



# Boletín Médico del Hospital Infantil de México

[www.elsevier.es/bmhim](http://www.elsevier.es/bmhim)



## TEMA PEDIÁTRICO

### Pasos para transformar una necesidad en una herramienta válida y útil para la detección oportuna de problemas en el desarrollo infantil en México



Antonio Rizzoli-Córdoba\* e Ismael Delgado-Ginebra

*Unidad de Investigación en Neurodesarrollo, Hospital Infantil de México Federico Gómez, México D.F., México*

Recibido el 5 de noviembre de 2015; aceptado el 7 de noviembre de 2015

Disponible en Internet el 23 de diciembre de 2015

#### PALABRAS CLAVE

Prueba de tamiz;  
Validación;  
Prueba EDI;  
Desarrollo infantil

#### KEYWORDS

Screening tool;  
Validation;  
CDE Test;  
Child development

**Resumen** Una prueba de tamiz es una herramienta cuya función es identificar individuos presuntamente enfermos en una población aparentemente sana al establecer el riesgo o sospecha de un problema de desarrollo. Debe tenerse precaución y ser muy cuidadoso al utilizar un instrumento de este tipo, por el riesgo de efectuar mediciones que no coincidan con la realidad y retrasar una intervención a quien lo requiere.

Antes de incorporar una prueba de este tipo como parte de la práctica clínica rutinaria, es necesario certificar que posee ciertas características que hacen meritoria su utilización. A este proceso de certificación se le denomina *validación*. El objetivo de este artículo fue describir los diferentes pasos que se llevaron a cabo desde la identificación de una necesidad de detección hasta la generación de una prueba de tamiz validada y confiable, utilizando como ejemplo el proceso del desarrollo de la versión modificada de la prueba Evaluación del Desarrollo Infantil (EDI) en México.

© 2015 Hospital Infantil de México Federico Gómez. Publicado por Masson Doyma México S.A. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

#### Steps to transform a necessity into a validated and useful screening tool for early detection of developmental problems in Mexican children

**Abstract** A screening test is an instrument whose primary function is to identify individuals with a probable disease among an apparently healthy population, establishing risk or suspicion of a disease. Caution must be taken when using a screening tool in order to avoid unrealistic measurements, delaying an intervention for those who may benefit from it.

Before introducing a screening test into clinical practice, it is necessary to certify the presence of some characteristics making its worth useful. This “certification” process is called

\* Autor para correspondencia.

Correos electrónicos: [antoniorizzoli@gmail.com](mailto:antoniorizzoli@gmail.com), [antoniorizzoli@hotmail.com](mailto:antoniorizzoli@hotmail.com) (A. Rizzoli-Córdoba).

validation. The main objective of this paper is to describe the different steps that must be taken, from the identification of a need for early detection through the generation of a validated and reliable screening tool using, as an example, the process for the modified version of the Child Development Evaluation Test (CDE or Prueba EDI) in Mexico.

© 2015 Hospital Infantil de México Federico Gómez. Published by Masson Doyma México S.A. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introducción

La construcción de una base sólida para un desarrollo saludable durante los primeros años de vida es requisito fundamental tanto para el bienestar individual como para la productividad económica, las comunidades exitosas y las sociedades civiles armoniosas<sup>1</sup>. La arquitectura del cerebro en desarrollo se construye a través de un proceso que inicia antes del nacimiento, continúa en la edad adulta y se establece como base, ya sea fuerte o frágil, para la salud, el aprendizaje y la conducta posterior<sup>2</sup>.

Durante los primeros 5 años de vida acontece una sucesión de periodos *sensibles*, cada uno relacionado con la formación de circuitos que se asocian con habilidades específicas. Así, las funciones cerebrales tienen un periodo crítico de máximo potencial de desarrollo en función de la edad. El nivel de desarrollo alcanzado depende de la interacción de los genes y las experiencias tempranas<sup>3</sup>. Con base en lo anterior surgió el concepto de desarrollo infantil temprano (DIT), que abarca desde el embarazo y hasta los 6 años de edad y es *un proceso de cambio en el que el niño aprende a dominar niveles siempre más complejos de movimiento, pensamiento, sentimientos y relaciones con los demás*. Se produce cuando el niño interactúa con las personas, los objetos y otros estímulos en su ambiente biofísico y social, y aprende de ellos<sup>4</sup>.

