Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), noviembre-diciembre 2026,
Volumen 10, Número 6.

https://doi.org/10.37811/cl_rcm.v10i6

NÓDULOS TIROIDEOS INDETERMINADOS: REVISIÓN SISTEMÁTICA DE BIOMARCADORES MOLECULARES PARA PREDECIR MALIGNIDAD Y REDUCIR CIRUGÍAS INNECESARIAS

INDETERMINATE THYROID NODULES: A SYSTEMATIC REVIEW OF MOLECULAR BIOMARKERS TO PREDICT MALIGNANCY AND REDUCE UNNECESSARY SURGERIES

José Alfredo Peñúnuri Domínguez
Hospital Ángeles del Pedregal

Mildred Galvez Gosebruch
Universidad Anáhuac México

Camila Cristina Lee
Universidad La Salle

Andrea Fernanda Hoyos Cabrera
Universidad Anáhuac México

Johana Michelle Torres Tello
Universidad Anáhuac Puebla

Angélica Alejandra Fernández Camarena
Universidad Lamar Guadalajara

Miguel Ángel Hernández Calderón
Universidad La Salle México

Diana Uribe García
Universidad Anáhuac

Laura Fernández Gómez Aguilar
Universidad Anáhuac

Cecilia Castillo Osorio
Grupo Educativo Siglo XXI

DOI: https://doi.org/10.37811/cl_rcm.v10i6.26976

Nódulos tiroideos indeterminados: revisión sistemática de biomarcadores moleculares para predecir malignidad y reducir cirugías innecesarias

José Alfredo Peñúnuri Domínguez
josepenunuri13@gmail.com
https://orcid.org/0000-0002-3113-5771
Hospital Ángeles del Pedregal

Mildred Galvez Gosebruch
mildredgosebruch@gmail.com
https://orcid.org/0009-0009-7009-9046
Universidad Anáhuac México

Camila Cristina Lee
dracamilalee@gmail.com
https://orcid.org/0009-0000-9205-971X
Universidad La Salle

Andrea Fernanda Hoyos Cabrera
andreafhc468@gmail.com
https://orcid.org/0009-0006-0072-3673
Universidad Anáhuac México

Johana Michelle Torres Tello
johanatello25@gmail.com
https://orcid.org/0009-0008-1507-6266
Universidad Anáhuac Puebla

Angélica Alejandra Fernández Camarena
angelicaale1908@gmail.com
https://orcid.org/0009-0004-0225-9793
Universidad Lamar Guadalajara

Miguel Ángel Hernández Calderón
tweety.calderon@gmail.com
https://orcid.org/0009-0000-4122-480X
Universidad La Salle México

Diana Uribe García
dianauribegarcia@hotmail.com
https://orcid.org/0009-0004-4880-6537
Universidad Anáhuac

Laura Fernández Gómez Aguilar
larua.gomezag02@gmail.com
https://orcid.org/0009-0008-2632-5088
Universidad Anáhuac

Cecilia Castillo Osorio
ces17cast@gmail.com
https://orcid.org/0009-0007-6762-030X
Grupo Educativo Siglo XXI

RESUMEN

Introducción: Los nódulos tiroideos con citología indeterminada plantean dificultades para distinguir lesiones benignas y malignas y seleccionar pacientes para cirugía. Los biomarcadores moleculares ofrecen información complementaria para la estratificación preoperatoria. **Objetivo:** Evaluar el rendimiento diagnóstico de los biomarcadores moleculares en nódulos Bethesda III y IV y su asociación con una menor utilización de cirugía diagnóstica. **Métodos:** Se realizó una síntesis descriptiva de los resultados verificables de seis publicaciones del corpus bibliográfico proporcionado, incluidos estudios primarios y metaanálisis. Se examinaron rendimiento diagnóstico, perfiles moleculares, utilización de cirugía y seguimiento. No se calcularon estimaciones agrupadas propias ni se sumaron muestras potencialmente solapadas. **Resultados:** En un estudio multicéntrico de 834 nódulos, Afirma GSC mostró valores predictivos negativos de 99% para Bethesda III y 98% para Bethesda IV, con seguimiento mediano de 23 meses y tasa global de falsos negativos de 2%. Un metaanálisis comunicó valores predictivos negativos de 96% para Afirma GSC y 93% para ThyroSeq v3 entre casos con confirmación quirúrgica. Otra síntesis encontró mayores odds de resección con manejo convencional que con evaluación molecular (OR: 2,258; IC95%: 1,548–3,293). En 200 nódulos GSC sospechosos y resecados, las alteraciones moleculares se asociaron con diferencias en malignidad, sin mayor agresividad histopatológica. **Conclusiones:** Las pruebas moleculares complementan la evaluación de nódulos indeterminados y se asocian con menor utilización de cirugía. Su interpretación depende de la población y del método de confirmación diagnóstica. La evidencia disponible no permite cuantificar uniformemente las cirugías innecesarias evitadas ni establecer la seguridad de la vigilancia a largo plazo.

