



CIRUGÍA ESPAÑOLA

www.elsevier.es/cirugia



Original

Rendimiento del sistema Bethesda en el diagnóstico citopatológico del nódulo tiroideo



Ismael Mora-Guzmán^a, José Luis Muñoz de Nova^{a,b,*}, Cristina Marín-Campos^a, José Antonio Jiménez-Heffernan^c, Juan Julián Cuesta Pérez^d, Marcos Lahera Vargas^e, Emma Torres Mínguez^a y Elena Martín-Pérez^{a,b}

^a Servicio de Cirugía General y del Aparato Digestivo, Hospital Universitario de La Princesa, Madrid, España

^b Departamento de Cirugía, Universidad Autónoma de Madrid, Madrid, España

^c Servicio de Anatomía Patológica, Hospital Universitario de La Princesa, Madrid, España

^d Servicio de Radiodiagnóstico, Hospital Universitario de La Princesa, Madrid, España

^e Servicio de Endocrinología, Hospital Universitario de La Princesa, Madrid, España

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Recibido el 29 de diciembre de 2017

Aceptado el 19 de febrero de 2018

On-line el 28 de marzo de 2018

Palabras clave:

Cáncer de tiroides

Bethesda

Punción-aspiración con aguja fina

Citología

RESUMEN

Introducción: La punción-aspiración con aguja fina es una pieza clave en la evaluación preoperatoria del nódulo tiroideo y el sistema Bethesda es el más aceptado para categorizar el análisis citológico. El objetivo del estudio es evaluar la validez del sistema Bethesda en la enfermedad nodular tiroidea para diagnosticar malignidad.

Métodos: Se incluye a los pacientes intervenidos de tiroides consecutivamente entre junio de 2010 y junio de 2017. Se realizó el análisis de la punción preoperatoria según el sistema Bethesda, correlacionando este dato con la histología definitiva para cada nódulo biopsiado. Los parámetros de prueba diagnóstica se calcularon como prueba de *screening* (verdadero positivo: categorías IV, V, VI) y como método para identificar malignidad (verdadero positivo: categorías V, VI).

Resultados: Se incluyó a 522 pacientes, de los que 184 (35,2%) presentaron un carcinoma en la histología definitiva; siendo el carcinoma papilar el más frecuente (84,2%). Los porcentajes de malignidad en el nódulo biopsiado para cada categoría Bethesda fueron: I, 0%; II, 1,5%; III, 6,4%; IV, 31%; V, 86,5% y VI, 100%. En el análisis como prueba de *screening*, se identificó una sensibilidad del 98,9%, especificidad del 84,4%, valor predictivo positivo del 69,6%, valor predictivo negativo del 99,5% y precisión diagnóstica global del 88,2%. En el análisis para detectar malignidad, los parámetros fueron: sensibilidad 98,6%, especificidad 97,6%, valor predictivo positivo 93,5%, valor predictivo negativo 99,5% y precisión diagnóstica global 97,9%. **Conclusiones:** El sistema Bethesda es un método sencillo y reproducible en la categorización citológica del nódulo tiroideo, una herramienta útil en el manejo y eficaz para identificar el riesgo de malignidad.

© 2018 AEC. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

* Autor para correspondencia.

Correo electrónico: jmunoz@salud.madrid.org (J.L. Muñoz de Nova).

<https://doi.org/10.1016/j.ciresp.2018.02.017>

0009-739X/© 2018 AEC. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Efficiency of the Bethesda System for Thyroid Cytopathology

ABSTRACT

Keywords:

Thyroid cancer
Bethesda
Fine-needle aspiration
Cytology

Introduction: Fine-needle aspiration biopsies are a key tool for preoperative assessment of thyroid nodules, and the Bethesda system is the preferred method to report cytological analysis. The purpose of this study is to assess the efficiency of the Bethesda system to identify the malignancy risk of thyroid nodules.

