

ORIGINAL

Validación de la versión en español de la escala de evaluación del trastorno por déficit de atención e hiperactividad (ADHD-RS-IV.es) en una muestra española[☆]



M. Vallejo-Valdivielso^{a,*}, C.A. Soutullo^{a,b,c}, P. de Castro-Manglano^{a,b,c}, J.J. Marín-Méndez^{a,c} y A. Díez-Suárez^{a,c}

^a Unidad de Psiquiatría Infantil y Adolescente, Departamento de Psiquiatría y Psicología Médica, Clínica Universidad de Navarra, Pamplona, Navarra, España

^b Unidad de Psiquiatría Infantil y Adolescente, Departamento de Psiquiatría y Psicología Médica, Clínica Universidad de Navarra, Madrid, España

^c Instituto de Investigación Sanitaria de Navarra (IDISNA), Navarra, Pamplona, España

Recibido el 14 de febrero de 2017; aceptado el 11 de mayo de 2017

Accesible en línea el 14 de julio de 2017

PALABRAS CLAVE

Trastorno por déficit de atención e hiperactividad; Evaluación; Escala Trastorno por déficit de atención e hiperactividad; *Attention-Deficit/Hyperactivity Disorder Rating Scale*

Resumen

Objetivos: El objetivo de este estudio es validar la versión en castellano del ADHD-RS-IV (ADHD-RS-IV.es) en una muestra española.

Métodos: A partir de una muestra total de 652 niños y adolescentes (rango edad: 6-17 años; media [DE] = 11,14 [3,27] años), se incluyó a 518 pacientes con TDAH (criterios DSM-IV-TR) y a 134 controles sanos. Para la evaluación de la estructura factorial, la validez y la fiabilidad de la escala se realizó un análisis factorial confirmatorio (CFA) mediante *structural equation modeling* en una matriz de correlaciones policóricas, y usando el método de máxima verosimilitud para la estimación. Se calcularon la validez discriminante y su valor predictivo mediante curvas *receiver operating characteristics*.

Resultados: La escala en castellano mostró una consistencia interna elevada, tanto para la escala total como para sus subescalas. El coeficiente alfa de Cronbach era 0,94 para la escala

[☆] Este trabajo no ha sido presentado en la Reunión Anual de la SEN. Sin embargo, algunos resultados preliminares han sido presentados como comunicación tipo póster en las siguientes reuniones científicas: 60.º Congreso AEPNYA-AACAP (San Sebastián, España, junio del 2016); 29th ECNP Congress of Applied and Translational Neuroscience (Viena, Austria, septiembre del 2016); 4th Eunethydis International Conference (Berlín, Alemania, octubre del 2016), y en el AACAP's 63rd Annual Meeting (Nueva York, EE. UU., octubre del 2016).

* Autor para correspondencia.

Correo electrónico: mvaldiviel@unav.es (M. Vallejo-Valdivielso).

KEYWORDS

Attention-Deficit/Hyperactivity Disorder; Assessment; Attention-Deficit/Hyperactivity Disorder rating scale; Attention-Deficit/Hyperactivity Disorder Rating Scale

total y $\geq 0,90$ para las subescalas. Los valores alfa ordinales eran 0,95 para la escala total y $\geq 0,90$ para las subescalas. El análisis CFA mostró un modelo de 2 factores (inatención e hiperactividad/impulsividad) intercorrelacionados. La escala ofrece buen poder discriminante (AUC = 0,97).

Conclusiones: La versión española del ADHD-RS-IV (ADHD-RS-IV.es) mostró una estructura bifactorial consistente con los modelos del DSM-IV-TR y DSM-5, y con el modelo propuesto por el autor de la escala original. Además, posee un alto poder discriminante, lo que lo convierte en un instrumento válido y fiable para medir la presencia y severidad de síntomas de TDAH en la población española.

© 2017 Sociedad Española de Neurología. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Validation of a Spanish-language version of the ADHD Rating Scale IV in a Spanish sample

Abstract

Objectives: The purpose of this study is to validate a Spanish-language version of the 18-item ADHD Rating Scale-IV (ADHD-RS-IV.es) in a Spanish sample.

Methods: From a total sample of 652 children and adolescents aged 6 to 17 years (mean age was 11.14 ± 3.27), we included 518 who met the DSM-IV-TR criteria for ADHD and 134 healthy controls. To evaluate the factorial structure, validity, and reliability of the scale, we performed a confirmatory factor analysis (CFA) using structural equation modelling on a polychoric correlation matrix and maximum likelihood estimation. The scale's discriminant validity and predictive value were estimated using ROC (receiver operating characteristics) curve analysis. **Results:** Both the full scale and the subscales of the Spanish-language version of the ADHD-RS-IV showed good internal consistency. Cronbach's alpha was 0.94 for the full scale and ≥ 0.90 for the subscales, and ordinal alpha was 0.95 and ≥ 0.90 , respectively. CFA showed that a two-factor model (inattention and hyperactivity/impulsivity) provided the best fit for the data. ADHD-RS-IV.es offered good discriminant ability to distinguish between patients with ADHD and controls (AUC = 0.97).

Conclusions: The two-factor structure of the Spanish-language version of the ADHD-RS-IV (ADHD-RS-IV.es) is consistent with those of the DSM-IV-TR and DSM-5 as well as with the model proposed by the author of the original scale. Furthermore, it has good discriminant ability. ADHD-RS-IV.es is therefore a valid and reliable tool for determining presence and severity of ADHD symptoms in the Spanish population.

© 2017 Sociedad Española de Neurología. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

El trastorno por déficit de atención e hiperactividad (TDAH) es un trastorno del neurodesarrollo que se caracteriza por la presencia de síntomas de inatención (IA) y/o hiperactividad-impulsividad (HI), persistentes y no correspondientes al nivel de neurodesarrollo normal del niño o adolescente¹. Estas 2 dimensiones (IA e HI) son resultado del estudio factorial de los 18 criterios diagnósticos del TDAH del DSM-IV-TR² y DSM-5¹. La evaluación completa de los síntomas de TDAH debe incluir una entrevista clínica con padres y paciente, e información sobre datos que recojan el impacto que los síntomas tienen en el funcionamiento global del niño, tanto en casa como en el colegio³⁻⁸.

