domina el debate sobre antidepresivos es de eficacia. Las diferencias de medias estandarizadas (SMD), el NNT y las diferencias en HAM-D provienen de ensayos diseñados para aislar el efecto farmacológico. La investigación de efectividad existe, y muestra que –en seguimientos largos y poblaciones reales– las tasas de remisión sostenida son aún más modestas que las observadas en los RCT clásicos. El ensayo pragmático más grande, el *Sequenced Treatment Alternatives to Relieve Depression* (STAR*D), mostró tasas de remisión

cercanas al 30% con el primer ISRS en condiciones de práctica clínica habitual, alcanzando un acumulado del 67% después de cuatro pasos sucesivos de tratamiento, aunque con tasas de recaída cercanas al 50 % en el seguimiento a doce meses¹⁴. Los datos muestran los límites de lo que el fármaco puede hacer por sí solo, tanto en la mejoría aguda como en la sostenibilidad de esa mejoría en el tiempo, y abren la pregunta de qué procesos clínicos acompañan al fármaco en los casos donde la recuperación se sostiene. Tanto el debate sobre eficacia como el de efectividad se concentran en el efecto del fármaco aislado, dejando afuera la dimensión del encuadre clínico que este artículo se propone trabajar.

Desplacemos entonces la discusión hacia la práctica clínica, donde se juega su sentido. La siguiente escena es una construcción ilustrativa, aunque cualquier psiquiatra reconocerá en ella un cuadro frecuente, y busca dar cuerpo concreto al argumento. Un lunes a última hora en el consultorio. Frente a nosotros hay una paciente que hace un mes no puede salir de la cama y fantasea con renunciar al trabajo, con separarse, con “desaparecer”. Es una persona con un hijo que la espera, un jefe que la presiona y una carga de vergüenza y cansancio que ninguna escala alcanza a registrar.

Después de algunas semanas de tratamiento, algo empieza a moverse. Sigue cansada y con momentos oscuros, pero vuelve a llegar a horario al trabajo, retoma una actividad que había abandonado y puede decir “no” donde antes se anulaba. Si tuviéramos que ponerle un titular a ese cambio, probablemente sería: “volvió a respirar”. La escena funciona como punto de partida para examinar qué queda afuera cuando el efecto terapéutico se mide exclusivamente por promedios.

Cambiamos ahora de escenario. Ya no estamos en el consultorio sino analizando un metanálisis, entre curvas, tamaños de efecto y desviaciones estándar. Allí, la misma intervención que permitió a esta paciente sostenerse un poco mejor aparece bajo otra etiqueta:

“SMD = 0,3. Efecto pequeño. Clínicamente no significativo”

Entre esas dos descripciones hay una brecha que merece atención. La crítica a los antidepresivos es necesaria, la evidencia tiene limitaciones metodológicas reales, el placebo es potente y la medicalización del sufrimiento tiene riesgos concretos. Pero si nos quedamos únicamente con ciertos umbrales estadísticos, terminamos diciendo algo que en la práctica suena absurdo, que una herramienta capaz de inclinar la balanza en la vida de muchas personas “no funciona”.

Retomando lo señalado en la introducción, conviene reforzar el argumento. Scott Alexander Siskind plantea que, si aplicáramos a toda la farmacología los mismos criterios con los que se juzga a los antidepresivos, gran parte de los fármacos que hoy consideramos estándar de cuidado quedarían descartados como “no significativos”¹⁵. Goodwin y Nutt han mostrado que el NNT de los antidepresivos es comparable al de muchos fármacos utilizados para condiciones médicas crónicas, lo que vuelve injustificada la asimetría con que se exige a los

antidepresivos un umbral de eficacia más alto que al resto de la farmacología². La pregunta, entonces, es qué estamos midiendo y qué dejamos afuera cuando reducimos la eficacia a un número.