Methods: Patients who underwent thyroid surgery between June 2010 and June 2017 were included. Samples were classified into 6 categories according to rates of malignancy associated with each diagnostic category. In order to investigate the correlation between categories, a statistical analysis compared the categories with pathology reports. Diagnostic indicators were calculated as a screening test (categories IV, V, VI as true-positive) and as a method to identify malignancy (V, VI as true-positive).

Results: In a series of 522 patients, we found 184 (35.2%) malignant tumours, papillary carcinoma being the most prevalent with 155 cases (84.2%). Malignant rates for diagnostic categories were: I, 0%; II, 1.5%; III, 6.4%; IV, 31%; V, 86.5%; VI, 100%. A robust correlation was identified between categories on statistical analysis. For the «screening test» analysis, sensitivity was 98.9%, specificity 84.4%, positive predictive value 69.6%, negative predictive value 99.5%, and diagnostic accuracy 88.2%. Analysing the accuracy to detect malignancy, values were: sensitivity 98.6%, specificity 97.6%, positive predictive value 93.5%, negative predictive value 99.5%, diagnostic accuracy 97.9%.

Conclusion: The Bethesda system is a clear and reliable approach to report thyroid cytology and therefore is an effective tool to identify malignancy risk and guide clinical management.

© 2018 AEC. Published by Elsevier España, S.L.U. All rights reserved.

Introducción

La enfermedad nodular tiroidea es muy frecuente en la población, sobre todo en mujeres y personas mayores, en las que puede afectar hasta al 60%^{1,2}. En las últimas décadas se ha producido un incremento significativo de la incidencia de cáncer de tiroides, sobre todo a costa de microcarcinomas^{3,4}. Así, en Estados Unidos la incidencia anual de cáncer de tiroides se ha triplicado, con un 40% para los microcarcinomas⁵.

Uno de los factores que se ha asociado a este hecho es el constante incremento en la realización de pruebas de imagen cervicales, principalmente de la ecografía tiroidea, lo que ha condicionado la identificación de un número creciente de nódulos tiroideos (NT) susceptibles de ser biopsiados mediante una punción-aspiración con aguja fina (PAAF)⁶. Con el objetivo de estandarizar la terminología que se empleaba para la descripción de la citología tiroidea, en 2007 se emitieron unas recomendaciones de consenso, conocidas como el sistema Bethesda (BS)⁷. Este sistema se basó en la creación de 6 categorías asociadas a un determinado riesgo de malignidad en cada una de ellas. Posteriores estudios han descrito los resultados tras adoptar las recomendaciones del BS, con una gran concordancia en la categorización de las muestras de PAAF⁸, si bien existe una limitación inherente al BS que es la variabilidad intra- e interobservador en el estudio citopatológico del NT⁹. Además, hay que tener en cuenta que el contexto más frecuente en la población general suele ser la existencia de un bocio multinodular: un único paciente puede presentar diversos nódulos subsidiarios de ser biopsiados mediante una PAAF

por sus características ecográficas, lo que aumenta la complejidad del proceso diagnóstico-terapéutico.

El objetivo del estudio es evaluar la validez del BS en la enfermedad nodular tiroidea para diagnosticar malignidad.

Métodos

La población en estudio estuvo compuesta por pacientes consecutivos intervenidos mediante cirugía tiroidea entre el 1 de junio de 2010 y el 30 de junio de 2017. Se incluyó solo a aquellos pacientes tratados mediante una primera intervención tiroidea cuya PAAF preoperatoria hubiera sido realizada en nuestro centro. Se excluyó a los pacientes con PAAF realizadas en otros centros, a los pacientes intervenidos sin PAAF preoperatoria o sin datos de PAAF disponibles según el BS (fig. 1).

El protocolo diagnóstico-terapéutico incluye en todos los pacientes una anamnesis, una exploración física, una analítica de la función tiroidea y una ecografía cervical. En los pacientes que cumplen criterios para realización de PAAF según las guías internacionales de la *American Thyroid Association*^{10,11}, dicho procedimiento se realiza bajo control ecográfico por un radiólogo asistido por un citólogo experto.

