

Validez del Ages & Stages questionnaires para predecir el desempeño cognitivo en los primeros años de educación escolar

Predictive value of Ages & Stages Questionnaires for cognitive performance at early years of schooling

Luisa Schonhaut B.^a, Marcela Pérez R.^a, Ana María Castilla F.^b, Sonia Castro M.^{a,b},
Patricia Salinas A.^a y Iván Armijo R.^c

^aServicio de Pediatría, Clínica Alemana, Facultad de Medicina, Universidad del Desarrollo, Santiago, Chile

^b"La Casa Azul", Centro Integral de Atención Psiquiátrica y Psicológica, Santiago, Chile

^cFacultad de Psicología, Universidad Gabriela Mistral, Santiago, Chile

Recibido el 20 de noviembre de 2015; aceptado el 18 de agosto de 2016

Resumen

Introducción: El *Ages and Stages questionnaires* (ASQ) fue validado en nuestro país para el tamizaje del desarrollo psicomotor. El objetivo es evaluar la validez del ASQ para predecir un menor rendimiento cognitivo en los primeros años de educación escolar. **Pacientes y Método:** Estudio de pruebas diagnósticas en una muestra de niños de nivel socioeconómico medio alto evaluados previamente con ASQ a los 8, 18 y/o 30 meses, y luego con la Escala de inteligencia para niños de Wechsler-tercera edición (WISC-III) entre los 6 y 9 años de edad. Cada evaluación con ASQ se contabilizó en forma independiente. La Escala de WISC-III fue estandarizada para la muestra, considerando rendimiento bajo cuando el puntaje total estaba bajo una desviación estándar de la media. **Resultados:** Se incluyeron 123 niños, correspondientes a 174 evaluaciones ASQ (42 de 8; 55 de 18; 77 de 30 meses de edad). El área bajo la curva ROC fue 80,7%, siendo superior a los 8 meses (98,0%) que a los 18 y 30 meses (78,1 y 79,3%, respectivamente). Considerando distintos criterios de corte con ASQ se obtuvo sensibilidad baja (27,8% a 50,0%), pero especificidad elevada (78,8% a 96,2%); el valor predictivo positivo varió entre 21,4% y 45,5%, mientras que el valor predictivo negativo fue 92,0-93,2%. **Conclusión:** El ASQ tiene baja sensibilidad pero excelente especificidad para predecir un menor rendimiento cognitivo durante los primeros años de educación escolar, siendo una buena alternativa para monitorizar el desarrollo psicomotor de niños que se atienden en el sector privado de salud de nuestro país.

Palabras clave:

Desarrollo psicomotor;
Tamizaje;
Ages and Stages questionnaires;
Aprendizaje;
Desarrollo cognitivo.

Abstract

Introduction: The Ages and Stages questionnaires (ASQ) has been recently validated in our country for developmental screening. The objective of this study is evaluate the validity of ASQ to predict low cognitive performance in the early years of schooling. **Patients and Method:** Diagnostic test studies conducted on a sample of children of medium-high socioeconomic level were evaluated using ASQ at least once at 8, 18 and/or 30 months old, and later, between 6 and 9 years old, reevaluated using the Wechsler Intelligence Scale for Children-third edition (WISC-III). Each ASQ evaluation was recorded independently. WISC-III was standardized, considering underperformance when the total score were under -1 standard deviation. **Results:** 123 children, corresponding to 174 ASQ assessments (42 of them were 8 months old, 55 were 18 months and 77 were 30 months of age) were included. An area under the ROC curve of 80.7% was obtained, showing higher values at 8 months (98.0%) compared to 18 and 30 months old (78.1 and 79.3%, respectively). Considering different ASQ scoring criteria, a low sensitivity (27.8 to 50.0%), but a high specificity (78.8 to 96.2%) were obtained; the positive predictive value ranged between 21 and 46%, while the negative value was 92.0-93.2%. **Conclusion** ASQ has low sensitivity but excellent specificity to predict a low cognitive performance during the first years of schooling, being a good alternative to monitor psychomotor development in children who attend the private sector healthcare in our country.

Keywords:

Psychomotor development; Screening; Ages and Stages questionnaires; Learning; Cognitive development.

