



## CRECIMIENTO Y DESARROLLO

# Validez concurrente de las cartillas de vigilancia para identificar alteraciones en el desarrollo del lactante



R. Rivera González<sup>a</sup>, Y. Villanueva Romero<sup>b</sup>, L. Amaro López<sup>b</sup>, C. Sánchez Pérez<sup>a,b,\*</sup>, M. Figueroa Olea<sup>a</sup>, K. M. Soler Limón<sup>a</sup>

<sup>a</sup>Laboratorio de Seguimiento del Neurodesarrollo del Instituto Nacional de Pediatría, México, D.F. México

<sup>b</sup>Maestría en Rehabilitación Neurológica, Universidad Autónoma Metropolitana Xochimilco, México, D.F. México

Recibido el 12 de junio de 2014; aceptado el 22 de agosto de 2014

### PALABRAS CLAVE

Pruebas de escrutinio del desarrollo;  
Validación de escalas;  
Vigilancia del desarrollo

### Resumen

**Introducción:** Las cartillas de vigilancia del desarrollo son una herramienta para evaluar el desarrollo durante los 2 primeros años de vida en el primer nivel de atención y son parte de las acciones del Sistema de Vigilancia y Promoción del Desarrollo Integral del Niño.

El objetivo del reporte es evaluar la validez concurrente por criterio externo de las conductas del desarrollo de las cartillas de vigilancia del Sistema de Vigilancia y Promoción del Desarrollo Integral del Niño respecto a 2 pruebas de desarrollo y proponer nuevos puntos de corte si el análisis de curvas ROC ofrece mejores indicadores de validez concurrente.

**Pacientes y métodos:** Se realizó un estudio transversal, prospectivo y descriptivo, comparando los resultados dicotomizados como normal o alterado de las cartillas respecto a 2 pruebas de desarrollo: "Escala de Desarrollo Infantil Bayley-II" y "Diagnóstico del desarrollo normal y anormal del niño", de Gesell-Amatruda.

**Resultados:** Se realizaron 2,702 evaluaciones a niños de 1 a 24 meses de edad, se estimaron los indicadores de validez con los criterios de calificación originales de la cartilla, mostrando sensibilidad de 75.99, especificidad 90.46 y valores predictivos positivos y negativos de 92.17 y 71.81, respectivamente. Los indicadores mejoraron a sensibilidad 84.43, especificidad 75.78, y valores predictivos positivos y negativos de 83.75 y 76.69. Por mes de edad, los indicadores de validez obtenidos fueron adecuados en 21 de los 24 meses analizados.

**Conclusiones:** Las cartillas de vigilancia son un adecuado instrumento para la detección de alteraciones y retrasos en el desarrollo durante los 2 primeros años de vida.

© 2015, Universidad Autónoma Metropolitana, Unidad Xochimilco. Publicado por Masson Doyma México S.A. Este es un artículo Open Access distribuido bajo los términos de la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

\*Autor para correspondencia: Av. Insurgentes Sur No. 3700-C, Col. Insurgentes Cuicuilco, Delegación Coyoacán, C.P. 04530, México, D.F., Tel.: 10840900 exts. 1720 y 1438. Correo electrónico: carmens@att.net.mx (C. Sánchez Pérez).

**KEYWORDS**  
 Development surveillance;  
 Development screening tests;  
 Validation

## Concurrent validity of Surveillance cards to identify child development disorders

### Abstract

**Introduction.** The development Surveillance Cards are tools to assess the development during the first two years of life in Primary Care, and are part of the activities of the Surveillance and Promotion of Integrated Child Development System (SIVIPRODIN). The objective of this report is to assess the concurrent validity, by external criteria, of the Development Behaviours of Surveillance SIVIPRODIN Cards regarding two tests of development, and propose new cut-offs if the ROC curve analysis provides better indicators of concurrent validity.

**Materials and Methods:** This is a cross-sectional, prospective, descriptive study, comparing the dichotomized normal and abnormal results of two development tests, the “Bayley Scale of Infant Development” and Gesell test “Developmental diagnosis-normal and abnormal child development”.

**Results.** A total of 2702 children 1-24 months old were evaluated. Validity indicators were estimated with cards with original qualification criteria, showing a sensitivity of 75.99, a specificity of 90.46, and positive and negative predictive values of 92.17 and 71.81, respectively. The indicators improved the sensitivity to 84.43, specificity to 75.78, and positive and negative predictive values to 83.75 and 76.14. By months of age, the validity indicators were adequate in 21 of the 24 months studied.

**Conclusions:** The Surveillance Cards are adequate tools for the detection of abnormalities and development delays during the first two years of life.

