

ORIGINAL BREVE

Concordancia interobservador en la detección de bocio por palpación en población escolar

A. González-González^{a,*}, A. Muñoz-Serrano^b, J.M. Tenías-Burillo^c, P. Falero-Gallego^b, E. García-Fuentes^d, C. Gutierrez-Repiso^d y R. Cañete^e

^a Sección de Endocrinología y Nutrición, Hospital General Universitario de Ciudad Real, Ciudad Real, España

^b Servicio de Pediatría, Hospital General La Mancha Centro, Alcázar de San Juan, Ciudad Real, España

^c Unidad de Investigación, Hospital General La Mancha Centro, Alcázar de San Juan, Ciudad Real, España

^d Servicio de Endocrinología y Nutrición, Hospital Regional Universitario Carlos Haya, CIBEROBN, Instituto de Investigación Biomédica de Málaga (IBIMA), Málaga, España

^e Unidad de Endocrinología Pediátrica, Hospital Universitario Reina Sofía (IMIBIC), Córdoba, España

Recibido el 29 de septiembre de 2012; aceptado el 23 de diciembre de 2012

Disponible en Internet el 18 de febrero de 2013

PALABRAS CLAVE

Bocio;
Concordancia;
Interobservador

Resumen

Objetivo: Estimar la concordancia interobservador en la detección de bocio por exploración física en población escolar.

Metodología: Se realizó un estudio transversal para la detección de bocio en una muestra representativa de 1.134 escolares de 6 a 12 años de 20 colegios del área de salud de La Mancha Centro. La exploración fue realizada de forma enmascarada por 2 únicos observadores. Se establecieron 5 grados de tamaño tiroideo (0, Ia, Ib, II y III). Se consideró bocio a partir del grado Ia. La concordancia se valoró en relación con variables como edad, sexo, índice de masa corporal, talla y día de exploración. Se utilizó el índice kappa como medida de concordancia.

Resultados: En los 1.097 escolares con doble exploración, en 96 (8,8%) se detectó bocio por el observador 1 y en 102 (9,3%) por el 2 ($p=0,58$). El grado de acuerdo interobservador en la identificación y en la graduación del bocio fue moderado (kappa 0,55; IC del 95%, 0,46-0,64) para la primera y sustancial (kappa ponderado 0,61; IC del 95%, 0,51-0,71) para la segunda. Fue algo mayor en las niñas y en los escolares de mayor edad, peso, talla e índice de masa corporal. La concordancia entre observadores fue relativamente estable a lo largo de la realización del estudio.

Conclusiones: La concordancia interobservador en la detección de bocio por palpación en nuestro estudio es moderada, aunque más baja en los niños más pequeños e invariable durante todo el periodo en el que se realizó.

© 2012 Asociación Española de Pediatría. Publicado por Elsevier España, S.L. Todos los derechos reservados.

* Autor para correspondencia.

Correo electrónico: abelg@sescam.jccm.es (A. González-González).

KEYWORDS

Goitre;
Agreement;
Interobserver

Interobserver agreement in detecting goitre by palpation in a school population**Abstract**

Objective: To estimate the agreement between observers on the detection of goitre by physical exploration in a school population.

Methodology: We performed a cross-sectional study to detect goitre in a representative sample of 1134 schoolchildren aged 6 to 12 years from 20 schools in the health area of La Mancha Centro. The examination was performed blinded by two observers. Five grades in thyroid size were established (0, Ia, Ib, II and III). Above grade Ia was considered as goitre. The agreement was assessed in relation to variables such as age, sex, body mass index, height, and day of examination. The weighted kappa was used to measure the agreement.

Results: In the 1097 schoolchildren with a dual examination, 96 (8.8%) cases of goitre were detected by observer 1, and 102 (9.3%) cases by observer 2, ($P=.58$). The degree of interobserver agreement in the identification and grading of goitre was moderate (kappa 0,55, 95%CI: 0,46 to 0,64) for the first, and substantial (weighted kappa 0,61; 95%CI: 0,51 to 0,71) for the second. The degree of agreement was somewhat higher in girls, older schoolchildren, increased weight, height, and body mass index. The interobserver agreement was relatively stable throughout the study.

Conclusions: The interobserver agreement in detecting goitre by palpation in our study is moderate, but is lower in younger children and stable for the duration of study.

© 2012 Asociación Española de Pediatría. Published by Elsevier España, S.L. All rights reserved.

