

Validación de un instrumento para la pesquisa de problemas del desarrollo en menores de 5 años en Uruguay

Gustavo Giachetto Larraz*, Anabella Santoro Ferrero†, María Mercedes Pérez Roca‡, Franklin Fernando Mass Mandagaran§

Resumen

Introducción: la detección temprana de problemas del desarrollo en niños permite aplicar intervenciones oportunas que mejoran el pronóstico y disminuyen la discapacidad. Esto requiere aplicar pruebas de pesquisa sensibles, sencillas y de bajo costo. Con este propósito, en Uruguay se desarrolló la Guía Nacional para la Vigilancia del desarrollo del Niño y la Niña menores de 5 años versión 2 (GNVD V2). Su utilización requiere conocer las propiedades psicométricas e índices fijos y variables.

Objetivo: validación concurrente de la GNVD V2 frente al Inventario de Desarrollo Battelle 4ª edición.

Método: se realizó un estudio transversal, observacional, en una muestra representativa de niños de 1 mes a 5 años de centros educativos de Montevideo. Se relacionaron los resultados de la aplicación de la GNVD V2 con los del Inventario de Desarrollo Battelle 4ª edición. Se calculó la sensibilidad, especificidad, valores predictivos positivo y negativo de la GNVD V2, según diferentes puntos de corte. Se estimó el beneficio clínico de su aplicación mediante las razones de verosimilitud o *likelihood ratios*. Se realizó análisis psicométrico.

Resultados: la validación se realizó con 341 niños. No haber fallado en ninguna de las 12 conductas ubicadas a la izquierda de las cuatro áreas correspondientes a su edad presenta sensibilidad 77%, especificidad 65%, valor predictivo positivo 42% y valor predictivo negativo 89%. Con dicho punto de corte, el LR (+) es 2,2 y el LR (-) 0,35 y la probabilidad de presentar un trastorno cuando no pasa la prueba 1,6 y cuando la pasa 0,26.

Conclusiones: la GNVD V2 posee propiedades psicométricas adecuadas e índices fijos similares a tests validados a nivel internacional que permiten recomendarlo como instrumento nacional de pesquisa.

Palabras clave: Desarrollo infantil
Validación
Prueba de pesquisa
Evaluación de desarrollo infantil

Key words: Child development
Validation
Screening test
Evaluation of child development

* Médico, Pediatra. Profesor Director de Clínica Pediátrica "C", Facultad de Medicina, Universidad de la República.

† Médica, Pediatra, Profesora Agregada Clínica Pediátrica "C", Facultad de Medicina, Universidad de la República.

‡ Directora División Protección Integral. Uruguay Crece Contigo. Ministerio de Desarrollo Social (MIDES). Médica, Pediatra. Asistente técnico, Área Programática de Salud de la Niñez, Ministerio de Salud (MS).

§ Licenciado en Estadística. Profesor asistente, Facultad de Ciencias Económicas y Administración, Universidad de la República.

Clínica Pediátrica "C", Facultad de Medicina (UDELAR); Uruguay Crece Contigo (MIDES); Programa de Salud de la Niñez (MS)

Correspondencia: Dr. Gustavo Giachetto Larraz, Clínica Pediátrica "C", Facultad de Medicina, Universidad de la República. Centro Hospitalario Pereira Rossell, Bulevar General Artigas 1550, CP 11600, Montevideo, Uruguay. Correo electrónico: ggiachet@gmail.com

Conflicto de intereses: el Proyecto de Investigación fue financiado por Uruguay Crece Contigo, MIDES. Recibió el segundo premio de la Academia Nacional de Medicina, 2018.

Recibido: 17/6/19

Aprobado: 25/9/19

Introducción

La primera infancia representa una etapa fundamental que influye de manera crítica durante el resto del ciclo vital. Problemas de elevada prevalencia mundial como enfermedades crónicas no transmisibles, enfermedades mentales, trastornos del aprendizaje e incluso criminalidad, se relacionan con factores de riesgo y daño en esta etapa⁽¹⁻³⁾.

Los países que invierten en programas de desarrollo de la primera infancia (PDI) de calidad poseen poblaciones más alfabetizadas, con mayor conocimiento básico sobre aritmética, mejores condiciones sanitarias y niveles inferiores de desigualdad. En América Latina y el Caribe, la inversión promedio es 0,4% del Producto Interno Bruto (PIB), la mitad que en los países de la OCDE (0,7%) y cuatro veces menos que en países de alto desempeño (1,7%)⁽⁴⁾. La inversión en la promoción del PDI presenta retribuciones en el transcurso de la vida superiores a la inversión inicial, y representa una estrategia fundamental para alcanzar los Objetivos de Desarrollo Sostenible 2030⁽⁵⁾.

Uruguay tiene una larga trayectoria en políticas públicas dirigidas a la primera infancia. El desarrollo exponencial de los Centros de Atención Integral a la Infancia y la Familia (Plan CAIF) desde 1988 y del Programa Uruguay Crece Contigo (UCC) desde el año 2012, ha contribuido a generar la base de un Sistema de Protección Integral a la Primera Infancia. A pesar de esto, la Encuesta Nacional de Desarrollo Infantil y Salud (ENDIS) muestra que aproximadamente el 25% de los niños y niñas menores de 4 años presenta algún tipo de rezago en el desarrollo, es decir, no alcanzan alguno de los hitos de desarrollo esperables para su edad. El rezago es mayor en los quintiles de menor ingreso⁽⁶⁾.

El Ministerio de Salud (MS) ha definido, entre los objetivos estratégicos orientados a “*mejorar el acceso y la atención de salud en el curso de vida*”, la disminución de la prevalencia de alteraciones del desarrollo en niños de 0 a 5 años y la atenuación de la severidad de las alteraciones diagnosticadas⁽⁷⁾.

Los trastornos del desarrollo constituyen un grupo heterogéneo de problemas de prevalencia creciente, que se presentan en la edad pediátrica, entre los que se encuentran déficits motores y neurosensoriales (auditivos y visuales), déficit intelectual, trastornos del espectro autista, trastornos del lenguaje y dificultades de aprendizaje, entre otros. Su detección temprana permite aplicar intervenciones eficaces que impactan positivamente en su evolución, disminuyen la discapacidad y favorecen la inclusión social⁽⁸⁻¹¹⁾. En este sentido, desarrollar acciones de pesquisa o tamizaje, que permitan identificar niños y niñas con riesgo de presentar posibles rezagos o

desvíos, representan un componente prioritario de los programas de salud^(12,13).