© 2015, Universidad Autónoma Metropolitana, Unidad Xochimilco. Published by Masson Doyma México S.A. This is an open access item distributed under the Creative Commons CC License BY-NC-ND 4.0 (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introducción

La vigilancia del desarrollo constituye una de las acciones implementadas por los servicios de salud y otros sectores formales e informales organizados para la atención de la salud y bienestar del niño con el fin de constatar el adecuado desarrollo y la presencia de condiciones favorables para su evolución.

Las herramientas y estrategias que se utilizan en la vigilancia del desarrollo deben ser evaluadas en cuanto su capacidad de identificar adecuadamente los casos que se encuentran en riesgo de presentar una alteración y requieren de acciones específicas para diagnosticarlos y atenderlos<sup>1</sup>.

Parte de las estrategias implican el uso de pruebas de tamizaje o escrutinio que permiten evaluar el desarrollo de una manera sencilla, con un entrenamiento básico y con materiales de fácil acceso. La Academia Americana de Pediatría recomienda que las herramientas deben corresponder a las características culturales de la población, y que la vigilancia debe realizarse de manera periódica, por ejemplo a los 3, 9, 18 y 36 meses de edad<sup>2</sup>. El test de Denver II es una de las pruebas más utilizadas en el ámbito pediátrico; sin embargo, no cuenta con validaciones suficientes, y aquellas que se han realizado en Estados Unidos muestran parámetros de sensibilidad y especificidad bajos<sup>3</sup>. El CAT-CLAMS es una herramienta emergente de amplio uso en EE. UU.; en Latinoamérica se han reportado estudios de edad de adquisición en población mexicana<sup>3</sup> que muestran no corresponder con las secuencias del desarrollo de los niños mexicanos. Romo-Pardo et al.<sup>4</sup> publicaron en 2012 una revisión sistemática de pruebas de tamizaje utilizadas en Latinoamérica y

Estados Unidos, donde se reportan valores variables de sensibilidad y especificidad de las distintas pruebas, y llama la atención que se reporten pocas pruebas elaboradas en América Latina y ninguna en México.

Vericat y Orden<sup>5</sup>, en una revisión del año 2006, reportan 10 pruebas de escrutinio, de las cuales 7 son elaboradas en países latinoamericanos (Argentina, Chile, Colombia, Brasil y México). De México se reporta la Escala de Desarrollo Integral del Niño, publicada por Atkin et al.<sup>6</sup>.

Rivera et al.<sup>7</sup> y Atkin et al.<sup>6</sup> han reportando anteriormente revisiones de pruebas para Latinoamérica y México, como la Valoración Neuroconductual del Desarrollo del Lactante (VANEDELA)<sup>8</sup>, como una prueba para identificar alteraciones del desarrollo con énfasis en las derivadas de lesión orgánica, y ha sido evaluada con diversas poblaciones de riesgo: alto, medio y bajo, arrojando buena sensibilidad y especificidad<sup>7</sup>.

También se reporta la Norma Oficial Mexicana<sup>9</sup> que retoma indicadores del CLAP-OPS, pero sin datos sobre su validez.

En los últimos años, se utiliza en centros de salud de México la evaluación propuesta por el Centro Nacional para la Salud del Infante y Adolescente, sobre la cual no hay datos de validez, y recientemente se ha publicado la Evaluación del Desarrollo Infantil (EDI)<sup>10</sup>, prueba de tamizaje validada en 438 niños menores de 5 años provenientes del Distrito Federal, Yucatán y Chihuahua, usando como estándar el Inventory de Desarrollo de Battelle-2 y el Bayley III.

Otros países de Latinoamérica cuentan con mayor avance en el diseño, validación e instalación de pruebas nacionales para la evaluación de su población infantil. El Test de Desarrollo Psicomotor (TEPSI) ha sido posiblemente la más fructífera por su diseminación y utilización en diversos reportes

