ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Araujo, María Caridad; Dormal, Marta; Rubio-Codina, Marta

Working Paper

La calidad de los programas de trabajo con familias y resultados en desarrollo infantil: El caso de Cuna Más en Perú

IDB Working Paper Series, No. IDB-WP-951

Provided in Cooperation with:
Inter-American Development Bank (IDB), Washington, DC

Suggested Citation: Araujo, María Caridad; Dormal, Marta; Rubio-Codina, Marta (2019) : La calidad de los programas de trabajo con familias y resultados en desarrollo infantil: El caso de Cuna Más en Perú, IDB Working Paper Series, No. IDB-WP-951, Inter-American Development Bank (IDB), Washington, DC,
https://doi.org/10.18235/0001366

This Version is available at:
https://hdl.handle.net/10419/234643

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

https://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DOCUMENTO DE TRABAJO DEL BID Nº IDB-WP-951

La calidad de los programas de trabajo con familias y resultados en desarrollo infantil:

El caso de Cuna Más en Perú

M. Caridad Araujo
Marta Dormal
Marta Rubio-Codina

Banco Interamericano de Desarrollo
División de Protección Social y Salud

Abril 2019

Banco Interamericano de Desarrollo

La calidad de los programas de trabajo con familias y resultados en desarrollo infantil:

El caso de Cuna Más en Perú

M. Caridad Araujo
Marta Dormal
Marta Rubio-Codina

Abril 2019

Catalogación en la fuente proporcionada por la
Biblioteca Felipe Herrera del
Banco Interamericano de Desarrollo
Araujo, Maria Caridad.
La calidad de los programas de trabajo con familias y resultados en desarrollo infantil: el caso de Cuna Más en Perú / M. Caridad Araujo, Marta Dormal, Marta Rubio-Codina.
p. cm. — (Documento de trabajo del BID ; 951)
Incluye referencias bibliográficas.
1. Home-based family services-Peru. 2. Parenting-Government policy-Peru. 3. Child care services-Government policy-Peru. 4. Child development-Government policy-Peru. 5. Child welfare-Government policy-Peru. I. Dormal, Marta. II. Rubio-Codina, Marta. III. Banco Interamericano de Desarrollo. División de Protección Social y Salud. IV. Título. V. Serie.
IDB-WP-951

http://www.iadb.org

Copyright © 2019 Banco Interamericano de Desarrollo. Esta obra se encuentra sujeta a una licencia Creative Commons IGO 3.0 Reconocimiento-NoComercial-SinObrasDerivadas (CC-IGO 3.0 BY-NC-ND) (http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode) y puede ser reproducida para cualquier uso no-comercial otorgando el reconocimiento respectivo al BID. No se permiten obras derivadas.

Cualquier disputa relacionada con el uso de las obras del BID que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la CNUDMI (UNCITRAL). El uso del nombre del BID para cualquier fin distinto al reconocimiento respectivo y el uso del logotipo del BID, no están autorizados por esta licencia CC-IGO y requieren de un acuerdo de licencia adicional.

Después de un proceso de revisión por pares, y con el consentimiento previo y por escrito del BID, una versión revisada de esta obra podrá reproducirse en cualquier revista académica, incluyendo aquellas referenciadas por la Asociación Americana de Economía a través de EconLit, siempre y cuando se otorgue el reconocimiento respectivo al BID, y el autor o autores no obtengan ingresos de la publicación. Por lo tanto, la restricción a obtener ingresos de dicha publicación sólo se extenderá al autor o autores de la publicación. Con respecto a dicha restricción, en caso de cualquier inconsistencia entre la licencia Creative Commons IGO 3.0 Reconocimiento-NoComercial-SinObrasDerivadas y estas declaraciones, prevalecerán estas últimas.

Note que el enlace URL incluye términos y condiciones adicionales de esta licencia.

Las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del Banco Interamericano de Desarrollo, de su Directorio Ejecutivo ni de los países que representa.

scl-sph@iadb.org

https://www.iadb.org/es/proteccionsocial

La calidad de los programas de trabajo con familias y resultados en desarrollo infantil: el caso de *Cuna Más* en Perú