Palabras clave: nódulo tiroideo, citología indeterminada, biomarcadores, diagnóstico molecular, neoplasias tiroideas

Indeterminate Thyroid Nodules: A Systematic Review of Molecular Biomarkers to Predict Malignancy and Reduce Unnecessary Surgeries

ABSTRACT

Introduction: Thyroid nodules exhibiting indeterminate cytological findings present significant diagnostic challenges in differentiating benign from malignant lesions and determining the appropriate indications for surgical intervention. Molecular biomarkers provide supplementary diagnostic information that may enhance preoperative risk stratification and inform clinical decision-making. **Objective:** To assess the diagnostic performance of molecular biomarkers in thyroid nodules classified as Bethesda categories III and IV and to determine their association with a reduction in the utilization of diagnostic surgical procedures. **Methods:** A descriptive synthesis was conducted using data from six publications, including primary studies and meta-analyses, selected from the provided literature. Diagnostic performance, molecular profiles, surgical utilization, and follow-up outcomes were assessed. No additional pooled estimates were calculated, and potentially overlapping study populations were not combined. **Results:** In a multicenter study including 834 thyroid nodules, Afirma GSC yielded negative predictive values of 99% and 98% for Bethesda categories III and IV, respectively, with a median follow-up of 23 months and an overall false-negative rate of 2%. A meta-analysis reported negative predictive values of 96% for Afirma GSC and 93% for ThyroSeq v3 in cases with surgical confirmation. A separate systematic review found that conventional management was associated with higher odds of surgical resection compared with molecular testing (OR, 2.258; 95% CI, 1.548–3.293). Among 200 Afirma GSC-suspicious nodules that underwent surgical resection, molecular alterations were associated with differences in malignancy rates but not with increased histopathological aggressiveness. **Conclusions:** Molecular testing provides additional diagnostic information in the evaluation of thyroid nodules with indeterminate cytology and is associated with reduced surgical utilization. The interpretation of these findings depends on the study population and the method used for diagnostic confirmation. Current evidence does not allow for consistent quantification of unnecessary surgeries avoided or definitive conclusions regarding the safety of long-term surveillance.

Keywords: thyroid nodule, indeterminate cytology, biomarkers, molecular diagnosis, thyroid neoplasms

Artículo recibido 10 octubre 2026
Aceptado para publicación: 17 septiembre 2026

INTRODUCCIÓN

Relevancia clínica y limitaciones de la citología.

A pesar de que los nódulos tiroideos son lesiones frecuentes y pueden presentarse hasta en 20-50% de la población, la mayoría corresponden a lesiones benignas, mientras que únicamente entre 5-15% presentan malignidad. [1,2] Se define como nódulo tiroideo aquella lesión localizada dentro de la glándula tiroides que puede distinguirse del parénquima tiroideo circundante mediante técnicas de imagen. [1] Actualmente, la evaluación de estas lesiones incluye métodos como la punción por aspiración con aguja fina (PAAF), cuyos hallazgos pueden clasificarse mediante distintos sistemas de reporte citopatológico. Esta evaluación permite estratificar el riesgo de malignidad y contribuir a la identificación de lesiones clínicamente relevantes, procurando al mismo tiempo evitar el sobrediagnóstico y sobretratamiento. [1,2]

Entre los sistemas utilizados para el reporte de la citología tiroidea destaca el Sistema Bethesda, constituido por seis categorías citológicas que abarcan desde muestras no diagnósticas y lesiones benignas hasta aquellas sospechosas o diagnósticas de malignidad. Sin embargo, una de las principales limitaciones de este sistema radica en los nódulos con resultados “indeterminados”, particularmente aquellos clasificados como Bethesda III y IV. En conjunto, estas categorías representan aproximadamente 20-30% de los diagnósticos, lo que requiere evaluaciones extras debido a un riesgo de malignidad que va desde el 5-25% y del 25-40% respectivamente. [1,2]

La presencia de un resultado citológico indeterminado genera, por tanto, incertidumbre en la toma de decisiones clínicas debido a que la citología por sí sola no permite establecer con certeza la naturaleza benigna o maligna de estas lesiones. Ante esta limitación, se han incorporado herramientas complementarias de estratificación, entre ellas las pruebas moleculares, con el objetivo de precisar el riesgo de malignidad y contribuir a la selección de la conducta terapéutica. Dentro de estas herramientas se han desarrollado diferentes pruebas moleculares aplicables a los nódulos tiroideos indeterminados, incluyendo estrategias orientadas a la confirmación o exclusión de la malignidad. [1,2]

Fundamento y alcance de biomarcadores

Los biomarcadores moleculares buscan reducir la incertidumbre diagnóstica de los nódulos tiroideos con citología indeterminada, especialmente en las categorías Bethesda III y IV. Funciona a partir de dos

estrategias: Las pruebas orientadas a descartar malignidad (*rule-out*) priorizan una sensibilidad y un valor predictivo negativo (VPN) elevados, de modo que un resultado benigno o negativo disminuya suficientemente la probabilidad de cáncer y permita considerar vigilancia en pacientes seleccionados. Las pruebas orientadas a identificar malignidad (*rule-in*), en cambio, buscan una mayor especificidad y valor predictivo positivo (VPP), donde un resultado positivo aumenta la probabilidad de neoplasia y apoya la decisión quirúrgica [2,4]. Esta distinción no es absoluta, ya que el VPP y el VPN dependen de la prevalencia de malignidad en la población donde se aplica la prueba [2,4].

Entre las plataformas actuales, Afirma Genomic Sequencing Classifier (GSC) analiza ARN mediante secuenciación de nueva generación, estudio transcriptómico y algoritmos de clasificación. En su validación clínica presentó una sensibilidad de 91%, especificidad de 68%, VPN de 96% y VPP de 47%, lo que refleja su capacidad para disminuir la probabilidad de malignidad ante un resultado benigno [2,3]. Por su parte, ThyroSeq v3 analiza ADN y ARN de 112 genes relacionados con cáncer tiroideo e identifica mutaciones, fusiones génicas, alteraciones en el número de copias y cambios en la expresión génica. En estudios de validación alcanzó una sensibilidad de 94%, especificidad de 82%, VPN de 97% y VPP de 66%, con una prevalencia de malignidad cercana al 28% [2,3].