Debido a que muchos de los problemas del desarrollo no son aparentes en etapas precoces y el juicio clínico de los pediatras y la observación de los padres o cuidadores no siempre es suficiente para detectar especialmente problemas leves o moderados, la detección oportuna requiere evaluar al niño con pruebas de pesquisa, tamizaje o *screening* objetivas, de fácil administración y bajo costo. Estas pruebas no son diagnósticas. Son útiles para realizar el seguimiento del desarrollo en general.

Estas pruebas deben tener una sensibilidad adecuada para detectar individuos probablemente enfermos en una población presuntamente sana⁽¹⁴⁾.

Desde 1996, la Organización Mundial de la Salud (OMS) promueve que cada país elabore y valide instrumentos de pesquisa adecuados a sus pautas culturales y psicométricamente válidos^(15,16). En el año 2007, el Programa Nacional de Salud de la Niñez del MS, con el apoyo del Centro Latinoamericano de Perinatología de la Organización Panamericana de la Salud (CLAP-OPS), inició el proceso de elaboración de un instrumento de pesquisa que culminó en el año 2010 con la aprobación de la Guía Nacional para la Vigilancia del Desarrollo del Niño y la Niña menores de 5 años. La guía, diseñada para ser aplicada en el primer nivel de atención, evalúa, en grupos de edad predefinidos, los logros en diferentes áreas del desarrollo: motriz grueso, coordinación, lenguaje y social. Adicionalmente, proporciona señales de alerta y un flujograma con orientaciones para la toma de decisiones.

La utilización de una prueba de pesquisa requiere estandarización y conocimiento acerca de sus propiedades psicométricas y de sus índices fijos (sensibilidad y especificidad)^(17,18). Esta información es esencial para evaluar la relación costo-beneficio que deriva de su implementación en programas de salud, así como para la toma de decisiones individuales. Utilizar un instrumento de pesquisa del desarrollo, sin conocer su sensibilidad y especificidad, es tan peligroso como utilizar cualquier otro instrumental o intervención médica sin esta información esencial.

El objetivo general de esta investigación fue la validación concurrente de la Guía Nacional para la Vigilancia del desarrollo del Niño y la Niña menores de 5 años versión 2 (GNVD V2) frente al Inventario de Desarrollo Battelle 4ª edición⁽¹⁹⁾. Los objetivos específicos fueron: 1) Realizar el análisis psicométrico de la GNVD V2 para determinar su validez interna. 2) Determinar la sensibilidad, especificidad, valores predictivos positivo y negativo de la GNVD V2 para los diferentes puntos de corte considerados.

La selección del Inventario de Desarrollo Battelle 4ª edición como patrón oro se basó en recomendaciones de expertos nacionales y extranjeros, su utilización en estudios internacionales similares, y en la experiencia nacional con su aplicación.

Material y método

Diseño

Se realizó un estudio transversal, observacional, entre octubre de 2016 y agosto de 2017. La población objetivo fueron los niños y niñas de 1 mes a 5 años y 11 meses que asistían a los CAIF, Centros de Cuidado Diurno del Instituto del Niño y Adolescente del Uruguay (INAU), centros educativos públicos y privados de la ciudad de Montevideo.

Se incluyeron niños y niñas previamente sanos que asistían a los centros mencionados, cuyos padres o tutores accedieron a participar del estudio previo consentimiento informado. Se excluyeron aquellos que al momento del estudio presentaban sospecha de retraso del desarrollo y enfermedades agudas que impedían la correcta evaluación (síndromes febriles, otitis, etc.) o presentaban enfermedades crónicas con retraso del desarrollo, enfermedades neurológicas, metabólicas y síndromes genéticos asociados a alteraciones del desarrollo. Los niños que luego de diez días de la aplicación de la GNVD V2 no completaron la evaluación diagnóstica, también se excluyeron.

Muestra

El diseño muestral fue probabilístico, bi-etápico y estratificado, proporcional a la cantidad de alumnos inscritos en dichos centros educativos durante 2015 y 2016. En la primera etapa, se seleccionaron centros educativos de cuatro estratos: públicos, CAIF, centros de cuidado diurno y privados. En la segunda, mediante muestreo sistemático, ordenado por edad y sexo, se seleccionó un número fijo de niños por centro. El tamaño muestral estimado fue 500 niños, considerando 25 centros educativos y 20 niños por centro, estimando una tasa de no respuesta de 25%.

Fases del estudio

El estudio se desarrolló en dos fases: I) Ajustes de diseño y II) Validación.

I) Ajustes de diseño

Se conformó un equipo de expertos nacionales que, previa revisión de diversos instrumentos de pesquisa disponibles y utilizados a nivel nacional e internacional, revisó los contenidos de cada uno de los ítems de la GNVD (denominación, definición, procedimiento de evaluación, materiales requeridos, criterios de logro,

orden de presentación, rangos de percentiles para su adquisición, instrumento de registro, instructivo). En base a sus recomendaciones se propuso un nuevo diseño en el que el conjunto de ítems de desarrollo estandarizados se graficó en barras correspondientes a los percentiles 25, 75 y 90 de la edad de su cumplimiento. Posteriormente, fue presentada a técnicos y profesionales referentes nacionales en desarrollo infantil de diferentes disciplinas (neuropediatría, psiquiatría infantil, fonoaudiología, psicología, maestra especializada, psicopedagogía, psicomotricidad y pediatría). Los aportes se recogieron para elaborar la versión final (GNVD V2).

II) Validación

Los niños con sus padres o cuidadores fueron citados a su centro educativo en dos instancias separadas: primero, para la administración de la GNVD V2 (pesquisa), y luego para la administración del Battelle (patrón oro). La aplicación del Battelle se realizó en forma independiente y ciega en un plazo no superior a diez días de la GNVD V2. La administración de la GNVD V2 fue realizada por un equipo de siete pediatras, docentes de Clínica Pediátrica de la Facultad de Medicina de la Universidad de la República, previamente capacitados. Previo al trabajo de campo, se analizó la variabilidad y concordancia interobservador en una muestra no probabilística de 39 niños de 0 a 48 meses. La concordancia observada entre los pediatras fue moderada a muy buena (Kappa 0,60 - 0,83). La administración de Battelle fue realizada por licenciados en psicomotricidad y psicología, con experiencia documentada y que participaron de un taller de nivelación. A aquellos niños con Battelle alterado, previo consentimiento informado de padres o tutores, se les ofreció completar la valoración por psiquiatra infantil o neuropediatra, según el tipo de alteraciones halladas. Para la definición de los problemas de desarrollo se utilizaron los criterios del Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-5). En todos los casos se realizó un informe de contrarreferencia a su pediatra tratante de modo de continuar la atención en su prestador de salud.