Hay varios cuestionarios siguiendo los criterios y la estructura factorial del DSM-IV-TR² para facilitar la recogida de la gravedad y la frecuencia de síntomas de TDAH por

parte de profesores y padres. Estos incluyen la ADHD-Rating Scale IV⁹, la ADHD Symptoms Rating Scale¹⁰, la Conners Rating Scale¹¹ y la SNAP-IV Rating Scale¹², entre otras. Para abordar los cambios aparecidos en el DSM-5¹, DuPaul y su equipo desarrollaron recientemente la ADHD Rating Scale-5³, avalando su validez y replicando la estructura bifactorial de los síntomas diana del TDAH, recogida tanto en el DSM-IV-TR² como el DSM-5¹. Tanto la utilidad clínica, como las consistentes propiedades psicométricas de la escala ADHD-RS-IV¹³ han sido replicadas en distintos estudios norteamericanos¹⁴, multinacionales¹⁵ y europeos¹⁶.

En España, Servera y Cardó publicaron en 2007 un estudio describiendo los datos normalizados del ADHD-RS-IV basados en una muestra de niños en edad escolar (6-12 años)¹⁷. Sin embargo, hasta donde sabemos, esta escala no ha sido validada en español. Los objetivos de este estudio son, primero, describir la frecuencia de síntomas de TDAH

tanto en una muestra de niños controles sanos sin TDAH, como en una muestra de pacientes con TDAH; y segundo, validar la versión en castellano de la escala de 18 ítems ADHD Rating Scale-IV (ADHD-RS-IV.es).

Sujetos, material y métodos

Sujetos

Realizamos un estudio naturalístico observacional de 652 participantes de edades entre los 6 y los 17 años de edad. El estudio se llevó a cabo a partir de 2 muestras separadas. Se incluyó en el estudio a 518 pacientes (no previamente tratados) de edades entre los 6 y 17 años en el momento del diagnóstico de TDAH (cualquier subtipo). Los sujetos fueron seleccionados a partir de los pacientes con diagnóstico de TDAH que habían acudido a la Unidad de Psiquiatría Infantil y Adolescente de la Clínica Universidad de Navarra entre enero del 2000 y abril del 2015. También se incluyó a 134 niños y adolescentes sin diagnóstico ni sospecha de TDAH, de edades entre los 6 y los 16 años. Los sujetos fueron seleccionados a partir de los pacientes que acudieron a lo largo del año 2014 al departamento de Pediatría de la Clínica Universidad de Navarra para consultas de niño sano.

Evaluación clínica de la muestra de pacientes con trastorno por déficit de atención e hiperactividad

Para la evaluación clínica y diagnóstica de TDAH se utilizaron los criterios diagnósticos del DSM-IV-TR² mediante una plantilla de la entrevista semiestructurada Kiddie-Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version (KSADS-PL)^{18,19}, realizada por psiquiatras especialistas en Psiquiatría del Niño y del Adolescente con experiencia (CAS, ADS, PdC). Como parte de la rutina diagnóstica, se recogieron, además, los cuestionarios completados por los padres del ADHD-RS-IV.es. Se excluyó a pacientes con TDAH que tuvieran un cociente intelectual igual o menor de 80 y/o con diagnóstico de trastorno del espectro autista. No se excluyó a los pacientes que presentaran comorbilidad con otros trastornos, como por ejemplo trastorno negativista desafiante, trastornos depresivos y/o trastornos de ansiedad.

Attention-Deficit/Hyperactivity Disorder Rating Scale (padres)

El ADHD-RS-IV⁹ es una escala de 18 ítems correspondientes para cada uno de los 18 síntomas que figuran en el diagnóstico DSM-IV-TR² de TDAH, similares a los del actual DSM-5¹. Se trata de una escala tipo Likert en la que cada ítem se califica de 0 a 3 puntos (nunca/rara vez, algunas veces, con frecuencia, con mucha frecuencia) cada uno. La puntuación total se calcula como la suma de las puntuaciones obtenidas en cada uno de los 18 ítems. Además de la puntuación total, se subdividen 2 puntuaciones atendiendo a la suma de síntomas de IA e HI. La versión en castellano del ADHD-RS-iv.es (véase la [tabla 1](#)) usada en este estudio fue traducida al castellano de España por el último autor (CAS) en el año 1999. Dicha traducción fue revisada por un

traductor independiente y la versión ha sido actualizada y adaptada sobre la base de las sugerencias y dificultades y dudas al cumplimentarlo por parte de padres, profesores y otros profesionales.

Todos los padres de los participantes (pacientes y controles sanos) completaron la escala ADHD-RS-IV.es, como parte del proceso diagnóstico en el caso de los pacientes, y de manera voluntaria y sin retribución alguna en el caso de los controles sanos. Los cuestionarios fueron administrados por miembros del equipo de la Unidad de Psiquiatría Infantil y Adolescente, así como del Departamento de Pediatría de la Clínica Universidad de Navarra.

Análisis estadístico

Se realizaron análisis estadísticos descriptivos con medidas de tendencia central y dispersión (media y desviación estándar [DE] para variables cuantitativas y proporciones e intervalos de confianza para las categóricas). Para la comparación de medias, se utilizó el test de la t de Student. Estos análisis se realizaron con el programa SPSS. Los subtipos DSM-IV-TR2 (combinado, predominio inatento o predominio hiperactivo/impulsivo) fueron analizados conjuntamente, sin diferenciar entre ellos. Se realizó el análisis de ítems, determinando las medias y desviaciones típicas de la muestra total como dividida por grupos (pacientes TDAH/controles sanos). También se determinó el coeficiente alfa de Cronbach y el alfa ordinal —según el método descrito por Zumbo et al. (2007) con la eliminación de un elemento²⁰.

