



REVISTA DE GASTROENTEROLOGÍA DE MÉXICO

www.elsevier.es/rgmx



ARTÍCULO DE REVISIÓN

Nuevas herramientas para la clasificación y puntuación de enfermedad por reflujo gastroesofágico

S. Siboni^{a,*}, M. Sozzi^a, P. Visaggi^b, R. De Maron^a y E. Asti^a

^a División de Cirugía General y de Urgencia, IRCCS Policlinico San Donato, Universidad de Milán, Milán, Italia

^b División de Gastroenterología, Departamento de Investigación Traslacional y Nuevas Tecnologías, Universidad de Pisa, Pisa, Italia

Recibido el 16 de abril de 2025; aceptado el 14 de mayo de 2025

PALABRAS CLAVE

Enfermedad por reflujo gastroesofágico; Endoscopia; Manometría de alta resolución; Estudio de monitoreo de reflujo

Resumen La enfermedad por reflujo gastroesofágico (ERGE) es una condición prevalente caracterizada por el flujo retrógrado de contenido gástrico al esófago, con un impacto significativo en la calidad de vida. Los abordajes diagnósticos tradicionales a menudo carecen de precisión debido al traslape de síntomas con otras condiciones. La presente revisión presenta la clasificación de la *American Foregut Society* (AFS), el puntaje de Milán, el puntaje pHoenix, el puntaje COuGH RefluX y el puntaje de Lyon, 5 nuevas herramientas diseñadas para mejorar la objetividad, reproducibilidad y la relevancia clínica del diagnóstico de ERGE.

La clasificación de la AFS refina la evaluación endoscópica de la integridad de la unión esofagogástrica (UEG) al incorporar parámetros medibles (longitud de hernia hiatal, diámetro de apertura hiatal y válvula «flap»), superando la subjetividad de la clasificación de Hill. El puntaje de Milán, derivado de la manometría de alta resolución (MAR), integra 4 parámetros (motilidad esofágica ineficaz, integral contráctil de UEG, morfología UEG y respuesta de elevación de pierna recta) para cuantificar la disrupción de la barrera antirreflujo (BAR).

El puntaje de pHoenix, desarrollado para el monitoreo prolongado remoto del pH, mide el tiempo de exposición a ácido (TEA) supino más profundamente, abordando las limitaciones del puntaje de DeMeester y el consenso de Lyon 2.0. El puntaje COuGH RefluX, un modelo de predicción clínica para síntomas laringofaríngeos, utiliza 6 parámetros (tos, obesidad, globo, hernia hiatal, regurgitación y sexo masculino) para estratificar la probabilidad de ERGE. Finalmente, el puntaje de Lyon integra datos endoscópicos y de impedancia de pH, categorizando pacientes en fenotipos (de sin ERGE a ERGE grave) y prediciendo desenlaces de tratamientos.

Estas herramientas, de manera colectiva, abordan los desafíos de diagnóstico al estandarizar las evaluaciones y mejorar la estratificación de los pacientes. Al reducir la ambigüedad diagnóstica y guiar la terapia personalizada, estas innovaciones prometen transformar el manejo

* Autor para correspondencia. División de Cirugía General y de Urgencia, IRCCS Policlinico San Donato. Piazza Malan 1, 20097, San Donato Milanese (Milano), Italia. Número telefónico: +39 0252774896

Correo electrónico: Stefano.siboni@grupposandonato.it (S. Siboni).

<https://doi.org/10.1016/j.rgmx.2025.05.005>

0375-0906/© 2025 Asociación Mexicana de Gastroenterología. Publicado por Masson Doyma México S.A. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Cómo citar este artículo: S. Siboni, M. Sozzi, P. Visaggi et al., Nuevas herramientas para la clasificación y puntuación de enfermedad por reflujo gastroesofágico, Revista de Gastroenterología de México, <https://doi.org/10.1016/j.rgmx.2025.05.005>

de ERGE, particularmente al seleccionar candidatos para intervenciones médicas o quirúrgicas de mayor impacto.

© 2025 Asociación Mexicana de Gastroenterología. Publicado por Masson Doyma México S.A. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

KEYWORDS

Gastro-esophageal reflux disease;
Endoscopy;
High-resolution manometry;
Reflux monitoring study

New tools for the grading and scoring of gastro-esophageal reflux disease

Abstract Gastroesophageal reflux disease (GERD) is a prevalent condition characterized by the retrograde flow of gastric contents into the esophagus, significantly impacting quality of life. Traditional diagnostic approaches often lack precision due to symptom overlap with other conditions. This review introduces the American Foregut Society (AFS) classification, Milan score, pHoenix score, COuGH RefluX score, and Lyon score, five novel tools designed to enhance the objectivity, reproducibility, and clinical relevance of GERD diagnosis.

The AFS classification refines the endoscopic assessment of esophagogastric junction (EGJ) integrity by incorporating measurable parameters (hiatal hernia length, hiatal aperture diameter, and flap valve), overcoming the subjectivity of the Hill classification. The Milan score, derived from high-resolution manometry, integrates four parameters (ineffective esophageal motility, EGJ-contractile integral, EGJ morphology, and straight leg raise response) to quantify anti-reflux barrier (ARB) disruption.

The pHoenix score, developed for prolonged wireless pH monitoring, weights supine AET more heavily, addressing limitations of the DeMeester score and Lyon 2.0 consensus. The COuGH RefluX score, a clinical prediction model for laryngopharyngeal symptoms, uses six parameters (cough, obesity, globus, hiatal hernia, regurgitation, male sex) to stratify GERD likelihood. Finally, the Lyon score integrates endoscopic and pH-impedance data, categorizing patients into phenotypes (from no GERD to severe GERD) and predicting treatment outcomes.

These tools collectively address diagnostic challenges by standardizing assessments and improving patient stratification. By reducing diagnostic ambiguity and guiding personalized therapy, these innovations hold promise for transforming GERD management, particularly in selecting candidates for escalated medical or surgical interventions.

© 2025 Asociación Mexicana de Gastroenterología. Published by Masson Doyma México S.A. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

La enfermedad por reflujo gastroesofágico (ERGE) es una condición común en la que contenido gástrico fluye de regreso al esófago, causando síntomas que pueden afectar de manera significativa la calidad de vida¹. La ERGE presenta un amplio rango de síntomas, que a menudo se categorizan en dos grupos: síntomas «típicos», como acidez y regurgitación, que tienen una gran probabilidad de estar ligados a ERGE, mientras que síntomas «atípicos», como tos, asma o ronquera, se pueden asociar con menor confiabilidad a la condición².

Para los pacientes que experimentan síntomas típicos de ERGE sin síntomas alarmantes, como pérdida de peso no deseado o disfagia, la *American Gastroenterological Association* recomienda probar con 8 semanas de terapia de inhibidor de bomba de protones (IBP), tomada una vez por día antes de una comida. Esta recomendación está sustentada por evidencia de nivel moderado y se considera una recomendación sólida³. La evaluación diagnóstica normalmente está indicada después de la prueba del IBP en caso de una respuesta inadecuada, o por adelantado si los síntomas presentes son atípicos⁴.

El diagnóstico de primera línea para ERGE es normalmente la endoscopia. Este procedimiento puede confirmar ERGE al identificar hallazgos específicos, como esofagitis grado Los Ángeles B, C o D, histológicamente confirmada por esófago de Barret, o una estenosis péptica. Además, la endoscopia proporciona también información valiosa sobre el grado de alteración de la unión esofagogastrica (UEG), evaluada por medio de la clasificación de Hill y puede detectar una hernia hiatal, la cual es una condición en la que el esfínter esofágico inferior (EEI) y el diafragma crural (DC) están separados, lo cual es un riesgo mayor para ERGE^{5,6}.

Si la endoscopia no proporciona evidencia concluyente de ERGE o si se está considerando la intervención quirúrgica, se vuelve necesaria la evaluación de la función esofágica. La manometría de alta resolución (MAR) es una herramienta clave en esta fase. Mientras que la MAR no permite un diagnóstico directo de ERGE, su papel es primordial, más allá de solamente guiar la colocación del catéter para evaluar el reflujo, pues ayuda a descartar otras condiciones que pueden simular ERGE, como la acalasia, otros trastornos de motilidad, o condiciones conductuales⁷⁻⁹. A lo largo de la década pasada, la MAR ha sido explorada extensamente para identificar variables y maniobras provocadoras que permi-

tan cuantificar la alteración de la barrera antirreflujo (BAR) y diferenciar pacientes ERGE de pacientes sanos¹⁰⁻¹⁵.

Para casos en los cuales la endoscopia no es concluyente, el monitoreo del reflujo se considera el estándar de oro para diagnosticar ERGE. A lo largo de los años, se han desarrollado diferentes criterios para interpretar los estudios de monitoreo de reflujo basados en catéter de 24h. En 1974, Johnston y DeMeester presentaron el puntaje de DeMeester, el cual evalúa 6 parámetros de reflujo y aún es utilizado ampliamente por cirujanos para seleccionar candidatos para cirugía de antirreflujo (CAR)¹⁶. Más recientemente, el consenso de Lyon 2.0 propuso referencias actualizadas: un tiempo de exposición a ácido (TEA) mayor del 6% indica definitivamente ERGE, un TEA entre 4 y 6% no es concluyente y TEA menor del 4% sugiere que no hay ERGE. Para casos no concluyentes (TEA 4-6%), se utilizan criterios adicionales para confirmar o descartar ERGE, incluidos un total de 80 episodios de reflujo o más, una media nocturna de impedancia basal (MNIB) menor de 1500ohms, o una asociación entre eventos de reflujo y síntomas^{4,17}.

Mientras que estos criterios y pruebas establecidos proporcionan una confianza sólida para el diagnóstico objetivo de ERGE, en la década pasada se han desarrollado varias herramientas novedosas para evaluar la gravedad de ERGE y la respuesta al tratamiento. Estas innovaciones tienen como objetivo hacer más preciso y menos invasivo el diagnóstico de ERGE, además de favorecer la identificación de pacientes que se podrían beneficiar de un tratamiento de mayor impacto. En el presente artículo revisamos estas nuevas herramientas, con enfoque en la evidencia actual, la practicidad de estas, el impacto potencial en la práctica clínica y sus limitaciones.

La clasificación de la *American Foregut Society*

Introducción

La endoscopia es la prueba diagnóstica de primera línea para síntomas gastrointestinales, ofreciendo una visualización directa del tracto gastrointestinal superior. Juega un rol crucial para excluir condiciones, como malignidades, atrofia gástrica y enfermedad por úlcera péptica, de la ruta diagnóstica de ERGE^{3,18}. El consenso de Lyon 2.0 fortalece más el rol crucial de la endoscopia al definir hallazgos específicos, como la esofagitis grado Los Ángeles B, C o D, el esófago de Barret, o la estenosis péptica, como diagnóstico de ERGE. Más allá del diagnóstico, la evaluación endoscópica de la UEG proporciona información esencial sobre factores anatómicos que contribuyen al reflujo. Para abordar limitaciones de las clasificaciones de la UEG anteriores, la *American Foregut Society* (AFS) ha presentado un novedoso sistema de clasificación para evaluar la integridad de la UEG, ofreciendo un abordaje más completo y estandarizado para la evaluación de ERGE¹⁹.