M. Caridad Araujo[i], Marta Dormal[i], Marta Rubio-Codina[i]

Resumen

Este estudio analiza la asociación entra la calidad de los programas de trabajo con familias y sus resultados sobre el desarrollo infantil. Se concentra en programas de acompañamiento familiar que operan a través de visitas domiciliarias, en los cuales un para profesional trabaja con el cuidador principal del niño para mejorar sus prácticas parentales, habilidades y conocimiento sobre desarrollo infantil. Específicamente, analiza el caso del *Programa Nacional Cuna Más* en Perú. La calidad de las visitas domiciliarias se mide a través de la Escala de Calificación de las Visitas Domiciliarias (*Home Visit Rating Scale, version A+2*, HOVRS-A+2) y los resultados en desarrollo infantil se miden a través de los Cuestionarios de Edades y Etapas (*Ages and Stages Questionnaires*, ASQ-3) y de las Escalas de Bayley (*Bayley Scales of Infant and Toddler Development*, Bayley-III). El análisis empírico es de naturaleza descriptiva. Los resultados muestran que los puntajes de la calidad de las visitas domiciliarias están correlacionados significativamente con medidas de desarrollo infantil, pero esta asociación no es consistente entre todas las medidas.

Códigos JEL: I12, J24, O15

Palabras clave: Visitas domiciliarias, calidad, desarrollo infantil, HOVRS-A+2, ASQ-3, Bayley-III

[i.] División de Salud y Protección Social, Banco Interamericano de Desarrollo, Washington DC, Estados Unidos.

1. Introducción

Las intervenciones de trabajo con familias y, en particular, aquellas que operan a través de visitas domiciliaras, han recibido atención de parte de los formuladores de política pública en América Latina y el Caribe (ALC) y en otras regiones del mundo en años recientes. Los programas de trabajo con familias son intervenciones que buscan cambiar los comportamientos y mejorar las prácticas parentales, las habilidades y el conocimiento sobre el desarrollo infantil a través del entrenamiento, el apoyo y la educación que brinda un agente para profesional (en adelante, nos referiremos a éste como 'el/la facilitador/a'), que trabaja directamente con la persona encargada del cuidado de los niños. Tres modalidades de atención son las más comunes: sesiones grupales, visitas domiciliarias, o trabajo durante los controles médicos (Berlinski y Schady 2015). La evidencia disponible, principalmente de ensayos de eficacia y de pilotos de pequeña escala rigurosamente evaluados, sugiere que estos programas pueden lograr impactos sostenidos en el desarrollo infantil de gran magnitud (Britto et al. 2016). Países como Perú, Brasil, y otros se han embarcado en el desafío de expandir la cobertura de la modalidad de visitas domiciliarias (en adelante, los 'programas de visitas domiciliarias') para atender a las familias en desventaja con niños en edades tempranas.

Como para cualquier servicio social, llevarlo a escala conlleva importantes desafíos de implementación. En las visitas domiciliarias, dado que es un área relativamente menos investigada, poco se conoce de los atributos de la calidad del servicio que es más crítico preservar durante el proceso de expansión. ¿Es suficiente con garantizar una dosis y duración mínima de las visitas? ¿Qué competencias deben manejar las facilitadoras para lograr, con éxito, cambiar los comportamientos de los padres y promover interacciones de alta calidad y estimulación psicosocial adecuada en el hogar? ¿Cuáles son los componentes clave de una visita domiciliaria exitosa que asegure impactos sostenibles en el desarrollo de los niños implementada a gran escala?

Paulsell et al. (2010) propone tres dimensiones de la calidad de las visitas domiciliarias: dosis, contenido y relaciones interpersonales. La dosis y el contenido pueden ser entendidos como elementos de la calidad *estructural*, mientras que la relaciones se refieren a lo que comúnmente se conoce como la calidad de *procesos*. La calidad de procesos se enfoca en aspectos dinámicos de la visita y en cómo se entrega el contenido de la intervención, en particular con respecto a las relaciones interpersonales y la cantidad y naturaleza de las interacciones que ocurren entre la facilitadora, el cuidador y el niño. La calidad estructural, por otra parte, se refiere a la presencia o ausencia de los recursos que facilitan estas interacciones durante la visita (por ejemplo, materiales, recursos, o experiencia, educación y salario de la facilitadora). Estas dos dimensiones de la calidad están altamente interrelacionadas. Por ejemplo, es más probable que la facilitadora tenga interacciones de alta calidad con el niño y el cuidador cuando ella cuenta con materiales de aprendizaje apropiados para apoyarse en su trabajo. Sin embargo, la medición de estos aspectos de la calidad difiere considerablemente: la calidad de procesos requiere la observación de expertos y la valoración e interpretación de las interacciones observadas. Por lo tanto, la medición

de la calidad de procesos es más complicada y demanda más tiempo que la de la calidad estructural. Esta última es más sencilla de cuantificar pues puede ser evaluada a través de una variedad de listas de chequeo.