**Tabla 1** Cartilla de Vigilancia del Desarrollo Conductas del Desarrollo

<b>Edad desarrollo</b>	<b>P*</b>	<b>Descripción</b>
Mes 1	1	Es capaz de buscar el pezón u objetos que rocen cerca de sus labios y luego succionarlos (chuparlos)
	2	Cuando se alimenta, su succión es fuerte, derrama poca leche por los labios durante toda la toma, no se atraganta ni se pone morado, no vomita
	3	Mueve doblando y estirando enérgicamente sus brazos y piernas
Mes 2	4	Sigue lentamente con la mirada el movimiento de las personas u objetos a la derecha y a la izquierda
	5	Reacciona a los sonidos, se aquiega o aumenta el movimiento de su cuerpo
	6	Al acostarlo boca abajo levanta su cabeza despegando la nariz
Mes 3	7	Inicia sonrisa social, se ríe cuando le hablan o lo ven
	8	Al cargarlo o sentarlo con ayuda ya sostiene la cabeza
	9	Observa objetos que se le colocan al frente
Mes 4	10	Acostado, gira la cabeza a cualquier lado buscando algo
	11	Sigue con los ojos los movimientos que hace con las manos
	12	Toma los objetos que se le ponen en su mano y se los lleva a la boca para chuparlos
Mes 5	13	Acostado, se pone de lado
	14	Alcanza objetos que se le ponen al frente
	15	Vocaliza, dice ooo/aaa o ah-gooo
	16	Voltea cuando se le habla
Mes 6	17	Contesta con sonidos como “agugu” cuando se le habla
	18	Sentado con ayuda, toma y levanta un objeto o juguete pequeño que está a su alcance
	19	Busca la cara de la mamá cuando se le esconde

**Desarrollo del niño: conductas esperadas de 7 a 12 meses de edad**

Mes 7 y 8	20	Se mantiene sentado sin ayuda y gira la cabeza hacia ambos lados
	21	Sostiene un objeto en cada mano
	22	Encuentra un juguete parcialmente escondido
	23	Acostado boca abajo, gira sobre su abdomen
	24	Estira los brazos cuando quiere que lo carguen
Mes 9 y 10	25	Reconoce su nombre cuando alguien le llama
	26	Puede entender “no”, “dame”, “ten”, “pon”, “mira”
	27	Tira objetos al piso y ve dónde cayó o escucha el sonido que producen al caer
	28	Junta el dedo pulgar al índice para tomar objetos pequeños
	29	Inicia el gateo y se desplaza o se queda de pie sujetándose de los muebles
	30	Puede rodarse para alcanzar un objeto
	31	Señala lo que quiere
Mes 11 y 12	32	Le gusta explorar, saca objetos de un recipiente o cajón
	33	Distingue quién es papá, mamá o ambos
	34	Inicia la imitación de algunos sonidos y gestos
	35	Puede beber de una taza aunque derrame o sin ayuda del vaso entrenador
	36	Jala la sabana o cobija para alcanzar un objeto que esta encima de ella o recurre a otra persona para obtenerlo
	37	Dice adiós con la mano, aplaude, manda besitos, hace ojitos
	38	Gatea o se pone de pie deteniéndose de los muebles

**Desarrollo del niño: conductas esperadas de 13 a 24 meses de edad**

Mes 13-15	39	Hace “solitos” o se mantiene parado al sostenerlo de una mano
	40	Indica lo que quiere con palabras o gestos
	41	Puede cambiar solo de posiciones con facilidad (se sienta, rueda, se acuesta)
	42	Identifica a los padres y puede llamarlos diciendo papá y mamá
	43	Busca un objeto de su interés previamente mostrado en los lugares que se lo esconden
	44	Después que se hace bailar una moneda o disco, toca al adulto para que lo repita o intenta hacerlo por sí solo
	45	Mete y saca objetos de un recipiente

**Tabla 1** Cartilla de Vigilancia del Desarrollo Conductas del Desarrollo (continuación)

Edad desarrollo	P*	Descripción
Mes 16-18	46	Camina solo
	47	Puede subir gateando un escalón
	48	Hace rayones
	49	Dice 2 a 3 palabras además de mamá y papá
	50	Imita quehaceres domésticos
	51	Bebe de taza sin ayuda
	52	Señala 2 partes de su cuerpo cuando se le pregunta
Mes 19-21	53	Arroja una pelota parado
	54	Al mostrarle un libro ayuda a cambiar las páginas
	55	Se agacha para recoger un objeto
	56	Dice 5 palabras además de mamá y papá
	57	Juega a darles de comer a sus juguetes
Mes 22-24	58	Patea la pelota sin sostenerse
	59	Corre
	60	Sube y baja escaleras sostenido de la mano descansando entre un escalón y otro (ambas cosas)
	61	Indica cuando se moja el pañal
	62	Señala 5 partes de su cuerpo
	63	Puede quitar la envoltura de un dulce o la cáscara de un plátano
	64	Es capaz de subirse a una silla grande solo

\* p indica puntaje acumulado, el puntaje en el cuadro negro indica el punto de corte en la propuesta original, el puntaje mínimo que deben sumar los niños de ese rango de edad para estimarse su desarrollo como normal.

de investigación de varios países. Es un test que evalúa niños de 2 a 5 años en las áreas de coordinación, motricidad y lenguaje<sup>5</sup>.