Introducción

La identificación de las dificultades del desarrollo psicomotor (DSM) es un eje central en los controles de supervisión de salud, puesto que da la oportunidad de derivar a los niños a una intervención oportuna. Considerando que la impresión clínica es poco precisa, en especial al evaluar niños que no tienen factores de riesgo^{1,2}, la Academia Americana de Pediatría recomienda el tamizaje del DSM mediante pruebas estandarizadas a los 9, 18 y 30 meses³, siendo la práctica similar en Atención Primaria de Salud de nuestro país⁴.

Existen distintas escalas para la evaluación del DSM, algunas deben ser aplicadas por un profesional capacitado, con el consiguiente inconveniente del consumo de tiempo del especialista y de la exposición del niño a una “situación extraña”, mientras que otras son completadas por los padres o cuidadores principales en el ambiente natural del niño³. La ventaja de estas últimas se relaciona con su bajo costo, fácil aplicación e involucramiento de los padres⁵⁻⁷.

Una de las escalas de reporte es el *Ages and Stages questionnaires* (ASQ), que fue desarrollada en la Universidad de Oregón en los años 80, y posteriormente actualizada y validada en varios países^{8,9}. En Estados Unidos ha ganado amplia popularidad en la última década¹⁰. El ASQ consiste en una serie de cuestionarios apropiados para distintos rangos etarios, que valora 5 dominios: comunicación, motricidad fina, motricidad gruesa, resolución de problemas y relaciones interpersonales.

De acuerdo a la validación original se recomienda considerar que un niño está en riesgo de déficit del

DSM y derivarlo a una evaluación más especializada en caso de tener un rendimiento inferior a -2 desviaciones estándar (DS) en al menos en una de las áreas del cuestionario⁸. No obstante, en investigaciones recientes hay discusión al respecto, ya que se ha reportado que al considerar un rendimiento inferior al punto de corte en 2 o más dominios aumenta la especificidad del ASQ¹¹⁻¹³, mientras que la eficiencia mejora al considerar el puntaje total^{13,14}. Independiente del criterio se ha descrito que al emplear un método de reporte de padres aumenta significativamente la pesquisa de posibles trastornos del desarrollo^{2,15}, siendo esperable hasta un 30% de sobrederivación de los niños evaluados^{8,16}. La debilidad de la evidencia actual radica en que la gran mayoría de los estudios son concurrentes, siendo importante diseñar investigaciones prospectivas que permitan definir un criterio de déficit que tenga real impacto clínico.

En nuestro país el ASQ fue validado en una amplia muestra comunitaria¹⁷. En una submuestra de niños de nivel socioeconómico medio alto, con desarrollo aparentemente típico, evaluamos la validez concurrente del ASQ, destacando valores de sensibilidad y especificidad del 73% y 81% respectivamente¹⁸. En dicha muestra el 14% de los niños tuvo un tamizaje positivo¹⁹, no obstante desconocemos cuál ha sido la trayectoria del desarrollo de aquellos niños.

El objetivo del presente manuscrito es evaluar la validez predictiva del ASQ para un menor rendimiento cognitivo en los primeros años de educación escolar, en una muestra de niños de nivel socioeconómico medio alto y, además, comparar la precisión diagnóstica relativa de distintos criterios del ASQ como predictores de rendimiento intelectual bajo el promedio.

Pacientes y Método

El presente estudio es de pruebas diagnósticas, realizado a partir de una muestra de oportunidad de niños de nivel socioeconómico medio alto, que fueron evaluados al menos en una ocasión con ASQ (a los 8, 18 y/o 30 meses) entre los años 2008 y 2011¹⁸, y luego reevaluados con la Escala de inteligencia para niños de Wechsler–tercera edición (WISC-III) entre los 6 y 9 años de edad. Cada una de las evaluaciones con ASQ se contabilizó de forma independiente para el análisis.

Criterios de inclusión

Haber sido previamente evaluados con las escalas de ASQ a los 8, 18 y/o 30 meses de edad, en ausencia de diagnóstico conocido de enfermedad neurológica, congénita o metabólica, o que el cuidador principal no hablara fluidamente el español. En los nacidos de menos de 37 semanas de edad gestacional la edad fue corregida en las evaluaciones de los 8 y 18 meses, a los 30 meses se consideró la edad cronológica.

Criterios de exclusión

Historia de afección intercurrente que pueda afectar el curso del desarrollo, como por ejemplo meningitis, enfermedad tumoral del SNC, accidente vascular, TEC severo diagnosticado posterior a la aplicación del ASQ. Además se excluyeron los niños que en el momento de la evaluación estaban viviendo fuera del país o que habían cumplido los 9 años en el momento de la evaluación.