Se realizó un análisis factorial confirmatorio (CFA) con STATA mediante *structural equation modeling* (SEM) en una matriz de correlaciones policóricas —así se recomienda cuando se trata de ítems con respuestas tipo Likert menores de 7 puntos²⁰— y usando el método de máxima verosimilitud. Como base para la comparación, se probó la adecuación de varios modelos: un modelo bifactorial, así como un modelo jerárquico para evaluar la posibilidad de un modelo bifactorial que incluyera un *general factor ADHD*, del que partieran los 2 factores (IA e HI), encontrado en un estudio previo²¹. Como medidas de bondad de ajuste incremental, además del test de la chi al cuadrado, se utilizaron el índice de Tucker-Lewis (*Tucker-Lewis index* [TLI]), y el índice de ajuste comparativo (*Comparative fit index* [CFI])^{22,23}. Se empleó el error cuadrático medio de aproximación (*Root-mean-square error of approximation* [RMSEA])²⁴ como medida de ajuste parsimonioso. Los valores inferiores a 2 en el cociente de la chi al cuadrado con los grados de libertad²⁵, iguales o superiores a 0,90 en el CFI y el TLI son indicadores de buen ajuste^{26,27}, e iguales o inferiores a 0,06 en el RMSEA²⁸.

Para determinar la fiabilidad de la escala se utilizó un comando programado en STATA, llamado *ordalpha*, basado parcialmente en el código de Joseph Coveney (2012) modificado²⁹, para obtener los coeficientes alfa ordinales de la escala total y sus subescalas. Estos cálculos se realizaron debido a que dichos coeficientes son preferidos en escalas tipo Likert por ser considerados mejores estimadores del grado de concordancia²⁰. No obstante, para facilitar la comparación con los resultados de otros estudios publicados en la literatura, con el programa SPSS se calcularon además los coeficientes alfa de Cronbach³⁰ de las 2 subescalas

Tabla 1 La versión en español del ADHD-RS-IV usada en el presente estudio

	Nunca o rara vez	Algunas veces	Con frecuencia	Con mucha frecuencia
IA1 No presta atención a las cosas o comete errores por descuido				
IA2 Tiene dificultad en mantener la atención en actividades o juegos				
IA3 No parece escuchar lo que se le está diciendo				
IA4 Tiene dificultad en completar/terminar tareas o no sigue instrucciones				
IA5 Tiene dificultad en organizarse (actividades, etc.)				
IA6 Evita actividades que requieren mucho esfuerzo mental (deberes, etc.)				
IA7 Pierde cosas necesarias para actividades o deberes				
IA8 Se distrae fácilmente por estímulos externos (por el ambiente...)				
IA9 Es olvidadizo de actividades diarias				
HA1 Mueve manos y pies o se mueve en su asiento				
HA2 Se levanta en clase o en otras situaciones cuando no debe				
HA3 Corre o se sube a sitios/cosas cuando se le dice que no lo haga				
HA4 Tiene dificultad para jugar en silencio				
HA5 Actúa como si estuviera activado por un motor. «No se le gastan las pilas»				
HA6 Habla excesivamente				
HA7 Responde abruptamente/impulsivamente antes de terminar la pregunta				
HA8 Tiene dificultad en esperar su turno en actividades de grupo				
HA9 Interrumpe a la gente o se mete en conversaciones o juegos				
Al tratarse de una escala Likert (0-3), el sistema de puntuación es el siguiente: «nunca o rara vez»: 0 puntos; «algunas veces»: 1 punto; «con frecuencia»: 2 puntos; «con mucha frecuencia»: 3 puntos.				
HA1-HA9: ítems correspondientes a los 9 síntomas de hiperactividad/impulsividad; IA1-IA9: ítems correspondientes a los 9 síntomas de inatención.				

y la escala total empleando toda la muestra. Además, para analizar la consistencia interna se utilizó también el índice de correlación media ítem-ítem, basado en una matriz de correlaciones de Pearson y en una matriz de correlaciones policóricas. Este índice es aconsejable como indicador de la homogeneidad de los ítems, especialmente en casos de escalas tipo Likert con un número situado entre 30 y 40 ítems. El índice de correlación media interítem (r) nos indica si los ítems son representativos de un único componente, recomendándose valores situados entre 0,15 y 0,50³¹.

Por último, se calculó la validez discriminante de la escala y su valor predictivo mediante curvas *receiver operating characteristics* (ROC). La validez discriminante es un indicador de la capacidad que tiene una escala para distinguir los diferentes subgrupos que pretende diferenciar. Se calculó el intervalo de confianza (IC) del 95% para la sensibilidad y la especificidad usando el método de Wilson³² programado en una macro de SPSS. El punto de corte óptimo se estableció en el índice de *Youden* máximo (J), es decir, el punto de máximo alejamiento de la curva ROC respecto a la diagonal ($J = \text{sensibilidad} + \text{especificidad} - 1$) utilizando siempre valores p de 2 colas.

Para los análisis estadísticos realizados en este estudio de validación se emplearon los programas SPSS v. 15 y STATA v.12.1.

Resultados

Seiscientos cincuenta y dos niños y adolescentes participaron en el estudio; 518 pacientes (muestra clínica)

diagnosticados por primera vez de TDAH (79,73% varones; edad [DE] = 11,40 [3,30] años) y 134 controles sanos (muestra no clínica) (47% varones; edad [DE] = 10,15 [2,97] años) fueron incluidos en el estudio.

Comparativa entre pacientes/controles sanos de los síntomas diana del trastorno por déficit de atención e hiperactividad

Las puntuaciones obtenidas por el grupo clínico fueron significativamente mayores que las obtenidas por el grupo no clínico ($p < 0,001$). Estas diferencias se hallaron en cada uno de los ítems de IA e HI por separado, así como en la puntuación total (tabla 2) (fig. 1).

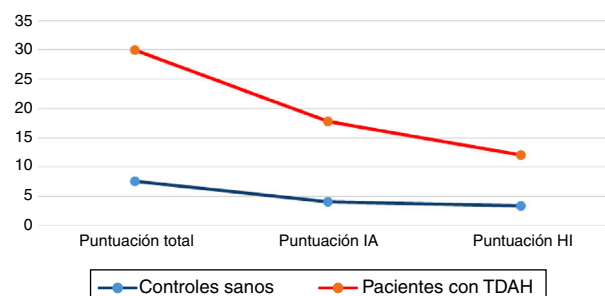


Figura 1 Comparación de puntuaciones totales (ADHD-RS-IV.es) y por subescalas (IA e HI) en la muestra clínica (pacientes con TDAH) y no clínica (controles sanos).