Lejarraga et al., entre 1988 y 1995, realizaron un estudio transversal de ámbito nacional en Argentina sobre el desarrollo psicomotor en 3,573 niños sanos entre los 0 y 5 años de edad, en el cual se estimaron centiles de edad de adquisición de 78 hitos de desarrollo, que conformaron la prueba nacional de pesquisa del desarrollo psicomotor<sup>11</sup>. En el mismo año, Pascucci et al. reportaron que el instrumento es válido para ser usado en el primer nivel de atención para el reconocimiento de niños con sospecha de sufrir trastornos del desarrollo<sup>12</sup>.

En Colombia, Ortiz Pinilla et al. desarrollaron la Escala Abreviada del Desarrollo en una muestra de 16,180 niños de las diversas regiones del país para la obtención de parámetros normativos. Esta escala tamiz permite detectar casos de alto riesgo, detención o retraso del desarrollo en niños menores de 5 años<sup>13</sup>.

El Sistema de Vigilancia y Promoción del Desarrollo Integral del Niño es una estrategia amplia que incluye dentro de sus acciones evaluar el desarrollo del niño de 1 a 24 meses de edad con 64 conductas del desarrollo, las cuales se integran en una cartilla junto con otros indicadores de riesgo psicosocial. Las conductas del desarrollo operan en el interior del sistema como una guía para la toma de decisiones en cuanto a necesidades diagnósticas de mayor complejidad, pero además, para ofrecer acciones adicionales de promoción del desarrollo y estimulación temprana<sup>7,14</sup>.

En las cartillas, los 64 indicadores se distribuyen en 13 rangos de edad mensuales los 6 primeros meses (19 ítems), bi-

mensuales entre los 7 y 12 meses (19 ítems) y trimestrales entre los 13 y 24 meses (26 ítems). Los indicadores fueron propuestos según la edad de presentación de los mismos en una muestra normativa de niños de bajo riesgo, y su sistema de aplicación y calificación implica que el niño realice todas las conductas correspondientes a su rango de edad y del rango anterior (tabla 1). Por lo mismo, las cartillas de vigilancia del desarrollo son susceptibles de ser evaluadas en cuanto a su capacidad de detectar alteraciones en el desarrollo<sup>7</sup>. Al ser considerada como una prueba de tamizaje del desarrollo, tiene la función de identificar a los niños con alta probabilidad de presentar una alteración o retraso del desarrollo en la población aparentemente sana; su aplicación establece la condición de riesgo o alteración, pero no confiere valor diagnóstico, debe ser fácil y rápida de aplicar, económicamente viable, confiable y válida (especificidad y sensibilidad > 0.70)<sup>4</sup>.

Estimar la utilidad de una prueba incluye varios procesos, uno de los cuales es el de validación concurrente, donde se busca conocer la medida en la que coinciden los resultados de dicha prueba con las evaluaciones diagnósticas habitualmente utilizadas (estándar de oro) que se aplican simultáneamente.

El objetivo de este trabajo es evaluar la validez concurrente por criterio externo de las conductas del desarrollo de las cartillas de vigilancia del Sistema de Vigilancia y Promoción del Desarrollo Integral del Niño respecto a la prueba Bayley II<sup>15</sup> y la valoración de Gesell-Amatruda<sup>16</sup>, y proponer nuevos puntos de corte si el análisis de curvas ROC ofrece mejores indicadores de validez concurrente por el uso de ambos criterios.

## Material y métodos

La validación de las conductas del desarrollo de las cartillas de vigilancia se realizó en una muestra de 2,702 niños procedentes de zonas urbanas del Distrito Federal, Estado de México, Morelos y Oaxaca. Cada mes de edad de los niños contó con una muestra de 70 a 110 niños, excepto el primer mes que se integró en una muestra de 56 niños. Por tratarse de un estudio de validación concurrente se admitió a niños de bajo riesgo y a niños con distinto perfil de riesgo psicosocial, el cual era desconocido por los evaluadores. Los niños con un evidente retraso en el desarrollo o que concurrían con un diagnóstico previo no fueron incluidos en el estudio. Algunos niños podían ser evaluados 2 (12) o 3 (8) veces si sus edades de evaluación correspondían con rangos distintos de edad semestral, pero las valoraciones eran consideradas por separado en virtud de que se trata de un diseño transversal.

Se calculó la sensibilidad, la especificidad, valores predictivos positivos y negativos y sus intervalos de confianza en 13 grupos de edad (mensual primeros 6 meses, bimensual 7 a 12 meses y trimestral de los 13 a 24 meses), utilizando los criterios de puntos alcanzados en la edad de valoración en cada edad (ejemplo 12, 19, 31 y 45 puntos a las edades de 4, 6, 9-10 y 13-15 meses como se indica en el tabla 1). Posteriormente, se estimaron mediante análisis de curvas ROC, nuevos puntos de corte para conformar una propuesta de puntuación que mejore los indicadores de validez, por mayor área bajo la curva y priorizando los criterios de sensibilidad y valor predictivo negativo<sup>17,18</sup>.