Antecedentes y limitaciones de herramientas anteriores

La clasificación de Hill ha sido considerada la herramienta principal desde 1996 para la evaluación endoscópica de la

válvula «flap» gastroesofágica, la cual es un componente crítico de la BAR⁵. A pesar de su relevancia, la clasificación de Hill no ha sido adoptada en la práctica de rutina debido a varias limitaciones. Se enfoca principalmente en la válvula «flap», con un mínimo de atención a la hernia hiatal. Además, se basa en evaluaciones subjetivas más que en parámetros medibles, y carece de una técnica o nomenclatura endoscópica estandarizada, lo cual reduce su confiabilidad. Por estas razones, la distinción entre grados I y II, ambos considerados normales, carece de significación clínica, mientras que la diferenciación entre grados II y III muestra una mala correlación con tasas de reflujo patológico. Estos defectos han limitado la utilidad de la clasificación de Hill tanto en la práctica clínica como en la investigación, destacando la necesidad de un sistema más robusto.

Descripción del puntaje

La clasificación de la AFS representa una mejora significativa al incorporar medidas objetivas de tres componentes clave de la BAR, resumidas en el acrónimo «LDF»: L (longitud) mide la longitud axial de la hernia hiatal, D (diámetro) evalúa el tamaño de la apertura hiatal, utilizando como referencia el diámetro endoscópico estándar de aproximadamente 1 cm, y F (flap [válvula]) evalúa la presencia (F+) o ausencia (F-) de la válvula «flap» esofagagástrica en el ángulo de His (fig. 1).

Este sistema asigna grados, de 1 (normal) a 4 (alteración anatómica grave), con el grado final determinado por el componente que exhibe la anomalía más significativa, reflejando el eslabón más débil en la BAR. Para asegurar la consistencia, la clasificación de la AFS proporciona una metodología clara, recomendando una insuflación prolongada (30-45s) y maniobras de rotación en la posición retroflexionada, para evaluar hernia hiatal deslizante. Estas técnicas estandarizadas pueden reducir la variabilidad interobservador y ayudar a evitar una subestimación de la afectación de la UEG, abordando una limitación clave de sistemas previos.

Evidencia y estudios de validación

Recientes estudios prospectivos han validado la clasificación de la AFS utilizando mediciones fisiológicas objetivas¹²⁻²⁰. En un estudio con 56 pacientes con sospecha de ERGE sometidos a endoscopia, MAR y estudio de monitoreo de reflujo, la clasificación AFS mostró una fuerte correlación tanto con reflujo patológico como con alteración de la UEG manométrica²⁰.

El estudio encontró un incremento progresivo de la prevalencia de TEA patológico (> 6%) en los grados AFS: 0% en grado I, 5.9% en grado II, 52% en grado III y 77.8% en grado IV (p < 0.001). Por otro lado, la clasificación de Hill demostró poca habilidad discriminatoria en la misma población, con tasas similares de reflujo patológico en los grados II y III (42.1% vs. 37.9%, p = 0.411), destacando sus limitaciones.

En el mismo estudio, cada componente de la clasificación AFS correlacionó con parámetros específicos de MAR, reflejando diferentes mecanismos de alteración de la UEG: el componente L correlacionó con morfología de la UEG y longitud intraabdominal del EEI, el componente D mostró

	Grado 1	Grado 2	Grado 3	Grado 4
				
	1	2	3	4
Longitud axial hiatal (L)	Nula (0 cm)	Nula (0 cm)	0-2cm	>2 cm
Apertura hiatal (D)	Ajustada 1 cm	Floja 1-2 cm	Abierta 2-3 cm	Apertura amplia > 3 cm
Válvula flap (F)	Presente, labio completo con forma de omega (F+)	Labio de válvula ausente, adelgazado y aplanado (F-)	Ausente (F-)	Ausente (F-)

Figura 1 La clasificación de la *American Foregut Society* de integridad de la unión esofagogástrica.

asociación significativa con la integral contráctil de la UEG (IC-UEG) y la presión basal del EEI. El componente F demostró correlación con la maniobra de Elevación de la Pierna Recta (EPR). Estos hallazgos confirman la relevancia fisiológica de los componentes de la clasificación AFS y su habilidad para estratificar pacientes de acuerdo con la gravedad de la afectación de la UEG. Otro artículo reciente demostró la superioridad de la clasificación AFS sobre la clasificación de Hill en términos de variabilidad interobservador y confirmó su mejor capacidad de predecir TEA²¹.

Implicaciones clínicas

A pesar de estar en las etapas tempranas de validación, la clasificación AFS ofrece un avance prometedor en la evaluación endoscópica de la UEG. Su protocolo preciso y reproducible, con parámetros medibles, minimiza la interpretación subjetiva, mejorando potencialmente la selección de pacientes para futuras investigaciones y reduce pruebas de diagnóstico innecesarias en pacientes con UEG intacta.

Limitaciones e investigaciones futuras

A pesar de sus avances, la clasificación AFS tiene ciertas limitaciones que requieren consideración. Su detallada metodología de evaluación puede requerir entrenamiento adicional para los endoscopistas y se requieren más estudios de variabilidad interobservador para asegurar la consistencia entre facultativos. Además, los procedimientos recomendados, como la insuflación prolongada y las maniobras provocadoras, extienden la duración de la endoscopia de rutina, lo cual podría resultar en complicaciones en contextos clínicos con sobrecarga. Por otro lado, los estudios de validación actuales también han sido realizados solamente en centros de referencia de tercer nivel para enfermedades esofágicas, lo cual limita potencialmente su aplicación en la práctica comunitaria. Finalmente, la clasificación AFS aún no aborda la evaluación de pacientes que se han sometido a

intervenciones antirreflujo, aunque la AFS propuso recientemente una evaluación separada completa para pacientes post-CAR²².

Se requiere de más estudios en diferentes contextos clínicos para confirmar la utilidad de la clasificación en escenarios del mundo real, fuera de centros de atención terciaria, para evaluar su valor predictivo a largo plazo, para guiar la selección de intervención y para validar su habilidad de predecir la ERGE patológica y estratificar a los pacientes efectivamente.

Conclusión

La clasificación AFS es un avance prometedor en la evaluación endoscópica de la integridad de la UEG, proporcionando un marco objetivo y estandarizado para evaluar pacientes con sospecha de ERGE. Aunque se requiere de mayor validación, en particular respecto a su generalización en los diferentes contextos de la práctica y su valor pronóstico, la clasificación tiene sustancial potencial para mejorar la práctica clínica y la investigación del manejo de ERGE.

El puntaje de Milán

Introducción

La MAR juega un papel auxiliar en la ruta diagnóstica de la ERGE, siendo utilizada principalmente para excluir trastornos de motilidad mayores que simulan síntomas de ERGE y para localizar con precisión el EEI para la colocación del catéter para el monitoreo de reflujo⁸. Sin embargo, a pesar de su rol secundario, artículos de consenso enfatizan la creciente importancia de la MAR para evaluar la alteración de la BAR y en la evaluación preoperatoria para CAR^{9,23-26}.

A diferencia de los métodos tradicionales de estándar de oro (monitoreo de pH y endoscopia), los cuales se enfocan en detectar reflujo de ácido y lesiones esofágicas, la MAR ofrece una capacidad única para cuantificar anomalías

funcionales subyacentes. La reciente introducción del puntaje de Milán potencia esta capacidad, ofreciendo un solo parámetro para evaluar la alteración de la BAR y predecir ERGE de manera objetiva.

Antecedentes y limitaciones de herramientas anteriores

La exploración de anomalías manométricas correlacionadas con medidas objetivas de ERGE en pacientes con síntomas gastrointestinales superiores está bien documentada²⁷. Uno de los padres de la cirugía esofágica moderna, el Dr. DeMeester, estudió ampliamente el potencial de la manometría convencional. Demostró el impacto de las características del EEI en la fisiopatología de ERGE, incluidas la longitud total e intraabdominal del EEI y su presión basal²⁸, la contracción del cuerpo²⁹, la hernia hiatal y la respuesta de la UEG a la presión abdominal elevada^{30,31}. Estos factores han sido adaptados a métricas de MAR, incluidas la Presión Integral de EEI (PIEEI)³², IC-UEG¹⁰, Motilidad Esofágica Ineficaz (MEI)³³, morfología de UEG^{11,34-37}, gradiente de presión toracoabdominal (GPTA)³⁸ y la maniobra de EPR^{12,13}.

En 2020, Masuda et al.¹⁴ propusieron un índice manométrico que combina la morfología UEG, PIEEI y GPTA. Sin embargo, el análisis de la característica operativa del receptor (ROC) dio un área bajo la curva (AUC) de 0.615, con sensibilidad del 56% y especificidad del 60.7% en el punto de corte óptimo, lo cual indicó una precisión de diagnóstico limitada y resaltó la necesidad de un sistema de puntaje más robusto.

Descripción del puntaje

El puntaje de Milán está diseñado para cuantificar la alteración de la BAR y para estimar el riesgo de ERGE objetiva

en pacientes que se someten a MAR y estudios de pH por síntomas persistentes. Integra 4 parámetros clave de MAR:

1. MEI: definida como >70% de degluciones ineficaces o $\geq 50\%$ de degluciones fallidas, como medida de tiempo de aclaramiento esofágico.
2. IC-UEG: calculado durante el periodo de referencia utilizando la herramienta integral contráctil distal (ICD) colocada sobre la UEG y ajustada para respiración y presión gástrica. Cuantifica la fuerza y duración de la contracción de la UEG, reflejando su competencia.
3. Morfología UEG: clasificada como tipo 1 en casos de sobreposición de EEI y DC, tipo 2 en casos de separación EEI-DC < 3 cm y tipo 3 con separación ≥ 3 cm. La morfología UEG identifica la presencia y tamaño de la hernia hiatal.
4. Respuesta EPR: realizada en posición supina como se describió anteriormente¹³, con elevación de las dos piernas si el aumento de la presión intraabdominal (PIA) es insuficiente con la elevación de una sola pierna. La EPR mide la habilidad de la UEG para contrarrestar el aumento de la PIA, lo cual es un factor fisiopatológico mayor de ERGE.

La prueba de Milán utiliza una fórmula matemática para medir estos parámetros, resultando en un puntaje final. El puntaje de Milán se calcula utilizando la herramienta calculadora en línea (www.milanscore.com) (fig. 2). Un valor ≥ 137 indica un riesgo del 50% para ERGE objetiva. Este abordaje estructurado ofrece una medida reproducible de la alteración de la BAR.