Tabla 2 Comparación de las puntuaciones en cada síntoma de TDAH (ADHD-RS-IV.es) en nuestra muestra de niños y adolescentes con (n = 518) y sin (n = 134) TDAH

Ítem	Puntuación en controles M (DE)	Puntuación en pacientes TDAH M (DE)	p
Subescala IA			
Ítem 1 (IA1): no presta atención/errores por descuido	0,63 (0,65)	2,07 (0,77)	< 0,001
Ítem 2 (IA2): dificultad en mantener atención en actividades o juegos	0,24 (0,52)	1,52 (0,96)	< 0,001
Ítem 3 (IA3): no parece escuchar lo que se le está diciendo	0,51 (0,59)	1,89 (0,88)	< 0,001
Ítem 4 (IA4): dificultad en completar/terminar tareas o no sigue instrucciones	0,30 (0,53)	2,12 (0,84)	< 0,001
Ítem 5 (IA5): dificultad en organizarse (actividades, etc.)	0,46 (0,60)	2,17 (0,93)	< 0,001
Ítem 6 (IA6): evita actividades que requieren mucho esfuerzo mental	0,43 (0,62)	2,27 (0,96)	< 0,001
Ítem 7 (IA7): pierde cosas necesarias para actividades o deberes	0,39 (0,60)	1,60 (1,05)	< 0,001
Ítem 8 (IA8): se distrae fácilmente por estímulos externos	0,73 (0,66)	2,41 (0,71)	< 0,001
Ítem 9 (IA9): es olvidadizo de actividades diarias	0,45 (0,60)	1,82 (0,94)	< 0,001
Puntuación subescala inatención (puntuación IA)	4,13 (3,57)	17,85 (5,48)	< 0,001
Subescala HI			
Ítem 10 (HA1): mueve manos y pies o se mueve en su asiento	0,54 (0,75)	1,79 (1,05)	< 0,001
Ítem 11 (HA2): se levanta en clase o en otras situaciones cuando no debe	0,24 (0,49)	1,41 (1,08)	< 0,001
Ítem 12 (HA3): corre o se sube a sitios/cosas cuando se le dice que no lo haga	0,37 (0,57)	1,16 (1,1)	< 0,001
Ítem 13 (HA4): tiene dificultad para jugar en silencio	0,22 (0,50)	0,98 (1,07)	< 0,001
Ítem 14 (HA5): actúa como si estuviera activado por un motor	0,37 (0,60)	1,26 (1,14)	< 0,001
Ítem 15 (HA6): habla excesivamente	0,62 (0,86)	1,30 (1,07)	< 0,001
Ítem 16 (HA7): responde abruptamente antes de terminar la pregunta	0,42 (0,63)	1,52 (1,05)	< 0,001
Ítem 17 (HA8): dificultad en esperar su turno en actividades de grupo	0,28 (0,53)	1,23 (1,05)	< 0,001
Ítem 18 (HA9): interrumpe a la gente o se mete en conversaciones o juegos	0,43 (0,58)	1,48 (1,08)	< 0,001
Puntuación subescala hiperactividad/impulsividad (puntuación HI)	3,43 (3,41)	12,11 (7,083)	< 0,001

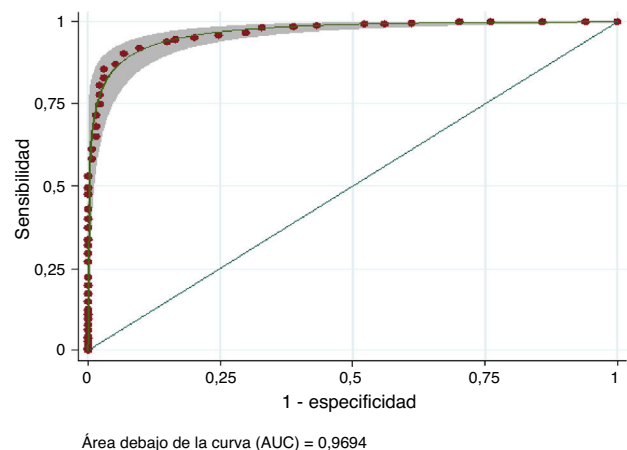
HI: hiperactividad/impulsividad; IA: inatención.

Validez discriminante de la versión española del ADHD-RS-IV.es: comparativa entre pacientes/controles sanos

El análisis mediante curva ROC muestra que la escala ADHD-RS-IV.es tiene una sensibilidad del 91,9% y una especificidad del 90,3% para diferenciar el grupo clínico del no clínico (fig. 2). En base al índice J, el punto de corte óptimo que definiría, de manera teórica y sobre la base de nuestros datos, con mayor exactitud el diagnóstico, son 17 puntos para la muestra general (17 puntos para chicos y 15 puntos para chicas; 19 puntos para niños y 16 para adolescentes [fig. 3] [tabla 2]).

Análisis factorial confirmatorio

La muestra de pacientes con TDAH y sin TDAH fue analizada conjuntamente, independientemente del subtipo DSM-IV-TR², ya que no existe razón probada en la literatura que justifique su análisis por separado²⁸. El modelo bifactorial (IA, HI) demostró un ajuste aceptable (CFI=0,92, TLI=0,91, RMSEA=0,07). El CFA mostró también que todos los ítems cargaban de manera significativa (r > 0,30 con p < 0,001) en los factores clínicamente correspondientes. Se confirma también la elevada correlación (r = 0,62)

**Figura 2** Curva ROC de la escala ADHD-RS-IV.es en la muestra total.

entre ambos factores (IA, HI), IC del 95% (0,57-0,68). La fiabilidad analizada mediante los coeficientes alfa ordinal y de Cronbach fueron elevados tanto para la escala total (0,93 y 0,95, respectivamente), como para la subescala de IA (0,93 y 0,95) y de HI (0,91 y 0,94) por separado.

El diagrama SEM completo se presenta en la figura 4.