## Instrumentos utilizados como estándar de oro

La Escala de Desarrollo Infantil Bayley II<sup>15</sup> es una de las pruebas de desarrollo más utilizadas para evaluar el desempeño del niño en la primera infancia; fue creada por Nancy Bayley en Estados Unidos en el año 1993. Existe una tercera versión de la escala de Bayley, pero su uso aún no es extenso en México. El Bayley II valora el desarrollo entre 1 mes y 3 años y medio de edad. La escala mental incluye ítems que evalúan capacidades como memoria, aprendizaje, percepción, clasificación y vocalización, y la escala psicomotora evalúa las actividades motoras gruesas y finas, incluyendo movimientos como el gateo, sentado, parado, caminar, correr, saltar y la coordinación sensorio-motora. En la actualidad es la prueba que más se ha utilizado como estándar de oro para la validación concurrente de pruebas de escrutinio del desarrollo. La prueba proporciona los resultados bajo criterios de normalización con media de 100 y desviación estándar de 15, no arroja un promedio o índice global, por lo que se tomaron los 2 valores de la escala psicomotora y mental por separado, con un punto de corte de 70, equivalente a -2 DS en cualquiera de las escalas para estimar el desarrollo como alterado.

La prueba de Gesell “Diagnóstico del desarrollo normal y anormal del niño”, propuesta por Arnold Gesell y Catherine Amatruda<sup>16</sup>, es una prueba que permite evaluar el desarrollo hasta los 42 meses de edad en 5 áreas: motor fino, motor grueso, adaptativo, personal social y lenguaje; no es una prueba estandarizada, su criterio de normalidad está determinado por la relación entre la edad de desarrollo y la edad cronológica. Para los efectos de la presente inves-

tigación se tomó como criterio el valor normalizado de la edad de desarrollo mostrado por los niños de bajo riesgo en cada mes de edad, considerándose como alteración el score Z < -2, que corresponde de manera global a un coeficiente de desarrollo entre 76 y 84 en cada área, al igual que en la prueba de Bayley II se consideró alterado el desarrollo del niño cuando la puntuación en días de desarrollo en cualquiera de sus áreas era menor -2 DS. La prueba también utiliza como estrategia el coeficiente general de desarrollo promediando los coeficientes de sus áreas, pero se estimó que el uso del mismo podía enmascarar alteraciones específicas.

El estándar de oro final para el presente reporte fue el resultado combinado de la aplicación de ambas pruebas, de tal forma que si cualquiera de las escalas de alguna de las pruebas arrojase un valor equivalente menor que -2 DS, se consideraba el resultado de la valoración como desarrollo alterado.

La aplicación de las cartillas y las pruebas de desarrollo fueron simultáneas, la calificación se realizó por separado por aplicadores estandarizados con confiabilidad interobservador > 90, participando 6 evaluadores de las pruebas de desarrollo y 12 evaluadores de las cartillas.

## Resultados

La estimación de la validez por criterio externo de las cartillas calificadas siguiendo la propuesta original y tomando como estándar la puntuación más baja alcanzada en las pruebas de desarrollo menor que -2 DS identificó valores de sensibilidad de 75.99 y especificidad de 90.46 (tabla 2) si se consideraban los puntos de corte inicialmente propuestos en el manual de las mismas. El análisis por grupo de edad presentó una tendencia a la alta especificidad y regular sensibilidad. Debe señalarse que en la forma original de calificación existían 13 grupos de edad y un punto de corte para cada uno de esos grupos. Para mejorar los puntos de corte para decidir la presencia de alteración, se realizó un análisis de curvas ROC para cada edad en meses, aportando puntos de corte adicionales para interpretar los resultados de la aplicación mes a mes.

**Tabla 2** Indicadores globales de validez concurrente de las cartillas del desarrollo tomando los puntos de corte originales

Pruebas de desarrollo				
	N	Retraso	Normal	Totales
Cartillas	col			
	fila			
Retraso	1225	104	1329	
	S 75,99	9,54		
Normal	387	986	1373	
	24,01	E 90,46		
	28,19	VP- 71,81		
	1612	1090	2702	

S: sensibilidad; E: especificidad; VP+: valor predictivo positivo; VP-: valor predictivo negativo.