Evidencia y estudios de validación

El puntaje de Milán fue desarrollado utilizando una cohorte multicentro de 295 pacientes y validada externamente en una cohorte separada de 233 pacientes³⁹. En análisis ROC mostró un AUC de 0.880 para predecir ERGE patológica,

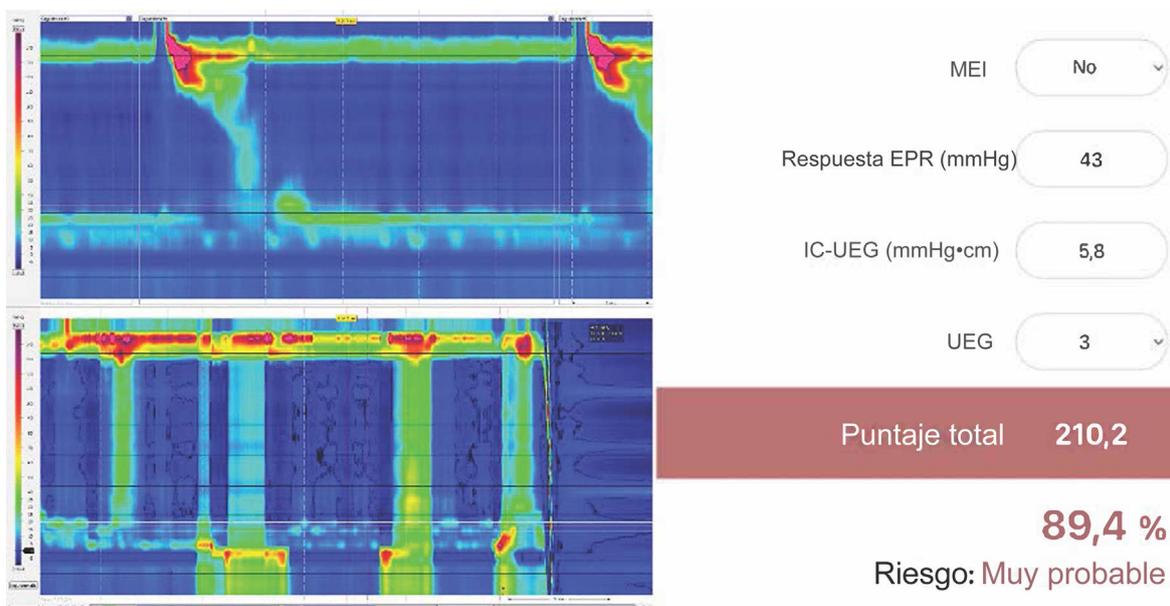


Figura 2 Ejemplo de trama de MAR y cálculo de puntaje de Milán utilizando el sitio web www.milanscore.com.

con discriminación y calibración sólidas (índice-c de Harrell corregido = 0.90, índice de calibración integrado 0.07) en la cohorte de validación. Estudios subsiguientes han explorado su aplicabilidad en diferentes escenarios clínicos. En un estudio publicado por el mismo grupo multicentro, la efectividad del puntaje de Milán fue confirmada en pacientes con morfología UEG tipo 2, el cual es un grupo difícil debido a la separación parcial del EEI-DC, logrando un AUC de 0.858 para identificar pacientes en riesgo de ERGE objetiva⁴⁰.

Otro estudio demostró su utilidad en pacientes con síntomas laringofaríngeos (SLF), con una sensibilidad del 57.1% y especificidad del 91.3%, lo cual sugiere potencial para agilizar las rutas de diagnóstico y mejorar la elección de candidatos para más pruebas⁴¹.

También, un estudio de un centro con 160 pacientes demostró la capacidad del puntaje de Milán para predecir desenlaces exitosos, posterior a CAR, ofreciendo una evaluación objetiva de la BAR postintervención. A pesar de sus limitaciones, como son su diseño retrospectivo, la falta de mediciones objetivas de desenlace y la inclusión de técnicas quirúrgicas variadas, este hallazgo destaca la promesa del puntaje de Milán en el contexto quirúrgico⁴².

Implicaciones clínicas

La fortaleza principal del puntaje de Milán radica en su habilidad para cuantificar el grado de alteración de la BAR. El carácter exhaustivo de la herramienta, integrando múltiples parámetros de MAR, la hace accesible para la mayoría de los pacientes que se someten a MAR. Al proporcionar tasas de riesgo de ERGE, también puede servir como una prueba anticipada para identificar pacientes que se podrían beneficiar de pruebas invasivas adicionales o excluir a pacientes de bajo riesgo de proceder a más evaluaciones.

Limitaciones e investigaciones futuras

A pesar de sus ventajas, el puntaje de Milán tiene algunas limitaciones que deben ser consideradas. La herramienta se basa en parámetros MAR, lo cual requiere de equipo especializado y experiencia que puede no estar disponible en todos los contextos clínicos.

La maniobra EPR, la cual es una piedra angular de la herramienta, representa algunos desafíos, ya que no todos los pacientes pueden realizarla de manera correcta, con un sesgo potencial de los resultados. La definición actual de una EPR eficaz (un incremento del 50% sobre la base en la presión intraesofágica durante la maniobra) se basa en opinión de expertos y no tiene una validación objetiva, lo cual requiere de la realización de más estudios. Además, el valor significativo de la EPR en el puntaje puede ser sobreestimado. Estudios de mayor alcance podrían refinar este balance o confirmar su poder predictivo. La valoración de los 4 componentes también se podría beneficiar de la integración con parámetros clínicos para mejorar la precisión.

Dada la naturaleza multifactorial de la ERGE y su cruce con otras condiciones, el puntaje de Milán no puede reemplazar un diagnóstico definitivo, sino que debería complementar la evaluación clínica endoscópica y los estudios de monitoreo de reflujo. También se debería explorar en futuras investigaciones el valor predictivo para desenlaces

de tratamientos quirúrgicos que validen completamente su utilidad clínica.

Conclusión

El puntaje de Milán representa un avance significativo en el diagnóstico de ERGE al ofrecer un método con fundamentos fisiológicos para estratificar el riesgo, con base en parámetros de MAR. Su evaluación exhaustiva de la función esofágica y la integridad de la BAR mejora la capacidad de identificar objetivamente ERGE, preparando el camino para un diagnóstico personalizado y estrategias terapéuticas. Aunque se requiere de mayor validación, en particular sobre su aplicación en diversos contextos clínicos y los desenlaces de tratamiento, el puntaje de Milán tiene potencial sustancial para mejorar el manejo de ERGE y permitir un diagnóstico y estrategias terapéuticas más personalizados.

El puntaje de pHoenix

Introducción

Hasta que se introdujo el puntaje de DeMeester (PDM) en 1974, el diagnóstico de ERGE se había basado en la valoración de síntomas, lo cual es un método a menudo impreciso debido a la sobreposición de pacientes y sujetos sanos^{16,43}. El PDM fue el primer intento de proporcionar una medida objetiva del TEA esofágico, estableciendo valores normativos para el diagnóstico de ERGE. Más recientemente, el consenso de Lyon 2.0 definió la ERGE objetiva con base en el TEA y proporcionó criterios para el monitoreo remoto de reflujo⁴. Sin embargo, ambos métodos tienen limitaciones, lo cual impulsó el desarrollo del puntaje de pHoenix para abordar estas limitaciones y mejorar la precisión diagnóstica⁴⁴.

Antecedentes y limitaciones de las herramientas anteriores

A lo largo de los años, el PDM se convirtió en la piedra angular del diagnóstico de ERGE, considerado el estándar de oro en la comunidad quirúrgica para seleccionar pacientes para CAR. El PDM es un puntaje compuesto que integra 6 parámetros: tiempo de reflujo en posición reclinada, en posición vertical y total, número total de episodios, número de episodios en 5 min y episodio más largo, midiendo cada uno con base en las desviaciones estándar obtenidas de controles saludables. Los parámetros con mayor variabilidad (p.ej., número de episodios, reflujo en posición vertical) contribuyen menos al puntaje, mientras que los parámetros que rara vez se encuentran en los controles (p.ej., episodios largos o reflujo en posición reclinada) tienen mayor peso. Este abordaje revolucionó el diagnóstico de ERGE al introducir métricas de TEA objetivas y ha sido correlacionada con hallazgos endoscópicos como la esofagitis y el esófago de Barrett⁴⁵⁻⁴⁸.

Sin embargo, el PDM presenta ciertas limitaciones: su cohorte de desarrollo fue pequeña, es necesario que los pacientes reporten con precisión periodos supinos y de alimentos y fue validada solamente para estudios de pH basados en catéter de 24h, lo cual lo hace susceptible de

variabilidad diaria y de carecer de umbrales para el monitoreo de reflujo prolongado.

A diferencia del PDM, el consenso de Lyon 2.0 define la ERGE patológica solamente por el TEA (>6%), introduciendo un rango límite (4-6%) que requiere de criterios adicionales y de establecer indicaciones específicas para un diagnóstico definitivo en pacientes sometidos a pruebas de monitoreo remoto de reflujo de 96 h. El consenso de Lyon 2.0 no considera otros parámetros de pH que pudieran afectar la gravedad de la ERGE, en particular el reflujo supino. Sin embargo, su umbral del 6% parece identificar ERGE de manera efectiva con excelente especificidad. Para abordar estas limitaciones, especialmente la falta de umbrales del PDM para el monitoreo prolongado y la ausencia de valoración del reflujo posicional del consenso de Lyon 2.0, es que Latorre-Rodríguez et al. desarrollaron el puntaje de pHoenix.

Descripción del puntaje

El puntaje de pHoenix es una medida compuesta derivada de una cohorte de pacientes con TEA 2-6% sometidos a monitoreo remoto de 48 h que utiliza la siguiente fórmula para medir el reflujo supino y vertical: $pHoenix = (\%TEA \text{ vertical} \times 0.991) + (\%TEA \text{ supino} \times 1.286)$.

El puntaje establece puntos de corte de 8.45 (superior) y 7.06 (inferior), definiendo un área gris para pacientes con puntajes en este rango. Su innovación radica en la asignación de un mayor peso al TEA supino, reflejando su mayor asociación con complicaciones de ERGE en comparación con el reflujo vertical.

La consideración de esta posición distingue al pHoenix del Lyon 2.0, que considera de igual manera cualquier tipo de reflujo, con base en el TEA total.

Evidencia y estudios de validación

La validación inicial del puntaje de pHoenix mostró resultados prometedores, con un AUC de 0.957 para la identificación de ERGE patológica, definida como PDM patológico⁴⁴.

La reducción en la proporción de pacientes clasificados como límite por los criterios de Lyon 2.0 del 77.2 al 13.2% ($p < 0.001$) es notoria, ofreciendo un diagnóstico claro para pacientes con TEA total no concluyente.

El puntaje demostró una buena sensibilidad y especificidad en los umbrales, con validación interna haciendo uso de sus propios componentes, confirmando su fortaleza.

Implicaciones clínicas

El puntaje de pHoenix ofrece una medida objetiva del reflujo con prioridad al TEA supino, el cual está ligado de manera más cercana a complicaciones de ERGE. A diferencia del PDM, define un área gris, lo cual es crucial, dada la naturaleza progresiva de la ERGE y establece un criterio de positividad claro para el monitoreo remoto del reflujo prolongado. En comparación con los criterios de Lyon 2.0, puede ser útil para identificar ERGE patológica en pacientes con TEA límite (4-6%).

Limitaciones e investigaciones futuras

A pesar de su potencial, el puntaje de pHoenix tiene algunas limitaciones. Su validación inicial fue realizada en un solo centro, lo cual limita su posible generalización a poblaciones diversas. Se requiere de estudios multicentro para confirmar su desempeño en contextos clínicos variados y con cohortes de mayor tamaño, lo cual también ayudaría a refinar los umbrales de diagnóstico.