DESARROLLO

Una traducción universal con letra chica

Para hacer comparables intervenciones de naturaleza distinta entre sí, la investigación clínica recurre a una medida común, el tamaño del efecto. Es una herramienta estadística que cuantifica la magnitud de una diferencia entre dos grupos en términos relativos, independientemente de la escala original con la que fue medida. En los ensayos sobre antidepresivos, la métrica más utilizada es el SMD, muy similar a la *d* de Cohen¹⁶. El SMD indica cuán lejos está el grupo tratado del grupo placebo, medido en unidades de desviación estándar. Cohen, en 1988, propuso una convención para interpretar estos valores¹⁷:

-0,2: efecto pequeño

-0,5: efecto moderado

-0,8: efecto grande

Esta regla es útil como referencia, pero es solo eso, una convención. No es una ley de la naturaleza, y usarla como un umbral rígido para decidir si algo “sirve” o “no sirve” es el origen de gran parte del malentendido.

Es importante aclarar desde el inicio que una magnitud de efecto pequeña no implica automáticamente ausencia de relevancia clínica. El tamaño del efecto describe una diferencia promedio entre grupos, y la significancia clínica, en cambio, depende del contexto, de la gravedad basal, de los objetivos del paciente y del impacto funcional del cambio. Un SMD de 0,3 puede ser irrelevante en un cuadro leve y, al mismo tiempo, clínicamente decisivo en una depresión grave con alto deterioro funcional. Confundir magnitud estadística con relevancia clínica es uno de los errores más frecuentes en este debate.

El argumento del efecto modesto

Esta es la posición que sostienen Kirsch^{4,5}, Moncrieff^{6,7}, Hengartner^{8,9} y Munkholm¹⁰. La mayoría de los ensayos miden la depresión con escalas como la de Hamilton (HAM-D), y durante años, organismos como el NICE británico establecieron tradicionalmente que una diferencia “clínicamente relevante” requería al menos 3 puntos en la HAM-D-17 o un SMD superior a 0,50¹⁸. La guía vigente de NICE (National Institute for Health and Care Excellence) (NG222, 2022) abandonó ese umbral específico y mantiene a los antidepresivos como tratamiento de primera línea para depresión moderada y grave, reconociendo su utilidad clínica más allá de los criterios estadísticos.

El metanálisis de Irving Kirsch sobre datos presentados a la Food and Drug Administration (FDA) de los Estados Unidos muestra que la diferencia promedio entre antidepresivo y placebo es de unos 1,8 a 2 puntos en la HAM-D, lo que se traduce en un SMD de -0,30^{4,5}. Con esos números y la vara de NICE, la conclusión parece inevitable y se ha convertido en un titular recurrente: el beneficio está por debajo de los criterios aceptados de significancia clínica. Leída en esos términos, la conclusión parecería

clara: los antidepresivos apenas superan al placebo. Vale la pena, sin embargo, examinar cuán bien calibrada está la herramienta con la que estamos midiendo.

DISCUSIÓN

Un estudio de Leucht y cols. hizo algo simple y revelador: comparó los tamaños del efecto de los psicofármacos con los de medicamentos de otras especialidades^{19,20}. El resultado fue una sorpresa para muchos: el efecto promedio de un antidepresivo sobre los síntomas depresivos es de la misma magnitud o mayor que el efecto de las estatinas en algunos de sus usos preventivos. En la revisión de Leucht, varios fármacos médicos de uso extendido presentan tamaños de efecto similares o incluso menores que los de los antidepresivos: las estatinas en prevención cardiovascular muestran un SMD $\approx 0,15$ y la aspirina en prevención secundaria un SMD $\approx 0,12$, mientras que los antidepresivos en tratamiento agudo se sitúan en torno a $0,30^{4,5,19}$. En promedio, los psicofármacos no son menos eficaces que muchos tratamientos médicos considerados estándar de cuidado¹⁹. Lo que está sobredimensionado, en realidad, es lo que esperamos que un fármaco haga por sí solo, en psiquiatría y en cualquier otra especialidad.

