

ORIGINAL

Propuesta para mejorar la sensibilidad del algoritmo diagnóstico de evaluación de la función tiroidea

Ruth Cano Corres*, Ana Argudo Ramírez, Javier Sánchez Álvarez, Pedro Alía Ramos y Ariadna Padró Miquel

Área de Bioquímica y Biología Molecular, Laboratori Clínic, Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat, Barcelona, España

Recibido el 5 de octubre de 2010; aceptado el 31 de enero de 2011

Disponible en Internet el 7 de abril de 2011

PALABRAS CLAVE

Tirotropina (TSH);
Función tiroidea;
Algoritmo diagnóstico

Resumen La medida en suero de la concentración de tirotropina y tiroxina es la base para la evaluación bioquímica de la función tiroidea. Con frecuencia, el intervalo de referencia de la tirotropina sirve como cribado inicial para valorar la necesidad de añadir la medida de tiroxina. Este trabajo se ha realizado con el objetivo de mejorar la sensibilidad diagnóstica del cribado.

Se seleccionaron todos los resultados de tirotropina y tiroxina solicitados de manera simultánea a pacientes de consultas externas: para la primera parte del estudio se usaron los del año 2008 ($n = 10.900$) y para la segunda parte, los de pacientes del año 2009 sin seguimiento en el año previo ($n = 5.367$). Se realizaron dos curvas ROC para delimitar el intervalo de decisión del algoritmo con una sensibilidad del 90% y se contabilizó el número de resultados falsos negativos obtenidos.

Los intervalos de tirotropina obtenidos en el primer y segundo estudio fueron (2,11-3,50) mIU.u./L y (2,04-3,41) mIU.u./L respectivamente. En ambos estudios la sensibilidad aumentó aproximadamente de un 70% de media con el intervalo de referencia a un 90% con el intervalo del algoritmo. El número de falsos negativos se redujo de 75 a 30 en el primer caso, y de 37 a 13 en el segundo.

La aplicación de un intervalo de tirotropina calculado para la evaluación de la función tiroidea, en pacientes ambulatorios con o sin seguimiento previo, supone un aumento en la sensibilidad diagnóstica, respecto al empleo del intervalo de referencia de tirotropina.

© 2010 AEBM, AEFA y SEQC. Publicado por Elsevier España, S.L. Todos los derechos reservados.

* Autor para correspondencia.

Correo electrónico: rcano@bellvitgehospital.cat (R. Cano Corres).

KEYWORDS

Thyrotropin (TSH);
Thyroid function;
Diagnostic algorithm

Proposal to improve the sensitivity of the diagnostic algorithm for thyroid function evaluation

Abstract The measurement of thyrotropin and thyroxine concentrations in serum is the basis of the biochemical evaluation of thyroid function. The reference interval of thyrotropin is frequently used as an initial screening to assess the need for thyroxine measurement. This study was carried out to obtain a different and more adjusted interval of thyrotropin, in order to improve the diagnostic sensitivity.

All of the results of thyrotropin and thyroxine requested at the same time on outpatients were selected: for the first part of the study, those from year 2008 ($n = 10,900$), and for the second part, those from 2009 with no follow-up in the previous year ($n = 5,367$). Two ROC curves were used to define the algorithm decision interval with a sensitivity of 90% and the number of false negative results was calculated.

The thyrotropin intervals obtained in the first and second studies were (2.11-3.50) mIU/L and (2.04-3.41) mIU/L, respectively. In both studies, the sensitivity increased approximately from an average of 70% to 90% of the confidence interval using the algorithm interval. The number of false negatives was reduced from 75 to 30 in the first case, and from 37 to 13 in the second case.

The application of a calculated thyrotropin interval to assess thyroid function in outpatients with or without prior monitoring, leads to an increase of the diagnostic sensitivity with regard to the use of the thyrotropin reference interval.

© 2010 AEBM, AEFA y SEQC. Published by Elsevier España, S.L. All rights reserved.