El enfoque inicial del estudio sobre pacientes con TEA entre el 2 y 6% puede haber introducido un sesgo de selección, potencialmente excluyendo a pacientes con patrones de reflujo claramente normales o patológicos. También, la precisión de los periodos supinos autorreportados durante el monitoreo de 48 h, a pesar del uso de diarios electrónicos, podría afectar la confiabilidad, ya que las inconsistencias en los reportes podrían sesgar la información relacionada con la posición.

En investigaciones futuras se debe incluir comparaciones directas con otros métodos diagnósticos de ERGE, como el monitoreo de la impedancia de pH y herramientas de evaluación de síntomas, para aclarar las ventajas del puntaje de pHoenix. También son esenciales los estudios prospectivos que evalúen los desenlaces de los pacientes después de terapia médica o quirúrgica, para validar su utilidad clínica. Finalmente, dado que la herramienta fue desarrollada utilizando monitoreo de reflujo remoto, su posible aplicación para estudios de pH con catéter de 24 h requiere de validación.

Conclusión

El puntaje de pHoenix representa un avance en el diagnóstico de ERGE por su incorporación de los patrones de reflujo supino en una medida compuesta ponderada. Su énfasis en el TEA supino y la propuesta de criterios de diagnóstico para el monitoreo remoto prolongado aborda limitaciones clave del PDM y de Lyon 2.0. Si bien se necesita una mayor validación en diversos contextos, el puntaje de pHoenix tiene el potencial de mejorar el diagnóstico de ERGE.

El puntaje COuGH Reflux

Introducción

Los síntomas laringofaríngeos (SLF), incluida la tos crónica, sensación de globo y ronquera, tienen un creciente impacto en la calidad de vida y desempeño social, lo cual lleva a un creciente número de consultas clínicas en gastroenterología y ORL⁴⁹. Determinar cuáles de estos pacientes están afectados por ERGE verdadero es aún complicado, ya que el espectro de SLF incluye un amplio rango de manifestaciones clínicas no específicas que pueden ser causadas por múltiples condiciones. La incertidumbre diagnóstica a menudo lleva a terapia empírica de IBP, lo cual puede ser ineficaz cuando la causa subyacente no es ERGE⁵⁰. Para evitar la sobremedicación de IBP, se deben realizar pruebas fisiológicas esofágicas anticipadamente^{4,51}, pero no siempre están disponibles en todos los contextos clínicos. El puntaje COuGH Reflux surgió como una herramienta de predicción clínica novedosa diseñada para abordar esta debilidad diag-

nóstica por medio de la estratificación de pacientes con SLF con base en su probabilidad de presentar ERGE⁵².

Antecedentes y limitaciones de herramientas anteriores

Tradicionalmente, el diagnóstico de ERGE siempre ha sido precedido de ensayos empíricos con IBP, que, aunque son convenientes, conducen a incertidumbre diagnóstica. El ensayo con IBP carece de especificidad y podría llevar a falsos positivos debido al efecto placebo o la mejora sintomática no relacionada con la supresión ácida⁵³. Este abordaje puede resultar en medicación de largo plazo innecesaria con costos asociados y potenciales efectos secundarios⁵⁴.

Aunque se han propuesto pruebas y parámetros alternativos^{55,56}, el diagnóstico definitivo se basa en la endoscopia y en el estudio de monitoreo de reflujo, preferentemente utilizando técnicas de monitoreo remoto prolongado⁵⁷. La endoscopia tiene sensibilidad limitada con pacientes con SLF, ya que hasta el 70% de los pacientes con SLF tienen enfermedad por reflujo no erosivo⁵⁸.

El monitoreo de reflujo ambulatorio es invasivo, costoso y a menudo no está disponible en sitios de atención primaria, además de que puede no ser bien tolerado por los pacientes.

Previo al puntaje COuGH RefluX, no existía una herramienta de predicción clínica específica validada para identificar ERGE en pacientes con SLF, ya que los cuestionarios y sistemas de puntaje no demostraban suficiente precisión diagnóstica^{59,60}.

Descripción del puntaje

El puntaje COuGH RefluX es un modelo de predicción clínico que incorpora 6 parámetros fácilmente evaluables: tos, sobrepeso/obesidad, sensación de globo, hernia hiatal, regurgitación y sexo masculino⁵². A cada parámetro se le asigna un valor específico con base en su poder predictivo relativo para ERGE (tabla 1).

El cálculo del puntaje es simple, solo requiere información clínica básica que se puede obtener durante una evaluación rutinaria del paciente sin pruebas especializadas, con la posible excepción de la hernia hiatal, la cual normalmente requiere endoscopia. El puntaje total va de 0 a 6 con un umbral bajo de 2.5 y un umbral alto de 5.0 para predecir ERGE.

Tabla 1 Cálculo de puntaje COuGH RefluX

Variabes	Puntos	Resultados
Tos	1.5	Puntaje ≤ 2.5 ERGE poco probable
Obesidad/Sobrepeso		
IMC 25-30	1.5	Puntaje 3.0-4.5
IMC ≥ 30	2.0	ERGE no concluyente
Globo	-1.0	
Hernia hiatal ≥ 1 cm	1	Puntaje ≥ 5
Regurgitación	1.5	ERGE probable
Sexo masculino	1.5	

ERGE: enfermedad por reflujo gastroesofágico; IMC: índice de masa corporal.

Evidencia y estudios de validación

El estudio original desarrolló y validó el puntaje COuGH RefluX en 856 pacientes (304 entrenamiento, 552 validación). La cohorte de entrenamiento estableció un modelo con un AUC de 0.68. Un umbral de 2.5 fue sensible al 82% y 5 fue específico al 79% para ERGE confirmada. En la cohorte de validación (AUC 0.67), la sensibilidad fue del 79% a 2.5 y especificidad del 81% a 5. Un umbral a 1.5 incrementó la sensibilidad al 93% mejorando su valor predictivo negativo, mientras que su desempeño disminuyó ligeramente al excluir la hernia hiatal (AUC 0.63).

Un siguiente estudio evaluó la aplicación práctica del puntaje COuGH RefluX en 232 pacientes, clasificados como poco probable (126 pacientes), no concluyente (74 pacientes) o probable (32 pacientes) para ERGE. Los autores encontraron tasas de ERGE progresivamente más elevadas (8%, 36.5%, 65.5%; $p < 0.001$). Tras la exclusión de esofagitis erosiva y esófago de Barrett, se observó una creciente tasa de respuesta a IBP en 196 pacientes en todos los grupos (12.4%, 45%, 73.3%; $p < 0.001$). El análisis multivariado identificó al puntaje COuGH RefluX y una MNIB más baja como predictores independientes de la respuesta a IBP (puntajes ≥ 5.0 , OR = 15.772; MNIB, OR = 0.915; $p < 0.001$)⁶¹.

Implicaciones clínicas

El puntaje COuGH RefluX proporciona una estratificación estandarizada de riesgo para ERGE en pacientes con SLF. Los pacientes en la categoría «ERGE poco probable» presentan características clínicas similares a la población sana, con una tasa baja de ERGE patológica, baja probabilidad de respuesta a IBP y MNIB más alta. La aplicación de esta herramienta tiene el potencial de evitar terapia innecesaria con IBP, reduciendo la sobremedicación.

En la categoría «ERGE probable», la herramienta proporciona sustento para iniciar la terapia con IBP con mayor seguridad en casos de evaluación médica inicial o para considerar pruebas fisiológicas invasivas en casos de SLF crónicos. La correlación documentada entre puntajes más elevados y mejor respuesta a IBP permite planear un tratamiento personalizado y establecer expectativas realistas con los pacientes.

Por otro lado, el puntaje COuGH RefluX puede ser una herramienta práctica en manos de especialistas de muchas áreas, ayudando a crear puntos en común, en particular entre gastroenterólogos, otorrinolaringólogos, neumólogos, cirujanos y médicos de atención primaria, quienes a menudo manejan a este grupo de pacientes.

Limitaciones e investigaciones futuras

A pesar de su prometedora utilidad, se deben reconocer varias limitaciones del puntaje COuGH RefluX. La herramienta se basa en parte en el reporte subjetivo de síntomas, lo cual introduce potencial variabilidad y sesgo de reporte. Las herramientas de evaluación de síntomas estandarizadas podrían mejorar la confiabilidad de este componente.

La detección de hernia hiatal, uno de los parámetros de la herramienta, requiere de endoscopia, la cual podría no estar disponible en todos los contextos clínicos. Esto podría

limitar su aplicabilidad en entornos de atención primaria sin fácil acceso a servicios endoscópicos. Se encontró que la herramienta tiene eficacia moderada, incluso sin la hernia hiatal, no obstante, este es un componente importante de la herramienta. Además, la pequeña cohorte de pacientes con síntomas laríngeos aislados no tuvo el poder estadístico suficiente para el modelado de subgrupos. Sin embargo, estos pacientes no son raros en la práctica clínica y ameritan validación específica.

Por otro lado, el puntaje COUGH RefluX no reemplaza la necesidad de juicio clínico o pruebas fisiológicas. Los valores moderados de AUC indican que, aunque la herramienta es valiosa y práctica como herramienta predictiva, no es un discriminador perfecto y debería ser interpretada dentro del contexto clínico amplio.

El puntaje de Lyon

Introducción

El consenso de Lyon 2.0 actualizó la definición moderna de enfermedad por reflujo en 2024 e introdujo el concepto de ERGE accionable, el cual se refiere a situaciones en las que un alto grado de confianza en el diagnóstico de ERGE es esencial, incluida una supresión de ácido de largo plazo posterior a un ensayo exitoso de IBP, pacientes con síntomas de ERGE refractarios a IBP o con síntomas atípicos, además de pacientes que requieren intervención escalada con manejo médico o tratamiento invasivo. Los criterios de monitoreo de reflujo endoscópicos y ambulatorios fueron definidos como concluyentes, de apoyo, limítrofe y no útiles para diagnosticar ERGE. Además, el Lyon 2.0 introdujo parámetros complementarios, como el número de episodios de reflujo (ERs), MNIB o asociación de síntomas de reflujo, para definir o refutar ERGE en casos con TEA limítrofe (4-6%), describir los umbrales a utilizar con monitoreo remoto prolongado y pruebas con IBP, además de agregar la definición de ERGE objetiva grado Los Ángeles B⁶².

El puntaje de Lyon se desarrolló posteriormente para integrar hallazgos de diferentes modalidades diagnósticas (estudio de impedancia de pH [MII-pH] y endoscopia), con base en el marco del Lyon 2.0. Este novedoso sistema de puntuación busca trasladar el concepto y evidencias que sustentan el Lyon 2.0 a la práctica clínica, condensando múltiples métricas de pruebas en una herramienta simple⁶³.

Antecedentes y limitaciones de herramientas anteriores

Los abordajes convencionales al diagnóstico de ERGE se han basado de manera importante en un solo parámetro o en métricas compuestas limitadas, cada una con limitaciones significativas. El PDM, a pesar de ser ampliamente utilizado, especialmente por la comunidad de cirujanos, se enfoca principalmente en parámetros de exposición a ácido, sin considerar datos endoscópicos o parámetros de impedancia¹⁶. De igual manera, la evaluación aislada de TEA puede resultar en una cantidad sustancial de casos no concluyentes (TEA 4-6%) y una mayor dependencia en la variabilidad diaria de monitoreo de reflujo con catéter.