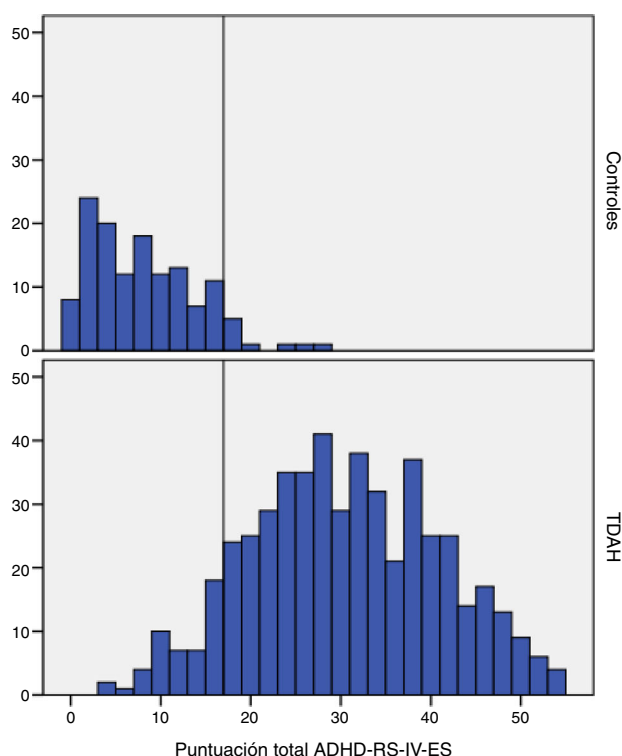


Figura 3 Puntuación total del ADHD-RS-IV.es: comparación de puntuaciones entre pacientes con TDAH y controles sanos. La línea vertical representa el punto de corte óptimo para la muestra total: 17 puntos.

Discusión

Este estudio evalúa la validez de la escala ADHD-RS-IV.es, y analiza las propiedades psicométricas de la escala en una muestra clínica y no clínica española. Teniendo en cuenta las características de nuestra muestra, no hay ningún estudio publicado hasta la fecha que haya realizado la validación de la escala ADHD-RS-IV en español en una muestra clínica y no clínica.

Tanto la homogeneidad como el tamaño muestral son aspectos que merecen consideración en el estudio de validación de una escala. A este respecto, no existe un acuerdo unánime sobre el número de sujetos mínimo que se han de incluir. El tamaño muestral que hemos utilizado en nuestro estudio satisface los criterios generalizados, ya que, en cuanto al valor absoluto, nuestra cifra de 652 sujetos cumpliría las indicaciones de Goldberg y Digman (1994), que recomiendan la inclusión de entre 500 y 1.000 participantes³³. Con respecto al número de sujetos reclutados por cada ítem del cuestionario, en nuestro caso la ratio es de 36:1, cumpliendo las recomendaciones más exigentes³⁴. Mediante el análisis descriptivo de los ítems, podemos contrastar su capacidad de poner de manifiesto las diferencias existentes entre el grupo clínico y no clínico. En nuestro estudio encontramos diferencias estadísticamente significativas ($p < 0,001$) en cada uno de los 18 ítems de la escala, así como en las puntuaciones de las subescalas (IA e HI) y en la puntuación total entre la muestra clínica y la no clínica. Esto demuestra que la escala ADHD-RS-IV.es, en nuestra muestra, ha tenido la capacidad de discriminar la

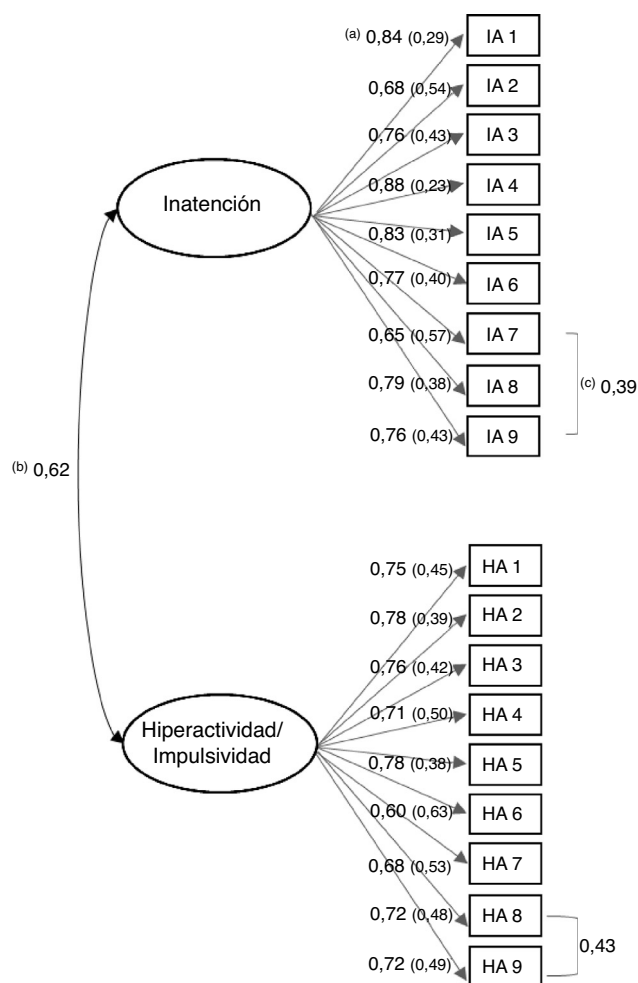


Figura 4 Diagrama SEM que muestra un modelo bifactorial, así como los valores de las correlaciones existentes entre las 2 subescalas entre las 2 dimensiones (subescalas) del ADHD-RS-IV.es.

En este diagrama, cada círculo corresponde con una dimensión, y cada rectángulo con un ítem (criterio diagnóstico). En este diagrama observamos cómo cada círculo está unido a los ítems que definen una misma dimensión y no están ligados a la otra (coeficiente de correlación media ítem-subescala). La ausencia de flecha que uniera un ítem con una dimensión significaría que la carga del factor con cualquiera de las dimensiones sería 0. La correlación entre las dimensiones y/o ítems están representadas con flechas de doble dirección (coeficiente de correlación interítem/intersubescala).

^a Coeficiente de correlación media ítem-subescala (error aleatorio).

^b Coeficiente de correlación intersubescala.

^c Coeficiente de correlación interítem.

presencia de diagnóstico. La media de puntuación total en nuestra muestra clínica fue de 29,96, ligeramente menor que la obtenida en otras muestras clínicas utilizadas en los estudios de validación europeos¹⁶ (35,8) y multinacionales¹⁵ (41,2). Esto subraya la importancia de obtener datos normativos de dicha escala en cada país.