Mientras el cerebro en maduración se especializa en asumir funciones cada vez más complejas, también se vuelve menos capaz de reorganizarse y adaptarse a nuevos retos. Una vez que el circuito se ha *instalado*, se irá estabilizando con la edad, dificultando su alteración posterior. Aunque las *ventanas de oportunidad* pueden permanecer abiertas por muchos años, modificar la conducta o construir nuevas habilidades sobre la base de circuitos cerebrales que no fueron debidamente instalados cuando se formaron por primera vez requiere de mucho más trabajo; además, se necesitan cantidades mayores de energía metabólica para compensar funciones de circuitos que no se desempeñan de la manera esperada<sup>5</sup>. Hacer las cosas de forma correcta desde la primera vez, en lugar de esperar y tratar de arreglarlas después, es menos costoso tanto para los individuos como para la sociedad en general<sup>2</sup>. Por esto, se debe vigilar la correcta formación de dichos circuitos.

La evaluación del desarrollo infantil es un proceso destinado a conocer y cuantificar el nivel de maduración alcanzado por un niño comparado con su grupo de edad<sup>6</sup>, lo que permite identificar alteraciones y establecer un perfil individualizado sobre las fortalezas y debilidades de los diferentes dominios evaluados. La detección temprana de problemas en el desarrollo es de suma importancia para el

bienestar de los niños y sus familias, ya que permite acceder a un diagnóstico y tratamiento oportunos<sup>7</sup>, y llevar a cabo acciones dentro de las *ventanas de oportunidad* para evitar la consolidación de circuitos inadecuados y favorecer que estos se reorganicen de una forma óptima. Estas acciones dan pie a que los niños y niñas puedan desarrollar habilidades cada vez más complejas acordes con su edad, y así alcanzar su máximo potencial.

El juicio clínico del personal de salud no es suficiente para realizar esta evaluación e identificar pacientes con retraso del desarrollo. De ahí la importancia de emplear herramientas estandarizadas para detectar a estos pacientes<sup>8</sup>. Una prueba de tamiz es una herramienta que tiene como función identificar individuos presuntamente enfermos en una población aparentemente sana al establecer el riesgo o sospecha de un problema de desarrollo. Se debe actuar cautelosamente al utilizar un instrumento, por el riesgo de efectuar mediciones que no coincidan con la realidad, y retrasar una intervención a quien sí lo requiere.

Antes de incorporar una prueba de este tipo como parte de la práctica clínica rutinaria es necesario *certificar* que tiene ciertas características que hacen meritoria su utilización. A este proceso de *certificación* se le denomina *validación de pruebas*<sup>9</sup>. Existen varios procesos que se han considerado como validación sin que en realidad lo sean. Los más frecuentes son efectuar únicamente traducciones, solamente realizar pruebas de concordancia o correlación con los resultados de la medición de otro instrumento o practicar tan solo pruebas de concordancia entre diferentes evaluadores<sup>10</sup>.

El objetivo de este artículo fue describir los pasos que se llevan a cabo para la validación de pruebas de tamiz ([tabla 1](#)), utilizando como ejemplo el proceso del desarrollo de la versión modificada, validación y evaluación en campo de la prueba EDI. Estos pasos deben responder una serie de preguntas que a continuación se exponen.

- ¿Existen pruebas que midan lo que se desea cuantificar de forma confiable?

La escala que se va a utilizar debe ser la mejor disponible. Por ello, es importante analizar previamente si existe un instrumento que cumpla con el objetivo que se desea alcanzar<sup>10</sup>. En el caso de pruebas de tamiz de desarrollo infantil en México, ante el interés de contar con un instrumento validado para la detección oportuna, se realizó una revisión sistemática basada en la información disponible en bases de datos y motores de búsqueda de publicaciones científicas<sup>11</sup>. No se encontró en la literatura científica información sobre alguna prueba de tamiz validada en México.

**Tabla 1** Descripción de las características o atributos que hacen meritoria la utilización de una prueba de tamiz y el proceso correspondiente en la validación de escalas\*

| Característica o atributo de la prueba                                                                             | Proceso de evaluación                      | Resultado esperado                                                                                               |
|--------------------------------------------------------------------------------------------------------------------|--------------------------------------------|------------------------------------------------------------------------------------------------------------------|
| La realidad existente está adecuadamente representada por la escala                                                | Validez de apariencia                      | La prueba parece medir lo que debe medir                                                                         |
| La escala refleja la estructura de dominios o de factores en los cuales fue dividida la realidad que se va a medir | Validez de contenido/constructo            | La prueba no deja factores sin medir ni mide dominios que no pertenecen al síndrome/enfermedad                   |
| La escala funciona de manera parecida a otros instrumentos para medir esa realidad que ya han sido validados       | Validez de criterio                        | La escala funciona de forma similar a otros instrumentos                                                         |
| La escala funciona bien bajo diferentes condiciones                                                                | Confiabilidad test-retest o interevaluador | La escala funciona bien cuando se aplica en diferentes oportunidades o cuando es aplicada por distintas personas |

\* A. R. Feinstein (ref.9)