Las limitaciones de estos abordajes para evaluar la integridad de la mucosa, la frecuencia de episodios de reflujo y múltiples modalidades de diagnóstico resultan en la incertidumbre del diagnóstico en respuesta a la terapia antirreflujo.

El consenso de Lyon 2.0 actualizado aborda varias de estas limitaciones al establecer un marco exhaustivo para el diagnóstico de ERGE que incorpora múltiples parámetros diagnósticos. El puntaje de Lyon pretende llenar el vacío entre criterios de diagnóstico sofisticados y una herramienta práctica para el uso clínico.

Descripción de la herramienta

El puntaje de Lyon incorpora múltiples parámetros de la endoscopia y el monitoreo de pH en un solo sistema exhaustivo de puntuación.

El puntaje de Lyon incluye dos parámetros endoscópicos y cuatro de MII-pH:

1. Esofagitis, clasificada de acuerdo con la clasificación de Los Ángeles como: grado A, una o más erosiones de mucosa < 5 mm; grado B, erosiones > 5 mm; grado C, roturas de la mucosa con cobertura de menos del 75% de la circunferencia esofágica; y grado D, \geq 75% del esófago⁶⁴.
2. Hernia hiatal, detectada con endoscopia o MAR.
3. TEA, ponderada con base en umbrales establecidos en el consenso de Lyon 2.0, con puntajes mayores para mayor exposición al ácido.
4. MNIB, una medida de integridad de mucosa. Los valores menores indican una mucosa esofágica comprometida, calculada manualmente como la media de tres mediciones nocturnas de 10 min cada una⁶⁵.
5. Número de Episodios de Reflujo, cuantificando la frecuencia de eventos de reflujo, identificados por medio de software y después verificados manualmente según el consenso de Wingate⁶⁶.
6. Asociación de Reflujo y síntomas, definida como un índice de síntomas > 50% o probabilidad de asociación de síntomas > 95%⁶⁷.

Cada parámetro tiene asignado un peso específico basado en su importancia relativa para el diagnóstico de ERGE y su correlación con desenlaces de tratamiento, agregando los valores ponderados para crear un puntaje compuesto (tabla 2).

Además, el puntaje de Lyon proporciona umbrales claros para categorizar pacientes en fenotipos diagnósticos: acidez funcional, defecto de permeabilidad, hipersensibilidad de reflujo, ERGE no concluyente, ERGE concluyente y ERGE grave.

Evidencia y estudios de validación

En el estudio de validación publicado por Gyawali et al. en 2024, la cohorte de desarrollo consistió en 281 pacientes con síntomas de ERGE, de los cuales 264 estaban en terapia con IBP y 18 se habían sometido a CAR. La mediana del puntaje de Lyon fue 6.5 (IQR 2.5-10.5). Un buen desenlace fue definido como un 50% de mejoría de síntomas utilizando una escala visual análoga (EVA) con un rango de 0 (sin sín-

Tabla 2 Cálculo y fenotipos del puntaje de Lyon

	Esofagitis	TEA	Episodios de reflujo	MNIB	Asociación de reflujo y síntomas	Hernia hiatal
Normal	Ninguno	<4.0%	<40	>2500	No	No
Anormal	LA-A	4-6%	>40	1500-2499	Sí	Sí
Concluyente	LA-B	>6.0%	NA	NA	NA	NA
Grave	LA-CLA-D	>10%	NA	NA	NA	NA
Fenotipo	Sin ERGE	Defecto de permeabilidad aislado	Hipersensibilidad a reflujo	ERGE límite	ERGE concluyente	ERGE grave
Puntaje de Lyon	0-0.5	0.5-2.0	1.0-3.0	≥ 3.0	≥ 5.0	≥ 10.0

ERGE: enfermedad por reflujo gastroesofágico; LA: clasificación de Los Ángeles; MNIB: media nocturna de impedancia basal; NA: no aplica; TEA: tiempo de exposición a ácido.

tomas a 100 (síntomas graves)⁶⁸. Las tasas de respuesta crecieron con evidencia de ERGE: del 7.1% en el fenotipo de acidez funcional, al 37.5% en ERGE límite, a $\geq 76.6\%$ en ERGE concluyente ($p < 0.001$). La AUC para predecir buenos desenlaces fue de 0.819. El puntaje de Lyon tuvo un mejor desempeño que el PDM (AUC 0.779, $p = 0.019$) para predecir desenlaces posteriores al tratamiento para ERGE, con un umbral óptimo de 6.25 (sensibilidad del 81.2%, especificidad del 73.4%).

Dos cohortes de validación diferentes fueron evaluadas. La europea (215 pacientes de dos centros en Italia) era dominante en casos de acidez, mientras que la asiática (258 pacientes de Taiwán) era dominante en pacientes con regurgitación y consistió en pacientes previamente no tratados. La media de los puntajes de Lyon fue 2.5 en la cohorte europea y 4.3 en la cohorte asiática, con AUC de 0.908 y 0.637, respectivamente ($p < 0.001$), para predecir respuesta al tratamiento. El puntaje de Lyon tuvo consistentemente un mejor desempeño que componentes individuales del PDM, especialmente en la cohorte asiática con predominancia en regurgitación. Con el umbral a 2.75 (sensibilidad 95%), identificó respuesta de síntomas en el 85.6% de la cohorte europea y el 77.9% en la cohorte asiática; con el umbral a 11.5 (especificidad 95%), solo el 14.4% de los pacientes de la cohorte europea y el 12.3% de la cohorte asiática no presentaron respuesta de síntomas.

Implicaciones clínicas

El puntaje de Lyon tiene el potencial de ofrecer ventajas clínicas significativas para el diagnóstico de ERGE. Integrando y ponderando métricas endoscópicas y de MII-pH, proporciona una herramienta exhaustiva para estratificar pacientes con síntomas de ERGE en fenotipos distintivos, que van de la acidez funcional a la ERGE grave. La capacidad del puntaje de Lyon de predecir la respuesta de síntomas a la terapia antirreflujo, con un rango de 7.1 en el fenotipo de acidez funcional a $\geq 75\%$ en ERGE concluyente, podría disminuir tratamientos innecesarios y mejorar los desenlaces terapéuticos en general, permitiendo tomar decisiones de tratamiento de manera personalizada. Los pacientes con puntajes bajos se pueden beneficiar de tratamientos conservadores, como cambios de estilo de vida o terapia médica mínima, mientras que los pacientes con puntajes altos podrían ser priorizados para opciones más invasivas, como CAR. La dependencia del puntaje en pruebas diagnósticas fácilmente accesibles y el potencial para automatizar mejoran su aplicación práctica, ofreciendo una herramienta de uso fácil, reproducible y de fácil acceso. En caso de ser validado de manera prospectiva, el puntaje de Lyon podría hacer el diagnóstico de ERGE más eficiente, ofreciendo una mejor selección de candidatos para escalar el tratamiento para ERGE.

Limitaciones e investigaciones futuras

A pesar de su prometedor desempeño, algunas limitaciones del puntaje de Lyon deben ser mencionadas. El diseño retrospectivo de los estudios de desarrollo introduce sesgos potenciales relacionados con la recolección de datos y la selección de pacientes. La validación prospectiva en con-

textos clínicos diversos es aún necesaria para confirmar la fortaleza y generalización de la herramienta.

El estudio antes mencionado se enfocó en pacientes con síntomas de ERGE típicos que estaban en terapia antisecretoria, por lo que la validación en presentaciones atípicas y cohortes en IBP es necesaria. La variabilidad de perfiles sintomatológicos en las cohortes de desarrollo y validación, como la predominancia de acidez funcional en la cohorte europea versus regurgitación en la cohorte asiática, sugiere el potencial de influencias étnicas o geográficas en el desempeño. Además, la dependencia del puntaje en métricas de impedancia de pH trae la cuestión de su aplicabilidad en contextos en los que se prefiere el monitoreo de pH remoto. Las investigaciones futuras deben priorizar ensayos multicentro prospectivos, para validar de manera externa la herramienta en poblaciones y fenotipos sintomatológicos diversos. Estudios que se enfoquen en cohortes terapéuticas específicas, comparativos de abordaje médico vs. quirúrgico e integraciones con otras herramientas (p.ej., puntaje de Milán o puntaje COuGH RefluX) podrían ofrecer una validación más robusta y mejorar su utilidad clínica.

Discusión

El diagnóstico y manejo de ERGE ha evolucionado significativamente en las décadas pasadas, impulsado por los avances en las tecnologías de diagnóstico y una mayor comprensión de su fisiopatología. Esta revisión destaca novedosas herramientas como la clasificación AFS, el puntaje de Milán, el puntaje de pHoenix, el puntaje COuGH RefluX y el puntaje de Lyon, las cuales buscan mejorar la precisión, objetividad y relevancia clínica de la evaluación para ERGE. Estas herramientas abordan dificultades para el diagnóstico de ERGE, como la subjetividad de métodos previos, limitaciones de evaluaciones con un solo parámetro y la necesidad de estratificación del tratamiento. En la [tabla 3](#) se muestra un resumen de la modalidad y desempeño de las herramientas mencionadas.

Todas las herramientas mencionadas arriba comparten la intención de proporcionar una evaluación más objetiva. Históricamente, el diagnóstico de ERGE se ha basado de manera importante en el reporte de síntomas y ensayos empíricos con IBP, pero este abordaje ha fracasado, dado el cruce significativo de síntomas con otras condiciones y efectos placebo. La introducción del PDM en 1974 marcó un cambio fundamental hacia la medición objetiva, al cuantificar el TEA y otros parámetros de reflujo¹⁶. Varias décadas después, el consenso de Lyon estableció puntos de comparación para TEA e incorporó métricas complementarias, como la MNIB y la asociación reflujo-síntoma⁴. Sin embargo, incluso cuando el umbral de TEA (6%) identifica ERGE de manera correcta, no consideró otros parámetros de monitoreo de reflujo (p.ej., reflujo supino) y careció de una herramienta simple y práctica para convertir sus conceptos a la práctica clínica. Para superar estas limitaciones, los puntajes de pHoenix y de Lyon han sido introducidos recientemente.

Debido a que la evaluación clínica, la endoscopia y la MAR son una parte integral de la evaluación para ERGE y las métricas tradicionales sufrían de subjetividad y carecían de precisión diagnóstica, diferentes grupos de investigación propusieron el puntaje COuGH RefluX, la clasificación AFS

y el puntaje de Milán, con la finalidad de estratificar de manera objetiva los SLF para probabilidad de ERGE, estandarizar la evaluación endoscópica de la UEG y cuantificar la afectación de la BAR^{19,39}.

En conjunto, estas herramientas reflejan el cambio de paradigma hacia una medicina de precisión en ERGE, con la finalidad de reducir la ambigüedad diagnóstica y optimizar la toma de decisiones terapéuticas.