Los niños con trastorno severo del desarrollo (que no tenían diagnóstico de enfermedad orgánica en el momento de la aplicación del ASQ), que no estaban escolarizados y/o en condiciones de participar de la evaluación, no fueron examinados con WISC-III, pero considerados con rendimiento bajo para el análisis final.

Los padres de los niños fueron contactados telefónicamente, y accedieron participar mediante un consentimiento informado. Luego completaron una ficha con información biodemográfica, antecedentes neonatales, académicos y referentes a posibles diagnósticos e intervenciones relacionadas con el desarrollo y aprendizaje de los niños. Todos los niños tenían al menos uno de los padres con educación superior y provenían de familias pertenecientes al quinto quintil de ingreso.

Instrumentos

*Ages and Stages questionnaire*⁸

Serie de 21 cuestionarios para niños de 2 meses hasta los 5 años. Cada cuestionario consta de 5 dominios, de 6 preguntas cada uno, que evalúan diferentes ámbitos del DSM, como son comunicación, motrici-

dad fina, motricidad gruesa, resolución de problemas y relaciones interpersonales. Cada pregunta se responde con las alternativas “sí” (10 puntos), “ocasionalmente” (5 puntos) o “todavía no” (0 puntos); los puntajes de cada dominio se suman, teniendo un total posible de 60 puntos por área. Al final hay una sección de 7 preguntas abiertas destinadas a indagar posibles preocupaciones de los padres, las que no fueron consideradas en el presente análisis. Los padres fueron quienes completaron el ASQ a las edades correspondientes. El cuestionario fue validado en nuestro país por Schonhaut y Armijo (2013 y 2015)^{17,18}.

*Escala de inteligencia para niños de Wechsler–tercera edición*²⁰

Instrumento clínico de administración individual para evaluar capacidad intelectual de niños de 6 a 16 años y 11 meses de edad, mediante 13 subpruebas, de las cuales 6 son verbales y 7 manuales. El rendimiento de los niños en las diversas pruebas se resume en 3 puntajes compuestos, que son el cociente intelectual verbal, de ejecución y el total. Este instrumento está adaptado y estandarizado en Chile por Ramírez y Rozas en el año 2007²¹. Para el presente análisis solo consideramos el cociente intelectual total.

El WISC-III fue aplicado por un grupo de psicólogas capacitadas en dicha evaluación y ciegas a los resultados de ASQ. Las evaluaciones fueron realizadas en una consulta o en el domicilio de los pacientes, y los puntajes asignados luego de una revisión consensuada por el equipo profesional participante. Finalmente se les entregó a los padres un informe del desempeño logrado por los niños.

Definición de riesgo de déficit del desarrollo psicomotor según el Ages and Stages questionnaires

Los puntajes totales del ASQ fueron estandarizados y llevados a puntaje Z en cada una de las edades evaluadas para así poder analizarlas en conjunto. Se revisaron distintos criterios de déficit: rendimiento < -2 desviaciones estándar (DS) en al menos un área del desarrollo; rendimiento < -2 DS en 2 o más áreas; puntaje total < -2 DS.

Definición de rendimiento bajo en la escala de inteligencia para niños de Wechsler

El puntaje obtenido en el test de WISC-III fue estandarizado para la muestra estudiada y se definió que el rendimiento era bajo cuando el puntaje total estaba bajo 1 DS de la media.

Análisis estadístico

Se analizó la normalidad de la distribución de los puntajes WISC-III y ASQ obtenidos usando la prueba de Shapiro-Wilk. Como indicador de consolidación

de desarrollo se utilizaron los puntajes estandarizados en la prueba WISC-III por todos los casos analizados, separados en 2 grupos: “rendimiento bajo” y “rendimiento normal” (el resto de los casos).

Se construyeron curvas ROC para determinar los puntajes en la prueba ASQ aplicados a las distintas edades de los niños, que permiten optimizar la precisión diagnóstica al momento de detectar casos con bajo rendimiento.