Por otro lado, las pruebas disponibles en América presentaban deficiencias o características diferentes que no las hacían las más adecuadas para su implementación en México. Como parte de la NOM-031-SSA1-1999, Para la atención a la salud del niño<sup>12</sup>, el punto 9.6.1 especifica que la valoración del desarrollo psicomotor del niño se debe realizar cada vez que el niño acuda a consulta para control de la nutrición y crecimiento, utilizando como referencia los límites de normalidad de conductas especificados en el apéndice F del mismo documento. No se encontró información disponible acerca de los valores de la prueba contenida en este apéndice (sensibilidad/especificidad) ni sobre el proceso de validación aparente o de criterio, además de que, como resultado de la prueba, solo se reporta si el niño fue identificado como con alarma. Debido a estos datos se concluyó que *“no existe un acuerdo sobre cómo evaluar el desarrollo físico, cognitivo, social y emocional de los niños en México. Los datos que se utilizan para medirlo son la matrícula de niños y niñas en educación preescolar o su inscripción en programas de cuidado”*<sup>4</sup>.

- Si se planea diseñar una nueva herramienta, ¿qué funciones clínicas debe cumplir y como estaría estructurada?

La herramienta que se requería en México debía cumplir con las siguientes funciones clínicas:

- a) Caracterización cualitativa del nivel de desarrollo de los niños. Identificación de cambios a través del tiempo que permitan realizar intervenciones adecuadas a las necesidades de los niños.
- b) Servir de base para una guía de manejo.

A falta de pruebas que cumplieran con estas funciones, la Dirección General del Programa Oportunidades (hoy PROSPERA), en 2010, solicitó a la Dra. Lourdes Schnaas, del Instituto Nacional de Perinatología, el diseño de la prueba de tamiz Evaluación del Desarrollo Infantil (EDI) en su versión original<sup>13</sup>. Esta prueba funcionaría como un instrumento para la detección oportuna de problemas del desarrollo en

niños y niñas desde un mes de vida y hasta un día antes de cumplir 5 años, utilizando un sistema de calificación en semáforo: verde (desarrollo normal), amarillo (rezago) o rojo (riesgo de retraso) (tabla 2)<sup>14</sup>.

## 2. Proceso de validación de una prueba de tamiz

Una vez que se ha diseñado o identificado un instrumento que pudiera ser de utilidad para cumplir el objetivo de identificar de forma oportuna la enfermedad o fenómeno de interés, la evaluación de la validez busca responder a las siguientes preguntas:

- ¿La escala parece medir lo que debe medir? ¿Refleja la estructura de dominios del fenómeno a evaluar?

Responder la primera pregunta es fundamental para establecer la aceptabilidad de la prueba en el escenario de aplicación, estableciendo con esto la *validez de apariencia*<sup>9</sup>. Esta validez depende del análisis de la *aparición* de la prueba y de sus ítems en cuanto a cómo buscan evaluar el fenómeno de interés (foco en el intercambio interpersonal, en la evidencia básica, coherencia biológica de los componentes, atención a colaboraciones personales). La respuesta a la segunda pregunta supone analizar los diferentes componentes de la prueba para establecer si el instrumento refleja el concepto que se pretende medir (omisiones importantes, inclusiones inapropiadas, peso de los componentes, escalas elementales satisfactorias y la calidad de la información básica), y con esto establecer la *validez de contenido*.

Para responder a estas preguntas se llevó a cabo un análisis minucioso de la versión original de la prueba EDI. Se detectaron áreas de oportunidad a partir de las cuales se desarrolló la versión modificada de la misma<sup>15</sup>. Para cada una de las preguntas se desarrollaron instrucciones y criterios precisos de calificación, se reorganizaron las preguntas en ejes con características comunes, se modificaron los criterios de calificación global para mayor congruencia y se añadió la modalidad observada en los ítems que se