Introducción

El papel de las hormonas tiroideas es esencial para el desarrollo y para la regulación del metabolismo. Las acciones más relevantes que ejercen son: estimular la síntesis proteica y la termogénesis, favorecer la glucólisis y la lipólisis, y aumentar la frecuencia y el gasto cardíaco. Su secreción se lleva a cabo por la glándula tiroidea y está regulada por un mecanismo que implica a la adenohipófisis y al hipotálamo. En las neuronas del núcleo paraventricular del hipotálamo se produce y secreta una hormona denominada tiroliberina (TRH), que llega hasta la adenohipófisis donde promueve la secreción de otra hormona, la tirotrópina (TSH). La TSH actúa sobre el tiroides, donde se produce la secreción de las hormonas tiroideas, triyodotironina o T3 y tiroxina o T4, que son las que ejercerán sus efectos sobre los diferentes órganos y tejidos. Altas concentraciones de T3 y T4 son capaces de inhibir la secreción de TSH y TRH cerrando un ciclo de retroalimentación negativa.

Entre las concentraciones plasmáticas de TSH y T4 existe una relación logarítmica-lineal, según la cual un cambio en una unidad de T4 conlleva un cambio de diez unidades de TSH. Por otra parte, la T4 es la que mejor refleja las concentraciones de hormona tiroidea disponible en suero, aunque la T3 sea la hormona biológicamente activa. En concreto se determina la concentración de T4 no unida a proteína, también denominada T4 libre, ya que proporciona mayor eficiencia diagnóstica en la detección de alteraciones tiroideas en los pacientes con anomalías en las proteínas de transporte¹.

Inicialmente, para la evaluación de la función tiroidea se determinaban las concentraciones de tirotrópina y tiroxina y las de triyodotironina sólo en casos de discrepancia entre ambas. La mejora en los métodos de medida de la tirotrópina² llevó a los especialistas a plantearse si la medida

de la concentración de tirotrópina en suero tenía la suficiente sensibilidad diagnóstica como para utilizarse como prueba inicial de cribado para la evaluación de la función tiroidea en la población general. Así se eliminaría la medida de T4 cuando fuese innecesaria.

Actualmente, en muchos laboratorios la determinación de la concentración de tiroxina está sujeta a que la concentración de tirotrópina se halle fuera de su intervalo de referencia³. Debido a la elevada variabilidad biológica interindividual de la tirotrópina⁴, su intervalo de referencia es muy amplio. Además, la tirotrópina pierde su valor diagnóstico en la enfermedad hipofisaria, ya que el hipotiroidismo central o secundario puede cursar con concentraciones de tiroxina bajas y de tirotrópina dentro de su intervalo de referencia⁵. La discordancia entre los valores de tiroxina y tirotrópina se debe a la reducida actividad biológica de la tirotrópina, que sigue siendo inmunológicamente activa^{6,7}.

Por ello, la decisión de medir la concentración de tiroxina basándose únicamente en el intervalo de referencia de la tirotrópina podría conllevar una baja sensibilidad. El presente estudio se ha realizado con el fin de obtener un intervalo de tirotrópina independiente del intervalo de referencia, que mejore su sensibilidad diagnóstica en el cribado de la población general para el estudio de la función tiroidea.

Material y métodos

La primera parte del estudio se dirigió a calcular un intervalo de decisión para la concentración de TSH que sirviera como base del algoritmo que se pretendía establecer. Para ello, se seleccionaron todos los resultados de los pacientes de consultas externas del Hospital Universitari de Bellvitge

a los que se les hubiese solicitado la determinación de tirotrópina y tiroxina no unida a proteína de manera simultánea durante el año 2008 ($n=10.900$). Los resultados de tirotrópina se separaron en dos grupos en función de si eran mayores o menores que el valor arbitrario de 2,42 mU/L (media del intervalo de referencia de la tirotrópina en nuestro laboratorio). Se realizaron dos curvas ROC, una para cada grupo, para delimitar los valores superior e inferior del intervalo de decisión del algoritmo con una sensibilidad del 90%. Los resultados de tirotrópina que se hallaron dentro de dicho intervalo, pero cuya concentración de tiroxina se situó fuera de su intervalo de referencia se consideraron falsos negativos. Se calculó el número de resultados falsos negativos obtenidos, y se verificó el algoritmo con datos de 2009 ($n=10.896$).