Por supuesto hay un contraargumento poderoso. La evidencia está plagada de sesgos de publicación y diseños metodológicos pobres, y los factores no específicos del ensayo –encuadre, expectativas, calidad del vínculo con el investigador– tienen un peso considerable sobre el resultado, comparable en muchos casos al del fármaco activo^{5,9,10}. Hengartner y Plöderl estiman que la diferencia mínima importante (MID) en la HAM-D-17 debería estar entre 3 y 5 puntos, por encima del promedio observado⁸. Hay que tomarse esto muy en serio, pero de reconocer estas fallas a saltar al nihilismo de la idea de que “no funcionan”²¹, hay un trecho más conceptual que numérico. La evidencia más matizada, como la que presentan Vöhringer y Ghaemi, apunta a una conclusión incómoda para los extremos. Los antidepresivos son claramente superiores al placebo en depresiones moderadas y graves, y probablemente inútiles en las leves¹¹. Son una herramienta más, con indicaciones precisas.

Sin embargo, la discusión no se agota en el plano metodológico. Su límite aparece cuando los promedios poblacionales se trasladan sin mediaciones al terreno de la práctica clínica, donde se encuentran con biografías singulares. En la realidad cotidiana del consultorio, en guardias, o salas de internación, atendemos pacientes con presentaciones clínicas, historias y comorbilidades muy distintas entre sí. Nuestra impresión clínica es que el “tamaño del efecto” de un antidepresivo bien indicado y bien acompañado suele ser mucho mayor que ese 0,3 promedio. ¿Por qué? Porque la clínica no es un RCT.

Hay tres razones para esa diferencia entre el promedio del ensayo y la clínica real. La primera es que el fármaco no actúa en el vacío. La pastilla es parte de un plan que puede incluir psicoterapia, cambios de hábitos y un acompañamiento clínico sostenido, de modo que lo que medimos como “efecto del fármaco” en la vida real es un cóctel complejo de farmacología, placebo, vínculo

y esperanza. La segunda es que no tratamos al “paciente promedio”. La mayoría de los RCT excluyen a pacientes con comorbilidades, trastornos de personalidad, riesgo suicida o consumo de sustancias, mientras que en el consultorio trabajamos con esa complejidad, incluido el riesgo agudo que los ensayos suelen dejar afuera por razones éticas y metodológicas. La tercera es que el sufrimiento no es lineal. Una reducción de 3 puntos en la HAM-D puede ser trivial para alguien con malestar leve, pero para quien está en la cima del pánico o la desesperanza, es el escalón que le permite volver a tomar aire, un salto cualitativo de la desesperación a sostener la vida cotidiana que ningún SMD captura.

Ningún número puede agotar lo que significa para alguien que pensaba en morir a diario empezar, simplemente, a sentir que vale la pena levantarse por la mañana.

En la clínica diaria, el efecto del fármaco rara vez aparece solo. Aparece dentro de un encuadre, atravesado por la confianza o la desconfianza que el paciente trae al consultorio. El profesional que dedica tiempo a explicar qué se puede esperar del fármaco, que cuida la información que ofrece y construye una alianza terapéutica sólida, está trabajando con algo más que una molécula. Está convocando un fenómeno propio, distinto del fármaco, que la literatura llama efecto placebo²².

La evidencia respalda esta lectura. El estudio del *NIMH Treatment of Depression Collaborative Research Program* mostró que la alianza terapéutica predice significativamente el resultado clínico en farmacoterapia activa, en placebo y en psicoterapia, operando como un factor relevante en todas las modalidades de tratamiento²³. Una observación reciente de Burke en *Lancet Psychiatry* señala que, en ensayos contemporáneos de dispositivos terapéuticos para depresión, las tasas de remisión en el grupo placebo han alcanzado el 41 %²². Burke propone entender ese fenómeno como un mecanismo activo que puede ser inducido éticamente en la práctica clínica a través de la calidad del encuadre y la relación terapéutica.

