

“En segundo lugar, las herramientas de tamizaje en desarrollo no son diagnósticas y sus resultados deben preceder a una evaluación más intensiva”<sup>9</sup>.

En cuanto a que aún se requieren estudios de reproducibilidad, estamos de acuerdo en que se requiere seguir realizando estudios con el INDIPCD-R con una muestra más amplia y con niños sanos y con diferentes patologías. Por este motivo decidimos dar a conocer estas observaciones a través de su prestigiada revista, para que quien desee aplicar dicho instrumento lo pueda hacer.

Por lo demás, agradecemos sus observaciones que nos impulsan a seguir cuidando la calidad del instrumento.

## Referencias

1. 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.
2. Bolaños C, Mina-Lara MM, Ramos-Ortega G. Indicadores de riesgo del perfil de conductas de desarrollo como instrumento para la detección temprana de retrasos y alteraciones en el desarrollo. *Acta Pediatr Mex.* 2015;36:135–46.
3. Bolaños C. Manual del perfil de conductas de desarrollo revisado. México D.F: Instituto de Terapia Ocupacional; 2005.
4. Bolaños C. Estudios de validación del perfil de conductas de desarrollo (PCD) como instrumento para evaluar el nivel de desarrollo de los niños de 0-4 años. *Bol Med Hosp Infant Mex.* 2003;60:79–89.
5. Bolaños MC, Golombek L, Márquez A, de la Riva M, Sánchez GCE, Gutiérrez GO, et al. Validez de correlación del perfil de conductas de desarrollo con la escala de desarrollo infantil Bayley II. *Acta Pediatr Mex.* 2006;27:190–9.
6. McCarthy AM, Wehby GL, Barron S, Aylward GP, Castilla EE, Javois LC, Goco N, Murray C. Application of neurodevelopmental screening to a sample of South American infants: The Bayley infant neurodevelopmental screener (BINS). *Infant Behav Dev.* 2012;35:280–94. <http://dx.doi.org/10.1016/j.infbeh.2011.12.003>
7. Weiss LG, Oakland T, Aylward G. The Bayley Infant Neurodevelopmental Screener (BINS): Different Test and Different Purpose. En: *Bayley-III clinical use and interpretation.* Burlington: Academic Press/Elsevier; 2010.
8. Meisels SJ, Provence S. Screening and assessment: Guidelines for identifying vulnerable children and their families. Washington, DC: National Center for Clinical Infant Programs; 1989.
9. Shahshahani S, Vameghi R, Azari N, Sajedi F, Kazemnejad A. Validity and reliability determination of Denver developmental screening test-II in 0-6 year-olds in Tehran. *Iran J Pediatr.* 2010;20:313–22.

Cristina Bolaños

*Instituto de Terapia Ocupacional, Ciudad de México, México*

*Correo electrónico: cbolanos@ito-edu.org.mx*

<http://dx.doi.org/10.1016/j.bmhmx.2016.05.004>

## Respuesta a carta al editor: Análisis comparativo de pruebas de tamiz para la detección de problemas en el desarrollo diseñadas y validadas en México



### Answer to letter to the editor: Comparative analysis of child development screening tools designed and validated in Mexico

Recibimos con agrado las observaciones referentes a nuestro artículo, ya que celebramos la discusión académica de alto nivel y esto nos da la oportunidad de ahondar sobre las motivaciones subyacentes de escribirlo, las cuales son las siguientes:

- a) Coadyuvar en la mejora continua de la calidad de la investigación en México.
- b) Promover el uso de listas de cotejo para mejorar el diseño, la metodología y la calidad de reporte de las publicaciones nacionales.
- c) Llamar la atención sobre el impacto que tiene la calidad del manuscrito, en el juicio subjetivo de que los clínicos podemos emitir sobre la utilidad de un método diagnóstico o terapéutico.

Existe una preocupación mundial sobre la reproducibilidad, confiabilidad y validez de la investigación publicada

y los sesgos a los que está expuesta, que ha llevado a la creación de distintas estrategias para reducirlos<sup>1</sup>.

Una de estas estrategias, apoyada por la Organización Mundial de la Salud (OMS) y la Organización Panamericana de la Salud (OPS) por lo eficiente y económico de su implementación, es la utilización de guías de reporte que nos sirvan de orientación para realizar investigaciones y publicaciones de mayor calidad metodológica. Entre ellas encontramos CONSORT para diseños experimentales, STROBE para observacionales, y STARD para validaciones, por mencionar algunos ejemplos. Estas guías son agrupadas para su fácil consulta en la página de la iniciativa *EQUATOR Network (Enhancing the QUALity and Transparency Of health Research)*<sup>2</sup>.