En la segunda parte del estudio, se intentó mimetizar la situación de primera visita o evaluación inicial de la función tiroidea. Con este fin, de todos los resultados de tirotrópina y tiroxina no unida a proteína solicitados simultáneamente durante los años 2008 y 2009, se seleccionaron sólo los de pacientes con un único resultado en 2009 y sin resultados durante 2008, es decir, sin seguimiento en el año previo ($n=5.367$). Se establecieron los límites del intervalo de decisión del algoritmo con una sensibilidad del 90%, de igual modo que en la primera parte del estudio.

Los intervalos de referencia de la tirotrópina y tiroxina no unida a proteína en nuestro laboratorio son 0,48-4,36 mU/L y 9,7-30,9 pmol/L respectivamente. Se trata de intervalos de referencia de producción propia obtenidos a partir de muestras de población sana representativas del área territorial del hospital.

Las medidas de tirotrópina y tiroxina no unida a proteínas en suero se realizaron en el analizador MODULAR E170 (Roche Diagnostics®) mediante inmunoanálisis electroquimioluminiscente heterogéneo con micropartículas recubiertas de estreptavidina.

Para el tratamiento de los datos se utilizó el programa Microsoft Excel y el programa estadístico Analyse-it.

Resultados

Estudio 1

En las figuras 1 y 2 se muestran las curvas ROC mediante las que se estimaron los valores discriminantes con una sensibilidad del 90% (tabla 1). El intervalo delimitado por estos valores es 2,11-3,50 mU/L. En 2008, de las 10.900 muestras analizadas, 2.592 se encontraron dentro de dicho intervalo y de ellas 30 fueron falsos negativos, con resultados por debajo del límite inferior del intervalo de referencia de la tiroxina: la mediana y el intervalo intercuartílico de la concentración de tiroxina de estos 30 casos fue 8,85 (8,25-9,42) pmol/L. Además, la aplicación del algoritmo supondría una disminución de un 24% (2.592/10.900) en el número de determinaciones de T4 solicitadas junto a la TSH en pacientes de consultas externas.

Con el fin de verificar el algoritmo, se aplicó a datos del 2009. Se obtuvieron 33 falsos negativos, que presentaron una mediana de la concentración de tiroxina de 9,2 pmol/L (8,6-9,5 pmol/L). Estos resultados son equiparables a los de 2008.

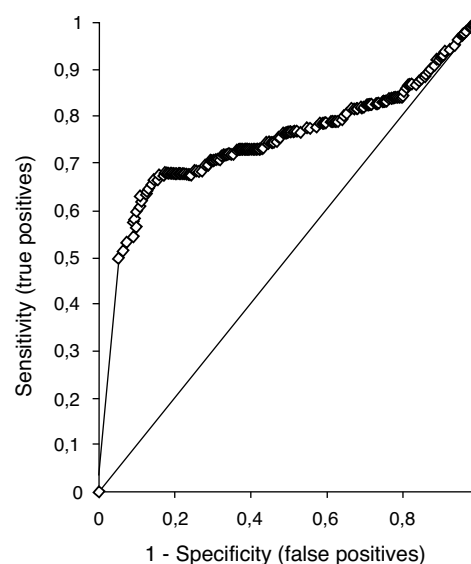


Figura 1 Curva ROC que determina el límite inferior del intervalo de decisión del algoritmo para el total de pacientes de consultas externas.

El número de falsos negativos que se hubiese obtenido aplicando el intervalo de referencia de TSH a los datos de 2008 en vez del obtenido en nuestro estudio, fue 75. Pese a que esto supondría una reducción de un 73% (7.951/10.900) en el número de determinaciones de T4 solicitadas conjuntamente con TSH en el ámbito ambulatorio, debe considerarse que la sensibilidad disminuye hasta el 75% de media (tabla 1).