Las muestras se clasifican según las recomendaciones del BS⁷, agrupadas en las 6 categorías descritas originalmente: I, insatisfactorio/no diagnóstico; II, benigno; III, atípicas de significado incierto/lesión folicular de significado incierto; IV, neoplasia folicular/sospecha de neoplasia folicular; V, sospechoso de malignidad y VI, maligno. La indicación para repetir una punción se limitó a aquellos casos con categorías

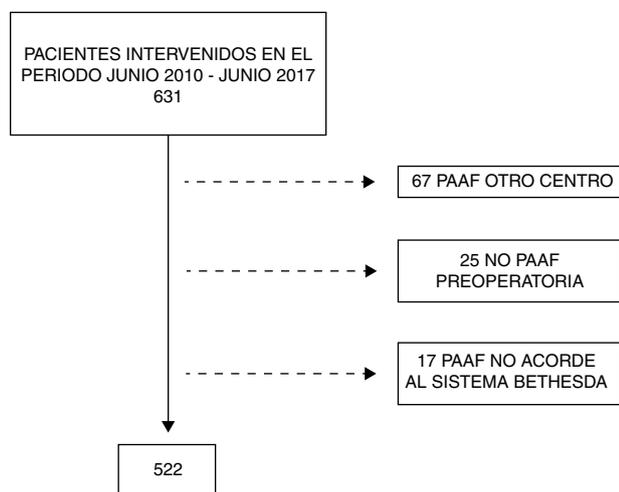


Figura 1 – Diagrama de pacientes incluidos en el estudio.

diagnósticas I y III, y ante punciones benignas, pero con un alto grado de sospecha clínico-radiológica.

Una vez realizada la punción, se indicó cirugía en aquellos pacientes con categorías IV, V y VI; en pacientes con categorías I persistentes tras repetir la punción, que presentaban un alto grado de sospecha clínico-radiológica; en pacientes con categoría III persistente tras repetir la punción o tras la punción inicial si existía un alto grado de sospecha y en pacientes con categoría II, pero que presentaban sintomatología atribuible a la nodularidad tiroidea, hiperfunción, crecimiento progresivo de los NT o si presentaban un tamaño > 4 cm en alguno de los NT. La técnica quirúrgica empleada en cada caso se basó en las características individuales del paciente, las categorías del BS y la localización de los NT. En general, se realizó una hemitiroidectomía ante la existencia de nodularidad unilateral o con nódulos contralaterales subcentimétricos ante categorías I-IV. Ante bocios multinodulares bilaterales sintomáticos, enfermedad de Graves o categorías V-VI, se optó por la realización de una tiroidectomía total.

Los datos de seguimiento de los pacientes y la correlación histológica final se realizó solo en pacientes con manejo quirúrgico. Si un paciente presentaba varias muestras de PAAF procedentes de diferentes NT, los resultados de cada punción y el correspondiente resultado histológico fue analizado por separado. Se realizó una revisión exhaustiva de cada NT evaluado, correlacionando minuciosamente la descripción de la ecografía que guió la PAAF (tamaño y localización) con los hallazgos de la pieza quirúrgica para confirmar la concordancia entre el NT biopsiado con su respectivo diagnóstico anatomopatológico definitivo.

En cuanto al diseño del estudio, se analizó una base de datos mantenida de forma prospectiva en la que se recogieron los datos diagnóstico-terapéuticos de todos los pacientes, en concreto, los datos demográficos, tamaño y localización ecográfica del NT, categoría diagnóstica de BS (en casos de múltiples punciones en el mismo paciente, se incluyó solo la categoría de mayor riesgo del BS), datos operatorios y datos anatomopatológicos.

Este estudio fue aprobado por el Comité Ético de Investigación Clínica de nuestro centro.

Análisis estadístico

El análisis estadístico se realizó mediante el programa SPSS® 23.0 para Windows (SPSS Inc., Chicago, Illinois, EE. UU.). Los resultados se expresaron como porcentajes para variables categóricas, y como media y desviación estándar para variables continuas, utilizando la mediana y el rango intercuartílico para variables con distribución asimétrica.