Si bien no existe el estándar absoluto en la interpretación de área bajo la curva (AUC en las siglas en inglés), es aceptado que un AUC igual a 0,5 representa un test sin valor predictivo; interpretándose valores entre 0,6 y 0,7 como aceptables; 0,7 a 0,9 como buenos y sobre 0,9 como excelentes²². Para comparar la precisión relativa obtenida entre las distintas edades de aplicación del ASQ para identificar casos con bajo rendimiento, comparando las AUC de las distintas curvas ROC obtenidas, se aplicó el método de DeLong²³.

Se revisaron las propiedades psicométricas del ASQ (sensibilidad, especificidad y valores predictivos) con distintos criterios de riesgo de déficit del desarrollo, definidos previamente.

Todos los análisis fueron realizados sobre la plataforma R²⁴, específicamente con la implementación del módulo PROC²⁵.

El estudio contó con la aprobación del Comité de Ética de la Investigación de la Facultad de Medicina de la Clínica Alemana-Universidad del Desarrollo.

Resultados

Se incluyeron en el análisis 123 niños con antecedente de haber recibido al menos una evaluación previa con ASQ. A 121 niños se les aplicó el test de WISC, y 2 niños con diagnóstico de trastorno del desarrollo no fueron evaluados, pero considerados en el análisis final, por lo que se analizó un total de 174 evaluaciones con ASQ. De esta muestra 73 habían nacido de término, 34 eran prematuros de 32 a 36 semanas de edad gestacional y 16 habían nacido antes de las 32 semanas de gestación (tabla 1).

Del total de las evaluaciones con ASQ 42 fueron realizadas a los 8 meses, 55 a los 18 meses y 77 evaluaciones fueron aplicadas a los 30 meses de edad. Siete niños tenían 3 evaluaciones, 37 tenían 2 evaluaciones y los restantes habían sido evaluados en una sola oportunidad (tabla 1).

El puntaje total obtenido en el test de WISC-III tuvo una distribución normal, con un promedio de 114,8 y desviación estándar 12,8 (test de normalidad de Shapiro-Wilk, $p = 0,19$). Quince niños cumplieron los criterios de rendimiento bajo en el WISC-III (13 con rendimiento inferior a -1 DS y 2 niños que no pudie-

ron ser evaluados), equivalente al 12,2% de la muestra.

En el análisis de la curva ROC se observó que los puntajes globales de ASQ produjeron un área bajo la curva de 80,7% (IC 95%: 71,8 a 89,5). En la figura 1 se muestran las curvas ROC a las edades en que se aplicó el ASQ, tomando como referencia el rendimiento bajo en el WISC-III; destaca una excelente precisión a los 8 meses, con un área bajo la curva ROC de 98,0% (IC 95%: 94,3 a 100), comparado con los 18 (78,1%) (IC 95%: 60,1-96,0) y los 30 meses (79,3%) (IC 95%: 67,3 a 91,4) en que la precisión fue buena. Al comparar la precisión de las curvas usando el Test de DeLong, encontramos diferencias significativas entre: 8 versus 18 meses ($D = -2,13$, $gl = 58,7$, $p = 0,037$); 8 versus 30 meses ($D = -2,90$, $gl = 90,0$; $p = 0,004$); pero no entre 18 versus 30 meses ($D = -0,12$; $gl = 99,16$; $p = 0,907$).

Al comparar los distintos criterios de riesgo de déficit del DSM, no encontramos importantes diferencias en las propiedades psicométricas del ASQ, considerando: puntaje total < -2 DS obtuvimos valores de sensibilidad del 33,3% y especificidad del 92,9%; cuando tomamos al menos 2 áreas bajo el punto de corte la sensibilidad fue del 27,8% y la especificidad del 96,2%, mientras que con el criterio de al menos un área bajo el punto de corte la sensibilidad subió a 50,0%, cayendo la especificidad a 78,8%. El valor predictivo positivo varió de 21,4% a 45,5%, y los porcentajes de sobre-

Tabla 1. Características biodemográficas de la muestra

	N	(%)
<i>Número de niños incluidos</i>	123	(100)
<i>Género</i>		
Masculino	70	(56,9)
Femenino	53	(43,1)
<i>Edad gestacional</i>		
RNT	73	(59,3)
PMT	34	(27,6)
PE	16	(13,0)
<i>Gemelar</i>		
Sí	33	(26,8)
No	90	(73,2)
<i>Número de evaluaciones con ASQ por niño</i>		
1	79	(64,2)
2	37	(30,0)
3	7	(5,7)
Total evaluaciones	174	(100)
<i>Edad evaluación</i>		
8 meses	42	(24,1)
18 meses	55	(31,6)
30 meses	77	(44,3)

PE: prematuros extremos (edad gestacional < 32 semanas de gestación); PMT: prematuros moderados y tardíos (32-36 semanas de edad gestacional); RNT: recién nacidos a término (37-42 semanas de gestación).