Introducción

La exploración física sigue siendo el método de elección de detección de bocio en estudios poblacionales¹ por su facilidad de implantación, la estandarización de sus resultados y su bajo coste. La ecografía, a pesar de ser más precisa, no ha logrado imponerse para su uso en estos estudios por la dificultad de estandarización de sus valores a todas las poblaciones², por su mayor coste económico y porque la palpación, en manos experimentadas, puede dar resultados que concuerdan satisfactoriamente con la medición ecográfica^{3,4}.

En los trabajos de investigación en los que participan varios observadores es esencial demostrar una buena concordancia entre los resultados obtenidos de ellos.

El presente estudio, preliminar a otro que medirá la prevalencia de la deficiencia de yodo en la población escolar de un área sanitaria, tiene como objetivo establecer el grado de concordancia en la exploración tiroidea por palpación entre los 2 exploradores participantes.

Sujetos y métodos

Se realizó un estudio transversal sobre una muestra representativa de los escolares de 6 a 12 años residentes en el área sanitaria de La Mancha Centro, que abarca poblaciones de Ciudad Real, Cuenca y Toledo.

Sobre una población de 13.896 niños repartidos en 73 centros escolares, se realizó un muestreo por conglomerados bietápico, con probabilidades proporcionales a los tamaños de la unidad de primera etapa. En la primera fase se seleccionaron 20 conglomerados (colegios) y en la segunda se realizó un muestreo aleatorio sistemático de 60 niños en cada centro seleccionado.

Los 2 únicos exploradores participantes en el estudio (AMS, PFG) son pediatras. Antes de iniciar el estudio ambos

unificaron criterios con la ayuda del investigador principal (AGG), endocrinólogo. Todos ellos con más de 10 años de experiencia clínica. La evaluación del tiroides se realizó, por inspección y palpación, de manera enmascarada por los 2 exploradores.

La clasificación del tamaño tiroideo se basó en la definición de bocio propuesta por Pérez et al.⁵ (tabla 1). De cada niño se escogió como medida del tamaño tiroideo la menor de entre las obtenidas cuando hubo duda entre los 2 evaluadores.

Las variables cuantitativas se resumieron mediante medidas de tendencia central y dispersión (media \pm desviación estándar [DE]) y la cualitativas con frecuencias absolutas y relativas, expresadas como porcentajes.

La concordancia ente observadores se estimó para el acuerdo en la identificación del bocio (presencia/ausencia) y el acuerdo en la graduación del mismo (0, Ia, Ib, II, III). Se estimó el grado de acuerdo con el índice kappa, utilizando la versión ponderada (con pesos cuadráticos de las distancias) para la concordancia en la graduación. La interpretación

Tabla 1 Clasificación del tamaño del tiroides por inspección y palpación

Grado	Tamaño
0	No palpable o lóbulo palpable, pero menor o igual que la falange distal del pulgar del individuo explorado
Ia	Nódulo palpable, pero mayor que la falange distal del pulgar del individuo explorado
Ib	Bocio palpable y visible solo con el cuello en extensión completa
II	Bocio visible con el cuello en posición normal
III	Bocio de gran tamaño, visible a 10 m de distancia

del índice kappa se hizo siguiendo los criterios de Landis y Koch⁶.

Calculamos el grado de acuerdo entre los observadores para diferentes estratos definidos por la edad, el sexo y algunas características antropométricas (peso, talla e índice de masa corporal [IMC]) del escolar. El posible impacto de cada variable sobre la concordancia la comprobamos contrastando la homogeneidad de los índices kappa estimados para cada categoría.

Los cálculos fueron realizados con el programa estadístico PASW 18.0 y EPIDAT 3.1 (Xunta de Galicia, OPS).

Resultados

Se reclutó a 1.134 alumnos procedentes de 20 centros de educación primaria del área estudiada.

En 1.097 escolares (589 niños [53,7%] y 508 niñas [46,3%], con una edad media \pm DE de 9,3 1,8 años) se realizó una exploración por 2 observadores de forma independiente, identificando y graduando la presencia de bocio.

De estos 1.097, en 96 (8,8%) se detectó bocio por el observador 1 y en 102 (9,3%) por el 2, aunque las diferencias no fueron estadísticamente significativas ($p=0,58$) (tabla 2). En 58 (5,3%) ambos observadores coincidieron en la detección de bocio y 140 niños (12,8%) fueron clasificados con bocio por

Tabla 2 Identificación del grado de bocio para cada observador

Grado	Observador 1	Observador 2
0	1001 (91,2%)	995 (90,7%)
Ia	77 (7,0%)	70 (6,4%)
Ib	18 (1,6%)	27 (2,5%)
II	1 (0,1%)	5 (0,5%)
III	0	0

uno u otro. El grado de concordancia entre observadores fue moderado (kappa 0,55; IC del 95%, 0,46 a 0,64).