En este artículo entendemos el efecto placebo como el poder del organismo de movilizar mecanismos psicobiológicos de curación a través de la creencia y el vínculo²². Existe con independencia del fármaco y opera en cualquier acto terapéutico cuando las condiciones lo permiten. La pregunta clínica relevante hoy es cómo aprovecharlo a favor del paciente, especialmente en cuadros graves donde toda herramienta cuenta. En la práctica, el efecto del fármaco y el del placebo bien convocado pueden potenciarse mutuamente, y esa potenciación produce resultados que un fármaco aislado difícilmente alcanza.

Todo esto se juega en el consultorio. Cuando nos sentamos frente a un paciente que sufre, lo que hacemos en los primeros minutos importa. El modo en que se lo mira, el tiempo que se le dedica, la información que se le ofrece sobre el fármaco, lo que se le promete y lo que no, y la calidad de la escucha que recibe son dimensiones que el ensayo no captura pero el resultado sí registra. Eso es parte fundamental del tratamiento. Hay psiquiatras que lo perciben desde la primera consulta y otros que tardan

más en advertirlo. La diferencia entre unos y otros no se ve en una escala de Hamilton ni en un metanálisis, pero se ve en el consultorio. Se ve en quién mejora y quién no, en quién sigue viniendo y quién abandona, en qué pacientes empiezan a recuperar la vida y cuáles se quedan flotando en la mejoría parcial. Esa es la dimensión clínica que los ensayos no capturan y que el quehacer profesional necesita reconocer.

Lo que un sistema de salud considera un efecto terapéutico “suficiente” no es una pregunta solamente técnica. Implica una decisión sobre qué nivel de sufrimiento la sociedad está dispuesta a tolerar antes de intervenir. A esto puede llamarse *política del sufrimiento*: el conjunto de decisiones implícitas, muchas veces invisibles para quien las toma, mediante las cuales se reparte el dolor evitable entre individuos y se establecen los umbrales para que ese dolor merezca o no respuesta clínica. Canguilhem mostró que definir lo normal y lo patológico implica siempre un juicio de valor²⁴. Adoptar un estándar hiperrestringido de relevancia clínica equivale, bajo apariencia de pureza metodológica, a aceptar tácitamente que el sufrimiento subumbral de muchas personas no merece intervención. Medicalizar todo malestar produce el efecto contrario y conduce a la trampa que denunciaba Illich, la expropiación de la capacidad de las personas para lidiar con su propia vulnerabilidad²⁵. Entre ambos extremos hay un trabajo clínico y conceptual permanente, que cada profesional realiza, consciente o no, cada vez que decide medicar, postergar la medicación o renunciar a ella.

Existe un camino más responsable que el purismo estadístico y la medicalización indiscriminada, que implica desconfiar tanto del entusiasmo como del nihilismo, elegir con mayor precisión a quién medicamos, reservando los antidepresivos para depresiones y ansiedades moderadas o graves, con claro deterioro funcional, y explicar con honestidad qué significa ese “efecto pequeño”.

En la práctica clínica, esa explicación podría sonar algo así: los estudios muestran que el fármaco ofrece, en promedio, una mejoría modesta. Está lejos de ser una cura, pero para muchas personas esa diferencia, aun siendo cuantitativamente discreta, alcanza para que la psicoterapia se sostenga, para volver al trabajo, para que la vida cotidiana deje de ser un peso insoportable. Conviene presentar el fármaco con realismo, como una herramienta entre otras dentro de un proceso terapéutico más amplio, y decidir junto al paciente si, en su situación particular, vale la pena probarla.