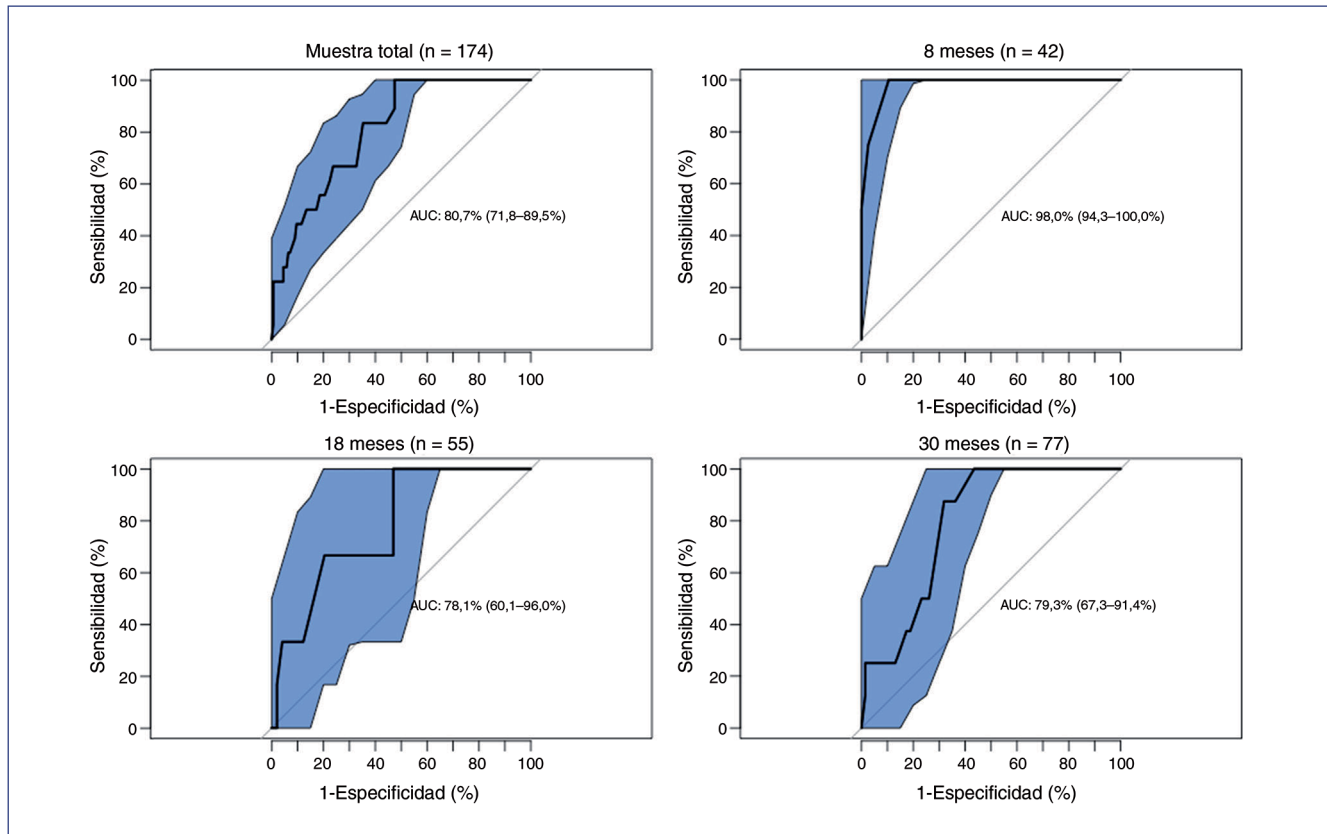


Figura 1. Curva ROC de la totalidad de la muestra y analizado según la edad de evaluación, tomando como referencia el rendimiento < -1 DS en la Escala de inteligencia para niños de Wechsler–tercera edición (WISC-III) estandarizada para la muestra. AUC: área bajo a curva ROC. Test de DeLong para comparar 2 curvas de ROC, 8 versus 18 meses (D = -2,13, gl = 58,7, p = 0,037); 8 versus 30 meses (D = -2,90; gl = 90,0; p = 0,004); 18 versus 30 meses (D = -0,12; gl = 99,16; p = 0,907).