**Tabla 2** Características generales de la versión original de la prueba EDI

| Característica                       | Descripción                                                                                                                                                                                                                                                                                                                                                    |
|--------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Organización general de la prueba    | 14 formatos de evaluación diferentes con base en la edad del niño                                                                                                                                                                                                                                                                                              |
| Rango de edad (meses)                | Uno a 59 meses (en niños que nacieron antes de las 37 semanas de gestación y hasta cumplir los dos años de edad se debe calcular la edad corregida para seleccionar las preguntas del grupo de edad correspondiente)                                                                                                                                           |
| Áreas del desarrollo evaluadas       | Se evalúan los diferentes dominios del desarrollo (cognitivo, personal-social, adaptativo, motor y comunicación) agrupado en las siguientes categorías: motor grueso, motor fino, lenguaje y social. En la evaluación de niños mayores a 36 meses se añade conocimiento. Se incluyeron además preguntas sobre factores de riesgo biológico y señales de alerta |
| Modalidad de aplicación de los ítems | Interrogado                                                                                                                                                                                                                                                                                                                                                    |
| Respuestas por ítem                  | Dicotómico: Sí/No                                                                                                                                                                                                                                                                                                                                              |
| Selección de los ítems               | La selección de los reactivos a evaluar fue realizada considerando que más del 75% de los niños de esa edad ya realizaran la conducta esperada, y de una forma secuencial en las diferentes pruebas                                                                                                                                                            |
| Número de ítems por prueba           | De 14 (un mes) a 36 (49 a 60 meses)                                                                                                                                                                                                                                                                                                                            |
| Resultado posible                    | a) Desarrollo normal (verde); b) Desarrollo normal con factores de riesgo (amarillo); c) Posible retraso en el desarrollo; d) Probable retraso en el desarrollo (rojo)                                                                                                                                                                                         |

EDI: Evaluación del Desarrollo Infantil.

consideraron pertinentes. Estas modificaciones se realizaron para cumplir de forma adecuada con los requerimientos de validez de apariencia y contenido (tabla 3).

Una vez lograda la validez de apariencia y contenido se buscó una *validación por consenso*. Para este fin, se organizó un panel de expertos que analizó cada uno de los ítems de la prueba EDI, comparando las dos versiones de la prueba utilizando una metodología de panel Delphi modificado. Este panel llegó a la conclusión de que la versión modificada de la prueba EDI era el instrumento de tamiz más adecuado para la evaluación de menores de 5 años en el contexto del país<sup>16</sup>. De forma adicional, se estableció que el Inventario de Desarrollo de Battelle 2.<sup>a</sup> edición en español (IDB-2), era el instrumento más adecuado para la corroboración diagnóstica<sup>17</sup>.

- ¿La prueba funciona de manera similar a otros instrumentos validados?

Para responder esta pregunta, es necesario que las pruebas sean sometidas a un proceso de *validación de criterio concurrente*. Este proceso busca conocer la medida en la que coinciden los resultados de dicha prueba con las evaluaciones diagnósticas consideradas *estándar de oro* o con aquellas habitualmente utilizadas y consideradas de referencia<sup>18</sup>. Para ello, se requiere de un estudio que cumpla con un diseño adecuado<sup>19</sup>, considerando la calidad de la evaluación y el riesgo de sesgo<sup>20</sup>.

Un estudio de validación permite establecer la capacidad de la prueba para las siguientes características: a) identificar como anormales a las personas que tienen la enfermedad (sensibilidad); b) identificar como normales a las personas que no tienen la enfermedad (especificidad); c) conocer el porcentaje del total de personas que obtienen

un resultado anormal en la prueba de tamiz en que se corrobora la enfermedad (valor predictivo positivo); y d) el porcentaje del total de personas con resultado normal en la prueba que realmente no tienen la enfermedad (valor predictivo negativo). En el caso de pruebas de tamiz, se recomienda que tengan valores de sensibilidad y especificidad mayores al 70%<sup>18</sup>.

Para conocer las propiedades de la prueba EDI en sus dos versiones (original y modificada), se llevó a cabo un estudio de validación de criterio concurrente<sup>21</sup> frente a dos instrumentos de evaluación del desarrollo considerados como estándar de referencia: IDB-2<sup>17</sup> y Bayley-III<sup>22</sup>. Como objetivo secundario de este estudio se realizó el análisis por componentes evaluados (dominios del desarrollo)<sup>23</sup> y para las diferencias por resultado anormal (amarillo vs. rojo)<sup>24</sup>. De acuerdo con el diseño del estudio y con los resultados obtenidos, y mediante un análisis comparativo de las pruebas de tamiz para problemas en el desarrollo diseñadas y validadas en México, se identificó que el estudio de validación de la prueba EDI era el que tenía menor riesgo de sesgo en los datos publicados<sup>25</sup>.

Una vez establecida la validez de la prueba, existen dos preguntas adicionales a considerar para poder llegar a contextualizar la confiabilidad de los resultados de la aplicación en campo.

- ¿En su aplicación en campo, se encuentran resultados similares a lo descrito en la validación de criterio?