Conviene tener presente, además, que los antidepresivos no actúan solo aliviando síntomas. Hay evidencia creciente de que inducen cambios neuroplásticos por mecanismos biológicos directos, aumento de BDNF (factor neurotrófico derivado del cerebro), neurogénesis hipocámpal y modulación de la plasticidad sináptica vía unión directa al receptor TrkB (receptor de tropomiosina quinasa B)^{26,27}. Estos mecanismos son distintos del efecto placebo, aunque pueden potenciarse en el encuentro clínico cuando la alianza, el encuadre y la calidad del cuidado activan dimensiones psicobiológicas adicionales. El sustrato neuroplástico, aun sin ser una cura por sí mismo,

abre una ventana de plasticidad cerebral que el placebo por sí solo no produce. Lo que ocurra dentro de esa ventana, el trabajo psíquico que la persona pueda sostener, con un vínculo terapéutico de calidad, es lo que decide si los cambios se traducen en una recuperación duradera o si los síntomas reaparecen al suspender el fármaco.

Finalmente, conviene reconocer al llamado “placebo” como expresión del poder del ritual, de la expectativa y, sobre todo, de la relación terapéutica. Lejos de oponerse a la farmacología, el cuidado del encuadre forma parte de la eficacia misma del tratamiento. Elegir la molécula adecuada importa, pero también importa cómo se prescribe, en qué contexto y con qué sentido compartido⁵.

Se trata, en definitiva, de sostener dos afirmaciones que conviven sin contradicción: los antidepresivos tienen en promedio un efecto modesto, y una parte relevante de ese efecto depende del placebo y del contexto. Para una persona concreta, en un momento concreto, ese incremento puede ser la diferencia entre seguir hundiéndose o empezar a remontar.

Entre el *paper* y la persona hay un espacio donde se arriesga lo decisivo. Ese espacio lo habita el juicio clínico, sostenido por una conversación honesta y por una ética del cuidado. Detrás de cada SMD hay una biografía concreta que intenta, como puede, recuperar alguna forma de vida propia.

Agradecimientos: Agradezco a todo el equipo de sala de internación de Psiquiatría del Hospital Italiano, Buenos Aires, y al equipo de Modra Psiquiatría Basada en Valores por las conversaciones y el sostén intelectual que hicieron posible este trabajo.

Declaración sobre el uso de inteligencia artificial (IA): durante la preparación de este manuscrito, el autor utilizó herramientas de inteligencia artificial como apoyo para tareas de edición y revisión del estilo. El autor revisó y validó el contenido final y asume la responsabilidad exclusiva por el contenido final de la publicación.

Conflictos de intereses: el autor declara no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: el autor declara que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet*. 2018;391(10128):1357-1366. [https://doi.org/10.1016/S0140-6736\(17\)32802-7](https://doi.org/10.1016/S0140-6736(17)32802-7).
2. Goodwin GM, Nutt D. Antidepressants; what's the beef? *Acta Neuropsychiatr*. 2019;31(1):59-60. <https://doi.org/10.1017/neu.2019.4>.
3. Hieronymus F, Jauhar S, Østergaard SD, et al. One (effect) size does not fit at all: interpreting clinical significance and effect sizes in depression treatment trials. *J Psychopharmacol*. 2020;34(10):1074-1078. <https://doi.org/10.1177/0269881120922950>.
4. Kirsch I, Deacon BJ, Huedo-Medina TB, et al. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med*. 2008;5(2):e45. <https://doi.org/10.1371/journal.pmed.0050045>.
5. Kirsch I. Placebo effect in the treatment of depression and anxiety. *Front Psychiatry*. 2019;10:407. <https://doi.org/10.3389/fpsy.2019.00407>.