Tabla 2. Análisis de las propiedades psicométricas en 174 evaluaciones con ASQ, con distintos criterios de definición de riesgo de déficit del desarrollo, tomando como referencia el rendimiento < -1 DS en la Escala de inteligencia para niños de Wechsler–tercera edición estandarizada para la muestra

Criterio ASQ	Sensibilidad	Especificidad	VPP	VPN	Sobre-referencia	Infra-referencia
Análisis puntaje total Z < -2	33,3 (14,4-58,9)	92,9 (87,4-96,3)	35,3 (15,3-61,3)	92,4 (86,7-95,8)	64,7 (32,9-82,5)	7,6 (4,4-14,0)
Al menos un área < -2 DS	50,0 (26,8-73,2)	78,8 (71,4-84,8)	21,4 (10,8-37,2)	93,2 (87,1-96,6)	78,5 (62,8-89,2)	6,8 (3,4-12,9)
Dos o más áreas < -2 DS	27,8 (10,7-53,6)	96,2 (91,4-98,4)	45,5 (18,1-75,4)	92,0 (86,4-95,5)	54,5 (24,6-81,9)	8,0 (4,5-13,5)

DS: desviación estándar; VPN: valor predictivo negativo; VPP: valor predictivo positivo.

referencia se encontraban entre 54,5% a 78,5%. Los valores predictivos negativos estuvieron entre 92,0% y 93,2% en los 3 criterios de análisis (tabla 2).

Discusión

En el presente estudio encontramos un área bajo la curva ROC de 80,7%, lo que traduce una buena capa-

acidad discriminativa global del ASQ para predecir un menor rendimiento cognitivo al inicio de la etapa escolar. Esto se manifestó en adecuados valores de especificidad (78,8% a 96,2%), pero baja sensibilidad (27,8% a 50,0%), independiente del criterio de riesgo utilizado.

Nuestros resultados contrastan con los estudios de Halbwachs et al., quienes reportaron una elevada sensibilidad (80%) y moderada especificidad (54%)

del puntaje global, en una muestra de prematuros²⁶, mientras que Kerstjens et al., basados en el criterio de un área bajo el punto de corte, obtuvieron 89% y 80% de sensibilidad y especificidad respectivamente para predecir necesidad de educación especial en un plazo de un año de seguimiento²⁷. No se han publicado otros estudios que muestren la validez predictiva del ASQ para predecir el menor rendimiento cognitivo a largo plazo.

En nuestro análisis destacaron los valores predictivos positivos bajos (21,4% a 45,5%), con alto porcentaje de sobre-referencia, siendo el criterio de 2 áreas bajo el punto de corte el que produjo una tasa de sobrederivación más cercana al valor esperado de 30%^{8,16}. El problema de la sobre-referencia es el alto costo económico y familiar que implica²⁸, factores a considerar en el momento de establecer políticas de salud pública. Por otro lado, las tasas de infra-referencia fueron inferiores al 10%, lo que es aceptable⁸.

A diferencia de lo reportado por Halbwachs con ASQ, y por Doyle con la escala de Bayley^{29,30}, no encontramos una tendencia a la mejoría de las propiedades psicométricas del ASQ a mayor edad de los niños. A los 8 meses se obtuvo una curva ROC con una excelente capacidad discriminativa (AUC 98%) comparada con los 18 y 30 meses en que el AUC fue buena (78,1% y 79,3%, respectivamente); considerando que el límite inferior del intervalo de confianza superó al 50% en todos los casos, podemos asumir que el ASQ identifica a los niños con posibles trastornos del desarrollo de forma superior al azar. Al respecto cabe destacar que a los 8 meses predominaron las dificultades del desarrollo motor grueso comparado con los demás dominios³¹, lo que demuestra cuán imbricadas están las distintas áreas del desarrollo.