Una vez identificadas las propiedades de la prueba en el estudio de la validación, se debe conocer si estas se mantienen en su aplicación en campo. Para conocer esto se realizó un estudio que incluyó niños de 16 a 59 meses de edad identificados con riesgo de retraso en la prueba EDI, en quienes

**Tabla 3** Características de la versión original de la prueba EDI y cambios en la versión modificada

| Característica                      | Componente        | Prueba EDI                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-------------------------------------|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                     |                   | Versión original <sup>a</sup>                                                                                                                                                                                                                                                                                                                                                 | Versión modificada <sup>b</sup>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| <i>Propósito y marco conceptual</i> | Función clínica   | <ul style="list-style-type: none"> <li>- Caracterización cualitativa del nivel de desarrollo infantil utilizando un sistema de semáforo</li> <li>- Ser capaz de identificar cambios en el nivel de desarrollo de los niños (a través de evaluaciones periódicas)</li> <li>- Ser la base para el desarrollo de acciones asociadas al resultado de la prueba</li> </ul>         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|                                     | Justificación     | <ul style="list-style-type: none"> <li>- Establece el cálculo de edad corregida, que disminuye los falsos positivos por edad inadecuada</li> <li>- Evalúa todos los dominios del desarrollo infantil</li> <li>- Incluye, además, dimensiones médicas que pueden ser la causa del problema y requieren manejo diferente (perímetro cefálico, movimientos anormales)</li> </ul> | <ul style="list-style-type: none"> <li>- A diferencia de la versión original (solo interrogada) incluye preguntas que deben ser observadas, lo que la vuelve más objetiva</li> <li>- Se reestructuran las preguntas en cinco ejes, para facilidad en la comprensión y calificación</li> <li>- Se desarrollan instrucciones específicas de aplicación y criterios de calificación para cada respuesta</li> <li>- Se concentran todas las pruebas en un solo manual de aplicación, evitando así el abastecimiento de 14 formatos diferentes</li> </ul> |
|                                     | Aplicabilidad     | El objetivo de la prueba es que sea posible aplicarla a todos los niños menores de cinco años del país                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| <i>Comprensibilidad</i>             | Simplicidad       | <ul style="list-style-type: none"> <li>- Todas las respuestas son dicotómicas (Sí/No)</li> <li>- Todas las áreas del desarrollo se califican igual</li> </ul>                                                                                                                                                                                                                 | - Se simplificaron los criterios de calificación                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|                                     | Oligovariabilidad | - Con dos o tres preguntas se evalúa un área del desarrollo mientras que para el dominio correspondiente en la prueba diagnóstica se ocuparían, al menos, tres veces más preguntas                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|                                     | Transparencia     | - La parte de señales de alerta es confusa porque para algunas, una es rojo, mientras que para otras preguntas es con dos o más                                                                                                                                                                                                                                               | - Se distribuyeron las preguntas agrupadas en señales de alerta de la versión original en tres ejes; aquellas preguntas relacionadas con la exploración neurológica se separaron en un eje independiente; las que con uno se convierte en rojo se agruparon en alarma; el resto, en alertas.                                                                                                                                                                                                                                                         |

Tabla 3 (continuación)

| Característica                            | Componente                              | Prueba EDI                                                                                                                                                                                            |                                                                                                                                                                 |
|-------------------------------------------|-----------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                           |                                         | Versión original <sup>a</sup>                                                                                                                                                                         | Versión modificada <sup>b</sup>                                                                                                                                 |
| <i>Replicabilidad</i>                     | Connotación biológica                   | -La versión original añade perímetro cefálico < percentil 10 o > percentil 90. Con ello, el 20% de la población obtendría un resultado anormal en la prueba sin tener problemas                       | Se modificó este criterio para ser $\pm 2$ DE                                                                                                                   |
|                                           | Claridad de las instrucciones           | No existen instrucciones para la administración de los ítems de la prueba                                                                                                                             | Se desarrolló un manual donde se explica, de forma completa, la forma de aplicación e interpretación de los ítems de la prueba                                  |
|                                           | Evaluación no sesgada                   | Esta versión establece solo el interrogatorio como fuente de obtención de la información, lo que da un sesgo tanto para la interpretación como para el cuidador para contestar si lo realiza o no     | -Se establecen dos modalidades de administración de ítems: interrogado y observado                                                                              |
| <i>Idoneidad de la escala de medición</i> | Comprensibilidad                        | -Todas las categorías tienen opciones exhaustivas, mutuamente excluyentes, exhaustivamente incluyentes y con valores reales                                                                           |                                                                                                                                                                 |
|                                           | Discriminación                          | -Al incluirse en la calificación los factores de riesgo biológico que son constantes para todas las edades, no se permitía evaluar cambios en el estado de la misma persona                           | -Se ajusta el sistema de calificación para que tanto los factores de riesgo biológico como las señales de alerta no modifiquen el resultado global de la prueba |
| <i>Validez de apariencia</i>              | Foco en el intercambio interpersonal    | -El hecho de que sea solo interrogado es una limitante en la calidad de la información obtenida                                                                                                       | -La corroboración de los aspectos a través de la observación de la realización de la conducta permite que sea más objetivo y preciso                            |
|                                           | Foco en la evidencia básica             | -Las preguntas para cada grupo de edad están acorde con lo que debe realizar el niño en ese periodo                                                                                                   |                                                                                                                                                                 |
|                                           | Coherencia biológica de los componentes | -Los elementos están agrupados en categorías que permiten una evaluación de aspectos similares en la misma área, además de que llevan una secuencia progresiva en complejidad a lo largo de la prueba |                                                                                                                                                                 |
|                                           | Atención a colaboraciones personales    | -Está basado en el entendimiento y cooperación de la madre o cuidador                                                                                                                                 | -Se establecen criterios para la evaluación de la conducta y <i>tips</i> para favorecer la evaluación                                                           |
|                                           |                                         |                                                                                                                                                                                                       |                                                                                                                                                                 |