El grado de acuerdo en la graduación del bocio fue mayor que en la identificación del mismo (kappa ponderado 0,61; IC95% 0,51 a 0,71).

El grado de acuerdo fue algo mayor en las niñas y en los escolares de mayor edad, peso, talla e IMC. La talla fue la variable que más se relacionó con el acuerdo en la identificación y graduación del bocio (tabla 3).

La concordancia entre observadores fue relativamente estable a lo largo de la realización del estudio, con un acuerdo que se mantuvo moderado, kappas entre 0,50 y 0,60, a partir del centenar de exploraciones realizadas (fig. 1).

Tabla 3 Niveles de concordancia en relación con la edad, el sexo y las variables antropométricas

	N.º	Concordancia identificación ^a	p	Concordancia graduación ^b	p
Sexo			0,27		0,29
Niños	589	0,50 (0,37-0,62)		0,56 (0,41-0,71)	
Niñas	508	0,59 (0,47-0,71)		0,66 (0,54-0,78)	
Edad			0,66		0,41
6-7 años	300	0,45 (0,25-0,66)		0,53 (0,30-0,76)	
8-9 años	362	0,53 (0,38-0,67)		0,52 (0,37-0,68)	
10-11 años	344	0,61 (0,47-0,74)		0,69 (0,54-0,84)	
12-14 años	91	0,52 (0,15-0,89)		0,68 (0,40-0,99)	
Masa corporal (kg/m²)			0,54		0,58
Q1 (12,6-15,9)	265	0,45 (0,11-0,79)		0,41 (0,08-0,74)	
Q2 (16,0-17,9)	267	0,44 (0,22-0,65)		0,66 (0,42-0,89)	
Q3 (18,0-20,4)	276	0,51 (0,34-0,68)		0,55 (0,37-0,73)	
Q4 (20,5-34,4)	273	0,60 (0,47-0,73)		0,63 (0,49-0,77)	
Peso (kg)			0,049		0,31
Q1 (16,8-26,9)	263	0,42 (0,09-0,74)		0,39 (0,07-0,71)	
Q2 (27,0-33,1)	272	0,44 (0,22-0,65)		0,46 (0,20-0,72)	
Q3 (33,2-41,8)	272	0,51 (0,35-0,68)		0,65 (0,48-0,82)	
Q4 (41,9-93,8)	276	0,61 (0,48-0,74)		0,65 (0,51-0,79)	
Talla (cm)			0,11		0,12
Q1 (110,1-127,5)	269	0,38 (0,10-0,65)		0,36 (0,09-0,64)	
Q2 (127,6-135,7)	272	0,42 (0,22-0,61)		0,55 (0,35-0,75)	
Q3 (135,8-145,8)	274	0,55 (0,40-0,71)		0,55 (0,37-0,72)	
Q4 (146,1-175,0)	269	0,67 (0,53-0,81)		0,72 (0,57-0,86)	

Resultados expresados como kappa (intervalo de confianza del 95%).

p: significación de la prueba de homogeneidad de kappas; Q1, Q4: cuartiles.

^a Índice kappa para el acuerdo entre observadores en la identificación (presencia/ausencia) de bocio.

^b Índice kappa ponderado para el acuerdo entre observadores en la graduación (0, Ia, Ib, II) del bocio.

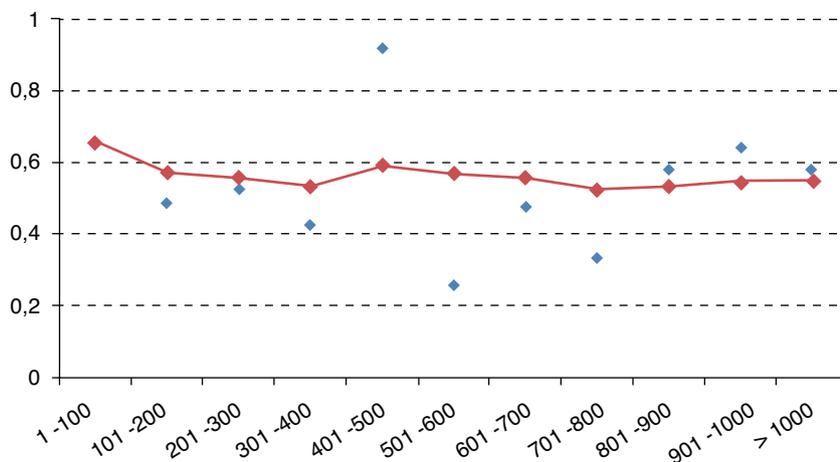


Figura 1 Evolución del acuerdo entre observadores a lo largo del estudio. Los puntos representan los índices de kappa para cada estrato de 100 exploraciones. La línea representa el índice de kappa para las exploraciones acumuladas.