Teniendo en cuenta el consumo total de reactivo de tiroxina no unida a proteína de nuestro hospital en el año 2008 —incluyendo también el gasto por controles de calidad, calibraciones y peticiones procedentes de pacientes ingresados— el ahorro en determinaciones de tiroxina aplicando el algoritmo con el intervalo de referencia habría sido

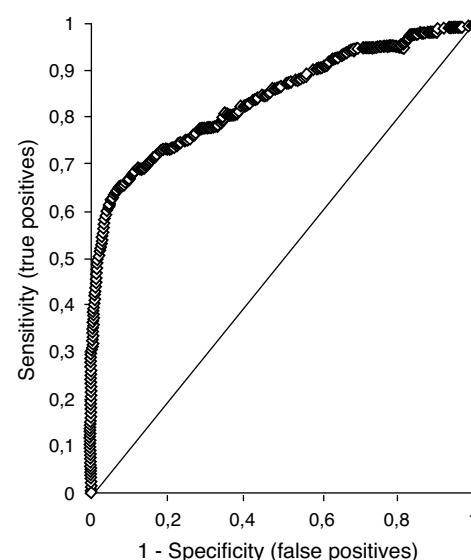


Figura 2 Curva ROC que determina el límite superior del intervalo de decisión del algoritmo para el total de pacientes de consultas externas.

Tabla 1 Comparación de la aplicación del intervalo calculado en el estudio 1 y del intervalo de referencia de TSH a los datos de pacientes de consultas externas del año 2008

	Concentración de tirotropina (mint.u./L)	Sensibilidad límite inferior	Sensibilidad límite superior	% Reducción de determinaciones de T4
Intervalo calculado	2,11-3,50	90%	90%	24%
Intervalo referencia	0,48-4,36	68%	82%	73%

% Reducción de determinaciones de T4: número de determinaciones de T4 solicitadas junto a TSH del ámbito ambulatorio no procedentes.

del 43% (7951/18373), mientras que utilizando los intervalos con una sensibilidad del 90% habría sido del 14% (2.592/18.373).

Estudio 2

Durante el año 2009 hubo 5.367 muestras sin seguimiento en el año previo. Siguiendo el mismo esquema metodológico del primer estudio, se calculó el intervalo del algoritmo con una sensibilidad del 90% mediante dos curvas ROC (figs. 3 y 4). El intervalo es 2,04-3,41 mint.u/L. En el año 2009, 1.539 muestras cumplieron las condiciones del algoritmo. De ellas, 13 fueron falsos negativos, con una mediana de la concentración de tiroxina no unida a proteína de 9,2 (8,6-9,4) pmol/L. La aplicación del algoritmo permitiría suprimir un 29% (1.539/5.367) de las determinaciones de tiroxina solicitadas junto a tirotropina a pacientes sin seguimiento previo.

Si se hubiese aplicado el intervalo de referencia de TSH, se habrían obtenido 37 falsos negativos. Aunque esto permitiría reducir en un 81% (4.330/5.367) el número de determinaciones de tiroxina, supondría una disminución de la sensibilidad de hasta un 71% de media (tabla 2).

Discusión

La aplicación de un intervalo de tirotropina de producción propia para la evaluación de la función tiroidea, en pacientes ambulatorios con o sin seguimiento previo, supone un aumento en la sensibilidad diagnóstica respecto al empleo del intervalo de referencia de tirotropina con este fin: de un 73% de media a un 90%. Por este motivo, aunque la disminución del número de determinaciones de T4 es mayor con la aplicación del intervalo de referencia de la tirotropina, el número de falsos negativos que origina es más del doble que con la aplicación del intervalo ajustado (75 frente a 30 en el primer caso y 37 frente a 13 en el segundo).

El aumento de una unidad de sensibilidad conlleva aproximadamente un coste adicional del 1,7% del presupuesto total destinado a la determinación de tiroxina no unida a proteína en nuestro centro.