Se valoró la correlación entre las diferentes categorías diagnósticas comparándolas entre sí respecto al resultado histológico final, para lo que se realizó un modelo logarítmico lineal (*likelihood ratio*) y un modelo chi-cuadrado, utilizando medidas simétricas de asociación. Los datos de malignidad empleados se calcularon asignando a cada nódulo biopsiado su correspondiente diagnóstico histológico final. Se consideraron diferencias estadísticamente significativas de forma bilateral con valores de $p < 0,05$. Se utilizó la correlación phi como medida del grado de asociación entre variables categóricas, cuyos valores oscilan entre +1 y -1. Según la fuerza de asociación: -1 indica una fuerte asociación negativa, +1 indica una fuerte asociación positiva y 0 indica ausencia de asociación.

Se calcularon los parámetros de prueba diagnóstica sensibilidad, especificidad, valores predictivos (valor predictivo positivo [VPP], valor predictivo negativo [VPN]) y precisión diagnóstica para detectar malignidad mediante 2 análisis. En el análisis como prueba de *screening* (análisis I) se consideró el resultado de la PAAF como indicación de cirugía por sospecha de malignidad (categorías del BS II vs. IV, V, VI). Según este análisis, los términos «positivo» o «negativo» constituyen la existencia o no de indicación quirúrgica de cara al análisis estadístico. Las categorías I y III se excluyeron de este análisis debido a que pueden implicar la repetición de PAAF. Se realizó un segundo análisis que midió la capacidad de la prueba para detectar malignidad (análisis II) ante punciones de alta sospecha (categorías V y VI) frente a punción de benignidad (categoría II).

Resultados

En el periodo del estudio se intervino a 631 pacientes, de los que se excluyeron para el estudio 67 pacientes con PAAF realizada en otro centro, 25 pacientes intervenidos sin PAAF preoperatoria y 17 pacientes en los que el informe de la PAAF no se realizó según el BS (fig. 1). Así, de los 522 pacientes incluidos, 433 (83%) fueron mujeres, con una edad media de $51,8 \pm 16$ años. La mediana del tamaño del NT evaluado preoperatoriamente fue de 2,5 cm (1,6-4). La citología más frecuente entre los pacientes intervenidos fue la categoría II (49%), con porcentajes muy similares de pacientes intervenidos con categorías III, IV, V y VI (14,9; 13,6; 7,1 y 11,5%, respectivamente). En 316 casos (60,5%) se realizó una tiroidectomía total; el resto fueron hemitiroidectomías con itsectomías (39,5%). Se asoció una disección del compartimento central en 66 casos (12,6%).

Respecto a los resultados histológicos, se identificaron 184 neoplasias malignas (35,2%), con el carcinoma papilar como el tumor más frecuente con 155 casos (84,2%), de los cuales 42 casos eran microcarcinomas incidentales (27,1% del

Tabla 1 – Porcentajes de malignidad de las categorías diagnósticas

Categoría diagnóstica	N.º de casos (%)	Riesgo de malignidad, N.º de casos (%)	Riesgo de malignidad en NT excluyendo miCPin, N.º de casos (%)	Riesgo de malignidad en NT evaluados N.º de casos (%)
I	17 (3,3)	6 (35,3)	6 (35,3)	0 (0)
II	259 (49,6)	37 (14,3)	12 (4,6)	4 (1,5)
III	78 (14,9)	18 (23,1)	9 (11,5)	5 (6,4)
IV	71 (13,6)	30 (42,3)	24 (33,8)	22 (31)
V	37 (7,1)	33 (89,2)	32 (86,5)	32 (86,5)
VI	60 (11,5)	60 (100)	60 (100)	60 (100)

miCPin: microcarcinoma papilar incidental (n = 42).