Discusión

La prevalencia de bocio es una medida indirecta de deficiencia nutricional de yodo en una población determinada⁷. Las 2 técnicas utilizadas en los estudios poblacionales para la medición del tamaño tiroideo, la exploración física y la ecografía presentan ventajas y desventajas una sobre la otra^{4,8,9}, pero la mayoría de los estudios utilizan la primera. Nosotros utilizamos una clasificación del tamaño tiroideo basada en un concepto de bocio propuesto en 1960 por Pérez et al.⁵, que considera patológico todo tiroides palpable con un lóbulo mayor que la falange distal del pulgar del individuo explorado, en lugar de una posterior (1993), más simple, que define como bocio todo tiroides detectado por palpación¹⁰ porque, aunque la segunda es más fácil de utilizar en estudios poblacionales, tiene menos especificidad y puede sobreestimar la prevalencia de bocio entre un 20 y un 30%⁸, y porque los estudios sobre prevalencia de bocio de zonas geográficas cercanas a la nuestra utilizaron clasificaciones similares. La prevalencia de bocio en nuestro estudio (8,8 y 9,3%, según el explorador) fue menor que la detectada en 1993 en 3 áreas diferentes de nuestra comunidad autónoma, que oscilaron entre el 18 y el 24%¹¹⁻¹³.

Un índice kappa entre 0,41 y 0,61 se considera moderado, y bueno por encima de 0,61⁴. El grado de concordancia interobservador en la identificación de bocio en nuestro estudio, que fue de 0,55, está en el rango medio de lo publicado^{4,8,14-16}. Una baja concordancia entre observadores influye sobre las estimaciones de la prevalencia de bocio¹⁴. Por esta razón, proponemos introducir la medición del grado de concordancia interobservador en los protocolos de estudio de bocio endémico. Además, como una buena concordancia interobservador no garantiza necesariamente la validez del procedimiento diagnóstico, también sería interesante incorporar en estos protocolos el grado de acuerdo entre la exploración física y ecográfica (más objetiva), al menos en un grupo reducido de sujetos.

El índice de concordancia se mantuvo estable a lo largo de la duración del estudio. Esta estabilidad puede deberse a que nuestros exploradores poseían ya una amplia experiencia en la palpación cervical al inicio del estudio y

probablemente no modificaron sus criterios diagnósticos a lo largo del mismo, algo que sí ocurre cuando los exploradores son inexpertos⁸.

El índice sí se modificó llamativamente con algunas características de los individuos analizados. La concordancia aumentó con la edad (salvo en el último cuartil), el IMC, el peso y, sobre todo, con la talla. Los índices de concordancia más bajos ($\leq 0,45$) fueron detectados en niños menores de 7 años, con un IMC $< 18 \text{ kg/m}^2$, con un peso menor de 33,2 kg y una talla inferior a 135,8 cm. Estos datos sugieren que la concordancia interobservador es paralela al tamaño de sujeto. Por otra parte, como además el volumen de tiroides aumenta progresivamente con la edad, el peso, la talla y el área de superficie corporal^{2,17}, nuestros resultados podrían interpretarse como que la concordancia interobservador depende del tamaño tiroideo, siendo las medidas antropométricas medidas indirectas de este. Esta impresión se ve corroborada por el hecho de haber detectado en las áreas con menor prevalencia de bocio, donde el tamaño del tiroides es menor, un menor índice de concordancia¹⁸; y esto podría explicar, en parte, el menor índice kappa hallado en nuestro estudio frente al de Peris Roig et al.¹⁴, a pesar de utilizar unos criterios de detección de bocio similares. Parece, pues, que la exploración cervical no es un método diagnóstico fiable de bocio en los niños más pequeños².

En conclusión, la concordancia interobservador en la detección de bocio por palpación en nuestro estudio es moderada, aunque más baja en los niños más pequeños y se muestra invariable durante todo el periodo en el que se realizó.

Financiación

Este trabajo ha sido financiado con la ayuda de un proyecto de investigación de la Fundación Socio-Sanitaria de Castilla-La Mancha (AN-2010/44).