Los resultados agrupados más recientes muestran que estas diferencias deben interpretarse con cautela. En el metaanálisis de Vardarli et al., Afirma GSC presentó una sensibilidad agrupada de 97% y especificidad de 50%, mientras que ThyroSeq v3 alcanzó 95% y 47%, respectivamente. La razón de verosimilitud negativa fue de 0.07 para GSC y 0.11 para ThyroSeq v3; a su vez, ThyroSeq v3 mostró un área bajo la curva ROC de 0.95, frente a 0.86 para GSC [4]. Sin embargo, estas cifras no permiten establecer una superioridad universal: los estudios incluidos no compararon ambas plataformas sobre las mismas muestras y el propio metaanálisis no encontró una diferencia significativa entre GSC y ThyroSeq v3 en su comparación mediante metarregresión [4].

Por ello, la elección de una prueba molecular debe responder a la pregunta clínica que se intenta resolver, más que a la búsqueda de una plataforma universalmente superior. El resultado molecular debe integrarse con la categoría citológica, los hallazgos ecográficos, la prevalencia local de malignidad y las características del paciente. En este contexto, tanto Afirma GSC como ThyroSeq v3 pueden contribuir

a disminuir el riesgo residual después de un resultado benigno o negativo y aportar información adicional cuando el resultado incrementa la sospecha de malignidad [2–4].

Brecha de conocimiento y objetivo de la revisión.

Las pruebas moleculares han mejorado la estratificación del riesgo de malignidad de nódulos tiroideos con citología indeterminada, sin embargo, es necesario diferenciar su precisión diagnóstica de su utilidad clínica. La precisión diagnóstica expresa su capacidad de distinguir lesiones benignas y malignas mediante indicadores como sensibilidad, especificidad y valores predictivos, mientras que, la utilidad clínica hace referencia a su contribución a las decisiones de manejo y los resultados del paciente. Por ello, un rendimiento favorable no demuestra por sí solo, que su implementación permita evitar cirugías innecesarias de forma segura [3].

Raghunathan et al. identifican como desafíos pendientes la fiabilidad de las pruebas y la definición de la duración óptima de vigilancia [3]. Por otro lado, Ahmadi et al. aportan evidencia sobre el rendimiento de Afirma GSC y los resultados de manejo de nódulos Bethesda III y IV, con un seguimiento de 23 meses [5]. Estos antecedentes justifican examinar hasta qué punto la información molecular deriva en un beneficio clínico y que limitaciones condicionan la interpretación de los resultados en pacientes que permanecen sin cirugía.

Como resultado, la evaluación de estas herramientas requiere considerar en conjunto la reducción de intervenciones quirúrgicas y la seguridad del seguimiento. La reducción en la utilización de cirugía no equivale directamente en una reducción del sobretratamiento, pues debe diferenciarse entre intervenciones evitadas e intervenciones pospuestas. Del mismo modo, es necesario examinar las malignidades detectadas después de resultados moleculares benignos o negativos, las cirugías realizadas durante la vigilancia y la duración e integridad del seguimiento [3,5].

Esta revisión tiene como objetivo sintetizar y evaluar críticamente la evidencia sobre la precisión diagnóstica y la utilidad clínica de los biomarcadores moleculares en nódulos tiroideos con citología indeterminada Bethesda III y IV, considerando su capacidad para predecir malignidad, su asociación con la reducción de cirugías diagnósticas potencialmente innecesarias y la seguridad del seguimiento de los pacientes con resultados benignos o negativos.

RESULTADOS

Características de las muestras

La evidencia primaria verificada comprendió el estudio multicéntrico de Ahmadi et al., con 834 nódulos indeterminados evaluados en tres centros estadounidenses; la serie mexicana, con 22 pacientes, de los cuales 18 recibieron pruebas moleculares; y el estudio retrospectivo de Peng et al., con 192 pacientes y 200 nódulos [5,6,14]. En Ahmadi et al., el tamaño mediano del nódulo fue de 2,1 cm y el seguimiento mediano de 23 meses. En Peng et al., la mediana de edad fue de 51 años y el 77,6% de los pacientes fueron mujeres. El número de pacientes correspondiente a los 834 nódulos de Ahmadi no está especificado en el resumen consultado.

Las síntesis secundarias incluyeron muestras diferentes: Vardarli et al. analizaron 53 estudios, con 71 conjuntos de datos y 6490 aspirados; Nguyen et al., 132 estudios y 66 448 nódulos, incluidos 30 292 nódulos Bethesda III/IV (45,6%); y Dowell et al. comunicaron 26 estudios en su resumen, con 1162 nódulos con confirmación quirúrgica evaluados mediante Afirma GSC y 1127 mediante ThyroSeq v3 [4,12,15]. Estas muestras se presentan por separado, sin sumarlas como una población única.

Rendimiento diagnóstico para malignidad

En Ahmadi et al., Afirma GSC mostró sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo de 95%, 81%, 50% y 99%, respectivamente, para Bethesda III; los valores correspondientes a Bethesda IV fueron 94%, 82%, 65% y 98% [5]. En los nódulos con predominio oncocítico, el valor predictivo negativo fue de 100% para Bethesda III y 98% para Bethesda IV, mientras que el valor predictivo positivo fue de 17% y 45%, respectivamente. Los denominadores de estos subgrupos no están disponibles en el resumen consultado.

En el metaanálisis de Vardarli et al., la sensibilidad agrupada de las cuatro plataformas fue de 0,95 (IC95%: 0,94–0,97) y la especificidad de 0,35 (IC95%: 0,28–0,43). ThyroSeq v3 presentó un área bajo la curva de 0,95 (IC95%: 0,93–0,96); Afirma GSC mostró una razón de verosimilitud negativa de 0,07 (IC95%: 0,02–0,19) [4]. Estos resultados corresponden al metaanálisis publicado, no a una estimación calculada para esta revisión.

Dowell et al. reportaron, entre los casos con confirmación quirúrgica, sensibilidad, especificidad y valor predictivo negativo de 94%, 42% y 96% para Afirma GSC, frente a 96%, 40% y 93% para ThyroSeq

v3. Al asumir que los resultados negativos no operados eran verdaderos negativos, los valores fueron 96%, 86% y 99%, y 97%, 83% y 99%, respectivamente [15].