Criterios para administración de la GNVD V2. Cada evaluador realizó una entrevista de antecedentes acorde al instructivo. Para aplicar la prueba se calculó la edad cronológica. En niños menores de 2 años, con edad gestacional al nacer inferior a 38 semanas, se estimó la edad corregida. En función de los logros en las conductas evaluadas en cada área se definieron dos categorías de resultados: el niño “pasa” o “no pasa”. Para ello, se trazó una línea vertical perpendicular a la edad del niño y se describieron las tres conductas más próximas a la izquierda de dicha línea (conductas “I”) y todas aquellas

que la cruzaban (conductas “C”). Se evaluó el número de niños que “pasaban” y “no pasaban” la GNVD V2 tomando diferentes puntos de corte según el número de conductas “I” y “C” no aprobadas para su edad. Independientemente del punto de corte, todos los casos con resultados “dudosos” se consideraron “no pasa”. Se seleccionó el punto de corte con mejor sensibilidad y especificidad. Este fue utilizado para estimar el resto de los indicadores de la prueba.

Criterios de evaluación del Inventario de Desarrollo Battelle 4ª edición. Se compró la licencia de la versión original, al momento de su administración se consideró la edad cronológica (corte día 30). Se definió “normal” una puntuación media de cociente intelectual (CI) 100 ± 15 y patológico una puntuación inferior a 85. Todos los padres fueron informados del resultado de esta prueba. Aquellos con CI global menor a 85 fueron derivados a los especialistas: pPsiquiatra infantil y neuropediatra. Los niños con CI global normal, pero con puntuaciones menores a 85 en algún área específica, recibieron recomendaciones.

Variables. Se registraron factores de riesgo biológicos, maternos y ambientales. Se aplicaron las siguientes definiciones operativas:

- Riesgo biológico, presencia de al menos uno de los siguientes: patología que requirió hospitalización en unidades de cuidado intensivo en el período neonatal, bajo peso al nacer (BPN) (peso al nacer menor a 2.500 g) y prematuridad (edad gestacional al nacer menor o igual a 37 semanas) y pequeño para la edad gestacional (PEG) (peso al nacer menor al percentil 10 para la edad gestacional).
- Riesgo materno, presencia de al menos uno de los siguientes: madre adolescente (edad materna menor a 18 años al momento del nacimiento), escolaridad menor a 12 años, patología mental, consumo problemático de sustancias.
- Riesgo ambiental, presencia de al menos una necesidad básica insatisfecha (NBI) evaluada por el índice elaborado y validado por la Facultad de Ciencias Sociales de la Universidad de la República (UDELAR)⁽²⁰⁾.

Aspectos éticos. El protocolo de estudio fue autorizado por las instituciones educativas participantes y aprobado por las comisiones de Investigación, Ética y Bioseguridad de la Facultad de Medicina de la UDELAR. Se solicitó consentimiento libre informado y escrito de padres o tutores, con posibilidad de abandonar el estudio en cualquier momento.

Análisis de datos. Se consideraron casos positivos para la GNVD V2 (pesquisa) aquellos que “no pasan”

dicha prueba, y casos positivos para Battelle (patrón oro o diagnóstico) los patológicos. Se establecieron cuatro resultados posibles: casos positivos por pesquisa y positivos por diagnóstico o verdaderos positivos (a); positivos por pesquisa y negativos por diagnóstico o falsos positivos (b); negativos por pesquisa y positivos por diagnóstico o falsos negativos (c), y negativos por pesquisa y por diagnóstico o verdaderos negativos (d). Para el cálculo de los índices de la prueba se aplicaron las siguientes fórmulas: *Sensibilidad* (Se) = $a/(a + c)$; *Especificidad* (Es) = $d/(b + d)$; *Valor predictivo positivo* (VPP) = $a/(a + b)$; *Valor predictivo negativo* (VPN) = $d/(c + d)$. Para estimar el beneficio clínico de la aplicación de la GNVD V2, independientemente de la prevalencia de los trastornos del desarrollo, se calcularon, a partir de la sensibilidad y especificidad, las razones de verosimilitud o *likelihood ratios* (LR): LR (+) = Tasa de verdaderos positivos/ Tasa de falsos positivos = $Se/1 - Es$; LR (-) = Tasa de falsos negativos/Tasa de verdaderos negativos = $1 - Se/Es$. En base al LR se calculó el riesgo de presentar un trastorno del desarrollo según el resultado de la prueba de pesquisa. Riesgo de presentar trastorno cuando “no pasa” GNVD: LR (+) x (1 – Prevalencia). Riesgo de presentar trastorno cuando “pasa” GNVD: LR (-) x (1 – Prevalencia). Para el procesamiento de los datos se utilizó el software de licencia libre R y el MPLUS 8.

El análisis psicométrico de la GNVD V2 fue realizado por especialistas del Instituto de Fundamentos y Métodos en Psicología (Facultad de Psicología de la UDELAR) mediante: 1) Correlaciones a nivel ítem. Debido a que la GNVD V2 utiliza ítems dicotómicos en las diferentes áreas, de acuerdo a las recomendaciones se analizaron las correlaciones tetracóricas interítem. Se consideró correlación baja $r 0,1 - 0,3$; moderada $r 0,4 - 0,6$ y fuerte $r 0,7 - 0,9$. Los valores negativos corresponden a relaciones inversamente proporcionales y los positivos directamente proporcionales. 2) Revisión de una dimensionalidad de las áreas. Debido a la cantidad de ítems de la escala y de participantes, se realizó un análisis de la unidimensionalidad de cada una de las áreas de la GNVD V2 por separado. Esto permitió utilizar una puntuación agregada en cada área y, a la vez, es un prerrequisito para el análisis con teoría de respuesta al ítem (TRI). Para ello, se realizó un análisis factorial confirmatorio con estimador de los mínimos cuadrados ponderados con medias y varianzas ajustadas (Weighted least squares with means and variances adjusted, WLSMV). 3) Análisis de los parámetros según teoría de respuesta al ítem. Esta teoría describe la relación entre los datos obtenidos con la aplicación de la GNVD V2 (“respuestas a los ítems” de la prueba) con determinadas variables latentes (características de los niños a quienes

se administró). Para ello, se consideró cada área como un único factor y se realizó un análisis de dos parámetros: dificultad y discriminación.