Tabla 2 – Análisis estadístico de las categorías del sistema Bethesda

Comparación de categorías diagnósticas	Chi-cuadrado ^a	phi ^b	LR ^c	GL	p
CD I vs. CD II vs. CD III vs. CD IV vs. CD V vs. CD VI	365,84	0,84	374,37	5	< 0,001
CD II vs. CD III vs. CD IV vs. CD V vs. CD VI	352,35	0,83	365,05	4	< 0,001
CD II vs. CD VI	294,44	-0,96	278,51	1	< 0,001
CD II vs. CD V	218,66	-0,86	148,52	1	< 0,001
CD II vs. CD IV	66,55	-0,45	52,83	1	< 0,001
CD II vs. CD III	5,46	-0,13	4,52	1	0,034
CD II vs. CD I	0,27	-0,03	0,51	1	1,000
CD III vs. CD VI	119,22	-0,93	153,70	1	< 0,001
CD III vs. CD V	74,74	-0,80	78,03	1	< 0,001
CD III vs. CD IV	15,13	-0,32	15,98	1	< 0,001
CD III vs. CD I	1,15	-0,11	2,03	1	0,581
CD IV vs. CD VI	66,15	-0,71	85,30	1	< 0,001
CD IV vs. CD V	29,97	-0,53	32,52	1	< 0,001
CD IV vs. CD I	7,02	-0,28	11,07	1	0,005
CD V vs. CD VI	8,55	-0,30	10,08	1	0,007
CD V vs. CD I	36,09	-0,82	43,69	1	< 0,001
CD VI vs. CD I	77,00	-1,00	81,29	1	< 0,001

CD: categoría diagnóstica; GL: grados de libertad; LR: likelihood ratio.

^a Modelo chi-cuadrado para evaluar asociación entre variables categóricas.

^b Coeficiente de correlación phi utilizado para evaluar la fuerza de asociación en variables categóricas.

^c Likelihood ratio (modelo log-lineal) utilizado para evaluar asociación en variables categóricas.

total de los carcinomas papilares y 8% del total de los pacientes intervenidos). El resto de las neoplasias identificadas fueron 19 carcinomas foliculares (10,3%), 8 carcinomas medulares (4,3%), un carcinoma anaplásico (0,5%) y un linfoma tiroideo (0,5%).

En cuanto a los porcentajes de malignidad en las diferentes categorías del BS, una vez excluidos los microcarcinomas incidentales, las tasas de malignidad para las categorías II, III, IV, V y VI fueron del 4,6; 11,5; 33,8; 86,5 y 100%, respectivamente (tabla 1). En la categoría I, la tasa de malignidad fue del 35,3%, pero en ningún caso se debió al nódulo biopsiado preoperatoriamente, mientras que en el global de la serie el 86,4% de los pacientes sí presentó el tumor sobre el nódulo que había sido biopsiado de forma preoperatoria. Así, las tasas de malignidad atribuible al nódulo biopsiado para las categorías II, III, IV, V y VI fueron del 1,5; 6,4; 31; 86,5 y 100%, respectivamente. Al analizar las diferencias entre los porcentajes de malignidad en cada una de las diferentes categorías, hemos encontrado una fuerte correlación en la práctica totalidad de las comparaciones (tabla 2). Únicamente no se detectaron diferencias estadísticamente significativas entre las categorías I y II ($p = 1,000$) y entre las categorías I y III ($p = 0,581$).

Tabla 3 – Parámetros de prueba diagnóstica del sistema Bethesda

Parámetro	Análisis I (%) ^a	Análisis II (%) ^b
Sensibilidad	98,9 (87/88)	98,6 (72/73)
Especificidad	84,4 (205/243)	97,6 (205/210)
VPP en CD VI	100 (43/43)	100 (43/43)
VPP en CD V	85,3 (29/34)	85,3 (29/34)
VPP en CD IV	31,3 (15/48)	—
VPP	69,6 (87/125)	93,5 (72/77)
VPN	99,5 (205/206)	99,5 (205/206)
Tasa de falsos negativos	1,1 (1/88)	1,4 (1/73)
Tasa de falsos positivos	15,6 (38/243)	2,4 (5/210)
Precisión diagnóstica	88,2 (292/331)	97,9 (277/283)

^a Considera casos en CD IV + V + VI verdaderos positivos y casos en CD II verdaderos negativos.