Para el correcto análisis de nuestros resultados debemos considerar que evaluamos niños que han crecido en un ambiente enriquecido culturalmente, y han recibido otras evaluaciones e intervenciones en su desarrollo, no analizadas en este estudio, lo que podría asociarse a un mejor rendimiento cognitivo y/o académico^{32,33}. El elevado rendimiento global de los niños, con baja prevalencia de problemas reales del desarrollo, podría explicar en gran medida los bajos valores predictivos positivos y la elevada sobre-referencia obtenida en nuestro estudio. Otra limitación se relaciona con el hecho de tratarse de una muestra de oportunidad, lo que no permite inferencia estadística. Además, podría cuestionarse el hecho de que solo medimos el desarrollo cognitivo sin evaluar otros dominios del desarrollo, como es la motricidad o el desempeño socioemocional, aspectos que escaparon a los objetivos del estudio, coincidiendo con otras publicaciones sobre este tema^{26,27}.

Como conclusión, podemos afirmar que, en la

muestra estudiada, el ASQ tiene baja sensibilidad pero excelente especificidad para predecir un menor rendimiento cognitivo durante los primeros años de educación escolar. Considerando el elevado porcentaje de sobre-referencia, recomendamos derivar a un niño para una evaluación más integral en caso de tener 2 o más áreas bajo el punto de corte; a los niños que tienen un área deficitaria se les podrá sugerir un plan de estimulación y una posterior reevaluación con ASQ. Nuestros resultados refuerzan la importancia de la evaluación temprana del desarrollo, siendo una buena alternativa el ASQ para monitorizar el DSM de niños que se atienden en el sector privado de salud de nuestro país.

Responsabilidades éticas

Protección de personas y animales: Los autores declaran que los procedimientos seguidos se conformaron a las normas éticas del comité de experimentación humana responsable y de acuerdo con la Asociación Médica Mundial y la Declaración de Helsinki.

Confidencialidad de los datos: Los autores declaran que han seguido los protocolos de su centro de trabajo sobre la publicación de datos de pacientes.

Derecho a la privacidad y consentimiento informado: Los autores han obtenido el consentimiento informado de los pacientes y/o sujetos referidos en el artículo. Este documento obra en poder del autor de correspondencia.

Financiación

Proyecto financiado por Concurso de proyectos internos Universidad del desarrollo (ID 201310091 61129950210) y Concurso Proyectos Sochipe 2012.

Conflicto de intereses

Los autores declaran no tener conflicto de intereses.

Agradecimientos

Agradecemos a las psicólogas Lorero González, Constanza Osorio, María Teresa Burich, Pía Margalet, Daniela Werner, Josefina Jirón y Bernardita Martínez por su apoyo, cariño y dedicación en la evaluación de los niños.