Resultados

La GNVD V2 se aplicó a 500 niños procedentes de 19 de los 25 centros seleccionados. No participaron del estudio seis centros por dificultades en la implementación del trabajo de campo. Se excluyeron del estudio 159 niños debido a imposibilidad de aplicar el Inventario Battelle en los plazos establecidos en el protocolo. La validación se realizó con 341 niños; 177 de sexo femenino (52%) y 164 masculino (48%). La distribución de niños por tipo de centro educativo fue 30,8% (n=105) en centros educativos privados; 28,2% (n=96) en centros educativos públicos; 26,4% (n=90) en CAIF y 14,7% (n=50) en centros diurnos.

La distribución por edad se observa en la tabla 1. El 47,5% (n=162) de los niños incluidos tenían una edad menor o igual a 36 meses y el 70% (n=239) menor o igual a 4 años.

El 46,3% (n= 158) de los niños no presentaba ningún factor de riesgo. Presentaba al menos un factor de riesgo biológico el 14,6% (n=50), un factor de riesgo materno el 14,1% (n=48) y una NBI 39% (n=133). La NBI más frecuente fue la vivienda (25%). Los niños procedentes de centros públicos, incluidos CAIF y centros diurnos, presentaron mayor riesgo materno y ambiental que aquellos procedentes de centros privados ($p<0,01$). En la tabla 2 se muestra la distribución de los factores de riesgo más relevantes por centro educativo.

La distribución de los puntajes globales de la aplicación del Inventario Battelle se muestra en la figura 1. El 25% (n=85) de los niños presentaron valores de CI patológicos. La proporción de niños con CI patológico por centros fue: CAIF 30%, centros diurnos 29%, jardines públicos 20% y privados 20%. En la tabla 3 se presentan las principales medidas de resumen del puntaje global, así como de sus componentes. En el área cognitiva el puntaje obtenido fue significativamente mayor en los niños procedentes de centros educativos privados. En el resto de las áreas no se observaron diferencias significativas entre los centros educativos. No se observó asociación significativa entre presencia de riesgo biológico (OR 1,61; IC95% 0,82-3,14) y materno (OR 1,33; IC95% 0,68-2,6) con Battelle patológico. En este estudio la presencia de riesgo ambiental aumentó la probabilidad de alteración en el Battelle (OR 1,95; IC95% 1,15-3,29).

Análisis psicométrico de la GNVD V2. 1) Correlaciones tetracóricas interítem. Los ítems presentan un comportamiento homogéneo entre ellos dentro de las

Tabla 1. Validación de un instrumento para la pesquisa de problemas del desarrollo en menores de 5 años en Uruguay. Distribución por edad (N=341).

Edad (meses)	n (%)
1 a 12	27 (7,9)
13 a 24	66 (19,4)
25 a 36	69 (20,2)
37 a 48	77 (22,6)
49 a 60	66 (19,4)
61 o más	36 (10,6)
Total	341 (100)

áreas. Las correlaciones son moderadas y altas para la mayoría de los ítems a la interna de las áreas, y moderadamente negativas entre ítems posicionados en los extremos de dificultad dentro de cada área (i.e.: los ítems para los niños de pocos meses correlacionan negativamente con los ítems destinados a los niños de 5 años). Las correlaciones entre ítems frecuentemente superan el punto de corte de $r = .95$, establecido como una correlación muy fuerte. Por dimensión, 15 de 182 correlaciones interítem en la dimensión social, 19 de 145 para motor, 13 de 333 para lenguaje, y 11 de 182 para coordinación superan dicho valor. 2) Revisión de una dimensionalidad de las áreas. En la tabla 4 se presentan los índices de ajuste del modelo y el rango de carga factorial para cada ítem. Como puede apreciarse, el par de índices CFI (índice de ajuste comparativo de Bentler) y TLI (Índice de Tucker - Lewis) superan el punto de corte recomendado por Hu & Bentler (1999), no así el error cuadrado de aproximación a las raíces medias (RMSEA). Adicionalmente, el rango de las cargas factoriales es elevado y significativo en todos los casos, por lo que se puede considerar adecuado el ajuste de los modelos. 3) Análisis de dos parámetros de teoría de respuesta al ítem. Para la dimensión social, las discriminaciones son adecuadas y las dificultades se ordenan de menor a mayor, excepto por el ítem “Inicia contacto social con otros niños”, que parece más sencillo de lo ubicado en la lista. Para el área motora, el ítem “Salta 2 m en un pie” presenta baja discriminación (menor a 0,7). La dificultad de los ítems está ordenada adecuadamente, excepto el ítem “Sube y baja las escaleras con ayuda”, que debería estar colocado en una posición anterior. En el área lenguaje se presenta un comportamiento adecuado en cuanto a discriminación. Sin embargo, en cuanto a dificultad, los ítems presentan valores no ordenados respecto a su posición en la escala, aunque dicha alteración es menor. En la

Tabla 2. Validación de un instrumento para la pesquisa de problemas del desarrollo en menores de 5 años en Uruguay. Distribución de los grupos de factores de riesgo más relevantes por centro educativo (N=341).

Factores de riesgo	Centro Educativo n (%)				Total n	p – valor (X ²)
	CAIF	Centro diurno	Centros educativos públicos	Centros educativos privados		
Biológicos	16 (17,8)	7 (14,0)	14 (14,6)	13 (12,4)	50 (14,6)	0,83
Prematurez	11 (12,2)	3 (6,0)	9 (9,4)	10 (9,5)	33	
PEG	8 (8,9)	3 (6,0)	4 (4,2)	8 (7,6)	23	
BPN	4 (4,4)	4 (8,0)	5 (5,2)	6 (5,7)	19	
Maternos	13 (14,4)	9 (18,0)	23 (24,0)	3 (2,9)	48 (14,1)	<0,01
Adolescente	9 (10,0)	5 (10,0)	12 (12,5)	3 (2,9)	29	
Consumo de sustancias	3 (3,3)	4 (8,0)	10 (10,4)	-	17	
Patología mental	1 (1,1)	4 (8,0)	4 (4,2)	-	9	
Ambiental (NBI)	35 (38,9)	24 (48,0)	45 (46,9)	29 (27,6)	133 (39,0)	<0,01
Vivienda	28 (31,1)	21 (42,0)	29 (30,2)	8 (7,6)	86	
Sin factores de riesgo	37 (43,0)	18 (37,5)	67 (76,1)	36 (46,4)	158 (46,3)	<0,01

dimensión de coordinación, las discriminaciones presentan valores adecuados. Las dificultades presentan valores adecuados, salvo por el ítem “Abotona y desabotona”, que debería ser reubicado.