Cada herramienta proporciona fortalezas específicas al panorama diagnóstico. La clasificación AFS mejora la clasificación Hill al introducir parámetros medibles (longitud, diámetro e integridad de válvula «flap») y un protocolo endoscópico estandarizado, reduciendo la subjetividad y mejorando la reproducibilidad. Su correlación con medidas fisiológicas, como parámetros de TEA y MAR, destaca su potencial para identificar pacientes con alteración significativa de la UEG, quienes se podrían beneficiar de una terapia de mayor impacto. De manera similar, el puntaje de Milán mejora la capacidad de la MAR para evaluar defectos funcionales de la BAR, ofreciendo una sola métrica predictora con excelente precisión diagnóstica para ERGE objetiva. Su capacidad para guiar la evaluación preoperatoria para CAR destaca su utilidad clínica en casos complejos.

El puntaje de pHoenix aborda limitaciones en el PDM y del Lyon 2.0 al hacer énfasis en el TEA supino, el cual es un factor de riesgo conocido para complicaciones de ERGE y establece umbrales para el monitoreo remoto prolongado. Su capacidad de reclasificar casos de TEA limítrofe (4-6%) en categorías diagnósticas más claras podría hacer más eficientes las decisiones de manejo terapéutico. El puntaje COuGH RefluX cubre una debilidad crítica al evaluar SLF, un conjunto de síntomas notablemente complicado, al ofrecer una herramienta simple con bases clínicas, con alta sensibilidad y especificidad a umbrales definidos. Finalmente, el puntaje de Lyon se destaca por su abordaje integrador, combinando datos endoscópicos y de impedancia de pH para estratificar pacientes en fenotipos accionables, con potencia predictiva demostrada para desenlaces de tratamiento (AUC 0.819).

A pesar de su promesa, estas herramientas tienen ciertas limitaciones que mitigan su adopción inmediata en la práctica de rutina, como su validación en centros de atención terciaria, lo cual es un hecho que plantea preguntas sobre su aplicación en la práctica comunitaria.

Otro hilo conductor común de estas herramientas es su etapa de desarrollo. La mayoría de ellas han sido validadas en contextos controlados o especializados, con poblaciones monocéntricas o retrospectivas y su desempeño en poblaciones diversas aún no ha sido explorado. Además, la naturaleza multifactorial de ERGE implica que no hay una sola herramienta que pueda capturar su complejidad, lo cual señala la necesidad de su uso complementario con el juicio clínico y diagnósticos establecidos.

A pesar de estas limitaciones, el impacto potencial de estas herramientas en la práctica clínica es sustancial, ya que ofrecen un abordaje más sofisticado y personalizado al diagnóstico de ERGE. Algunas pruebas diagnósticas no son fácilmente disponibles en varios países, lo cual hace que la evaluación clínica y la evaluación endoscópica de primera línea sean cruciales para la estratificación de los pacientes⁶⁹. Bajo esta perspectiva, el puntaje COuGH RefluX y la clasificación AFS durante la endoscopia podrían ayudar a seleccionar pacientes que presenten alto riesgo de ERGE y

Tabla 3 Resumen del contexto, modalidad y desempeño de las herramientas

Herramienta	Contexto	Aplicación	Modalidad	Umbral	Desempeño
Clasificación AFS	Sospecha de ERGE	Evaluación endoscópica de integridad de UEG	Endoscopia	Patológico: tipo III-IV AFS	Correlación con TEA patológico (> 6%): 0% (grado I) 5.9% (grado II) 52% (grado III) 77.8% (grado IV) AUC 0.880 Sensibilidad 80.4% Especificidad 85.2%
Puntaje de Milán	Síntomas de ERGE típico y atípico Evaluación postoperatoria	Evaluación de alteración de BAR	MAR	137	AUC 0.957 para predecir DeMeester patológico
Puntaje de pHoenix	Síntomas de ERGE	Diagnóstico de ERGE en casos de TEA limítrofe (2-6%) con monitoreo prolongado	Monitoreo remoto de pH	Umbral bajo: 7.06 Umbral alto: 8.45	AUC 0.957 para predecir DeMeester patológico
Puntaje COuGH RefluX	Síntomas de ERGE atípicos (SLF)	Estratificación de riesgo para ERGE en pacientes con síntomas laringofaríngeos	Cuestionario	Bajo: ≤ 2.5 Intermedio: 3-4.5 Alto: ≥ 5	AUC 0.67 Sensibilidad 79% Especificidad 81%
Puntaje de Lyon	Síntomas típicos de ERGE	Diagnóstico y fenotipos de ERGE, prediciendo desenlaces de tratamiento	Multi-modal (endoscopia + MII-pH)	6.25	AUC 0.819 Sensibilidad 81.2% Especificidad 73.4%

AFS: *American Foregut Society*; BAR: barrera antirreflujo; AUC: área bajo la curva; UEG: unión esofagogástrica; ERGE: enfermedad por reflujo gastroesofágico; MAR: manometría de alta resolución; MII-pH: Impedancia de pH intraluminal multicanal; SLF: síntomas laringofaríngeos.

alteración profunda de UEG para pruebas fisiopatológicas y reducir pruebas innecesarias en pacientes con menor riesgo o barreras intactas. El puntaje COuGH RefluX destaca como una herramienta particularmente adecuada para contextos con recursos limitados debido a su sustento en parámetros clínicos que pueden ser evaluados durante evaluaciones de rutina, con la excepción de la detección de la hernia hiatal, la cual puede requerir endoscopia. En contextos de atención primaria o en áreas con bajos recursos donde la endoscopia no está disponible, el desempeño de la herramienta sin hernia hiatal es aún moderado, lo cual sugiere su utilidad potencial como instrumento de triaje. Sin embargo, su dependencia en el reporte subjetivo de síntomas podría introducir variabilidad, requiriendo de cuestionarios estandarizados para mejorar la confiabilidad. Por otro lado, la clasificación AFS ofrece una evaluación endoscópica estandarizada de la integridad de la UEG pero representa algunos desafíos en la práctica cotidiana debido a la necesidad de especialistas capacitados. Para mejorar la disponibilidad, se podrían explorar protocolos simplificados de capacitación o evaluaciones endoscópicas guiadas por telemedicina.

Es bien sabido que la respuesta al IBP y los desenlaces positivos de CAR dependen en gran medida de un diagnóstico objetivo de ERGE^{70,71}. El enfoque del puntaje de pHoenix en el reflujo supino podría refinar la escalada del tratamiento, particularmente en pacientes con síntomas nocturnos o TEA limítrofe, mientras que la naturaleza exhaustiva y los fenotipos del puntaje de Lyon podrían ofrecer evidencia objetiva

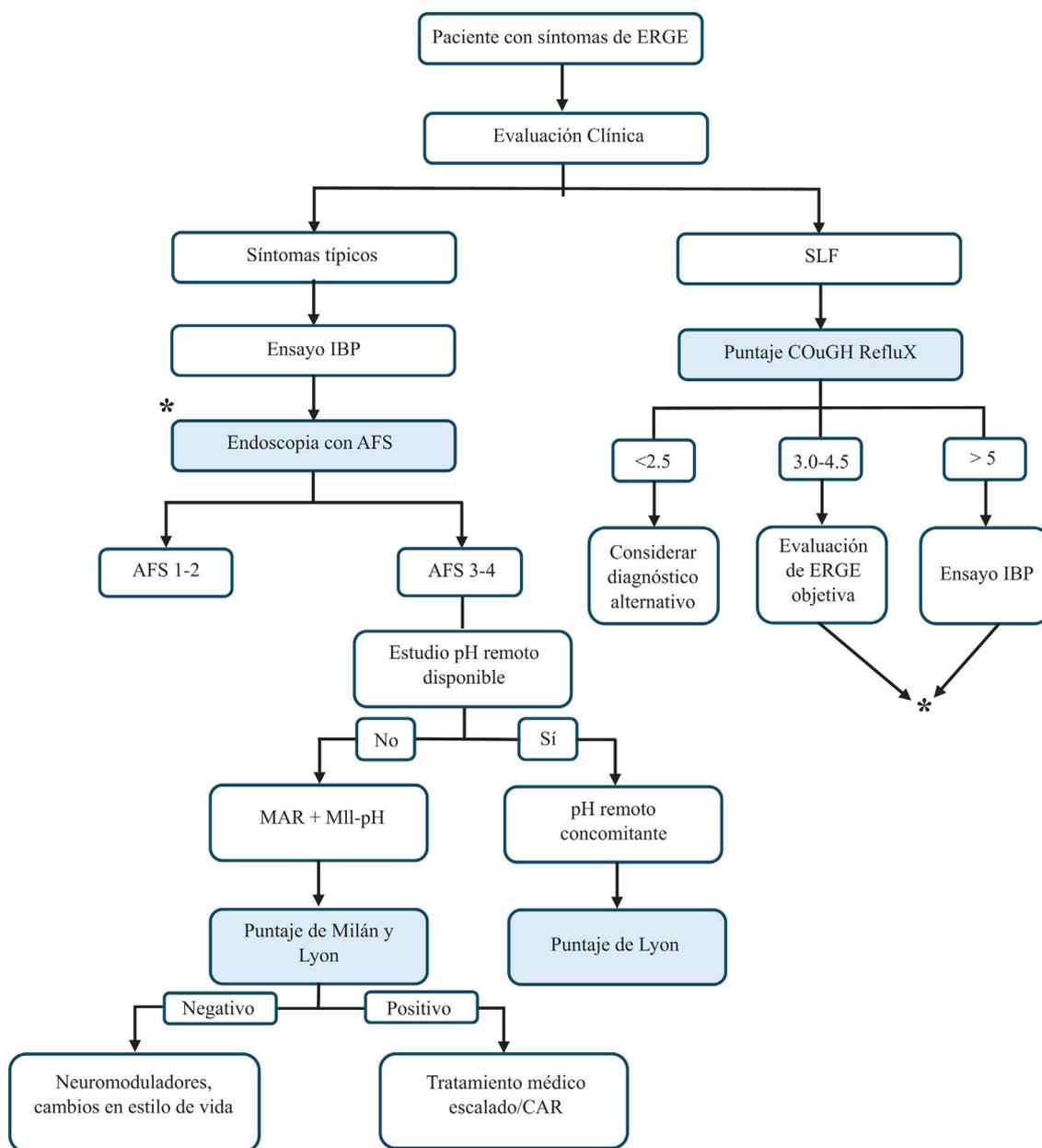


Figura 3 Propuesta de integración de estas novedosas herramientas a la vía de diagnóstico para ERGE. AFS: *American Foregut Society*; CAR: cirugía antirreflujo; ERGE: enfermedad por reflujo gastroesofágico; IBP: inhibidor de bomba de protones; MAR: manometría de alta resolución; MII-pH: impedancia de pH intraluminal multicanal; SLF: síntomas laringofaríngeos.

para seleccionar pacientes para cirugía (ERGE demostrada y grave) u opciones terapéuticas menos invasivas (neuromoduladores, IBP).

Finalmente, la capacidad del puntaje de Milán para cuantificar la alteración de la BAR e identificar casos de ERGE grave lo hacen una prueba particularmente apropiada y útil previo a la CAR.