**Tabla 3** (continuación)

| Característica              | Componente                         | Prueba EDI                                                                                                                                                                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-----------------------------|------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                             |                                    | Versión original <sup>a</sup>                                                                                                                                                                           | Versión modificada <sup>b</sup>                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| <i>Validez de contenido</i> | Omissiones importantes             | -A diferencia de las demás pruebas de tamiz, es la única que incluye aspectos de exploración neurológica, que pueden ser la causa de los problemas y cuya identificación favorece una atención oportuna |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|                             | Inclusiones inapropiadas           | -En las preguntas del perímetro cefálico se incluía un rango muy amplio de resultado                                                                                                                    | -Se acotó el valor de corte para considerar anormal el perímetro cefálico                                                                                                                                                                                                                                                                                                                                                                                                               |
|                             | Peso de los componentes            | -Existen en la misma categoría (señales de alerta) ítems que con uno solo se modifica el resultado, otros que con dos, etcétera                                                                         | -Se agrupan los ítems con criterios de calificación similares en ejes independientes                                                                                                                                                                                                                                                                                                                                                                                                    |
|                             | Escalas elementales satisfactorias | -No se estipula el criterio para cada ítem de la prueba                                                                                                                                                 | -Se establecieron criterios claros para la respuesta de cada ítem                                                                                                                                                                                                                                                                                                                                                                                                                       |
|                             | Calidad de la información básica   | -Las instrucciones son muy generales y no se explican a detalle los procesos necesarios para aplicar correctamente la prueba (cálculo de edad corregida y medición perímetro cefálico)                  | -Se especifican las instrucciones para la evaluación de cada ítem y se explican a detalle los procesos más complicados (cálculo de edad corregida y medición de perímetro cefálico)                                                                                                                                                                                                                                                                                                     |
| <i>Facilidad de uso</i>     |                                    | La prueba estaba diseñada para ser utilizada en 14 formatos independientes, lo que complicaba el abastecimiento de cada uno                                                                             | -Se agrupan los 14 formatos en un mismo manual donde, además a través de pestañas, se puede consultar como evaluar los ítems de la prueba<br>-Es de fácil aplicación y puede ser aplicada por personas con bajo nivel de escolaridad<br>-Los formatos de aplicación se encuentran en un manual; solo se requiere la reproducción de la hoja de respuesta que va al expediente<br>-Los materiales para la aplicación pueden ser elaborados a partir de cosas disponibles en la localidad |

EDI: Evaluación del Desarrollo Infantil; DE: desviación estándar.

<sup>a</sup> Rizzoli-Córdoba y colaboradores (ref.21).

<sup>b</sup> Rizzoli-Córdoba y colaboradores (ref.15).

se buscó establecer la confirmación diagnóstica de retraso mediante la aplicación de IDB-2. En este estudio se encontró que el 93.2% de los niños identificados con riesgo de retraso por la prueba EDI tuvieron diagnóstico de retraso en al menos un dominio evaluado por la prueba diagnóstica<sup>26</sup>. Este resultado fue similar al descrito en el estudio de la validación de criterio (93.8%)<sup>21</sup>, con lo que se corroboró la confiabilidad de los resultados al aplicar la prueba en campo.

- ¿La escala funciona bien cuando se aplica en diferentes momentos o cuando es aplicada por distintas personas?

Si un mismo paciente es evaluado al mismo tiempo por dos evaluadores diferentes, los resultados de la evaluación deberían ser similares; esto es lo que se mide con la *confiabilidad interevaluador*. La concordancia en el resultado de la prueba al evaluar al mismo paciente en diferentes

ocasiones en un lapso corto de tiempo es lo que se denomina *confiabilidad test-retest*. En un estudio que tuvo como objetivo evaluar un modelo de supervisión que permitiera identificar la calidad de la aplicación de la prueba EDI a nivel poblacional, se encontró una concordancia interevaluador del 86.1% a través de la aplicación simultánea de la prueba como estudio de sombra; y una concordancia en la re-aplicación de la prueba entre el supervisor y el aplicador del 88.1%. Estas diferencias en el resultado estuvieron asociadas con la incorrecta administración de alguno de los ítems por personal específico, lo que refuerza la necesidad de acompañamiento y verificación de la correcta aplicación en campo para obtener resultados confiables<sup>27</sup>.