El poder de discriminación de un ítem se puede estimar analizando el coeficiente de correlación ítem-total

Tabla 3 Puntos de corte óptimos y percentiles correspondientes con los datos normativos de la muestra/submuestras

	Valor punto de corte				Datos normativos							Punto de corte de la muestra total ^a y percentiles correspondientes a los datos normativos de la muestra total	Punto de corte de las submuestras ^b y percentiles correspondientes a los datos normativos de las distintas submuestras
		Sensibilidad (%)	Especificidad (%)	AUC	IC del 95%	n	M (DE)	p90	p93	p95	p98		
Muestra total	17 ^a	91,9	90,3	0,969	0,96 - 0,98	134	7,57 (5,83)	15,5	16,55	18	23,90	17 (p93-p95)	
Chicas	15 ^b	92,4	83,1	0,961	0,94 - 0,98	71	7,42 (5,67)	16	16,96	18	21,68		15 (p85-p90)
	17 ^a	87,4	88,7									17 (p93-p95)	
Chicos	17 ^b	93,1	92,1	0,969	0,95 - 0,99	63	7,73 (6,04)	15	16,52	17,8	27,44		17 (p93-p95)
	17 ^a	93,1	92,1									17 (p93-p95)	
Niños	19 ^b	90,8	92,5	0,973	0,96 - 0,99	84	9,25 (6,05)	17	18	19,5	26,6		19 (p93-p95)
	17 ^a	95,5	86									17 (p90-p93)	
Adolescentes	16 ^b	91,2	95,1	0,986	0,97 - 0,99	50	4,74 (4,14)	11	12,43	13,9	15		16 (>p98)
	17 ^a	87,8	100									17 (>p98)	

AUC: área debajo de la curva (*area under the curve*); IC del 95%: intervalo de confianza del 95%; M (DE): media (desviación estándar).

^a Punto de corte óptimo establecido en el índice de Youden máximo (sensibilidad + especificidad – 1) en la muestra total.

^b Punto de corte óptimo establecido en el índice de Youden máximo (sensibilidad + especificidad – 1) en las submuestras (chicos, chicas; niños, adolescentes).

corregido (r) entre la puntuación del ítem y la obtenida en la escala total, así como entre la puntuación del ítem y la subescala correspondiente. Los valores iguales o superiores a 0,25-0,30 se consideran adecuados³⁵. Tal y como se observa en la figura 4, todos los coeficientes (cargas factoriales) se encuentran por encima del rango 0,25-0,30, demostrando que los ítems son representativos de un único factor y dando cuenta de la robustez de la escala. La presencia de correlación elevada ($r > 0,40$) entre ítems o factores implica que ambos están midiendo aspectos muy similares. En nuestro modelo encontramos que entre ambos factores existe una correlación moderadamente elevada ($r = 0,62$), demostrando que ambos factores definen conjuntamente una misma entidad diagnóstica. Además, se observan correlaciones positivas elevadas entre distintos pares de ítems: IA7 («pierde cosas») y IA9 («es olvidadizo»), $r = 0,39$; HA8 («dificultad para guardar turno») y HA9 («interrumpe a otros»), $r = 0,41$. La presencia de estos índices podría indicar que la eliminación de uno de ellos podría mejorar la fiabilidad de la escala. Sin embargo, en este caso, tras la realización del análisis eliminando los ítems, la calidad del ajuste no mejora, por lo que el modelo no se modifica.

Con respecto a la estructura factorial de la escala ADHD-RS-IV.es, los resultados obtenidos en el análisis factorial exploratorio y confirmatorio demuestran que el modelo bifactorial es el que mejor se ajusta a la estructura de la escala (CFI=0,92, TLI=0,91, RMSEA=0,07, chi al cuadrado = 4,04), con valores de CFI y TLI > 0,90, que indican buena bondad de ajuste, aunque algunos autores²³ señalan que deberían considerarse valores más cercanos a 0,95. Un valor RMSEA menor de 0,06 indica buen ajuste y valores entre 0,06 y 0,08 indican un ajuste moderado²³. En nuestro análisis podemos comprobar que el valor de la chi al cuadrado entre los grados de libertad es mayor de 2, superior al valor recomendado²⁵. Sin embargo, usando el método de máxima verosimilitud, este valor puede verse alterado en muestras grandes, por lo que no puede utilizarse como único indicador de ajuste en los análisis con SEM, pero sí es conveniente incluirlo en los análisis confirmatorios dado que otros índices se basan en este³³. De este modo, y a pesar del elevado valor de la chi al cuadrado de nuestro modelo, los excelentes valores del resto de las medidas de bondad de ajuste (CFI, TLI y RMSEA) respaldan sin lugar a dudas la replicación de la estructura factorial de la escala original⁹. A pesar de esto, también se sometió a contraste el modelo jerárquico, que mostró un adecuado ajuste, contrariamente a lo encontrado en un estudio previo por Ullebø et al. en 2012 en el que observaron que la escala se ajustaba a un modelo bifactorial a partir de un *General factor ADHD*²¹.

La fiabilidad de un instrumento nos indica la precisión con la que mide aquello que pretende medir, estimándose mediante el alfa de Cronbach³⁰ y, con más precisión en el caso de escalas Likert, mediante el alfa ordinal²⁹. Cuanto mayor es el valor alfa, mayor será la consistencia interna de la escala²⁹. En nuestro estudio, los valores (*Cronbach's/ordinal*) podrían considerarse excelentes tanto para la escala total (0,94/0,95) como para las 2 subescalas de la IA (0,937/0,95) y la HI (0,91/0,94), siendo similares a los demostrados en estudios de validación previos¹⁴⁻¹⁶ (tabla 3).