Las investigaciones futuras se deberán enfocar en validaciones prospectivas multicentro para establecer la fortaleza de estas herramientas en poblaciones diversas y escenarios clínicos. La integración de parámetros clínicos (p.ej., gravedad de síntomas, IMC) con métricas fisiológicas podría mejorar la precisión diagnóstica, como se observa en el uso de sexo y obesidad en el puntaje COUGH RefluX. También, el desarrollo de estudios comparativos que pongan a prueba estas herramientas unas contra otras, o en combinación, será clarificador de sus ventajas relativas y desempeño combinado.

Con base en el desempeño y las características de cada herramienta, proponemos un algoritmo diagnóstico para integrar estas nuevas herramientas en la ruta diagnóstica para ERGE, como se observa en la [tabla 3](#). Los estudios de desenlace a largo plazo son críticos para validar estos puntajes como herramientas útiles para predecir la respuesta al tratamiento de terapia médica o CAR.

Conclusiones

La clasificación AFS, el puntaje de Milán, el puntaje de pHoenix, el puntaje COUGH RefluX y el puntaje de Lyon representan avances significativos en el diagnóstico de ERGE, ofreciendo herramientas objetivas y estandarizadas para mejorar la precisión y personalizar el tratamiento. La clasificación AFS y el puntaje COUGH RefluX agilizan las evaluaciones endoscópicas y clínicas, respectivamente, ayudando con la selección de pacientes en la atención primaria. El puntaje de Milán y el puntaje de pHoenix refinan las evaluaciones preoperatorias y de TEA, mientras que el puntaje de Lyon integra datos multimodales para predecir desenlaces de tratamiento. La ruta diagnóstica propuesta ([fig. 3](#)) integra estas herramientas para guiar a los practicantes clínicos desde la evaluación inicial de síntomas a las pruebas de confirmación y escalada de tratamiento, asegurando un manejo de ERGE más eficiente y personalizado. Conforme la validación avance y las barreras prácticas sean abordadas, estas innovaciones tendrán el potencial del manejo de ERGE, mejorando la precisión diagnóstica, reduciendo el sobretratamiento y optimizando los desenlaces para los pacientes.

Financiación

No se recibió patrocinio de ningún tipo para llevar a cabo este artículo.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Referencias

1. Kahrilas PJ, Jonsson A, Denison H, et al. Impact of regurgitation on health-related quality of life in gastro-oesophageal reflux disease before and after short-term potent acid suppression therapy. *Gut*. 2014;63:720–6, <http://dx.doi.org/10.1136/gutjnl-2013-304883>.
2. Calabrese F, Pasta A, Bodini G, et al. Applying Lyon consensus criteria in the work-up of patients with extra-oesophageal symptoms - A multicentre retrospective study. *Aliment Pharmacol Ther*. 2024;59:1134–43, <http://dx.doi.org/10.1111/apt.17934>.
3. Katz PO, Dunbar KB, Schnoll-Sussman FH, et al. ACG clinical guideline for the diagnosis and management of gastroesophageal reflux disease. *Am J Gastroenterol*. 2022;117:27–56, <http://dx.doi.org/10.14309/ajg.0000000000001538>.
4. Gyawali CP, Yadlapati R, Fass R, et al. Updates to the modern diagnosis of GERD: Lyon consensus 2.0. *Gut*. 2024;73:361–71, <http://dx.doi.org/10.1136/gutjnl-2023-330616>.
5. Hill LD, Kozarek RA, Kraemer SJ, et al. The gastroesophageal flap valve: in vitro and in vivo observations. *Gastrointest Endosc*. 1996;44:541–7, [http://dx.doi.org/10.1016/s0016-5107\(96\)70006-8](http://dx.doi.org/10.1016/s0016-5107(96)70006-8).
6. Fass R. Epidemiology and pathophysiology of symptomatic gastroesophageal reflux disease. *Am J Gastroenterol*. 2003;98:S2–7, [http://dx.doi.org/10.1016/s0002-9270\(03\)00009-1](http://dx.doi.org/10.1016/s0002-9270(03)00009-1).
7. Patel A, Posner S, Gyawali CP. Esophageal high-resolution manometry in gastroesophageal reflux disease. *JAMA*. 2018;320:1279–80, <http://dx.doi.org/10.1001/jama.2018.8694>.
8. Savarino E, de Bortoli N, Bellini M, et al. Practice guidelines on the use of esophageal manometry - A GISMA-D-SIGE-AIGO medical position statement. *Dig Liver Dis*. 2016;48:1124–35, <http://dx.doi.org/10.1016/j.dld.2016.06.021>.
9. Gyawali CP, Roman S, Bredenoord AJ, et al. Classification of esophageal motor findings in gastro-esophageal reflux disease: Conclusions from an international consensus group. *Neurogastroenterol Motil*. 2017;29:e13104, <http://dx.doi.org/10.1111/nmo.13104>.
10. Nicodème F, Pipa-Muniz M, Khanna K, et al. Quantifying esophagogastric junction contractility with a novel HRM topographic metric, the EGJ-Contractile Integral: normative values and preliminary evaluation in PPI non-responders. *Neurogastroenterol Motil*. 2014;26:353–60, <http://dx.doi.org/10.1111/nmo.12267>.
11. Tolone S, De Bortoli N, Marabotto E, et al. Esophagogastric junction contractility for clinical assessment in patients with GERD: a real added value? *Neurogastroenterol Motil*. 2015;27:1423–31, <http://dx.doi.org/10.1111/nmo.12638>.
12. Rogers BD, Rengarajan A, Ali IA, et al. Straight leg raise metrics on high-resolution manometry associate with esophageal reflux burden. *Neurogastroenterol Motil*. 2020;32:e13929, <http://dx.doi.org/10.1111/nmo.13929>.
13. Siboni S, Kristo I, Rogers BD, et al. Improving the diagnostic yield of high-resolution esophageal manometry for GERD: the “Straight Leg-Raise” international study. *Clin Gastroenterol Hepatol*. 2023;21:1761–70, <http://dx.doi.org/10.1016/j.cgh.2022.10.008>, e1.
14. Masuda T, Mittal SK, Kovacs B, et al. Simple manometric index for comprehensive esophagogastric junction barrier competency against gastroesophageal reflux. *J Am Coll Surg*. 2020;230:744–55, <http://dx.doi.org/10.1016/j.jamcollsurg.2020.01.034>, e3.
15. Akimoto S, Singhal S, Masuda T, et al. Esophagogastric junction morphology and distal esophageal acid exposure. *Dig Dis Sci*. 2016;61:3537–44, <http://dx.doi.org/10.1007/s10620-016-4331-y>.
16. Johnson LF, Demeester TR. Twenty-four-hour pH monitoring of the distal esophagus. A quantitative measure of gastroe-

- sophageal reflux. *Am J Gastroenterol.* 1974;62:325–32. PMID: 4432845.
17. Visaggi P, Mariani L, Svizzero FB, et al. Clinical use of mean nocturnal baseline impedance and post-reflux swallow-induced peristaltic wave index for the diagnosis of gastro-esophageal reflux disease. *Esophagus.* 2022;19:525–34, <http://dx.doi.org/10.1007/s10388-022-00933-6>.
 18. Kahrilas PJ, Shaheen NJ, Vaezi MF, et al. American Gastroenterological Association Medical Position Statement on the management of gastroesophageal reflux disease. *Gastroenterology.* 2008;135:1383–91, <http://dx.doi.org/10.1053/j.gastro.2008.08.045>, e5.
 19. Nguyen NT, Thosani NC, Canto MI, et al. The American Foregut Society white paper on the endoscopic classification of esophagogastric junction integrity. *Foregut.* 2022;2:339–48, <http://dx.doi.org/10.1177/26345161221126961>.
 20. Siboni S, Sozzi M, Kersik A, et al. Does the American Foregut Society endoscopic classification of esophagogastric junction integrity predict acid reflux and esophagogastric junction disruption? *Foregut.* 2024, <http://dx.doi.org/10.1177/26345161241230940>.
 21. Swei E, Almario JA, Abou-Diwan E, et al. American Foregut classification is superior to Hill classification in evaluation of the gastroesophageal junction. *Foregut.* 2024;4:404–10, <http://dx.doi.org/10.1177/26345161241293522>.
 22. Nguyen NT, Bell R, Abu-Dayyeh B, et al. American Foregut Society white paper on the endoscopic assessment of the gastroesophageal valve after anti-reflux surgery. *Foregut.* 2024, <http://dx.doi.org/10.1177/26345161241300480>.
 23. Pauwels A, Boecxstaens V, Andrews CN, et al. How to select patients for antireflux surgery? The ICARUS guidelines (international consensus regarding preoperative examinations and clinical characteristics assessment to select adult patients for antireflux surgery). *Gut.* 2019;68:1928–41, <http://dx.doi.org/10.1136/gutjnl-2019-318260>.
 24. Jobe BA, Richter JE, Hoppo T, et al. Preoperative diagnostic workup before antireflux surgery: an evidence and experience-based consensus of the Esophageal Diagnostic Advisory Panel. *J Am Coll Surg.* 2013;217:586–97, <http://dx.doi.org/10.1016/j.jamcollsurg.2013.05.023>.
 25. Visaggi P, Ghisa M, Barberio B, et al. Gastro-esophageal diagnostic workup before bariatric surgery or endoscopic treatment for obesity: position statement of the International Society of Diseases of the Esophagus. *Dis Esophagus.* 2024;37:doae006, <http://dx.doi.org/10.1093/dote/doae006>.
 26. Salvador R, Pandolfino JE, Costantini M, et al. The role of high-resolution manometry before and following antireflux surgery: the Padova consensus. *Ann Surg.* 2025;281:124–35, <http://dx.doi.org/10.1097/SLA.0000000000006297>.
 27. Siboni S, Bonavina L, Rogers BD, et al. Effect of increased intra-abdominal pressure on the esophagogastric junction: a systematic review. *J Clin Gastroenterol.* 2022;56:821–30, <http://dx.doi.org/10.1097/MCG.0000000000001756>.
 28. DeMeester TR, Wernly JA, Bryant GH, et al. Clinical and in vitro analysis of determinants of gastroesophageal competence. A study of the principles of antireflux surgery. *Am J Surg.* 1979;137:39–46, [http://dx.doi.org/10.1016/0002-9610\(79\)90008-4](http://dx.doi.org/10.1016/0002-9610(79)90008-4).
 29. Joelsson BE, DeMeester TR, Skinner DB, et al. The role of the esophageal body in the antireflux mechanism. *Surgery.* 1982;92:417–24. PMID: 7101132.
 30. DeMeester TR, Lafontaine E, Joelsson BE, et al. Relationship of a hiatal hernia to the function of the body of the esophagus and the gastroesophageal junction. *J Thorac Cardiovasc Surg.* 1981;82:547–58. PMID: 7278346.
 31. Wernly JA, DeMeester TR, Bryant GH, et al. Intra-abdominal pressure and manometric data of the distal esophageal sphincter Their relationship to gastroesophageal reflux. *Arch Surg.* 1980;115:534–9, <http://dx.doi.org/10.1001/archsurg.1980.01380040156028>.
 32. Hoshino M, Sundaram A, Mittal SK. Role of the lower esophageal sphincter on acid exposure revisited with high-resolution manometry. *J Am Coll Surg.* 2011;213:743–50, <http://dx.doi.org/10.1016/j.jamcollsurg.2011.09.002>.
 33. Gyawali CP, Zerbib F, Bhatia S, et al. Chicago Classification update (V4.0): technical review on diagnostic criteria for ineffective esophageal motility and absent contractility. *Neurogastroenterol Motil.* 2021;33:e14134, <http://dx.doi.org/10.1111/nmo.14134>.
 34. Yadlapati R, Kahrilas PJ, Fox MR, et al. Esophageal motility disorders on high-resolution manometry: Chicago Classification version 4.0©. *Neurogastroenterol Motil.* 2021;33:e14058, <http://dx.doi.org/10.1111/nmo.14058>.
 35. Tolone S, de Cassan C, de Bortoli N, et al. Esophagogastric junction morphology is associated with a positive impedance-pH monitoring in patients with GERD. *Neurogastroenterol Motil.* 2015;27:1175–82, <http://dx.doi.org/10.1111/nmo.12606>.
 36. Luther J, Zarro S, Sagaram M, et al. Intermittent hiatus hernia on high-resolution manometry associates with abnormal reflux burden similar to persistent hiatus hernia. *Am J Gastroenterol.* 2023;118:2071–4, <http://dx.doi.org/10.14309/ajg.0000000000002353>.
 37. Akimoto S, Singhal S, Masuda T, et al. Classification for esophagogastric junction (EGJ) complex based on physiology. *Dis Esophagus.* 2017;30:1–6, <http://dx.doi.org/10.1093/dote/dox048>.
 38. Masuda T, Mittal SK, Kovacs B, et al. Thoracoabdominal pressure gradient and gastroesophageal reflux: insights from lung transplant candidates. *Dis Esophagus.* 2018;31, <http://dx.doi.org/10.1093/dote/doy025>.
 39. Siboni S, Sozzi M, Kristo I, et al. The Milan score: a novel manometric tool for a more efficient diagnosis of gastroesophageal reflux disease. *United European Gastroenterol J.* 2024;12:552–61, <http://dx.doi.org/10.1002/ueg2.12565>.
 40. Ferrari D, Siboni S, Sozzi M, et al. The Milan Score predicts objective gastroesophageal reflux disease in patients with type 2 esophagogastric junction. *Neurogastroenterol Motil.* 2025:e14987, <http://dx.doi.org/10.1111/nmo.14987>.
 41. Siboni S, Sozzi M, Visaggi P, et al. The Milan Score is an effective manometric tool to predict gastroesophageal reflux in patients with laryngopharyngeal symptoms. *Neurogastroenterol Motil.* 2025:e70015, <http://dx.doi.org/10.1111/nmo.70015>.
 42. Sozzi M, Siboni S, Cusmai L, et al. A new manometric tool to evaluate the efficacy of laparoscopic anti-reflux surgery: first postoperative application of the Milan score. *Ann Esophagus.* 2024;7:25, <http://dx.doi.org/10.21037/aoe-24-32>.
 43. Johnson LF, DeMeester TR. Development of the 24-hour intraesophageal pH monitoring composite scoring system. *J Clin Gastroenterol.* 1986;8:52–8, <http://dx.doi.org/10.1097/00004836-198606001-00008>.
 44. Latorre-Rodríguez AR, Mittal SK, Simmonds H, et al. pHoenix score: development and validation of a novel approach to decrease the number of inconclusive GERD diagnoses. *Surg Endosc.* 2024;38:6880–93, <http://dx.doi.org/10.1007/s00464-024-11105-1>.
 45. Schwameis K, Lin B, Roman J, et al. Is pH testing necessary before antireflux surgery in patients with endoscopic erosive esophagitis? *J Gastrointest Surg.* 2018;22:8–12, <http://dx.doi.org/10.1007/s11605-017-3554-3>.
 46. Xu D-T, Feng G-J, Zhao L-L, et al. Relationship among symptoms, mucosal injury, and acid exposure in gastroesophageal reflux disease. *Chin Med J (Engl).* 2013;126:4430–4. PMID: 24286401.
 47. Schlottmann F, Andolfi C, Herbella FA, et al. GERD: presence and size of hiatal hernia influence clinical presentation, esophageal function, reflux profile, and degree of mucosal injury. *Am Surg.* 2018;84:978–82. PMID: 29981634.