Para la detección oportuna de problemas en el DIT se deben plantear, como primera línea, las pruebas de tamiz que parezcan medir lo que deben medir (validez de apariencia), incluyan todos los dominios del desarrollo a ser evaluados (validez de constructo), que al compararse con las evaluaciones diagnósticas consideradas *estándar de oro* tengan valores de especificidad (porcentaje del total de personas con desarrollo normal por el estándar de oro que la prueba de tamiz identifique como normales) y sensibilidad (porcentaje de las personas enfermas por el estándar de oro que la prueba de tamiz identifique como anormales) superiores al 70% (validez de criterio)<sup>18</sup>; que tenga resultados similares cuando se aplique en diferentes oportunidades o por distintas personas (confiabilidad test-retest o evaluador); dado que el desarrollo infantil es un proceso de cambio, que detecte estas diferencias con el paso del tiempo (sensibilidad al cambio); y que sean rápidas y de fácil administración, para que puedan ser utilizadas por personal no especialista en las unidades de atención primaria (utilidad)<sup>28</sup>.

La evidencia disponible sobre la prueba EDI corrobora el cumplimiento de cada uno de los puntos antes señalados, por lo que es considerada actualmente la prueba más adecuada para la detección oportuna de problemas del desarrollo en México<sup>16</sup>, y es la prueba recomendada por el Centro Nacional de la Infancia y la Adolescencia para la evaluación del desarrollo infantil en las unidades de atención primaria del país<sup>29</sup>. El análisis de la información obtenida a través de su aplicación en campo permitirá identificar las áreas de oportunidad, y corroborar si los resultados de los diferentes pasos de la validación son consistentes al ser aplicada como parte de una política pública.

## Financiamiento

Para este escrito no se obtuvo financiamiento.

## Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

## Agradecimientos

Dr. Onofre Muñoz Hernández (tutor de Doctorado); Dr. Juan Garduño Espinosa, Dr. Camilo de la Fuente Sandoval, Dr. Ricardo Pérez Cuevas (Comité Tutorial de Doctorado); Dra. Lucina Isabel Reyes Lagunes, Dra. Hortensia Reyes Morales, Dra. Guadalupe Silvia García de la Torre, Dr. Miguel

Ángel Villasís Kever (integrantes del jurado del examen de grado); Dra. Rosa Elba Leyva Huerta, Mtra. María de los Ángeles Torres Prieto, Lic. Magnolia Silva Méndez (Coordinación de Posgrado de la Universidad Nacional Autónoma de México); Dr. Gabriel O'Shea Cuevas, Dr. Daniel Aceves Villagrán, Lic. Joaquín Carrasco Mendoza, Mtra. Fátima Adriana Antillón Ocampo (Comisión Nacional de Protección Social en Salud); Dr. Mariano Ramírez Degollado, Amapola Adell Grass, Dr. Daniel Carrasco Daza (In memoriam).

## Referencias

- Shonkoff JP, Phillips DA, editores. Committee on Integrating the Science of Early Childhood Development. From Neurons to Neighborhoods. The Science of Early Childhood Development. Washington, D.C: National Academy Press; 2000. p. 588.
- Shonkoff JP, Karakowsky Y. Neurobiología del desarrollo en la primera infancia. Fundación para una sociedad sostenible. En: Santibañez-Martínez L, Calderón Martín del Campo DE, editores. Los invisibles —Las niñas y los niños de 0 a 6 años—. México D.F: Mexicanos Primero Visión 2030; 2014. p. 23–46. A.C.
- Nelson CA. Neural plasticity and human development: the role of early experience in sculpting memory systems. Dev Sci. 2000;3:115–36.
- Myers R, Martínez A, Delgado MA, Fernández JL, Martínez A. Desarrollo Infantil Temprano en México. Diagnóstico y recomendaciones. Washington, D.C: División de Protección Social y Salud, Banco Interamericano de Desarrollo; 2013. Disponible en: <http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=37427911>.
- Knudsen E. Sensitive periods in the development of the brain and behavior. J Cogn Neurosci. 2004;16:1412–25.
- Costas MC. Evaluación del desarrollo en atención temprana. Rev Interuniv Form Profr. 2009;23:39–55.
- 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.
- Mackrides PS, Ryherd SJ. Screening for developmental delay. Am Fam Physician. 2011;84:544–9.
- Feinstein AR. Clinimetrics. London: Yale University Press; 1987. p. 272.
- Sánchez R, Echeverry J. Validación de escalas de medición en salud. Rev Salud Publica. 2004;6:302–18.
- 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.
- NORMA Oficial Mexicana NOM-031-SSA2-1999, Para la atención a la salud del niño. Disponible en: <http://www.salud.gob.mx/unidades/cdi/nom/031ssa29.html>.
- O'Shea-Cuevas G, Rizzoli-Córdoba A, Aceves-Villagrán D, Villagrán-Muñoz VM, Carrasco-Mendoza J, Halley-Castillo E, et al. Sistema de Protección Social en Salud para la detección y atención oportuna de problemas en el desarrollo infantil en México. Bol Med Hosp Infant Mex. 2015;72:429–37.
- Rizzoli-Córdoba A, O'Shea-Cuevas G, Aceves-Villagrán D, Mares-Serratos BB, Martell-Valdez L, Vélez Andrade V. Evaluación del Desarrollo Infantil en México. En: Santibañez-Martínez L, Calderón-Martín del Campo D, editores. Los Invisibles —Las