6. Moncrieff J, Kirsch I. Empirically derived criteria cast doubt on the clinical significance of antidepressant-placebo differences. *Contemp Clin Trials*. 2015;43:60-62. <https://doi.org/10.1016/j.cct.2015.05.005>.
7. Moncrieff J. The myth of the chemical cure: a critique of psychiatric drug treatment. New York: Palgrave Macmillan; 2009.
8. Hengartner MP, Plöderl M. Estimates of the minimal important difference to evaluate the clinical significance of antidepressants in the acute treatment of moderate-to-severe depression. *BMJ Evid Based Med*. 2022;27(2):69-73. <https://doi.org/10.1136/bmjebm-2020-111600>.
9. Hengartner MP. Evidence-biased antidepressant prescription: overmedicalisation, flawed research, and conflicts of interest. Cham: Springer; 2022.
10. Munkholm K, Paludan-Müller AS, Boesen K. Considering the methodological limitations in the evidence base of antidepressants for depression: a reanalysis of a network meta-analysis. *BMJ Open*. 2019;9(6):e024886. <https://doi.org/10.1136/bmjopen-2018-024886>.
11. Vöhringer PA, Ghaemi SN. Solving the antidepressant efficacy question: effectsizesinmajordepressive disorder. *Clin Ther*. 2011;33(12):B49-B61. <https://doi.org/10.1016/j.clinthera.2011.11.019>.
12. Frank E, Prien RF, Jarrett RB, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence. *Arch Gen Psychiatry*. 1991;48(9):851-855. <https://doi.org/10.1001/archpsyc.1991.01810330075011>.
13. Rush AJ, Kraemer HC, Sackeim HA, et al. Report by the ACNP Task Force on response and remission in major depressive disorder. *Neuropsychopharmacology*. 2006;31(9):1841-1853. <https://doi.org/10.1038/sj.npp.1301131>.
14. Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry*. 2006;163(11):1905-1917. <https://doi.org/10.1176/ajp.2006.163.11.1905>.
15. Alexander S. All medications are insignificant in the eyes of God and traditional effect size criteria. *Astral Codex Ten*. 2023 May 30 [citado 2026 mar 30]. Disponible en: https://open.substack.com/pub/astralcodexten/p/all-medications-are-insignificant?utm_campaign=post-expanded-share&utm_medium=web.
16. Sullivan GM, Feinn R. Using effect size-or why the P aalue is not enough. *J Grad Med Educ*. 2012;4(3):279-282. <https://doi.org/10.4300/JGME-D-12-00156.1>.
17. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
18. National Institute for Health and Care Excellence. Depression in adults: treatment and management. London: NICE; 2022. NICE guideline [NG222].
19. Leucht S, Hierl S, Kissling W, et al. Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. *Br J Psychiatry*. 2012;200(2):97-106. <https://doi.org/10.1192/bjp.bp.111.096594>.
20. Leucht S, Fennema H, Engel R, et al. What does the HAMD mean? *J Affect Disord*. 2013;148(2-3):243-248. <https://doi.org/10.1016/j.jad.2012.12.001>.
21. Stegenga J. Medical nihilism. Oxford: Oxford University Press; 2018.
22. Burke MJ. A fundamental change is needed for appraising placebo responses in psychiatry. *Lancet Psychiatry*. 2023;10(5):316-317. [https://doi.org/10.1016/S2215-0366\(23\)00068-8](https://doi.org/10.1016/S2215-0366(23)00068-8).
23. Krupnick JL, Sotsky SM, Simmens S, et al. The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: findings in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *J Consult Clin Psychol*. 1996;64(3):532-539. <https://doi.org/10.1037//0022-006x.64.3.532>.
24. Canguilhem G. The normal and the pathological. New York: Zone Books; 1991.
25. Illich I. Medical nemesis: the expropriation of health. New York: Pantheon Books; 1976.
26. Casarotto PC, Giryh M, Fred SM, et al. Antidepressant drugs act by directly binding to TRKB neurotrophin receptors. *Cell*. 2021;184(5):1299-1313. e19. <https://doi.org/10.1016/j.cell.2021.01.034>.
27. Castrén E, Monteggia LM. Brain-derived neurotrophic factor signaling in depression and antidepressant action. *Biol Psychiatry*. 2021;90(2):128-136. <https://doi.org/10.1016/j.biopsych.2021.05.008>.