Estratificación según alteraciones moleculares

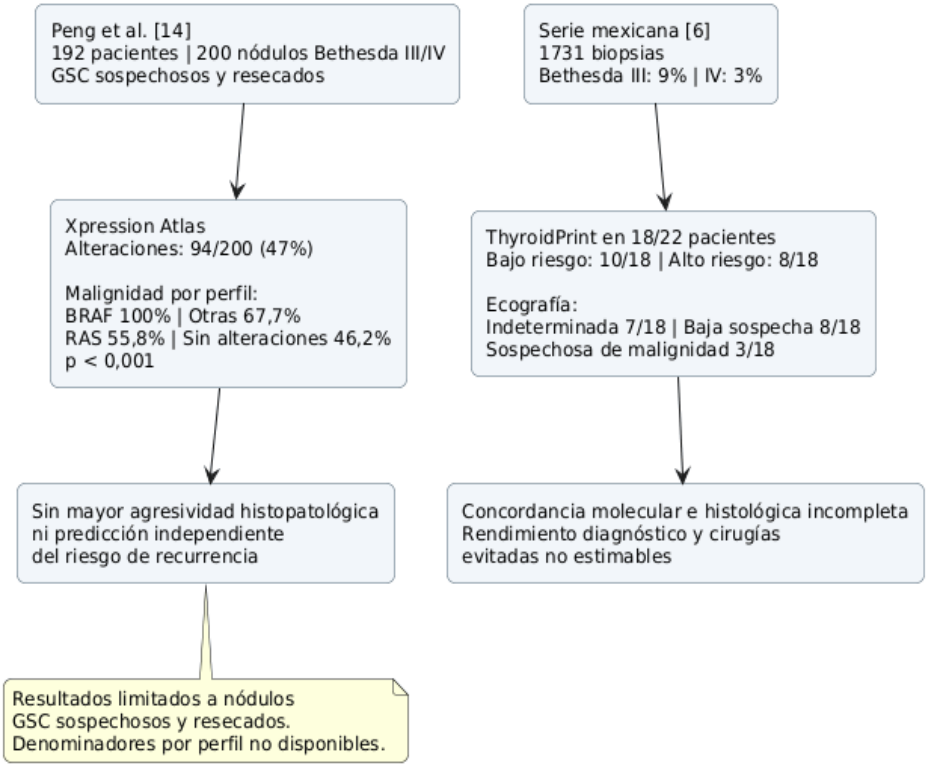
Peng et al. evaluaron 192 pacientes con 200 nódulos Bethesda III/IV, todos con resultado sospechoso mediante Afirma GSC y confirmación histopatológica posterior a la resección. Xpression Atlas detectó alteraciones en 94/200 nódulos (47,0%). La frecuencia de malignidad difirió entre los perfiles moleculares: 100% en los nódulos con alteraciones de BRAF (IC95%: 71,5–100,0%), 67,7% con otras alteraciones (IC95%: 48,6–83,3%), 55,8% con alteraciones de RAS (IC95%: 41,3–69,5%) y 46,2% sin alteraciones detectadas (IC95%: 36,5–56,2%; $p < 0,001$) [14]. Los denominadores de estos subgrupos no estaban disponibles en el resumen consultado.

La mayor frecuencia de malignidad en los nódulos con alteraciones no se acompañó de mayores frecuencias de extensión extratiroidea, invasión linfática o angioinvasión. Asimismo, el análisis multivariable no identificó ningún subgrupo molecular como predictor independiente de la categoría de riesgo de recurrencia de la American Thyroid Association de 2025 [14]. Estos resultados corresponden exclusivamente a nódulos GSC sospechosos y resecados; no describen el riesgo de todos los nódulos indeterminados. journals.sagepub.

En la serie mexicana se identificaron 1731 biopsias por aspiración, de las cuales 9% corresponden a Bethesda III y 3% a Bethesda IV. Se realizó ThyroidPrint en 18/22 pacientes con lesiones indeterminadas: 10/18 obtuvieron clasificación molecular de bajo riesgo y 8/18, de alto riesgo. La evaluación ecográfica identificó 7/18 casos indeterminados, 8/18 de baja sospecha y 3/18 sospechosos de malignidad [6].

Los autores comunicaron decisiones terapéuticas en cinco pacientes, todos Bethesda III, y hallazgos de carcinoma papilar en las tiroidectomías descritas; otros cuatro pacientes estaban programados para cirugía. Sin embargo, el resumen no presenta una correspondencia individual completa entre clasificación molecular, intervención e histopatología. Por ello, no permite establecer cuántos carcinomas correspondieron a cada categoría molecular ni calcular sensibilidad, especificidad, valores predictivos o cirugías evitadas [6].

Figura 1. Estratificación molecular de nódulos tiroideos



Utilización de cirugía y seguimiento

Nguyen et al. encontraron mayores odds de resección con manejo convencional que con pruebas moleculares (OR: 2,258; IC95%: 1,548–3,293; p < 0,001). Las pruebas de primera generación también se asociaron con mayores odds de cirugía que las de segunda generación (OR: 1,709; IC95%: 1,122–2,601; p = 0,013) [12]. Estas estimaciones describen utilización de cirugía; no proporcionan un número absoluto de intervenciones innecesarias evitadas.

Ahmadi et al. comunicaron una tasa global de falsos negativos de 2% y crecimiento durante la vigilancia en 22% de los nódulos con resultado GSC benigno, con seguimiento mediano de 23 meses [5]. El resumen no especifica los numeradores y denominadores correspondientes ni permite cuantificar complicaciones quirúrgicas, mortalidad o malignidad diagnosticada tardíamente.