La publicación de nuestro artículo es guiada por este espíritu, y por ello su objetivo es “comparar la calidad del reporte de validación publicado y el riesgo de sesgo entre las pruebas de tamiz desarrolladas y validadas en México”<sup>3</sup>. Por esto, no se emiten opiniones sobre la utilidad comparativa de las pruebas -lo que requeriría invariablemente un diseño experimental-, sino sobre el contenido de los reportes de validación, así como un juicio subjetivo basado en listas de cotejo validadas y reconocidas mundialmente sobre el riesgo de sesgo que los datos presentados tienen.

Brevemente nos permitimos responder a las observaciones emitidas:

1. Se menciona: “En la tabla 1. Descripción general de las pruebas de tamiz comparadas en el estudio, en la

- columna "Aspectos evaluados...". En respuesta a ello agradecemos el señalamiento del error, ya que debe decir en dicha columna "Aspectos variados sin agrupar por áreas del desarrollo".
2. Con respecto al señalamiento del rango de edad en la misma tabla, este fue obtenido de la lectura del artículo de la validación en su página 138<sup>4</sup>: "Los INDIPCD-R quedaron integrados inicialmente en dos formatos: el formato "A" con 44 reactivos (que se utiliza con niños de 0 a 2 años) y el "B" con 41 reactivos (para niños de 2 a 4 años)", lo que resulta confuso con la posterior descripción de seis rangos de edad y de una normalización realizada con una muestra completamente distinta a la que se refiere el artículo. Por lo tanto, en el análisis comparativo, en la columna de rango de edad debería decir "De seis a 48 meses (Evaluados en 6 grupos)". Consideramos de importancia que la autora corrija las contradicciones que originan el malentendido en su artículo de validación.
  3. Con respecto al señalamiento de los errores contenidos en la tabla 2, estos datos se obtienen íntegros del artículo analizado<sup>4</sup>, donde en el resumen se especifica: "Materiales y métodos: estudio transversal comparativo con 145 infantes provenientes de una clínica y dos estancias infantiles (CENDI...". Si bien se menciona en renglones posteriores la normalización de la escala con 347 niños, dicha normalización no es la validación publicada y por tanto no es sujeto del artículo de revisión que los autores escribieron. Posteriormente se observa que la muestra total a la que se le aplicaron las dos pruebas y que por tanto fue analizada para la validación es de 83 niños<sup>4</sup>.
  4. En cuanto a: "La aseveración que establecen en la tabla 4: "Cuestionario para evaluar riesgo de sesgos en estudios de precisión diagnóstica QUADAS, que menciona que el INDIPCD-R presenta un alto riesgo de sesgo, es errónea"<sup>3</sup>, estamos en desacuerdo, ya que lo confuso y contradictorio del escrito analizado<sup>4</sup> refiere una muestra de validación de 145 niños, pero después menciona que los datos de sensibilidad y especificidad se obtuvieron de solo 83 niños; después menciona una normalización de la cual no se publican datos de 347 niños, lo cual fue parte de lo que nos llevó a afirmar que existe un alto riesgo de sesgo en cuanto a la selección de su muestra, basándonos solo en la calidad del reporte de validación. En este sentido, la única afirmación categórica en la que coincidimos que podría haber duda es la que se refiere a la validez del estándar de referencia, pues PCD-R fue validada contra una prueba completamente independiente y creada por otro grupo de personas, ampliamente reconocida como estándar de oro en el medio<sup>5</sup>.
  5. Con respecto al cuestionamiento "el INDIPCD-R no puede compararse contra sí mismo, los expertos no consideraron que la escala de desarrollo Perfil de Conductas de Desarrollo (PCD-R), que fue utilizada como estándar de oro, es una prueba independiente de los INDIPCD-R", consideramos que la aseveración hecha por nosotros: "La INDIPCD-R presenta un alto riesgo de sesgo de su prueba índice y estándar de oro, ya que se valida contra sí misma"<sup>3</sup> es exagerada y debía haber sido fraseada como: "La INDIPCD-R presenta un alto riesgo de sesgo de su prueba índice y estándar de oro, ya que se valida

tomando como estándar de oro una prueba de la cual se deriva y es fruto del trabajo de los mismos autores".