- niñas y los niños de 0 a 6 años—. México D.F: Mexicanos Primero Visión 2030; 2014. p. 117–32. A.C.
15. Rizzoli-Córdoba A. Diseño de una metodología de evaluación del desarrollo infantil en México (Tesis presentada para obtener el grado de Doctor en Ciencias). México D.F: Universidad Nacional Autónoma de México; 2015.
  16. Comisión Nacional de Protección Social en Salud. Manual para la Evaluación de Menores de Cinco Años con Riesgo de Retraso en el Desarrollo. México, D.F.: Secretaría de Salud; 2013. 88p. Disponible en: <http://www.himfg.edu.mx/descargas/documentos/EDI/ManualparaEvaluacionaMenoresde5conRiesgoenelDesarrollo.pdf>.
  17. Newborg J. Battelle Developmental Inventory, Spanish. Second edition Itasca, IL: Riverside Publishing; 2005.
  18. Rydz D, Shevell MI, Majnemer A, Oskoui M. Developmental screening. *J Child Neurol*. 2005;20:4–21.
  19. Bossuyt P, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003;326:41–4.
  20. Whiting PF, Rutjes AWS, Reitsma JB, Bossuyt PM, Kleijmen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003;3:25.
  21. 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.
  22. Bayley N. Bayley Scales of Infant and Toddler Development® (Bayley-III®). Third edition Texas: Psychological Corporation; 2005.
  23. Rizzoli-Córdoba A, Schnaas-Y-Arrieta L, Ortega-Ríosvelasco F, Rodríguez-Ortega E, Villasís-Keever MA, Aceves-Villagrán D, et al. Child Development Evaluation Test analysis by field improves detection of developmental problems in children. *Bol Med Hosp Infant Mex*. 2014;71:154–62.
  24. Rizzoli-Córdoba A, Ortega-Ríosvelasco F, Villasís-Keever MA, Pizarro-Castellanos M, Buenrostro-Márquez G, Aceves-Villagrán D, et al. Reliability for detection of developmental problems using the semaphore from the Child Development Evaluation test: Is a yellow result different from a red result. *Bol Med Hosp Infant Mex*. 2014;71:277–85.
  25. Orcajo-Castelán R, Sidonio-Aguayo B, Alcacio-Mendoza JA, López-Díaz GL. Análisis comparativo de pruebas de tamiz para la detección de problemas en el desarrollo diseñadas y validadas en México. *Bol Med Hosp Infant Mex*. 2015;72:364–75.
  26. Rizzoli-Córdoba A, Campos-Maldonado MC, Vélez-Andrade VH, Delgado-Ginebra I, Baqueiro-Hernández CI, Villasís-Keever MA, et al. Evaluación diagnóstica del nivel de desarrollo en niños identificados con riesgo de retraso en la prueba Evaluación del Desarrollo Infantil. *Bol Med Hosp Infant Mex*. 2015;72:397–408.
  27. Villasís-Keever MA, Rizzoli-Córdoba A, Delgado-Ginebra I, Mares-Serratos BB, Martell-Valdez LG, Sánchez-Velázquez O, et al. Diseño de un modelo de supervisión para la aplicación de la prueba de Evaluación del Desarrollo Infantil en las unidades de atención primaria en México. *Bol Med Hosp Infant Mex*. 2015;72:385–96.
  28. Glascoe FP, Squires J. Issues with the new developmental screening and surveillance policy statement. *Pediatrics*. 2007;119:861–2.
  29. Centro Nacional para la Salud de la Infancia y la Adolescencia. Desarrollo Infantil Temprano. Lineamientos Técnicos. México D.F.: Secretaría de Salud; 2015.