Por otra parte, estimamos la capacidad discriminante de la escala mediante el análisis de curvas ROC, que permiten

Tabla 4 Comparación de nuestros resultados con otros estudios

	Consistencia interna (coeficientes alfa de Cronbach)		
	IA	HI	Total
DuPaul et al., 1998	0,93	0,91	
Zhang et al., 2005	0,724	0,825	0,795
Döpfner et al., 2006	0,81	0,87	0,88
Vallejo et al.	0,937	0,91	0,94

HI: subescala de hiperactividad/impulsividad de la escala ADHD-RS-IV.es; IA: subescala de inatención de la escala ADHD-RS-IV.es.

estimar la sensibilidad y la especificidad para los distintos puntos de corte de la escala y dan una idea global de la capacidad de discriminación de la escala: «el área bajo la curva ROC estima la capacidad que tiene una prueba para distinguir o discriminar entre enfermos y no enfermos; cuanto mayor sea el área bajo la curva, más capacidad tendrá la prueba para discriminar entre unos y otros»³⁶. El punto de corte óptimo se estableció en 17 puntos para la escala total, con una elevada sensibilidad (92%) y especificidad (99%), mostrando un elevado valor del área bajo la curva (AUC=0,97). Esto demuestra que la escala ADHD-RS-IV.es muestra un buen poder discriminante. Estos valores también se encontraron en los subanálisis realizados atendiendo a género y edad (tabla 4). Este punto de corte (17), obtenido tanto para la muestra total como para las submuestras, corresponde al percentil entre 93 y 95 de los datos normativos extraídos de nuestra pequeña muestra control (tabla 4). En otras palabras, usando este punto de corte, el 93% de la muestra total (clínica y no clínica), cumpliría criterios de TDAH. Al realizar los análisis por subgrupos (varones/mujeres; niños/adolescentes) se encontraron percentiles similares. Además, calculamos los puntos de corte óptimos para cada submuestra (tabla 2), obteniendo un punto de corte menor en el caso de las chicas (15) y mayor en el caso de los chicos (19).

En este estudio cabe señalar algunas limitaciones importantes para interpretar y generalizar sus resultados. En primer lugar, no diferenciamos si el informante era el padre o la madre, tal y como sí recogen algunos estudios previos. En segundo lugar, dado que únicamente fueron los padres quienes cumplieron la escala, los datos obtenidos en este estudio no podrían extrapolarse a la interpretación de dicha escala en profesores. Por último, dada la falta de datos normativos de la escala ADHD-RS-IV.es en la población infantil y adolescente¹⁷, lo evaluamos en una pequeña muestra control de 134 sujetos sanos. Por esta razón, serían necesarios estudios que incluyan muestras más amplias de controles sanos serían necesarias para comprobar estos datos normativos preliminares. Aunque nuestro estudio indica que las puntuaciones de la escala ADHD-RS-IV.es completadas por los padres son capaces de discriminar el grupo clínico del no clínico, cabe recordar que el diagnóstico nunca debe realizarse sobre la base de ningún cuestionario, sino que debe basarse siempre en la entrevista clínica con entrevistas a los padres y al paciente, cuestionarios, información de los profesores y exploraciones complementarias.

A pesar de estas limitaciones, concluimos que la escala ADHD-RS-IV.es presenta una alta fiabilidad y validez, equiparables a la escala original, presenta la misma estructura factorial que la original y propuesta en DSM-IV-TR² y DSM-5¹, y muestra propiedades psicométricas clínicamente aceptables, convirtiéndola en un instrumento válido y fiable para evaluar la intensidad y frecuencia de los síntomas de TDAH en población clínica.

Conflicto de intereses

María Vallejo-Valdivielso ha recibido financiación departamental para estudios de investigación (no personal), de la Fundación Caja Navarra (CAN), de la Sociedad Vasco-Navarra de Psiquiatría (SVNP), del Gobierno de Navarra y de la Sociedad Española de Psiquiatría Infantil y Adolescente (AEPNYA). También ha recibido financiación para distintos cursos de formación continuada por parte de Shire, Janssen y Lundbeck.

César A. Soutullo ha recibido financiación departamental para estudios de investigación (no personal) de la Fundación Caja Navarra (CAN), Eli Lilly, Lundbeck, Shire y TEVA. Ha servido en consejos asesores de: Fundación Alicia Koplowitz, Editorial Médica Panamericana, Eli Lilly, European Network on Hyperkinetic Disorder (EUNETHYDIS), Instituto de Salud Carlos III (FIS), NeuroTech Solutions Ltd, Ministerio de Sanidad (Guías de Práctica Clínica en TDAH y Depresión), Rubió y Shire. Ha impartido ponencias en cursos de formación médica continuada (no sobre un fármaco) para Eli Lilly, Shire, Universidad Internacional Menéndez Pelayo y la Universidad Internacional de La Rioja (UNIR). Ha recibido además derechos autor por parte de: DOYMA, Editorial Médica Panamericana, EUNSA y Mayo Ediciones.

Pilar de Castro-Manglano ha recibido financiación departamental para estudios de investigación (no personal) de la Fundación Caja Navarra (CAN), Eli Lilly, Lundbeck y Shire. Ha servido en consejos asesores de: Fundación Alicia Koplowitz, Editorial Médica Panamericana y Eli Lilly. Ha impartido ponencias en cursos de formación médica continuada (no sobre un fármaco) para Shire.

Juan J. Marín-Méndez ha recibido financiación departamental para estudios de investigación (no personal) de la Fundación Caja Navarra (CAN), del Instituto Carlos III, del Gobierno de Navarra, de la Fundación QPEA y de Shire. También ha recibido financiación para distintos cursos de formación continuada por parte de Shire, Eli Lilly, Rovi, Roche y Pfizer.

Azucena Díez-Suárez ha recibido financiación departamental para estudios de investigación (no personal) de la Fundación Caja Navarra (CAN), Otsuka Pharmaceuticals, Lundbeck y Shire. Ha servido en consejos asesores de: Fundación Alicia Koplowitz, Editorial Médica Panamericana y Eli Lilly. Ha impartido ponencias en cursos de formación médica continuada (no sobre un fármaco) para Shire.