Características de la GNVD V2. En la tabla 5 se muestran los índices fijos (Se y Es) y variables (VPP y VPN) obtenidos de la administración de la GNVD V2, según diferentes puntos de corte. No haber fallado en ninguna de las 12 conductas “I” de las cuatro áreas es el punto de corte con mayor sensibilidad (76%) y la falla en una conducta “I” en alguna de las áreas evaluadas el que presenta mayor especificidad (89%). Se tomó como punto de corte aquel con mayor sensibilidad, es decir “no haber fallado en ninguna de las 12 conductas “I” de las cuatro áreas” (Se=76%, Es=65%, VPP=42%, VPN=89%). En la tabla 6 se muestran los datos que permitieron calcular estos parámetros. Tomando dicho punto de corte, el LR (+) es 2,2 y el LR (-) 0,35, el riesgo de presentar un trastorno cuando “no pasa” la GNVD 1,6 y de presentar un trastorno cuando pasa 0,26. Los índices fijos y variables de la GNVD V2 no se modifican con la presencia de al menos un factor de riesgo biológico, materno o una NBI (tabla 7). Al igual que lo observado con la administración Battelle, la presencia de riesgo biológico (OR 1,84, IC95% 0,97-3,47) y materno (OR 0,82, IC95% 0,44-1,52) no se asoció con falla en la GNVD V2. Sí se observó asociación con la presencia de riesgo ambiental (OR 2,56, IC95% 1,57-4,16).

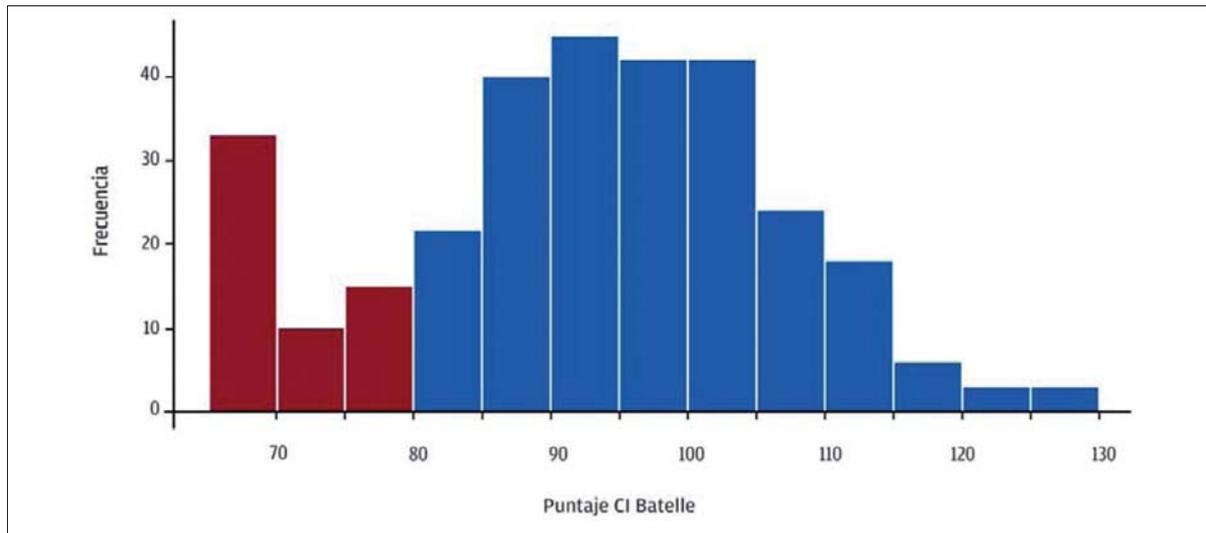
Discusión

Un aspecto importante a considerar, antes de establecer los criterios de riesgo en un test de pesquisa de problemas de desarrollo, es el tamaño y tipo de muestra en el que se realizaron los estudios de validación. Esta validación se realizó en una muestra de niños procedentes de centros educativos públicos y privados de la ciudad de Montevideo, con participación de niños y niñas de diferentes estratos socioeconómicos. A pesar de las limitaciones metodológicas relacionadas con la toma de una muestra limitada a Montevideo, una pérdida de niños superior a la estimada (31,8%), y la composición final de la muestra (menor número de niños en los extremos etarios), lo que afecta el poder o potencia del estudio, el diseño al azar y el número incluido resulta una fortaleza. Es importante señalar que en varios de los estudios de validación de instrumentos de pesquisa ampliamente utilizados en la práctica clínica en diferentes países, el tamaño de la muestra es cercano a 100^(21,22).

Otro aspecto relevante es la distribución de factores de riesgo en la muestra utilizada en la validación. La prevalencia de problemas de desarrollo en la población puede influir en los resultados de la validación⁽²³⁻²⁵⁾. En poblaciones con mayor riesgo biológico o ambiental que el observado en la muestra estudiada, es probable que la prevalencia de los trastornos del desarrollo sea más elevada. En esas condiciones los resultados de la validación pueden ser diferentes a los obtenidos y deben ser

Tabla 3. Validación de un instrumento para la pesquisa de problemas del desarrollo en menores de 5 años en Uruguay. Principales medidas de resumen del puntaje global y sus componentes (N=341).

	Total	Personal/Social	Motor	Comunicación	Adaptativo	Cognitivo
Mínimo	65	65	65	65	65	65
Q1	84,5	86	84	79	88	85
Mediana	95	96	95	89	96	95
Media	93,2	94,1	93,5	88,7	96,8	93,8
Q3	103	103	102	99	99	103
Máximo	128	131	126	128	128	126
SD	14,2	14,2	13,6	14,1	16,1	14,3

**Figura 1.** Validación de un instrumento para la pesquisa de problemas del desarrollo en menores de 5 años en Uruguay. Distribución del puntaje de coeficiente intelectual de Battelle (n=341).

considerados cuando se extrapolan a la población general, donde la prevalencia es menor. En general, la proporción de falsos positivos es menor y el valor predictivo mayor en los estudios realizados en poblaciones de alto riesgo. En este trabajo, la proporción de niños con al menos un factor de riesgo biológico o materno fue similar. Para ambos instrumentos aplicados, la presencia de riesgo biológico y materno no se asoció con resultados alterados, aunque sí se observó asociación con la presencia de riesgo ambiental evaluado por al menos una NBI. Si bien la evaluación del riesgo ambiental por este método presenta importantes limitaciones, ya que no se evalúa directamente los aspectos vinculares ni las pautas de crianza, estos resultados concuerdan en líneas gene-

rales con diversos estudios disponibles que señalan el impacto de los factores de riesgo medioambientales en la prevalencia de trastornos del desarrollo^(1,3,26).