Síntesis de resultados

Las publicaciones verificadas mostraron una elevada capacidad para obtener resultados negativos en ausencia de malignidad, con estimaciones dependientes de la plataforma y del método de confirmación diagnóstica. En los 834 nódulos estudiados por Ahmadi et al., Afirma GSC presentó valores predictivos

negativos de 99% para Bethesda III y 98% para Bethesda IV, con seguimiento mediano de 23 meses y una tasa global de falsos negativos de 2% [5]. En el metaanálisis de Dowell et al., los valores predictivos negativos entre casos con confirmación quirúrgica fueron de 96% para Afirma GSC y 93% para ThyroSeq v3; ambos alcanzaron 99% al asumir como verdaderos negativos los resultados negativos no operados [15].

La utilización de cirugía fue menor en los grupos evaluados mediante pruebas moleculares. Nguyen et al. comunicaron mayores odds de resección con manejo convencional que con evaluación molecular (OR: 2,258; IC95%: 1,548–3,293), así como con pruebas de primera frente a segunda generación (OR: 1,709; IC95%: 1,122–2,601) [12]. Estas estimaciones no equivalen a un número absoluto de cirugías innecesarias evitadas.

La estratificación molecular también mostró diferencias en la frecuencia de malignidad según la alteración detectada, sin una asociación paralela con mayor agresividad histopatológica en la cohorte de Peng et al. [14]. La serie mexicana aportó resultados de clasificación molecular en 18 pacientes, pero sin información suficiente para estimar su rendimiento diagnóstico [6]. No se calculó una estimación conjunta propia ni se sumaron las muestras de las publicaciones, dado que combinan pacientes, nódulos y aspirados, y no se ha descartado el solapamiento entre cohortes.

Figura 2. Síntesis de resultados

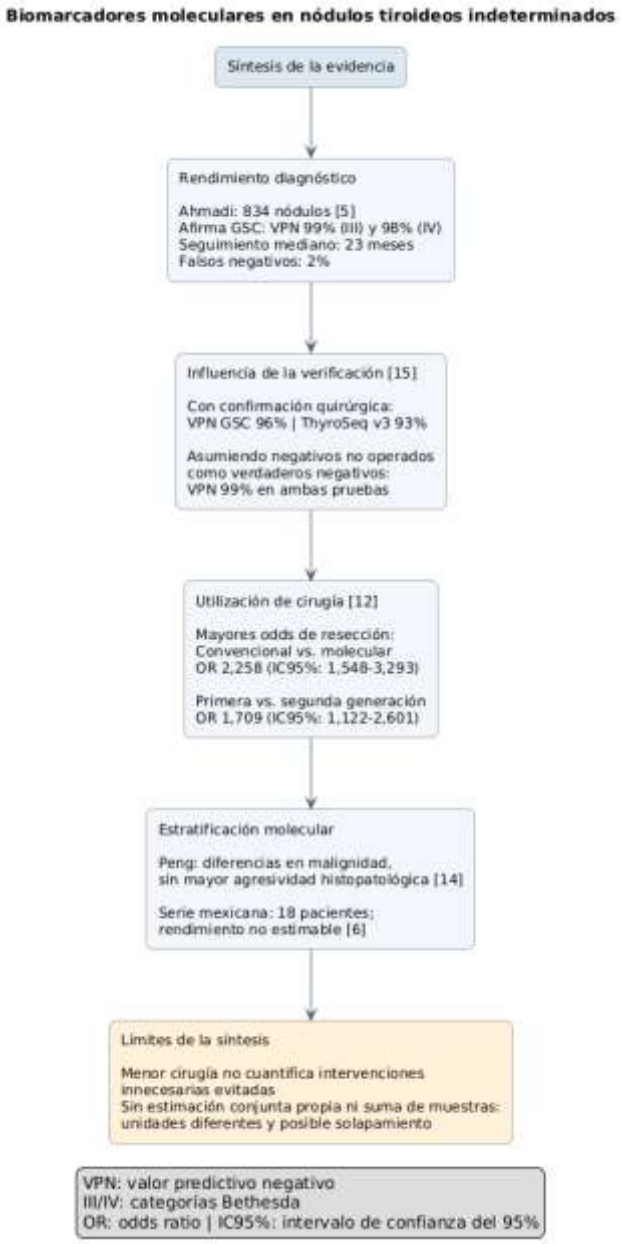


Tabla 1. Desenlaces y técnicas principales

Técnica o comparación	Muestra	Desenlace principal	Referencia
Afirma GSC	834 nódulos; 3 centros	Bethesda III: Se 95%, Sp 81%, VPP 50%, VPN 99%. Bethesda IV: Se 94%, Sp 82%, VPP 65%, VPN 98%.	Ahmadi et al. [5]
Afirma GSC y ThyroSeq v3	1162 y 1127 nódulos con confirmación quirúrgica, respectivamente	Se/Sp/VPN: GSC 94%/42%/96%; ThyroSeq v3 96%/40%/93%. Con negativos no operados asumidos verdaderos negativos: VPN 99% en ambas pruebas.	Dowell et al. [15]
Afirma GEC/GSC y ThyroSeq v2/v3	53 estudios; 6490 aspirados	Se agrupada 0,95 (IC95% 0,94–0,97); Sp 0,35 (0,28–0,43). AUC ThyroSeq v3: 0,95 (0,93–0,96).	Vardarli et al. [4]
Pruebas moleculares frente a manejo convencional	132 estudios; 30 292 nódulos Bethesda III/IV	Mayores odds de cirugía con manejo convencional: OR 2,258 (IC95% 1,548–3,293). Primera frente a segunda generación: OR 1,709 (1,122–2,601).	Nguyen et al. [12]
Afirma Xpression Atlas	192 pacientes; 200 nódulos resecados GSC sospechosos	Alteraciones: 94/200. Malignidad: BRAF 100%; RAS 55,8%; otras 67,7%; sin alteraciones 46,2% ($p < 0,001$).	Peng et al. [14]
ThyroidPrint	18/22 pacientes evaluados molecularmente	Bajo riesgo: 10/18; alto riesgo: 8/18. Concordancia completa con histopatología no disponible.	Avalos-Jiménez et al. [6]
Vigilancia tras Afirma GSC	Cohorte de 834 nódulos; mediana 23 meses	Tasa global de falsos negativos 2%; crecimiento en 22% de los nódulos GSC benignos. Denominadores de ambos desenlaces no disponibles en el resumen.	Ahmadi et al. [5]