^b Considera casos en CD V + VI verdaderos positivos y casos en CD II verdaderos negativos.

Con relación al rendimiento del BS, al analizar su utilidad como prueba de screening (análisis I: categoría II vs. IV + V + VI) encontramos una sensibilidad para detectar malignidad del

98,9%, con una especificidad del 84,4%, un VPP del 69,6%, un VPN del 99,5% y una precisión diagnóstica global del 88,2% (tabla 3). Este análisis, en punciones altamente sospechosas (análisis II: categoría II vs. V + VI), aumentó la precisión global de la prueba hasta el 97,9% (sensibilidad 98,6%, especificidad 97,6%, VPP 93,5% y VPN 99,5%).

Discusión

El método más utilizado para la descripción y categorización de las muestras de PAAF tiroidea es el BS^{8,10}. Este se basa en 6 categorías para cada una de las cuales existe un riesgo estimado de cáncer de tiroides¹². Nuestro trabajo busca revisar este riesgo, confrontando los hallazgos citológicos con el único *gold standard* posible: el estudio histológico definitivo en el paciente intervenido mediante cirugía tiroidea. Las tasas de malignidad en las que se ha basado el estudio estadístico no tienen en cuenta los microcarcinomas incidentales, ya que la mayoría de ellos tendrán un curso clínico indolente. Además, hemos considerado solo la existencia de malignidad sobre el nódulo biopsiado, ya que buscamos definir la capacidad del estudio citológico para identificar la malignidad, no la de la selección ecográfica del nódulo que biopsiar. A pesar de esto, queremos reseñar que solo 20 pacientes (3,8%) presentaron un tumor >1 cm en uno de los nódulos no biopsiados.

Nuestro análisis muestra la existencia de una fuerte correlación para cada una de las categorías del BS a la hora de identificar malignidad, con diferencias entre casi todas ellas (tabla 2). Esto podría justificar el mantenimiento de las 6 categorías, tal y como ha sucedido en la reciente revisión del BS¹³. Los porcentajes de malignidad observados para cada categoría se encontraron, en general, dentro de los límites descritos^{9,10}.

Solo un 3,3% de los pacientes intervenidos presentaban una categoría I, cifra muy por debajo de los datos ofrecidos por otros autores (7-26%^{8,14}). Creemos que esta cifra se basa en que todas las PAAF se han realizado guiadas por ecografía, evitando la aspiración, al tiempo que un citólogo evaluaba *in situ* la calidad del material obtenido. El porcentaje de malignidad para el nódulo biopsiado dentro de esta categoría fue del 0%, lo cual representa un porcentaje ideal si se tiene en cuenta que los resultados de esta categoría inciden en la importancia de la obtención de material satisfactorio para el análisis citológico.

La tasa de malignidad asociada a la categoría III fue 6,4%, similar a la propuesta originalmente⁷. Si bien trabajos posteriores presentaron tasas de hasta el 48% para esta categoría, esto puede atribuirse a la selección de pacientes quirúrgicos y a la inclusión de neoplasias incidentales en el análisis¹⁵. Para tratar de valorar esta amplia variación de malignidad dentro de la categoría III, se recomienda valorar el cociente entre pacientes categoría III y categoría VI¹⁶, cuyo valor ideal deberá encontrarse entre 1 y 3. Valores por encima de 3 indicarían una sobreutilización de la categoría III, mientras que valores inferiores a 1 se deberían a una baja utilización de esta categoría, con el consecuente riesgo de pérdida de sensibilidad para detectar malignidad. En nuestra serie, este cociente fue 1,3, lo que nos acerca a la parte más eficiente de la horquilla recomendada.

Respecto a los parámetros de prueba diagnóstica del BS, al valorarlo como prueba de *screening*, es decir, aquella punción que indica una intervención quirúrgica (tabla 3), se objetivó una sensibilidad del 98,9% y un VPN del 99,5%. Estos datos son similares a los detallados en el trabajo de Bongiovanni et al.⁹ en cuanto a la sensibilidad, si bien nuestro VPN mejora ampliamente el promedio indicado en dicho estudio (99,5 vs. 47%). Este último dato es de especial importancia, ya que el principal objetivo de la PAAF preoperatoria es descartar la existencia de malignidad con el objeto de disminuir el número de cirugías innecesarias por este motivo.