Es tarea de cada hospital determinar la relación coste-eficiencia del algoritmo, decidiendo cuánto puede disminuir la sensibilidad del algoritmo con el consiguiente aumento del ahorro, pero sin comprometer los estándares de calidad que tenga establecidos. En todo caso, el estudio de la sensibilidad del algoritmo es imprescindible, puesto que la decisión de cuánto se restringe el intervalo debe basarse en este dato, con independencia de si coincide o no con el intervalo

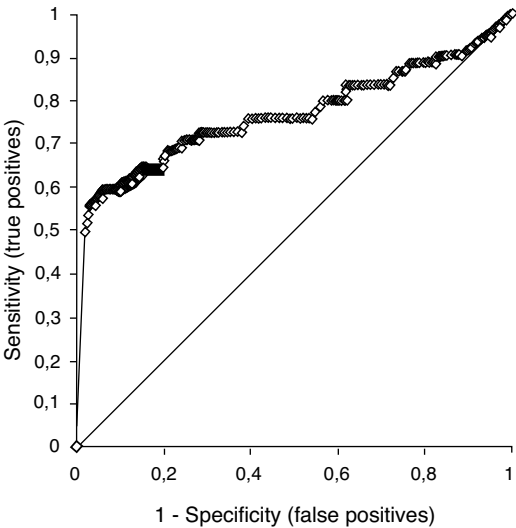


Figura 3 Curva ROC que determina el límite inferior del intervalo de decisión del algoritmo para pacientes ambulatorios sin seguimiento tiroideo previo.

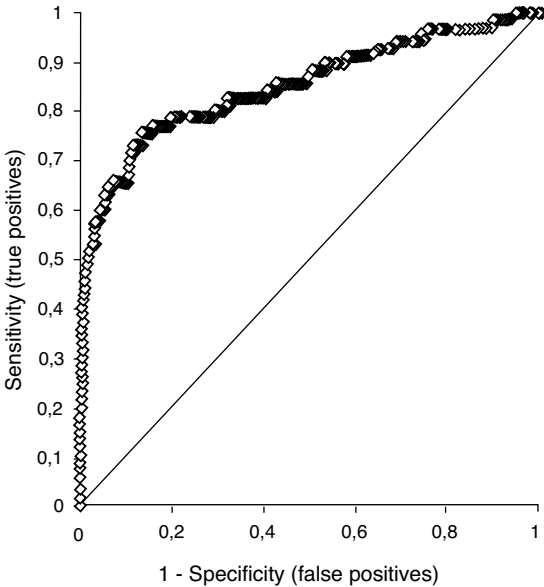


Figura 4 Curva ROC que determina el límite superior del intervalo de decisión del algoritmo para pacientes ambulatorios sin seguimiento tiroideo previo.

Tabla 2 Comparación de la aplicación del intervalo calculado en el estudio 2 y el intervalo de referencia de TSH a los datos de pacientes de consultas externas sin seguimiento previo del año 2009

	Concentración de tirotrópina (mIU./L)	Sensibilidad límite inferior	Sensibilidad límite superior	% Reducción de determinaciones de T4
Intervalo calculado	2,04-3,41	90%	90%	29%
Intervalo referencia	0,48-4,36	59%	83%	81%

% Reducción de determinaciones de T4: número de determinaciones de T4 solicitadas junto a TSH del ámbito ambulatorio no procedentes.

de referencia. Con la aplicación de un algoritmo calculado similar a los propuestos en los estudios 1 y 2 se reduciría anualmente alrededor de un 25% las determinaciones de tiroxina de consultas externas solicitadas conjuntamente con tirotrópina en nuestro hospital, conociendo exactamente la sensibilidad diagnóstica del algoritmo. Esto supone una optimización de los recursos del laboratorio para la evaluación de la función tiroidea.