Se: sensibilidad; Sp: especificidad; VPP/VPN: valor predictivo positivo/negativo; AUC: área bajo la curva; OR: odds ratio; IC95%: intervalo de confianza del 95%. Las estimaciones agrupadas proceden de los metaanálisis citados.

Tabla 2. Estudios principales

Estudio	Diseño y ámbito	Muestra y técnica	Datos relevantes
Vardarli et al. 2024 [4]	Revisión sistemática y metaanálisis diagnóstico	53 estudios; 71 conjuntos de datos; 6490 aspirados. Afirma GEC/GSC y ThyroSeq v2/v3.	Bethesda III/IV; NIFTP considerada maligna en el análisis.
Ahmadi et al. 2024 [5]	Estudio retrospectivo multicéntrico; 3 centros de Estados Unidos.	834 nódulos Bethesda III/IV. Afirma GSC. Número de pacientes no disponible en el resumen.	Tamaño mediano 2,1 cm; seguimiento mediano 23 meses; cirugía o vigilancia ecográfica ≥1 año.
Avalos-Jiménez et al. 2026 [6]	Prospectivo, observacional y descriptivo, según los autores; México	1731 BAAF en 2 centros; 22 pacientes seleccionados, 18 con ThyroidPrint.	Bethesda III 9% y IV 3% de las BAAF. Base de datos 2022–2024; edad y sexo no disponibles en el resumen.
Nguyen et al. 2026 [12]	Revisión sistemática y metaanálisis global	132 estudios; 66 448 nódulos totales; 30 292 Bethesda III/IV (45,6%).	Compara manejo convencional, pruebas moleculares y generaciones de plataformas; desenlaces de resección y malignidad.

Peng et al. 2026 [14]	Retrospectivo unicéntrico; marzo 2020–septiembre 2025	192 pacientes; 200 nódulos Bethesda III/IV, GSC sospechosos y resecados; Xpression Atlas.	Mediana de edad 51 años; mujeres 77,6%. Subgrupos BRAF, RAS, otras alteraciones y sin alteraciones.
Dowell et al. 2025 [15]	Revisión sistemática y metaanálisis diagnóstico	26 estudios según el resumen; 1162 nódulos quirúrgicos GSC y 1127 ThyroSeq v3.	Distingue confirmación quirúrgica y negativos no operados asumidos verdaderos negativos.

BAAF: biopsia por aspiración con aguja fina; NIFTP: neoplasia tiroidea folicular no invasiva con características nucleares de tipo papilar. Los tamaños de muestra no se suman: las unidades difieren y no se ha descartado solapamiento entre publicaciones

DISCUSIÓN

El valor clínico de los biomarcadores moleculares en los nódulos tiroideos indeterminados depende de su capacidad para modificar una decisión terapéutica bajo incertidumbre. Un resultado negativo puede favorecer la vigilancia en pacientes seleccionados, pero la reducción de cirugía adquiere relevancia únicamente si se acompaña de seguimiento suficiente para detectar malignidad posteriormente. Por ello, precisión diagnóstica, utilización de cirugía y seguridad del manejo conservador deben evaluarse como desenlaces relacionados, aunque no intercambiables [3,5].

La comparación entre Afirma GSC y ThyroSeq v3 exige considerar cómo se establece el diagnóstico de referencia. Dowell et al. mostraron que las estimaciones cambian al incorporar nódulos negativos no operados como verdaderos negativos [15]. Restringir el análisis a casos resecados introduce sesgo de verificación, porque la indicación quirúrgica selecciona una población con características diferentes de quienes permanecen en vigilancia. En cambio, asumir benignidad en todos los casos no operados puede sobreestimar el rendimiento cuando el seguimiento es insuficiente. Estas diferencias dificultan atribuir superioridad a una plataforma y hacen necesario interpretar cada estimación según su método de confirmación diagnóstica.

La clasificación histopatológica también condiciona las comparaciones. Vardarli et al. consideraron la neoplasia tiroidea folicular no invasiva con características nucleares de tipo papilar —NIFTP— dentro del desenlace maligno [4]. Esta decisión modifica el significado de un resultado molecular positivo: la identificación de una lesión neoplásica susceptible de resección no implica necesariamente la detección de carcinoma invasivo. Las diferencias entre estudios pueden reflejar, por tanto, tanto el comportamiento de las pruebas como la definición del desenlace. Separar carcinoma, NIFTP y lesiones benignas

permitiría precisar qué enfermedad identifica cada plataforma y qué beneficio clínico aporta su detección.