En la segunda parte del análisis de los parámetros de prueba diagnóstica del BS hemos considerado su capacidad para asegurar la existencia de malignidad (categorías V y VI), obteniendo en este caso una especificidad del 97,6% y un VPP del 93,5%, con una precisión global del 97,9%. Estos datos permiten definir al BS como una herramienta muy fiable cuando se trata de confirmar la existencia de malignidad.

Una de las limitaciones del presente estudio son las inherentes al análisis citológico. Los patólogos deben mantenerse alerta ante la posibilidad de más errores en el análisis de lesiones quísticas, bocio multinodular o solapamiento de lesiones con características citomorfológicas similares, como la presencia de células foliculares reactivas o lesiones con células de Hürthle. Estos hallazgos aparecen sobre todo en el contexto de las categorías I y III, precisamente las implicadas en las únicas comparaciones entre categorías del BS que no mostraron diferencias significativas. No obstante, consideramos que el bajo número de pacientes que conformaron la categoría I limita la capacidad para detectar diferencias.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

BIBLIOGRAFÍA

- Burman KD, Wartofsky L. Clinical practice. Thyroid nodules. *N Engl J Med.* 2015;373:2347-56.
- Guth S, Theune U, Aberle J, Galach A, Bamberger CM. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *Eur J Clin Invest.* 2009;39:699-706.
- Davies L, Morris LG, Haymart M, Chen AY, Goldenberg D, Morris J, et al. American Association of Clinical Endocrinologists and American College of Endocrinology disease state clinical review: The increasing incidence of thyroid cancer. *Endocr Pract.* 2015;21:686-96.
- Vaccarella S, Franceschi S, Bray F, Wild CP, Plummer M, Dal Maso L. Worldwide thyroid-cancer epidemic? The increasing impact of overdiagnosis. *N Engl J Med.* 2016;375:614-7.
- Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin.* 2014;64:9-29.
- Ross DS. Predicting thyroid malignancy. *J Clin Endocrinol Metabol.* 2006;91:4253-5.
- Ali SZ, Cibas ES. The Bethesda system for reporting thyroid cytopathology. Definitions criteria and explanatory notes. Nueva York: Springer. 2010.

8. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda system for reporting thyroid cytopathology: A meta-analysis. *Acta Cytol.* 2012;56:333-9.
9. Cibas ES, Baloch ZW, Fellegara G, Livolsi VA, Raab SS, Rosai J, et al. A prospective assessment defining the limitations of thyroid nodule pathologic evaluation. *Ann Intern Med.* 2013;159:325-32.
10. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al., 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer. The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid.* 2016;26:1-133.
11. Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid.* 2009;19:1167-214.
12. Ali SZ. Thyroid cytopathology: Bethesda and beyond. *Acta Cytol.* 2011;55:4-12.
13. Cibas ES, Ali SZ. The 2017 Bethesda system for reporting thyroid cytopathology. *Thyroid.* 2017;27:1341-6.
14. Theoharis CG, Schofield KM, Hammers L, Udelsman R, Chhieng DC. The Bethesda thyroid fine-needle aspiration classification system: Year 1 at an academic institution. *Thyroid.* 2009;19:1215-23.
15. Iskandar ME, Bonomo G, Avadhani V, Persky M, Lucido D, Wang B, et al. Evidence for overestimation of the prevalence of malignancy in indeterminate thyroid nodules classified as Bethesda category III. *Surgery.* 2015;157:510-7.
16. Krane JF, Vanderlaan PA, Faquin WC, Renshaw AA. The atypia of undetermined significance/follicular lesion of undetermined significance malignant ratio: A proposed performance measure for reporting in the Bethesda system for thyroid cytopathology. *Cancer Cytopathol.* 2012;120:111-6.