Este estudio se planteó únicamente en pacientes de seguimiento ambulatorio, ya que son ampliamente conocidas las alteraciones que se dan en las concentraciones de tirotrópina y tiroxina en los pacientes ingresados. Esta situación es conocida como enfermedad general no tiroidea⁸.

Se decidió llevar a cabo el segundo estudio siguiendo la hipótesis de que la aplicación del algoritmo en la valoración inicial de la función tiroidea, supondría la obtención de un intervalo más amplio y próximo al intervalo de referencia, que el obtenido con todos los pacientes de consultas externas. Sin embargo, encontramos resultados muy parecidos en ambos enfoques. Una limitación del segundo estudio es que sólo se pudo hacer una simulación de la situación de valoración inicial de la función tiroidea. Por otro lado, cabe destacar el elevado número de pacientes que participaron (n=5.367).

La verificación del algoritmo aplicado a todos los pacientes de consultas externas con los datos del año 2009, permite suponer que el intervalo de decisión del algoritmo permanecerá vigente mientras no se modifique el intervalo de referencia de la concentración de tiroxina en suero.

En diversos artículos publicados se utiliza el intervalo de referencia de tirotrópina para discernir cuando es necesario medir la concentración de tiroxina no unida a proteína^{5,9}. Wardle et al⁵ emplearon el intervalo de referencia de TSH para clasificar a los pacientes según su función tiroidea y como en nuestro caso, encontraron también falsos negativos debido a la presencia de enfermedad hipofisaria. Riesco et al⁹, partían de unos valores de decisión que en un inicio se aplicaban en su laboratorio. Realizaron un estudio aplicando el intervalo de referencia más estrecho que el delimitado por los valores iniciales, y concluyeron que es mejor aplicar los valores de referencia, pero no explican de qué manera se obtuvieron los valores iniciales ni la sensibilidad diagnóstica en ninguno de los casos. En contraposición, el presente trabajo propone establecer de antemano la sensibilidad diagnóstica para elaborar un intervalo discriminante.

En los dos estudios que aquí se presentan se obtuvieron falsos negativos (TSH dentro del intervalo del algoritmo y T4 fuera de su intervalo de referencia), 30 en el primero y 13 en el segundo, que se correspondieron a situaciones pato-

lógicas concretas: epilepsia tratada con anticonvulsivos, neoplasias, esclerosis múltiple e hipotiroidismo secundario.

Respecto a los antiepilepticos, se ha descrito que la carbamacepina aumenta la velocidad del metabolismo de las hormonas tiroideas por inducción del citocromo hepático P450. La fenitoína desplaza la unión de hormonas tiroideas a proteínas, sobre todo de la globulina enlazante de tiroxina, dando lugar a una disminución del 20-40% de tiroxina, y en menor grado de T3¹⁰. En el caso de las neoplasias, se ha visto que algunos de los tratamientos empleados, como los inhibidores de tirosina quinasa, producen hipotiroidismo y disfunción tiroidea¹¹. Y finalmente, en un estudio realizado en población española se demostró que los pacientes con esclerosis múltiple no tratada presentaban mayor prevalencia de alteraciones tiroideas que la población general¹². También se ha asociado el tratamiento con interferón tipo I, empleado en la esclerosis múltiple y en tumores sólidos, al desarrollo de alteraciones tiroideas¹³.

De los 75 casos de falsos negativos que se hubieran hallado en caso de aplicar el intervalo de referencia en el algoritmo, un 80% pertenecen a los cuatro grupos previamente descritos, mientras que el 20% restante (n=15) presentan un contexto clínico muy variado.