La asociación entre evaluación molecular y menor utilización de cirugía constituye un hallazgo relevante, pero no demuestra por sí sola que todas las intervenciones evitadas fueran innecesarias [12]. La decisión de operar también depende de características clínicas, ecográficas y preferencias del paciente [3]. Además, la comparación entre generaciones de pruebas puede estar influida por cambios contemporáneos en los criterios de intervención y vigilancia. La interpretación más sustentada es que las pruebas moleculares contribuyen a seleccionar pacientes para cirugía, mientras que la magnitud del beneficio atribuible exclusivamente a la prueba requiere comparaciones que controlen estas diferencias. Los hallazgos de Peng et al. permiten distinguir riesgo de malignidad y agresividad tumoral [14]. La asociación de determinados perfiles con carcinoma no se acompañó de mayor agresividad histopatológica ni permitió identificar predictores independientes de la categoría de recurrencia. Por ello, la información molecular destinada a establecer probabilidad diagnóstica no debe trasladarse automáticamente a decisiones sobre extensión quirúrgica. Asimismo, la ausencia de alteraciones en Xpression Atlas no equivale a un resultado GSC benigno: la cohorte incluyó exclusivamente nódulos GSC sospechosos y resecados. Esta selección limita la extrapolación de sus resultados a pacientes candidatos a vigilancia.

La serie mexicana aporta información contextual sobre la aplicación de ThyroidPrint, aunque su muestra molecular de 18 pacientes y la ausencia de correspondencia completa entre prueba e histopatología impiden estimar con precisión su rendimiento [6]. Su contribución debe entenderse como experiencia inicial de implementación, sin asumir equivalencia con plataformas evaluadas en cohortes mayores. La validación local requiere documentar prevalencia de malignidad, características de los pacientes, resultados no concluyentes, decisiones posteriores y seguimiento, variables necesarias para valorar la transferencia de resultados entre entornos asistenciales.

Limitaciones de la evidencia

La certeza de esta síntesis está limitada por la combinación de estudios primarios y metaanálisis, el posible solapamiento de cohortes y la extracción incompleta de algunas publicaciones. Tampoco se dispone todavía de una evaluación formal del riesgo de sesgo de esta revisión; las discrepancias internas

identificadas en Dowell et al. requieren conciliación antes de cerrar el manuscrito. Las investigaciones futuras deberían priorizar comparaciones directas, confirmación diagnóstica claramente definida y seguimiento prolongado de los resultados negativos. Su desenlace clínico central debe ser la reducción de intervenciones evitables acompañada de una evaluación explícita de malignidad diagnosticada tardíamente, complicaciones y calidad de vida.

CONCLUSIONES

Rendimiento diagnóstico y selección del manejo

Las pruebas moleculares, particularmente afirma GSC y ThyroSeq v3, complementan la evaluación de los nódulos tiroideos con citología indeterminada Bethesda III y IV al aportar información adicional sobre la probabilidad de malignidad. La evidencia revisada respalda su capacidad para disminuir la incertidumbre diagnóstica, especialmente ante resultados benignos o negativos; sin embargo, su interpretación debe integrarse con los hallazgos clínicos, ecográficos y citológicos, así como con la prevalencia de malignidad de la población evaluada [4, 15].

En pacientes seleccionados, un resultado molecular benigno o negativo puede apoyar la elección de vigilancia clínica y ecográfica en lugar de cirugía diagnóstica inmediata. Este resultado no excluye por completo la malignidad ni elimina la necesidad de seguimiento y reevaluación ante cambios sospechosos. Asimismo, las diferencias en la selección de pacientes y en los métodos de confirmación diagnóstica condicionan las estimaciones de rendimiento e impiden establecer una superioridad universal entre ambas plataformas. Por ello, su principal aporte consiste en atribuir a una selección individualizada del manejo [4, 15].

Utilidad clínica, limitaciones y evidencia futura

La utilidad clínica de las pruebas moleculares en nódulos tiroideos indeterminados radica en su potencial para optimizar las decisiones terapéuticas y disminuir la utilización de cirugía diagnóstica. La evidencia revisada muestra una menor frecuencia de intervención quirúrgica en pacientes evaluados mediante estas herramientas frente al manejo convencional; sin embargo, esta asociación no permite establecer que todas las cirugías evitadas fueran innecesarias ni demostrar una mejoría del pronóstico a largo plazo [12]. Asimismo, la identificación de alteraciones moleculares asociadas con mayor probabilidad de

malignidad no implica necesariamente una mayor agresividad tumoral, por lo que sus resultados deben interpretarse dentro del contexto clínico individual [14].

A pesar de estos beneficios potenciales, la evidencia presenta limitaciones relacionadas con la heterogeneidad metodológica, los sesgos de verificación y la duración del seguimiento de los pacientes con resultados moleculares negativos [5,15]. Por ello, se requieren estudios prospectivos con comparaciones directas entre plataformas y seguimiento prolongado que permitan determinar su impacto sobre las intervenciones quirúrgicas evitables, la detección tardía de malignidad, las complicaciones y la calidad de vida. Estas investigaciones serán fundamentales para establecer con mayor precisión la seguridad y el beneficio clínico de incorporar las pruebas moleculares en el manejo de los nódulos tiroideos indeterminados.

REFERENCIAS BIBLIOGRÁFICAS

1. Grani G, Sponziello M, Filetti S, Durante C. Thyroid nodules: diagnosis and management. Nature Reviews Endocrinology [Internet]. 2024 Aug 16;20(12):715–28. Available from: https://pubmed.ncbi.nlm.nih.gov/39152228/
2. Fumagalli C, Serio G. Molecular testing in indeterminate thyroid nodules: an additional tool for clinical decision-making. Pathologica [Internet]. 2023 Sept;115(4):205–16. Available from: https://old.pathologica.it/article/view/887
3. Raghunathan R, Praw SS, Livhits M. Molecular testing for indeterminate thyroid nodules: past, present, and future. Current Opinion in Endocrinology, Diabetes & Obesity [Internet]. 2023 Aug 2;30(5):231–7. Available from: https://www.ovid.com/jnls/co-endocrinology/fulltext/10.1097/med.0000000000000829~molecular-testing-for-indeterminate-thyroid-nodules-past
4. Vardarli I, Tan S, Görges R, Krämer BK, Herrmann K, Brochhausen C. Diagnostic accuracy of Afirma gene expression classifier, Afirma gene sequencing classifier, ThyroSeq v2 and ThyroSeq v3 for indeterminate (Bethesda III and IV) thyroid nodules: a meta-analysis [Internet]. Vol. 13, Endocrine connections. England; 2024. p. e240170. Available from: https://pubmed.ncbi.nlm.nih.gov/38771544/