El análisis por grupos de edad mensual con nuevos puntos de corte mostró valores adecuados ( $> 70$ ) de sensibilidad en 21 de los 24 meses evaluados: 2-19 y 22-24 meses, pero a los 18, 22 y 23 meses con regular especificidad, y en el primer mes con especificidad baja; en todas las demás 21 edades, la especificidad fue buena.

Modificados los puntos de corte, la sensibilidad global de las cartillas subió a 84.43, pero la especificidad bajó a 75.78, el valor predictivo positivo fue de 83.75 y el valor predictivo negativo fue de 76.69 (tabla 3). En la tabla 4 se muestran los resultados de la estimación de los indicadores de validez concurrente mes a mes a partir de los ajustes estimados por el análisis de curvas ROC, proponiéndose nuevos puntos de corte con adecuados valores de sensibilidad en 21 de las 24 edades y regulares (60 a 70) en 3 edades, y en el primer mes, la sensibilidad fue  $< 60$ . En 20 de las 24 edades, la sensibilidad también fue buena, y en 17 de las edades se observaron valores predictivos negativos adecuados, y en 5 edades fueron regulares y solo en 2 edades fueron bajos (tabla 4).

**Tabla 3 Indicadores globales de validez concurrente de las cartillas del desarrollo tomando los puntos de corte sugeridos por el análisis de curvas ROC**

**Pruebas de desarrollo**

Cartillas	N	Retraso	Normal	Totales
	% col			
	% fila			
Retraso	1,361	264	1,625	
S 84.43		24.22		
VP+ 83.75		16.25		
Normal	251	826	1,077	
	15.57	E 75.78		
	23.31	VP- 76.69		
	1,612	1,090	2,702	

S: sensibilidad; E: especificidad; VP+: valor predictivo positivo; VP-: valor predictivo negativo.

**Tabla 4 Valores de los indicadores de validez concurrente por mes de edad tomando como criterio los puntos de cohorte del análisis de curvas ROC**

Edad meses	Sensibilidad	Especificad	VP+	VP-	Punto de corte alterado si es menor que:
1	53.85	55.00	43.75	64.71	4
2	89.19	70.37	80.49	82.61	7
3	80.88	83.87	91.67	66.67	8
4	87.06	91.67	94.87	80.00	11
5	87.63	84.31	91.40	78.18	15
6	88.89	81.63	91.43	76.92	18
7	90.48	91.67	96.94	76.74	23
8	73.86	85.06	87.84	77.23	24
9	83.49	78.95	91.92	62.50	30
10	86.05	75.76	82.22	80.65	32
11	93.14	90.91	95.96	85.11	37
12	80.52	70.97	77.50	74.58	39
13	89.90	83.33	92.13	67.31	44
14	89.61	85.29	86.84	84.43	47
15	75.00	80.65	73.33	81.97	48
16	85.25	95.59	94.55	87.84	51
17	70.21	75.00	66.00	78.46	54
18	83.67	65.91	73.21	78.38	58
19	81.13	70.00	82.69	67.74	59
20	69.05	70.45	69.05	70.45	59
21	64.29	78.57	66.67	76.74	60
22	86.67	62.50	81.25	71.43	62
23	78.05	65.00	82.05	59.09	63
24	70.00	73.68	87.50	48.28	63

VP+: valor predictivo positivo; VP-: valor predictivo negativo.

De manera general, los valores de sensibilidad y especificidad son adecuados en todos los meses, excepto los meses 20 a 23, donde la sensibilidad o la especificidad son regulares, pero no ambas, y en el primer mes tanto sensibilidad como especificidad son bajas (tabla 4).

## Discusión

Los resultados iniciales muestran indicadores de validez concurrente adecuados, con mejor especificidad que sensibilidad global, pero un sesgo hacia la alta especificidad y baja sensibilidad en algunas edades, lo que de forma general se traduce en que con los puntos de corte originales a los niños les fue más fácil resolver las demandas presentadas en las cartillas para quedar en el grupo de desarrollo normal de las cartillas que en el correspondiente a las pruebas de desarrollo. Aunque la sensibilidad se muestra global en la tabla 2, es el resultado de aplicar 13 puntos de corte originales.