Por otra parte, es necesario que los estudios de validación se realicen en muestras que contemplen la probabilidad de hallar diferentes trastornos, como ocurre en poblaciones de alto riesgo. Esto resulta especialmente importante al estimar la sensibilidad y especificidad de una prueba, debido a que la prevalencia de pacientes con un determinado problema del desarrollo puede ser limitada⁽²¹⁾. Por lo tanto, un hallazgo que refuerza la validez de la muestra es la prevalencia de trastornos del desarrollo y el tipo de alteraciones observadas mediante Battelle (25%), similar a lo comunicado en la primera ronda de la

Tabla 4. Validación de un instrumento para la pesquisa de problemas del desarrollo en menores de 5 años en Uruguay. Índices de ajustes para los modelos unidimensionales de cada componente de la GNVDV2.

Área	Chi 2	gl	RMSEA	RMSEA 90% IC	CFI	TLI	Rango de cargas (STDY)**
Social	1.205,68*	170	0,132	0,125 a 0,139	0,96	0,96	0,73 a 0,99
Motor	976,21*	135	0,133	0,126 a 0,141	0,96	0,96	0,53 a 0,99
Lenguaje	1.973,48*	324	0,121	0,116 a 0,126	0,97	0,97	0,69 a 0,99
Coordinación	1.298,74*	170	0,138	0,131 a 0,145	0,96	0,95	0,72 a 0,99

* $p < 0,001$. ** Todas las cargas son significativas al nivel $p < 0,001$.
STDY: estandarización STDY

Tabla 5. Validación de un instrumento para la pesquisa de problemas del desarrollo en menores de 5 años en Uruguay. Sensibilidad, especificidad, valores predictivo positivo y negativo de la GNVDV2, según puntos de corte considerados.

Puntos de corte	S	E	VPP	VPN
Ninguna falla en conductas "I" de las cuatro áreas	0,76	0,65	0,42	0,89
Falla en una conducta "I" de cualquier área	0,56	0,89	0,64	0,85
Falla en dos conductas "I" de cualquier área	0,34	0,95	0,71	0,80
Falla en una conducta "I" y en una "C"	0,55	0,89	0,66	0,84
Sin fallas en conductas "I" y una falla en una conducta "C"	0,75	0,70	0,47	0,89
Sin fallas en conductas "I" y dos fallas en conductas "C"	0,74	0,73	0,50	0,89
Sin fallas en conductas "I" y tres fallas en conductas "C"	0,69	0,75	0,46	0,87

Conducta "I": conducta a la izquierda de la línea de referencia. Conducta "C": conducta que cruza línea de referencia.

Encuesta Nacional de Desarrollo Infantil y Salud (ENDIS 2014). En esta encuesta, mediante la aplicación del instrumento Ages and Stages Questionnaire (ASQ-3), en el total de niños de la muestra de Montevideo, el 83% de los niños se ubica en el rango de normalidad⁽⁸⁾.

Este estudio de validación incluyó no solo la caracterización de la prueba en términos de sus índices fijos (sensibilidad y especificidad) y variables (valores predictivos), sino el análisis de sus propiedades psicométricas. Es importante señalar que en revisiones sistemáticas sobre validación de pruebas de pesquisa, la mayoría no incluye este tipo de análisis en el proceso de validación. En el análisis psicométrico, la GNVD V2 muestra valores aceptables de ajuste unidimensional, el par de índices CFI y TLI presentan valores superiores a 0,90. En el análisis mediante teoría de respuesta al ítem la GNVD V2 presenta buenos valores de discriminación y

dificultad. Estos resultados son muy útiles para realizar recomendaciones de mejora en su diseño. Así, algunos ítems pueden ser reorganizados en términos de su ubicación en la progresión por edad dentro de cada área. Por otro lado, varios ítems presentaron poca variabilidad entre ellos, esto es, ante una respuesta positiva en uno la probabilidad de respuesta positiva en el otro es muy alta, con varianzas compartidas superiores a 90% en los casos de correlaciones superiores a 0,95. Esto permite identificar algunos ítems candidatos a ser eliminados, lo que permitiría reducir la longitud de la GNVD V2 y promover una administración aun más sencilla.

Este estudio permite conocer la sensibilidad y especificidad de diferentes puntos de corte de la GNVD V2. Considerando que el objetivo de la misma es la pesquisa de trastornos del desarrollo en menores de 5 años, el punto de corte con mayor sensibilidad para definir que pasa o aprueba la GNVD V2 es "no haber fallado en nin-

Tabla 6. Validación de un instrumento para la pesquisa de problemas del desarrollo en menores de 5 años en Uruguay. Sensibilidad, especificidad, valores predictivo positivo (VPP) y negativo (VPN) tomando como punto de corte ninguna falla en conductas "I" de las cuatro áreas (N=341).

		Battelle		Total
		Patológico	Normal	
GNVD	No pasa	65	90	155
	Pasa	20	166	186
Total		85	256	341

Sensibilidad 76% (IC95% 67-85); especificidad 65% (IC95% 59-71); VPP 42% (IC95% 34-50); VPN 89% (IC95% 84-93).

guna de las 12 conductas "I" de las cuatro áreas" (Se 0,77; Es 0,65). Por lo tanto, se puede afirmar que la GNVD V2 posee una sensibilidad adecuada para ser utilizada en programas nacionales de pesquisa del desarrollo⁽²⁷⁾. La GNVD V2 posee una sensibilidad similar a la Prueba Nacional de Pesquisa (PRUNAPE 0,80), Parents' Evaluation of Developmental Status (PEDS 0,74-0,79), a las últimas versiones del Denver II Developmental Screening (0,56-0,83) e inferior a Battelle Developmental Inventory Screening (0,72-0,93)^(21,28).

La sensibilidad y especificidad, a pesar de definir la validez de la prueba diagnóstica independientemente de la prevalencia de la enfermedad en la población, no proporcionan información relevante para la toma de decisiones clínicas ante un determinado resultado. Es por ello que adquieren relevancia los LR o razones de verosimilitud. Estos son valores inherentes a la prueba e independientes de la prevalencia de la enfermedad que reflejan su capacidad para cambiar una probabilidad pretest a una nueva probabilidad postest. Los $LR > 1$ aumentan la probabilidad de que el trastorno esté presente, a mayor valor de LR, mayor probabilidad. Por el contrario, $LR < 1$ disminuyen la probabilidad de que el trastorno esté presente, a menor el LR, mayor es la disminución de la probabilidad. En este caso, aunque los valores de LR hallados por sí solo parecen indicar cambios modestos en la probabilidad posprueba, su contribución a la práctica clínica puede ser importante si se la utiliza en forma combinada con una correcta valoración clínica, incluyendo evaluación de factores de riesgo y examen clínico⁽²⁹⁾.