Referencias

- Sheldrick RC, Merchant S, Perrin EC. Identification of developmental-behavioral problems in primary care: A systematic review. *Pediatrics*. 2011;128:356-63.
- Hix-Small H, Marks K, Squires J, Nickel R. Impact of implementing developmental screening at 12 and 24 months in a pediatric practice. *Pediatrics*. 2007;120:381-9.
- American Academy of Pediatrics. Council on children with disabilities, section on developmental behavioral pediatrics, Bright Futures Steering Committee, Medical Home Initiatives For Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*. 2006;118:405-20.
- Normas técnicas de evaluación y estimulación del desarrollo psicomotor en el niño y la niña menor de 6 años. Ministerio de Salud. Santiago Chile 2004 [consultado 13 Nov 2015]. Disponible en: <http://www.crececontigo.gob.cl/wp-content/uploads/2013/06/Normas-tecnicas-estimulacion-y-evaluacion-desarrollo.pdf>
- Sices L, Drotar D, Keilman A, Kirchner HL, Roberts D, Stancin T. Communication about child development during well-child visits: Impact of parents' evaluation of developmental status screener with or without an informational video. *Pediatrics*. 2008;122:e1091-9.
- Schonwald A, Horan K, Huntington N. Developmental screening. Is there enough time? *Clinical Pediatrics*. 2009;48:648-655.
- Schonhaut L, Armijo I, Millán T, Herreros J, Hernández K, Salgado AM, et al. Comparación entre la evaluación tradicional desarrollo psicomotor versus una prueba autoadministrada. *Rev Chil Pediatr*. 2010;81:498-505.
- Squires J, Bricker D. Ages & Stages questionnaires. Third edition. (ASQ-3): A parent-completed child-monitoring system. Stanford: Paul Brookes Publishing Company; 2009.
- Schonhaut L, Armijo I. Aplicabilidad del Ages and Stages questionnaire (ASQ) para el tamizaje del desarrollo psicomotor. *Rev Chil Pediatr*. 2014;85:12-21.
- Radecki L, Sand-Loud N, O'Connor KG, Sharp S, Olson LM. Trends in the use of standardized tools for developmental screening in early childhood: 2002-2009. *Pediatrics*. 2011;128:14-9.
- Limbos MM, Joyce DP. Comparison of the ASQ and PEDS in screening for developmental delay in children presenting for primary care. *J Dev Behav Pediatr*. 2011;32:499-511.
- Kapci EG, Kucuker S, Uslu RI. How applicable are Ages and Stages questionnaires for use with Turkish children? *Top Early Child Spec Educ*. 2010;30:176-88.
- Flamant C, Branger B, Nguyen S, et al. Parent-Completed developmental screening in premature children: A valid tool for follow-up programs. *PLoS One*. 2011;6:4-9.
- Hornman J, Kerstjens JM, de Winter AF, Bos AF, Reijneveld SA. Validity and internal consistency of the Ages and Stages questionnaire 60-month version and the effect of three scoring methods. *Early Hum Dev*. 2013;89:1011-5.
- Guevara JP, Gerdes M, Localio R, et al. Effectiveness of developmental screening in an urban setting. *Pediatrics*. 2013;131:30-7.
- Glascoe FP, Byrne KE. The accuracy of three developmental screening tests. *J Early Interv*. 1993;17:368-79.
- Armijo I, Schonhaut L, Cordero M. Validation of the Chilean version of the Ages and Stages questionnaire (ASQ-CL) in community health settings. *Early Hum Dev*. 2015;91:671-6.
- Schonhaut L, Armijo I, Schonstedt M, Alvarez J, Cordero M. Validity of the ages and stages questionnaires in full term and preterm infants. *Pediatrics*. 2013;131:e1468-75.
- Schonhaut L, Salinas P. Evaluación del desarrollo psicomotor en las consultas pediátricas: experiencia en Clínica Alemana. *Contacto Científico*. 2012;2:48-55 [consultado 13 Nov 2015] Disponible en: <http://contactocientifico.alemana.cl/ojs/index.php/cc/article/view/46>
- Wechsler D. Test de inteligencia para niños WISC-III. Manual (traducción de Ofelia Castillo). Buenos Aires. Argentina: Paidós; 1997.
- Ramírez V, Rosas R. Estandarización del WISC-III en Chile: descripción del test, estructura factorial y consistencia interna de las escalas. *Psykhé (Santiago)*. 2007;16:91-109.
- Swets A. Measuring the accuracy of diagnostic systems. *Science*. 1988;240:1285-93.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*. 1988;83:7-845.
- The R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing [consultado 22 Jun 2016]. Disponible en: <http://www.R-project.org/>
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.
- Halbwachs M, Muller JB, Nguyen The Tich S, et al. Usefulness of parent-completed ASQ for neurodevelopmental screening of preterm children at five years of age. *PLoS One*. 2013;8:e71925.
- Kerstjens JM, Bos AF, ten Vergert EMJ, de Meer G, Butcher PR, Reijneveld SA. Support for the global feasibility of the Ages and Stages questionnaire as developmental screener. *Early Hum Dev*. 2009;85:443-7.
- Cadman D, Chambers LW, Walter SD, Ferguson R, Johnston N, McNamee J. Evaluation of public health preschool child developmental screening: the process and outcomes of a community program. *Am J Public Health*. 1987;77:45-51.
- Halbwachs M, Muller J-B, Nguyen The Tich S, et al. Predictive Value of the parent-completed ASQ for school difficulties in preterm-born children < 35 weeks' GA at five years of age. *Neonatology*. 2014;106:311-6.
- Doyle LW, Davis PG, Schmidt B, Anderson PJ. Cognitive outcome at 24 months is more predictive than at 18 months for IQ at 8-9 years in extremely low birth weight children. *Early Hum Dev*. 2012;88:95-8.
- Schonhaut L, Schonstedt M, Álvarez J, Salinas P, Armijo I. Desarrollo psicomotor en niños de nivel socioeconómico medioalto. *Rev Chil Pediatr*. 2010;81:122-7.
- Guralnick MJ. Effectiveness of early intervention for vulnerable children: A developmental perspective. *Am J Ment Retard*. 1998;102:319-45.
- Anderson LM, Shinn C, Fullilove MT, et al. The effectiveness of early childhood development programs. *Am J Prev Med*. 2003;24:32-46.