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Agradecimientos

A Angel Arias Arias, de la Unidad de Investigación de La Mancha Centro, por su labor en la confección y depuración de la base de datos.

Bibliografía

1. WHO, UNICEF, ICCIDD. Assessment of Iodine Deficiency Disorders and Monitoring their elimination. A guide for managers. WHO/NHD/01.1 2001.
2. Foo LC, Zulfiqar A, Nafikudin M, Fadzil MT, Asmah AS. Local versus WHO/International Council for Control of Iodine Deficiency Disorders-recommended thyroid volume reference in the assessment of iodine deficiency disorders. *Eur J Endocrinol.* 1999;140:491-7.
3. Nordmeyer JP, Simons M, Wenzel C, Scholten T. How accurate is the assessment of thyroid volume by palpation? A prospective study of 316 patients. *Exp Clin Endocrinol Diabetes.* 1997;105:366-71.
4. Zimmermann M, Saad A, Hess S, Torresani T, Chaouki N. Thyroid ultrasound compared with World Health Organization 1960 and 1994 palpation criteria for determination of goiter prevalence in regions of mild and severe iodine deficiency. *Eur J Endocrinol.* 2000;143:727-31.
5. Pérez C, Scrimshaw NS, Muñoz JA. Técnicas de las encuestas sobre bocio endémico. En: *El bocio endémico. Monografía de la OMS*, 44. WHO: Ginebra; 1961. p. 399-414.
6. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-74.
7. Soriguer F, García-Fuentes E, Rojo G, Santiago P, Olveira G, Garriga MJ, et al. Protocolo para el estudio de trastornos debidos a la deficiencia nutricional de yodo. *Endocrinol Nutr.* 2005;52:105-24.
8. Peterson S, Sanga A, Eklöf H, Bunga B, Taube A, Gebre-Medhin M, et al. Classification of thyroid size by palpation and ultrasonography in fields surveys. *Lancet.* 2000;355:106-10.
9. Soriguer FJC, Millón MC. Diagnóstico de bocio en los estudios poblacionales de déficit de yodo: ¿palpación o ecografía? *Endocrinol Nutr.* 2000;47:203-4.
10. WHO. Report of a Joint WHO/UNICEF/ICCIDD Consultation on indicators for assessing iodine deficiency disorders and their control programmes. Geneva: World Health Organization; 1992. p. 22-9.
11. Sanabria C, Rodríguez-Arno J, De Paz V, Parejo MD, Gómez-Pan A, Escobar del Rey F. Detección de bocio endémico en la población escolar de la zona rural de la provincia de Toledo. *Endocrinol Nutr.* 1993;40:239-43.
12. Calvo de Haro MA, Escobar del Rey F. Bocio endémico y déficit de yodo en la población escolar de la provincia de Guadalajara. *Endocrinol Nutr.* 1993;40:233-8.
13. Donnay Candil S, Requejo Castillo R, Navarro Collado T, Martínez Orgado J, Mora Bañales F, Martín Escobar E, et al. Prevalencia de bocio y estado nutricional de yodo en la población escolar de la Sierra de Cuenca. *Endocrinología.* 1993;40:228-32.
14. Peris Roig B, Merchante Alfaro AA, Tenias Burillo JM, Atienzar Herráez N, Calvo Rigual F, López García MJ. Interobserver agreement in detecting goitre in a school population between 6 and 14 years old. *Rev Esp Salud Publica.* 2009;83:339-43.
15. Martínez-Salgado H, Castañeda-Limones R, Lechuga-Martín del Campo D, Ramos-Hernández RI, Orozco-López M, Rivera-Dommarco J, et al. Iodine deficiency and other potential goitrogens in the persistence of endemic goiter in Mexico. *Gac Med Mex.* 2002;138:149-56.
16. Gómez R, Millón MC, Soriguer F, Mancha I, Garriga MJ, Muñoz R, et al. Comparación de diferentes criterios de diagnóstico poblacional de la deficiencia de yodo (DDY). Prevalencia de bocio endémico en la Axarquía (Málaga). *Endocrinol Nutr.* 2000;47:260-6.
17. Vitti P, Martino E, Aghini-Lombardi F, Rago T, Antonangeli L, Maccherini D, et al. Thyroid volume measurement by ultrasound in children as a tool for the assessment of mild iodine deficiency. *J Clin Endocrinol Metab.* 1994;79:600-3.
18. WHO/ICCIDD. Recommended normative values for thyroid volume in children age 6-15 years. *Bulletin of the World Health Organization.* 1997;75:95-7.