No se dispone hasta el momento de un instrumento de pesquisa de desarrollo que reúna las características psicométricas y validez ideales. Si bien la GNVD V2 no ha sido comparada con otros instrumentos de pesquisa, no se dispone de evidencia científica que justifique utili-

Tabla 7. Validación de un instrumento para la pesquisa de problemas del desarrollo en menores de 5 años en Uruguay. Sensibilidad, especificidad, valores predictivo positivo (VPP) y negativo (VPN), según puntos de corte de mayor sensibilidad y factores de riesgo.

Índices de la GNVD	Factores de riesgo	
	No	Por lo menos uno
Sensibilidad	0,76	0,78
Especificidad	0,65	0,68
VPN	0,89	0,86
VPP	0,42	0,46

zar un instrumento en lugar de otro, ya que todos tienen fortalezas y debilidades en cuanto a su método y tiempo de aplicación, materiales, características del estudio de validación y dimensiones evaluadas. Por lo tanto, es posible que su aplicación deba ser complementada con otros instrumentos que evalúen aspectos vinculares, pautas de crianza, elementos de alerta de trastornos del espectro autista (TEA)^(7,28).

Conclusiones

La GNVD V2 desarrollada en Uruguay posee propiedades psicométricas adecuadas y sus índices fijos permiten utilizarla como instrumento de pesquisa de alteraciones del desarrollo en niños menores de 5 años. Los resultados de este estudio brindan información valiosa a tomadores de decisión en políticas públicas para utilizar el punto de corte más adecuado para la detección de trastornos del desarrollo en programas de pesquisa. Esta información resulta imprescindible para los equipos de salud del primer nivel de atención, que son los responsables de su aplicación y, por lo tanto, quienes deben conocer los alcances y limitaciones del instrumento. En el campo de la investigación aporta información inédita en la bibliografía de nuestro país, que puede ser utilizada en futuros estudios poblacionales. El desafío es incorporar las recomendaciones de ajuste psicométrico en una versión actualizada, capacitar a los equipos en su aplicación y desarrollar un sistema informático con alarmas que permitan tomar decisiones oportunas y evaluar intervenciones a corto y largo plazo.

Abstract

Introduction: early detection of developmental problems in children enables the application of timely interventions that improve prognosis and reduce disabilities.

To that end, the application of sensitive, simple and low-cost screening procedures is required, what led to the creation of the “National Guidelines for the Surveillance of Early Childhood Development (under 5 years old children)” (GNVD V2). Use of these guidelines relies on awareness of psychometric properties as well as fixed rates and variables.

Objective: concurrent validation of GNVD V2 with Battelle Developmental Inventory, Fourth Edition.

Method: we conducted a transversal, observational study in a representative sample of children between 1 month and 5 years old from different schools in Montevideo. Subsequently, results from the application of GNVD V2 were compared to the Battelle Developmental Inventory, Fourth Edition. Sensitivity, specificity and predictive positive and negative values of the GNVD V2 were calculated for the different cut-off points. The clinical benefit if its application was estimated by means of likelihood ratios. A psychometric analysis was performed.

Results: validation was done with 341 children. No failure occurring in any of the behaviors to the left of the four areas corresponding to age represents 77% sensitivity, 65% specificity, 42% positive predictive value and 89% negative predictive value. Given such cut-off point, LR+ is 2.2 and LR- is 0.35, and the likelihood of there being a disorder when not passing the test is 1.6 and 0.26 when passing it.

Conclusions: GNVD V2 has the right psychometric properties and fixed rates that are similar to globally validated tests that allow for it to be recommended as the national screening instrument.

Resumo

Introdução: a detecção precoce de problemas de desenvolvimento em crianças permite aplicar intervenções oportunas que melhoram o prognóstico e diminuem as deficiências. Para isso é necessário aplicar provas de pesquisa sensíveis, simples e de baixo custo. Com este propósito, desenvolveu-se no Uruguai a Guia Nacional para la Vigilancia del desarrollo del Niño y la Niña menores de 5 años Versión 2 (GNVD V2). Para utilizá-la é necessário conhecer as propriedades psicométricas e os índices fixos e variáveis.

Objetivo: validação concorrente da GNVD V2 comparada com o Inventário de Desenvolvimento Battelle 4ª edição.

Métodos: realizou-se um estudo transversal, observacional, em uma amostra representativa de crianças com idades entre 1 mês e 5 anos de centros educativos de Montevideo. Os resultados da aplicação da GNVD V2 foram comparados com os do Inventário de Desenvolvimento Battelle 4ª edição. A sensibilidade, especificidade, valores preditivos positivos e negativos da GNVD V2 segundo di-

ferentes pontos de corte foram calculados. O benefício clínico de sua aplicação mediante as razões de verossimilhança ou “likelihood ratio” (LR) foi estimado. Uma análise psicométrica foi realizada.

Resultados: a validação foi realizada com 341 crianças. Não haver mostrado falhas em nenhuma das 12 condutas a esquerda das quatro áreas correspondentes à sua idade apresenta sensibilidade 77%, especificidade 65%, valor preditivo positivo 42% e valor preditivo negativo 89%. Com este ponto de corte, o LR (+) é 2.2 e o LR (-) 0.35 e a probabilidade de apresentar um transtorno quando não passa a prova 1.6 e quando passa 0.26.

Conclusões: a GNVD V2 possui propriedades psicométricas adequadas e índices fixos similares a testes validados a nível internacional que permitem recomendá-lo como instrumento nacional de pesquisa.