6. En respuesta a la afirmación: "Decidimos utilizar el PCD-R como estándar de oro por ser una prueba que cuenta con estudios de validez y confiabilidad en población abierta mexicana" y al ejemplo mencionado de Bailey como estándar de oro en la validación de BINSS<sup>6</sup>, consideramos que el prestigio de un grupo de investigadores no los exenta de cuestionamientos, y que en ambos casos existe un claro conflicto de interés, que si bien no es malo *per se*, debió haber sido reportado.
7. Por último, se afirma que nuestro escrito utiliza de forma indistinta las palabras "diagnóstico y tamizaje". Esta aseveración no tiene fundamento, pues el título especifica que en el artículo sólo se comparan pruebas de tamizaje y no de diagnóstico, incluyéndose en el comparativo solo las primeras. Los autores tenemos muy clara la diferencia entre los dos conceptos.

El uso de la frase: "listas de cotejo de calidad de reporte de estudios de validación de precisión diagnóstica" está completamente justificado debido a que si bien los instrumentos de tamizaje y diagnóstico son distintos en su objetivo, constructo, creación y uso clínico, no son diferentes en su metodología de validación. Así, reportan exactamente las mismas áreas de concordancia entre la prueba evaluada y el estándar de oro que los autores hayan elegido (sensibilidad, especificidad, valor predictivo negativo, valor predictivo positivo) y están expuestos a los mismos errores y limitaciones metodológicas que pueden inducir sesgos.

Agradecemos la lectura cuidadosa de nuestro artículo y, en un espíritu de colaboración donde todos pugnamos por la mejoría de la calidad de la investigación en México, esperamos que esta carta impulse a la autora a realizar las correcciones en la redacción del artículo usado en la comparación y que inducen al malentendido, con la finalidad de que el INDIPCD-R alcance el reconocimiento, el impacto y la utilización que justamente se merece.

## Referencias

1. Higgins J, Green S. Introduction to sources of bias in clinical trials. Cochrane Handbook for Systematic Reviews of Interventions. En: The Cochrane initiative. Chichester, UK: John Wiley & Sons, Ltd; 2011. Accesado el 18 de abril del 2016. Disponible en: [http://handbook.cochrane.org/chapter\\_8/8.4.introduction.to\\_sources\\_of\\_bias\\_in\\_clinical\\_trials.htm](http://handbook.cochrane.org/chapter_8/8.4.introduction.to_sources_of_bias_in_clinical_trials.htm)
2. Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. BMC medicine. 2010;8:1-6, <http://dx.doi.org/10.1186/1741-7015-8-24>
3. 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.
4. Bolaños C, Mina-Lara MM, Ramos-Ortega G. Indicadores de riesgo del perfil de conductas de desarrollo como instrumento para la detección temprana de retrasos y alteraciones en el desarrollo. Acta Pediatr Mex. 2015;36:135-46.
5. Bolaños MC, Golombek L, Márquez A, de la Riva M, Sánchez GCE, Gutiérrez GO, et al. Validez de correlación del perfil de conductas

de desarrollo con la escala de desarrollo infantil Bayley II. *Acta Pediatr Mex.* 2006;27:190–9.

6. McCarthy AM, Wehby GL, Barron S, Aylward GP, Castilla EE, Javois LC, et al. Application of neurodevelopmental screening to a sample of South American infants: The Bayley infant neurodevelopmental screener (BINS). *Infant Behav Dev.* 2012;35:280–94, <http://dx.doi.org/10.1016/j.infbeh.2011.12.003>

Rodrigo Orcajo-Castelán<sup>a,c,d,\*</sup> y Beatriz Sidonio-Aguayo<sup>b</sup>

<sup>a</sup> *UMCAE Hospital Regional de Psiquiatría Unidad Morelos, Ciudad de México, México*

<sup>b</sup> *Departamento de Rehabilitación Pediátrica, Hospital Infantil de Mexico "Federico Gómez", Ciudad de México, México*

<sup>c</sup> *Principles and Practice of Clinical Research Course, Harvard T.H. Chan School of Public Health, Cambridge, Massachusetts, EUA*

<sup>d</sup> *Centre for Global Mental Health, King's College London/London School Of Hygiene and Tropical Medicine, Londres, Reino Unido*

\* Autor para correspondencia.

*Correo electrónico:*

[rodrigo\\_aquilino.orcajo\\_castelan@kcl.ac.uk](mailto:rodrigo_aquilino.orcajo_castelan@kcl.ac.uk)  
(R. Orcajo-Castelán).

<http://dx.doi.org/10.1016/j.bmhmx.2016.05.005>