En una situación ideal, se podría conseguir una sensibilidad cercana al 100% siempre que se evitara la aplicación del algoritmo a los pacientes que son causa de falsos negativos. Sin embargo, al aplicar el intervalo más restrictivo, se observó que las peticiones de dichos pacientes provenían de 10 servicios distintos de consultas externas, mientras que si ampliáramos el intervalo hasta el de referencia, habría 17 servicios solicitantes distintos. Si a esto se une la habitual escasez o incluso carencia de los datos clínicos que acompañan las solicitudes, resulta del todo impracticable intentar alcanzar una sensibilidad del 100% ni seleccionando por servicio ni por orientación clínica. Por consiguiente, a excepción del servicio de endocrinología, por su evidente experiencia en la materia, el algoritmo debería ser aplicado a todos los pacientes de consultas externas. En dicho caso, con nuestra propuesta la sensibilidad bajaría únicamente hasta el 90%, siendo la mediana de la concentración de tiroxina no unida a proteína de estos falsos negativos, muy próxima al límite inferior de su intervalo de referencia.

En resumen, dado que el laboratorio no dispone de las herramientas necesarias para sospechar los falsos negativos de antemano, el conocimiento de la sensibilidad del algoritmo es imprescindible para la toma de cualquier decisión acerca del equilibrio entre ahorro y sensibilidad.

Empleando el intervalo de referencia de tirotrópina para identificar cuando es necesaria la medida de la concentración de tiroxina no unida a proteína en la población general,

se obtiene un elevado número de falsos negativos. Por ello, proponemos que cada laboratorio estime un intervalo de tirotropina de elaboración propia con la sensibilidad diagnóstica deseada, para reducir el número de falsos negativos al coste económico que considere asumible.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

1. Nelson JC, Wilcox RB. Analytical performance of free and total thyroxine assays. *Clin Chem*. 1996;42:146–54.
2. García C, Rodríguez J, Berlanga E. Exploración bioquímica de la función tiroidea. Ed. Comité de Publicaciones de la Sociedad Española de Bioquímica Clínica y Patología Molecular. 2008. p. 39-40.
3. Ladenson PW, Singer PA, Ain KB, Bagchi N, Bigos ST, Levy EG, et al. American Thyroid Association Guidelines for Detection of Thyroid Dysfunction. *Arch Intern Med*. 2000;160:1573–5.
4. Nagayama I, Yamamoto K, Saito K, Kuzuya T, Saito T. Subject-based reference values in thyroid function tests. *Endocr J*. 1993;40:557–62.
5. Wardle CA, Fraser WD, Squire CR. Pitfalls in the use of thyrotropin concentration as a first-line thyroid-function test. *Lancet*. 2001;357:1013–4.
6. Beck-Peccoz P, Persani L. Variable biological activity of thyroid-stimulating hormone. *Eur J Endocrinol*. 1994;131:331–40.
7. Beck-Peccoz P, Amr S, Menezes-Ferreira MM, Faglia G, Weintraub BD. Decreased receptor binding of biologically inactive thyrotropin in central hypothyroidism. Effect of treatment with thyrotropin-releasing hormone. *N Engl J Med*. 1985;312:1085–90.
8. Warner MH, Beckett GJ. Mechanisms behind the non-thyroidal illness syndrome: an update. *J Endocrinol*. 2010;205:1–13.
9. Riesco Prieto M, Bennassar Barceló A, Pérez Esteban G, Ribes Vallés JL. Evaluation of a diagnostic protocol of thyroid dysfunction. *Aten Primaria*. 1997;20:355–60.
10. Verrotti A, Scardapane A, Manco R, Chiarelli F. Antiepileptic drugs and thyroid function. *J Pediatr Endocrinol Metab*. 2008;215:401–8.
11. Torino F, Corsello SM, Longo R, Barnabei A, Gasparini G. Hypothyroidism related to tyrosine kinase inhibitors: an emerging toxic effect of targeted therapy. *Nat Rev Clin Oncol*. 2009;6:219–28.
12. Munteis E, Cano JF, Flores JA, Martínez-Rodríguez JE, Miret M, Roquer J. Prevalence of autoimmune thyroid disorders in a Spanish multiple sclerosis cohort. *Eur J Neurol*. 2007;14:1048–52.
13. Monzani F, Caraccio N, Dardano A, Ferrannini E. Thyroid autoimmunity and dysfunction associated with type I interferon therapy. *Clin Exp Med*. 2004;3:199–210.