Con las modificaciones realizadas a partir del análisis de curvas ROC, la sensibilidad global de la prueba mejoró de 75.99 a 84.43 conservándose valores adecuados de especificidad (75.78); sin embargo, en algunas edades la especificidad fue regular (60 a 70) pero aceptables según los parámetros utilizados en pruebas similares<sup>3-7</sup>. Una forma de resolver el hecho de privilegiar la sensibilidad a costa de la especificidad es generar un rango intermedio de 1 o 2 puntos en los que esas puntuaciones inestables se presenten como un resultado “dudoso” quedando para ellos implementar medidas que se deriven de la apreciación de otras condiciones que rodean el proceso de desarrollo y su vigilancia como la crianza, accesibilidad a servicios de salud y condiciones de riesgo o protección psicosocial, como la escolaridad de los padres y acceso a programas de estimulación<sup>7,19</sup> y de esa forma se pueden establecer parámetros de mayor seguridad para el establecimiento de la normalidad y la alteración. En una exploración preliminar de los datos de este estudio, aproximadamente un 12% de la población presentarían esas puntuaciones intermedias a las que eventualmente en la interpretación podrían calificarse como dudosos en los resultados de la aplicación, y para los cuales una decisión puede ser tomada en una reevaluación 2 a 4 semanas después. No presentamos esos rangos en el presente reporte, pues conlleva modificar drásticamente los criterios de calificación originales por la aparición de un posible tercer resultado.

El que en algunas edades se originasen valores regulares de especificidad con los nuevos puntos de corte por incrementarse el número de casos detectados, generándose junto a una mayor proporción de verdaderos positivos una cantidad mayor de falsos positivos, con una reducción de la especificidad, puede deberse a la necesidad de ítems adicionales próximos a esas edades, pues la variación entre una edad y otra reducida en puntos, por ejemplo entre los 18 y 19 meses, es de un punto, y entre los 22 y 24 la diferencia es de 2 puntos. Ante esta situación, es posible que la incorporación de indicadores adicionales en esas edades permita diferenciar con mayor precisión el desarrollo normal del alterado.

Cuando se configura una prueba deben de atenderse 2 supuestos, que los niños a la edad en la que se les explora

realicen un número de conductas y que esas conductas correspondan a un orden o secuencia que las bases teóricas y empíricas indican<sup>19,20</sup>.

Considerar solo el orden puede llevar a limitar la evaluación por la ubicación de los ítems en los grupos de edad en que se exploran, sin indagar otras posibilidades de ejecución en edades de la prueba mas avanzadas. Sin embargo, durante la aplicación de la cartilla fue posible explorar de manera continua las ejecuciones del niño hasta que no le fuese posible resolver completamente una edad, por lo que el establecimiento de puntuaciones sumadas permitió que los aplicadores se centren durante la aplicación en lo que podía hacer el niño y no en la forma en que este ajustaba al ordenamiento que tienen las conductas en el formato de las cartillas. De esta forma, nosotros replanteamos que la indicación para la aplicación no debe ser cumplir con los indicadores de su edad y la anterior, sino además, sumar otras conductas que pueda realizar de edades posteriores del formato aunque solo cumpla parcialmente las que corresponden con su propia edad, pero desde el punto de vista clínico no debe ignorarse que la falla de una conducta en su edad o en la edad anterior puede reflejar alguna alteración correspondiente al área específica que representa dicha conducta.

Como lo recomiendan otras pruebas como el CAT-CLAMS, se debe iniciar con el bloque de edad en la que todas las conductas son positivas y continuar hasta el bloque de edad en que todas resulten negativas<sup>21</sup>.

Aunque como se menciona líneas arriba, una medida que se desprende del presente reporte es aplicar los ítems de forma continua, aun en esos casos el orden en que se disponen los ítems, respecto a la forma en que se presentan en una población de referencia contribuye a lograr una mejor estimación del desarrollo del niño, ya que de existir conductas que tienden a presentarse en edades previas a las que se sugiere en la prueba puede generar falsos negativos en los rangos de edad cercanos, mientras que de existir conductas que solo son posibles de realizar por la población en edades posteriores a las que se propone, se tiende a generar falsos positivos. Sin embargo, la estimación de las edades de presentación de los ítems es una tarea pendiente que es posible implementar si se analiza únicamente las adquisiciones de niños de bajo riesgo en un reporte futuro.

Consideramos que las cartillas de vigilancia son un instrumento útil para detectar alteraciones en el desarrollo en los 2 primeros años de vida, aunque su uso en los meses 1 y 21 debe hacerse con reserva por la baja sensibilidad, y en los meses 22 y 23 deben tomarse en cuenta valores regulares de especificidad, lo que deriva únicamente en un incremento de falsos positivos. Sin embargo, los valores generales de sensibilidad y especificidad son superiores a los mostrados por otros instrumentos analizados por Romo en un trabajo reciente<sup>4</sup>, considerando a nuestro juicio que una limitación en la mayoría de esos reportes es que no se analizan estos parámetros de validez concurrente mes a mes, generándose la falsa impresión de que la prueba se comporta de manera homogénea en las distintas edades, cuando es sabido que las pruebas de desarrollo se van transformando según la edad en la que se aplique. Por lo mismo, las cartillas de vigilancia del Sistema de Vigilancia y Promoción del Desarrollo Integral del Niño son una opción para vigilar el desarrollo de manera sencilla, mostrando valores similares o superiores a los mostrados por otras pruebas<sup>3-7,11-13</sup>.