Bibliografía

1. **WHO, UNICEF, World Bank Group.** Advancing early childhood development: from science to scale. An executive summary for The Lancet's Series. Lancet 2016. Disponible en: https://www.thelancet.com/pb-assets/Lancet/stories/series/ecd/Lancet_ECD_Executive_Summary.pdf [Consulta: 17 diciembre 2017].
2. **Irwin LG, Siddiqi A, Hertzman C.** Desarrollo de la primera infancia: un potente equalizador. Informe Final. Vancouver: HELP, 2007:82p. Disponible en: http://www.who.int/social_determinants/publications/early_child_dev_ecdkn_es.pdf [Consulta: 24 marzo 2018].
3. **Center on the Developing Child at Harvard University.** From best practices to breakthrough impacts: a science-based approach to building a more promising future for young children and families. Cambridge, MA: Center on the Developing Child, 2016. Disponible en: https://pediatricsociale.fondationdrjullien.org/wp-content/uploads/2016/07/cdc_rapport_complet_2016.pdf [Consulta: 27 marzo 2018].
4. **Funaro R.** Poco dinero para los más pequeños. (Blog: Ideas que cuentan, octubre 2015). Disponible en: <https://blogs.iadb.org/Ideasquecuentan/2015/10/28/poco-dinero-para-los-mas-pequenos/> [Consulta: 16 diciembre 2018].
5. **Heckman JJ.** Invest in early childhood development: reduce deficits, strengthen the economy. Disponible en: https://heckmanequation.org/assets/2013/07/F_HeckmanDeficitPieceCUSTOM-Generice_052714-3-1.pdf [Consulta: 27 diciembre 2017].
6. **Uruguay. Instituto Nacional de Estadística.** Salud, nutrición y desarrollo en la primera infancia en Uruguay: primeros resultados de la ENDIS. Grupo de Estudios de Familia. Montevideo: INE, UdelaR, OPP, MSP, MIDES, 2015. Disponible en: <http://www.ine.gub.uy/documents/10181/35704/SALUD%2C+NUTRICI%C3%93N+Y+DESARROLLO+EN+LA+PRIMERA+INFANCIA+EN+URUGUAY+PRIMEROS+RESULTADOS+DE+LA+ENDIS/7be3f504-ebb9-4427-bb5d-cb4d9f242a7b> [Consulta: 8 diciembre 2018].

7. **Uruguay. Ministerio de Salud Pública.** Objetivos Sanitarios Nacionales 2020. Disponible en: https://www.paho.org/uru/index.php?option=com_docman&view=download&alias=531-objetivos-sanitarios-web2017&category_slug=publicaciones-comunicacion&Itemid=307 [Consulta: 24 marzo 2019].
8. **Palfrey JS, Singer JD, Walker DK, Butler JA.** Early identification of children's special needs: a study in five metropolitan communities. *J Pediatr* 1987; 111(5):651-9.
9. **Blackman JA.** Developmental screening infants, toddlers, and preschoolers. En: Levine MD, Carey WB, Crocker AC. En: *Developmental Behavioral Pediatrics*. 2a ed. Philadelphia, PA: WB Saunders, 1992:617-23.
10. **Blackman JA, Healy A, Ruppert ES.** Participation by pediatricians in early intervention: impetus from Public Law 99-457. *Pediatrics* 1992; 89(1):98-102.
11. **Bennett F, Guralnick M.** Effectiveness of developmental intervention in the first five years of life. *Pediatr Clin North Am* 1991; 38:1513-28.
12. **Committee on practice and ambulatory medicine, Bright futures periodicity Schedule workgroup.** 2017 Recommendations for preventive pediatric health care. *Pediatrics* 2017; 139(4):e20170254.
13. **Weitzman C, Wegner L; Section on developmental and behavioral pediatrics, Committee on psychosocial aspects of child and family health, Council on early childhood and Society for Developmental Behavioral Pediatrics.** Promoting optimal development: screening for behavioral and emotional problems. *Pediatrics* 2015; 135(2):384-95.
14. **Wilson JM, Jungner G.** Principles and practice of screening for disease. Geneva: World Health Organization, 1968. (Public health papers, 34).
15. **Lansdown RG, Goldstein H, Shah PM, Orley JH, Di G, Kaul KK, et al.** Culturally appropriate measures for monitoring child development at family and community level: a WHO collaborative study. *Bull World Health Organ* 1996; 74:283-90.
16. **Colombia. Ministerio de Salud.** Escala abreviada de desarrollo. Manual de instrucciones. Bogotá: Guadalupe, 1991.
17. **Mokkink LB, Prinsen CA, Bouter LM, de Vet HC, Terwee CB.** The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther* 2016; 20(2):105-13. Disponible en: <http://www.redalyc.org/articulo.oa?id=235046202001> [Consulta: 8 agosto 2017].
18. **Streiner DL, Norman GR, Cairney J.** Health measurement scales: a practical guide to their development and use. 4 ed. Oxford: Oxford University, 2008.
19. **Newborg J.** Battelle Developmental Inventory. 2ª ed. (Normative update BDI-2 NU). Itasca, IL: Riverside, 2005.
20. **Calvo JJ, coord, Borrás V, Cabella W, Carrasco P, De los Campos H, Koolhaas M.** Atlas sociodemográfico y de la desigualdad del Uruguay. Las necesidades básicas insatisfechas a partir de los Censos 2011. Montevideo: Trilce, 2013: 78 p.
21. **Romo-Pardo B, Liendo-Vallejos S, Vargas-López G, Rizzoli-Córdoba A, Buenrostro-Márquez G.** Pruebas de tamizaje de neurodesarrollo global para niños menores de 5 años de edad validadas en Estados Unidos y Latinoamérica: revisión sistemática y análisis comparativo. *Bol Med Hosp Infant Mex* 2012; 69(6):450-62.
22. **Schonhaut L, Salinas P, Armijo I, Schonstedt M, Álvarez J, Manríquez M.** Validación de un cuestionario autoadministrado para la evaluación del desarrollo psicomotor. *Rev Chil Pediatr* 2009; 80(6):513-9.
23. **Cochrane AL, Holland WW.** Validation of screening procedures. *Br Med Bull* 1971; 27(1):3-8.
24. **Hall DMB, ed.** Health for all children: a programme for child health surveillance. Oxford: Oxford University, 1989.
25. **DSM IV.** Manual diagnóstico y estadístico de los trastornos mentales. 4 ed. Barcelona: Masson, 1995.
26. **Shonkoff JP, Garner AS.** The lifelong effects of early childhood adversity and toxic stress. *Pediatrics* 2011; 129(1):e232-46.
27. **Rydz D, Srouf M, Oskoui M, Marget N, Shiller M, Birnbaum R, et al.** Screening for developmental delay in the setting of a community pediatric clinic: a prospective assessment of parent-report questionnaires. *Pediatrics* 2006; 118(4):e1178-86.
28. **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(1):405-20.
29. **Silva Fuente-Alba C, Molina Villagra M.** Likelihood ratio (razón de verosimilitud): definición y aplicación en Radiología. *Rev Argent Radiol* 2017; 8(3):204-8.

Contribución de autores

Los autores participaron en igual medida en la elaboración del artículo
 Gustavo Giachetto, <https://orcid.org/0000-0003-3775-4773>
 Anabella Santoro, <https://orcid.org/0000-0003-1761-0211>
 Fernando Massa, <https://orcid.org/0000-0002-2922-4097>
 Mercedes Pérez Rocha, <https://orcid.org/0000-0003-3575-3814>