48. Neto RML, Herbella FAM, Schlottmann F, et al. Does DeMeester score still define GERD? *Dis Esophagus*. 2019;32:doy118, <http://dx.doi.org/10.1093/dote/doy118>.
49. Lechien J-R, Mouawad F, Bobin F, et al. Review of management of laryngopharyngeal reflux disease. *Eur Ann Otorhinolaryngol Head Neck Dis*. 2021;138:257–67, <http://dx.doi.org/10.1016/j.anorl.2020.11.002>.
50. Kahrilas P, Anastasiou F, Bredenoord AJ, et al. Proton pump inhibitors: rational use and use-reduction. *The Windsor Workshop. Dig Dis*. 2024;42:211–20, <http://dx.doi.org/10.1159/000538399>.
51. Ribolsi M, Luca-Guarino MP, Balestrieri P, et al. The results from up-front esophageal testing predict proton pump inhibitor response in patients with chronic cough. *Am J Gastroenterol*. 2021;116:2199–206, <http://dx.doi.org/10.14309/ajg.000000000001389>.
52. Krause AJ, Kaizer AM, Carlson DA, et al. Validated clinical score to predict gastroesophageal reflux in patients with chronic laryngeal symptoms: COuGH Reflux. *Clin Gastroenterol Hepatol*. 2024;22:1200–9, <http://dx.doi.org/10.1016/j.cgh.2024.01.021>, e1.
53. Spantideas N, Drosou E, Bougea A, et al. Proton pump inhibitors for the treatment of laryngopharyngeal reflux. A systematic review. *J Voice*. 2020;34:918–29, <http://dx.doi.org/10.1016/j.jvoice.2019.05.005>.
54. Francis DO, Rymer JA, Slaughter JC, et al. High economic burden of caring for patients with suspected extraesophageal reflux. *Am J Gastroenterol*. 2013;108:905–11, <http://dx.doi.org/10.1038/ajg.2013.69>.
55. Fuchs HF, Müller DT, Berlth F, et al. Simultaneous laryngopharyngeal pH monitoring (Restech) and conventional esophageal pH monitoring—correlation using a large patient cohort of more than 100 patients with suspected gastroesophageal reflux disease. *Dis Esophagus*. 2018;31, <http://dx.doi.org/10.1093/dote/doy018>.
56. Kurylo CM, Eastwood D, Blumin JH, et al. Correlation of esophageal mean nocturnal baseline impedance with markers of laryngopharyngeal reflux. *Laryngoscope*. 2023;133:1927–32, <http://dx.doi.org/10.1002/lary.30421>.
57. Krause AJ, Greytak M, Kaizer AM, et al. Diagnostic yield of ambulatory reflux monitoring systems for evaluation of chronic laryngeal symptoms. *Am J Gastroenterol*. 2024;119:627–34, <http://dx.doi.org/10.14309/ajg.000000000002557>.
58. Vardar R, Varis A, Bayrakci B, et al. Relationship between history, laryngoscopy and esophagogastroduodenoscopy for diagnosis of laryngopharyngeal reflux in patients with typical GERD. *Eur Arch Otorhinolaryngol*. 2012;269:187–91, <http://dx.doi.org/10.1007/s00405-011-1748-y>.
59. Belafsky PC, Postma GN, Koufman JA. Validity and reliability of the reflux symptom index (RSI). *J Voice*. 2002;16:274–7, [http://dx.doi.org/10.1016/s0892-1997\(02\)00097-8](http://dx.doi.org/10.1016/s0892-1997(02)00097-8).
60. Geeratragool T, Maneerattanaporn M, Prapruetkit J, et al. Association between laryngopharyngeal reflux clinical scores and esophageal multichannel intraluminal impedance pH monitoring interpretation according to Lyon Consensus 2.0. *Dis Esophagus*. 2025;38:doae098, <http://dx.doi.org/10.1093/dote/doae098>.
61. Wong M-W, Hsiao S-H, Wang J-H, et al. Practical application of the COuGH Reflux score for patients with laryngopharyngeal symptoms: real-world assessment in predicting gastroesophageal reflux disease and proton-pump inhibitor response. *Am J Gastroenterol*. 2024, <http://dx.doi.org/10.14309/ajg.000000000003240>.
62. Visaggi P, Del Corso G, Gyawali CP, et al. Ambulatory pH-impedance findings confirm that grade B esophagitis provides objective diagnosis of gastroesophageal reflux disease. *Am J Gastroenterol*. 2023;118:794–801, <http://dx.doi.org/10.14309/ajg.000000000002173>.
63. Gyawali CP, Marchetti L, Rogers BD, et al. The Lyon score: a novel reflux scoring system based on the Lyon Consensus 2.0 that associates with treatment outcome from antireflux therapy. *Am J Gastroenterol*. 2024, <http://dx.doi.org/10.14309/ajg.000000000003083>.
64. Lundell LR, Dent J, Bennett JR, et al. Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification. *Gut*. 1999;45:172–80, <http://dx.doi.org/10.1136/gut.45.2.172>.
65. Frazzoni M, Savarino E, de Bortoli N, et al. Analyses of the post-reflux swallow-induced peristaltic wave index and nocturnal baseline impedance parameters increase the diagnostic yield of impedance-pH monitoring of patients with reflux disease. *Clin Gastroenterol Hepatol*. 2016;14:40–6, <http://dx.doi.org/10.1016/j.cgh.2015.06.026>.
66. Gyawali CP, Rogers B, Frazzoni M, et al. Inter-reviewer variability in interpretation of pH-impedance studies: the Wingate Consensus. *Clin Gastroenterol Hepatol*. 2021;19:1976–8, <http://dx.doi.org/10.1016/j.cgh.2020.09.002>, e1.
67. Roman S, Gyawali CP, Savarino E, et al. Ambulatory reflux monitoring for diagnosis of gastro-esophageal reflux disease: update of the Porto consensus and recommendations from an international consensus group. *Neurogastroenterol Motil*. 2017;29:1–15, <http://dx.doi.org/10.1111/nmo.13067>.
68. Patel A, Sayuk GS, Gyawali CP. Parameters on esophageal pH-impedance monitoring that predict outcomes of patients with gastroesophageal reflux disease. *Clin Gastroenterol Hepatol*. 2015;13:884–91, <http://dx.doi.org/10.1016/j.cgh.2014.08.029>.
69. Davis TA, Gyawali CP. Refractory gastroesophageal reflux disease: diagnosis and management. *J Neurogastroenterol Motil*. 2024;30:17–28, <http://dx.doi.org/10.5056/jnm23145>.
70. Dhar A, Maw F, Dallal HJ, et al. Side effects of drug treatments for gastro-oesophageal reflux disease: current controversies. *Frontline Gastroenterol*. 2020;13:45–9, <http://dx.doi.org/10.1136/flgastro-2019-101386>.
71. Patti MG, Allaix ME, Fisichella PM. Analysis of the causes of failed antireflux surgery and the principles of treatment: a review. *JAMA Surg*. 2015;150:585–90, <http://dx.doi.org/10.1001/jamasurg.2014.3859>.