## Financiación

Fondo Sectorial de Investigación en Salud y Seguridad Social CONACYT. Proyecto SALUD-2011-1-162227.

## Referencias

1. Figueiras C. Manual para la vigilancia del desarrollo infantil en el contexto AIEPI Washington: OPS; 2006.
2. 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. Fe de erratas en: *Pediatrics*. 2006;118:1808-9.
3. Sánchez Pérez MC, Mandujano Valdés MA, Martínez I, Muñoz-Ledo P, Rivera González R. Los procedimientos de tamizaje para la evaluación y el seguimiento del desarrollo infantil. *Rev Cien Clín.* 2004;5:11-20.
4. 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;69:450-62.
5. Vericat A, Orden AB. Herramientas de screening del desarrollo psicomotor en Latinoamérica. *Rev Chil Pediatr*. 2010;81:392-401.
6. Atkin LC. Paso a paso. Cómo evaluar el crecimiento y desarrollo de los niños. México: UNICEF, Editorial Pax; 1987.
7. Rivera R. Vigilancia del desarrollo en el primer nivel de atención. En: Rivera R, Sánchez C (Eds.). Vigilancia del desarrollo integral del niño. México: Editores de Textos Mexicanos; 2009. p. 9-15.
8. Sánchez MC, Benavides H, Mandujano M, Rivera R, Martínez I, Alvarado G. Valoración Neuroconductual del Desarrollo del Lactante (VANEDELA). 3.<sup>a</sup> ed. México: Universidad Autónoma Metropolitana-Xochimilco; 2007.
9. Norma Oficial Mexicana NOM-031-SSA2-1999 para la Atención a la Salud del Niño. México: Diario Oficial de la Federación, 12 de febrero 2001.
10. Rizzoli-Córdoba A, Schnaas-Arrieta L, Liendo-Vallejos S, Buenrostro-Márquez G, Romo-Pardo B, Carreón-García J, et al. Validación de un instrumento para la detección oportuna de problemas de desarrollo en menores de 5 años en México. *Bol Med Hosp Infant Mex*. 2013;70:195-208.
11. Lejarraga H, Pascucci MC, Krupitzky S, Kelmansky D, Bianco A, Martínez E, et al. Psychomotor development in Argentinean children aged 0-5 years. *Paediatr Perinat Epidemiol*. 2002; 16:47-60.
12. Pascucci MC, Lejarraga H, Kelmansky D, Álvarez M, Boullón M, Breiter P, et al. Validación de la prueba nacional de pesquisa de trastornos de desarrollo psicomotor en niños menores de 6 años. *Arch Argent Pediatr*. 2002;100:374-84.
13. Ortiz Pinilla N. Escala Abreviada de Desarrollo-EAD. Bogotá: Editorial del Ministerio de Salud de Colombia; 1991. p. 3-41.
14. Sánchez C, Rivera González R, Figueroa Olea M. Promoción del desarrollo infantil temprano. Orientación y estrategias. México: Editores de Textos Mexicanos; 2009.
15. Bayley N. Bayley Scales of Infant Development Manual. Nueva York: The Psychological Corporation; 1993. p. 67-77.
16. Gesell A, Amatruda C. Diagnóstico del desarrollo Normal y Anormal del niño. México: Paidós; 1992. p. 29-35.
17. Martínez Arias R. Psicometría: Teoría de los test psicológicos y educativos. 2.<sup>a</sup> ed. España: Síntesis Psicología; 2005.
18. Sattler J. Evaluación infantil. México: Manual Moderno; 2003. p. 161-4.
19. Luinge MR, Post WJ, Wit HP, Goorhuis-Brouwer SM. The ordering of milestones in language development for children from 1 to 6 years of age. *J Speech Lang Hear Res*. 2006;49:923-40.
20. Rivera González R, Sánchez C, Corral Guille I, Figueroa Olea M, Soler Limón K, Martínez Vázquez I, et al. Edad de presentación de los reactivos del Test de Denver II en niños de 0 a 4 años de edad del Estado de Morelos. *Salud Mental*. 2013;36:459-70.
21. Accardo PJ, Capute AJ. The Capute Scales. Cognitive Adaptive Test/Clinical Linguistic & Auditory Milestone Scale (CAT/CLAMS): Baltimore: Brookes Publishing Company; 2005.