5. Ahmadi S, Kotwal A, Bikas A, Xiang P, Goldner W, Patel A, et al. Outcomes of Cytologically Indeterminate Thyroid Nodules Managed With Genomic Sequencing Classifier [Internet]. Vol. 109, The Journal of clinical endocrinology and metabolism. United States; 2024. p. e2231–9. Available from: https://pubmed.ncbi.nlm.nih.gov/38415829/

6. Jiménez NIA, Hernández BES, Chapa IM, Ariza NMI, Domínguez AG. Análisis molecular de nódulos tiroideos indeterminados en México: experiencia de 5 años en un centro hospitalario de referencia. Patología Revista Latinoamericana. 2026;64:1. Available from: https://openurl.ebsco.com/contentitem/doi:10.24245%2Fpatrl.v64id.110082?sid=ebsco:plink:crawler&id=ebsco:doi:10.24245%2Fpatrl.v64id.110082

7. Solis I. Enfrentamiento actual del nódulo tiroideo. Revista Médica Clínica Las Condes. 2025 July;36(4):260–6. Available from: https://www.sciencedirect.com/science/article/pii/S0716864025000744

8. Martínez MER. Nódulo tiroideo. ANALES DE OTORRINOLARINGOLOGÍA MEXICANA. 2025;70(3). Available from: https://www.revisionporpares.com/index.php/AORL/article/view/10419

9. Martínez Gómez AV, Pacheco Guevara KA, Castro Guachilema MA, Vega Tenesaca YM, Cabrera Valdiviezo EE, Alcivar Ramos GM, et al. Revisión bibliográfica del diagnóstico y manejo actual del nódulo tiroideo. Florence: Interdisciplinary Journal of Health and Sustainability. 2024 Sept 30;2(2):e24001. Available from: https://florence.sapienzaeditorial.com/index.php/home/article/view/21

10. Morales P. Biomarcadores de metilación del ADN para el diagnóstico precoz del cáncer de tiroides. Csic.es. Universidad de Oviedo; 2023. Available from: https://digital.csic.es/handle/10261/340352

11. ARRANZ JIMÉNEZ R. Marcadores radiológicos prequirúrgicos en pacientes intervenidos por nódulos tiroideos. Sistematización actualizada y clasificación estandarizada. Produccioncientifica.ucm.es. 2023. Available from: https://produccioncientifica.ucm.es/documentos/64f22ec3255f420df2d33fcd?lang=es_419

12. Nguyen TPX, Bychkov A, Jung CK, Kakudo K, Rana C. Impact of molecular testing on surgical decision-making in indeterminate thyroid nodules: a global meta-analysis across test generations. Endocr Pathol. 2026;37:22. doi:10.1007/s12022-026-09917-z. https://link.springer.com/article/10.1007/s12022-026-09917-z

13. Azaryan I, Maxwell C, Tran DH, Sipos JA, Endo M. Molecular testing for the management of indeterminate thyroid nodules. Endocr Relat Cancer. 2026;33(2):e250258. doi:10.1530/ERC-25-0258. https://journals.bioscientifica.com/erc/article-abstract/33/2/e250258/75538/Molecular-testing-for-the-management-of?redirectedFrom=fulltext

14. Peng TD, Hughes EG, Kim J, Tseng CH, SmookePraw S, Wu JX, Yeh MW, Livhits MJ. Afirma Xpression Atlas alteration status is associated with malignancy risk in indeterminate thyroid nodules. Thyroid. 2026 Aug 28. doi:10.1177/10507256261484715.

15. https://journals.sagepub.com/doi/10.1177/10507256261484715

16. Dowell N, Begum S, Muzaffar J, Boelaert K, Nieto H. Comparing the diagnostic accuracy of Afirma GSC to ThyroSeq V3 in cytologically indeterminate thyroid nodules [Internet]. Vol. 14, European thyroid journal. England; 2025. p. e250296. Available from: https://pubmed.ncbi.nlm.nih.gov/41324437/

17. Valderrabano, P., & McIver, B. (2017). Evaluation and Management of Indeterminate Thyroid Nodules: The Revolution of Risk Stratification Beyond Cytological Diagnosis. Cancer control: journal of the Moffitt Cancer Center, 24(5), 1073274817729231. https://doi.org/10.1177/1073274817729231

18. Livhits MJ, Zhu CY, Kuo EJ, Nguyen DT, Kim J, Tseng CH, et al. Effectiveness of Molecular Testing Techniques for Diagnosis of Indeterminate Thyroid Nodules: A Randomized Clinical Trial. JAMA Oncology [Internet]. 2021;7(1):70–77. Available from: https://jamanetwork.com/journals/jamaoncology/fullarticle/2773499

19. Kim NE, Raghunathan RS, Hughes EG, Longstaff XR, Tseng CH, Li S, et al. Bethesda III and IV Thyroid Nodules Managed Nonoperatively After Molecular Testing With Afirma GSC or

ThyroSeq v3. The Journal of Clinical Endocrinology & Metabolism [Internet]. 2023;108(9):e698–e703. Available from: https://academic.oup.com/jcem/article/108/9/e698/7095610

20. Lévesque F, Payne RJ, Beaudoin D, Boucher A, Fortier PH, Massicotte MH, et al. Publicly Funded Molecular Testing of Indeterminate Thyroid Nodules: Canada's Experience. The Journal of Clinical Endocrinology & Metabolism [Internet]. 2025;110(4):e1031–e1037. Available from: https://academic.oup.com/jcem/article/110/4/e1031/7680219