Bibliografía

1. American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders, 5th edition (DSM-5). Washington, DC: American Psychiatric Association; 2013.
2. American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders, fourth edition, text revised (DSM-IV-TR). Washington, DC: American Psychiatric Association; 2000.
3. DuPaul GJ, Power TJ, Anastopoulos AD, Reid R, McGoey KE, Ikeda MJ. Teacher ratings of attention deficit hyperactivity disorder symptoms: Factor structure and normative data. *Psychol Assess.* 1997;9:436–44.
4. DuPaul GJ, Reid R, Anastopoulos AD, Lambert MC, Watkins MW, Power TJ. Parent and teacher ratings of attention-deficit/hyperactivity disorder symptoms: Factor structure and normative data. *Psychol Assess.* 2016;28:214–25.
5. Anastopoulos AD, Shelton TL. Assessing attention-deficit/hyperactivity disorder. New York: Kluwer Academic/Plenum Press; 2001.
6. Barkley RA, editor. Attention-deficit/hyperactivity disorder: A handbook for diagnosis and treatment. 4th ed. New York: Guilford Press; 2015.
7. DuPaul GJ, Reid R, Anastopoulos AD, Power TJ. Assessing ADHD symptomatic behaviors and functional impairment in school settings: Impact of student and teacher characteristics. *Sch Psychol Q.* 2014;29:409–21.
8. Pelham WE Jr, Fabiano GA, Massetti GM. Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *J Clin Child Adolesc Psychol.* 2005;34:449–76.
9. DuPaul GJ, Power TJ, Anastopoulos AD, Reid R. ADHD Rating Scale-IV: Checklists, norms and clinical interpretation. New York: Guilford Press; 1998.
10. Holland ML, Gimpel GA, Merrell KW. Innovations in assessing ADHD: Development, psychometric properties, and factor structure of the ADHD Symptoms Rating Scale (ADHD-SRS). *J Psychopathol Behav Assess.* 1998;20:307–32.
11. Conners CK. Multi-Health Systems. 3rd ed. Toronto: Conners; 2008.
12. Swanson J, Nolan W, Pelham WE. The SNAP-IV Rating Scale. 1992 [consultado 14 Feb 2017]. Disponible en: <http://www.ADHD.net/snap-iv-instructions.pdf>
13. DuPaul GJ, Anastopoulos AD, Power TJ, Reid R, McGoey KE, Ikeda MJ. Parent ratings of ADHD symptoms: Factor structure, normative data, and psychometric properties. *J Psychopathol Behav Assess.* 1998;20:83–102.
14. Power TJ, Doherty BJ, Panichelli-Mindel SM, Karustis JL, Eiraldi RB, Anastopoulos AD, et al. Integrating parent and teacher reports in the diagnostic assessment of ADHD. *J Psychopathol Behav Assess.* 1998;20:57–81.
15. Zhang S, Faries DE, Vowles M, Michelson D. ADHD Rating Scale IV: Psychometric properties from a multinational study as a clinician-administered instrument. *Int J Methods Psychiatr Res.* 2005;14:186–201.
16. Döpfner M, Steinhausen HC, Coghill D, Dalsgaard S, Poole L, Ralston SJ, et al. Cross-cultural reliability and validity of ADHD assessed by the ADHD Rating Scale in a pan-European study. *Eur Child Adolesc Psychiatry.* 2006;15(Suppl 1):S46–55.
17. Servera M, Cardo E. ADHD Rating Scale-IV en una muestra escolar española: datos normativos y consistencia interna para maestros, padres y madres. *Rev Neurol.* 2007;45:393–9.
18. Kaufman J, Birmaher B, Brent D, Rao U, Ryan N. Schedule for affective disorders and schizophrenia for school aged children — Present and lifetime version (K-SADS-PL). Pittsburgh: University of Pittsburgh Press; 1996.
19. Ambrosini PJ. Historical development and present status of the schedule for affective disorders and schizophrenia for school-age children (KSADS). *J Am Acad Child Adolesc Psychiatry.* 2000;39:49–58.
20. Zumbo BD, Gadermann AM, Zeisser C. Ordinal versions of coefficients alpha and theta for Likert ratings scales. *J Mod App Stat Meth.* 2007;6:21–9.

21. Ullebø AK, Breivik K, Gillberg C, Lundervold AJ, Posserud MB. The factor structure of ADHD in a general population of primary school children. *J Child Psychol Psychiatry*. 2012;53:927–36.
22. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull*. 1990;107:238–46.
23. Byrne BM. Structural equation modeling with AMOS: Basic concepts, applications and programming. Mahwah: Erlbaum; 2001.
24. Steiger JH, Lind JC. Statistically-based tests for the number of common factors. Paper presented at the annual spring meeting of the Psychometric Society. Iowa City: IA; 1980.
25. Kline RB. Principles and practice of structural equation modeling. New York: Guilford Press; 1998.
26. Boomsma A. Reporting analyses of covariance structures. *Struct Equ Modeling*. 2000;7:461–83. http://dx.doi.org/10.1207/S15328007SEM0703_6
27. Browne MW, Cudeck R. Alternative ways of accessing model fit. En: Bollen KA, Long JS, editores. *Testing structural equation models*. Newbury Park: Sage; 1993. p. 136–62.
28. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6:1–55. <http://dx.doi.org/10.1080/107055199095401>
29. Coveney, J. (2012). Calculate alpha after polychoric factor analysis [consultado 26 Jun 2017]. Disponible en: <http://www.stata.com/statalist/archive/2012-02/msg00696.html>
30. Cronbach LF. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297–334.
31. Clark LA, Watson D. Constructing validity: Basic issues in objective scale development. En: Kazdin AE, editor. *Methodological issues & strategies in clinical research*. 3.ª ed. Washington: APA; 2003. p. 207–31.
32. Newcombe RG. Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Stat Med*. 1998;17:857–72.
33. Goldberg LW, Digman JM. Revealing structure in the data: Principles of exploratory factor analysis. En: Strack S, Lorr M, editores. *Differentiating normal and abnormal personality*. New York: Springer; 1994. p. 216–42.
34. Hair JF, Anderson RE, Tatham RL, Black WC. *Multivariate data analysis with readings*. 4.ª ed. Englewood Cliffs: Prentice Hall; 1995.
35. Nunally JC, Bernstein IJ. *Teoría psicométrica*. Madrid: McGraw-Hill; 1995.
36. De Irala J, Martínez-González, Sánchez Vilelga A. Epidemiología clínica. Diagnóstico precoz. En: De Irala-Estévez J, Martínez-González MA, Seguí-Gómez M, editores. *Epidemiología aplicada*. Barcelona: Ariel; 2008. p. 393–432.