Varios instrumentos han sido utilizados para medir la calidad de las visitas domiciliarias. En su mayoría, estos instrumentos hacen énfasis en medir la calidad general de las visitas domiciliarias y la de la relación entre la facilitadora y el cuidador. Los instrumentos empleados en estudios publicados son:

- Las series de Escalas de Calificación de las Visitas Domiciliarias (*Home Visit Rating Scales series,* HOVRS, HOVRS-A, HOVRS-A+ y HOVRS-A+2; Roggman et al. 2006, Roggman et al. 2014). Consiste en sietes escalas, organizadas en dos dominios, el primero enfocado en las estrategias que usa el facilitador para efectivamente realizar la visita y la segunda, que mide el nivel de involucramiento del cuidador y el niño y la naturaleza de sus interacciones.
- El Formulario de Contenido de las Visitas Domiciliarias y Formulario de Características (*Home Visit Content and Characteristics Form,* HVCCF; Boller et al. 2009). Se usa como complemento al HOVRS, pues documenta el rango de características estructurales de las visitas domiciliarias, la dosis y el contenido.
- El Sistema de Calificación COACH (*COACH rating system,* Dishion et al. 2010) es un sistema de observación diseñado para cuantificar la fidelidad en la entrega del contenido, con foco en cinco dimensiones: precisión conceptual, observación y receptividad a las necesidades de la familia, estructura de la visita, enseñanza cuidadosa y apropiada, y motivación.

Un elemento común en la administración y calificación de todos estos instrumentos es que requieren observadores rigurosamente capacitados y con amplio conocimiento del protocolo de calificación de cada uno de ellos.

Schodt et al. (2015) realiza una revisión de la literatura sobre la medición de la calidad de las visitas domiciliarias y de los pocos estudios que exploran la asociación entre la calidad de las visitas domiciliaras y los resultados sobre el desarrollo infantil. Estos estudios, descritos brevemente a continuación, son descriptivos: presentan correlaciones entre la calidad de las visitas domiciliarias, la calidad del ambiente en el hogar y el desarrollo del niño. No cuentan con un diseño que permita establecer una relación causal entre las visitas de alta calidad y el desarrollo infantil.

Vogel et al. (2015) es parte del estudio *Baby FACES*, una evaluación longitudinal descriptiva del programa *Early Head Start* en los Estados Unidos. Los autores caracterizaron las visitas domiciliarias utilizando el HOVRS-A y el HVCCF y exploraron su asociación con los resultados de desarrollo infantil a la edad de 2 años. Ellos documentaron una asociación positiva y significativa entre el segundo dominio de HOVRS-A, Efectividad del facilitador, y entre la escala de Involucramiento del niño con el puntaje de regulación emocional medida con la Escala de Calificación del Comportamiento del Bayley (*Bayley Behavior Rating Scale*). El orden de magnitud de estas asociaciones está entre 0,13 y 0,14. Consistentemente, los autores también reportaron una asociación negativa de -0,17 entre la calidad de la visita domiciliaria y

los problemas de comportamiento reportados por las propias facilitadoras empleando la Evaluación Socioemocional Corta para Bebés y Párvulos (*Brief Infant Toddler Social Emotional Assessment*). No obstante, ellos documentaron correlaciones negativas y significativas entre la calidad de la visita domiciliaria y la motricidad gruesa, motricidad fina y resolución de problemas, medidas con el ASQ-3. El tamaño de estas asociaciones está entre -0,14 y -0,17. No se encontraron correlaciones significativas entre las medidas de la calidad de las visitas domiciliarias y la calidad del ambiente en el hogar medida por la escala HOME.

Smith et al. (2013) analizó los datos del programa Chequeo Familiar (*Family Check-Up*) en Estados Unidos. Este modelo brinda herramientas a los padres para manejar los problemas de comportamiento de los niños. Los autores usaron el sistema de clasificación COACH para caracterizar la fidelidad en la entrega del contenido y explorar el efecto indirecto de la calidad de las visitas domiciliarias sobre los resultados de los niños. Específicamente, encontraron que las facilitadoras que eran más competentes en adherirse al contenido del programa tenían mayor capacidad de involucrar a los padres en la sesión y esto se tradujo en una mejora pequeña pero significativa en las habilidades parentales del uso de las técnicas de manejo positivo del comportamiento promovidas por el programa (0,06 de una desviación estándar, DE). Un año después, los niños también mostraron menos problemas de comportamiento, medidos por la escala de externalización de la Lista de Chequeo de Comportamiento Infantil (*Child Behavior Checklist*; -0,24 SD).

Roggman et al. (2016) examinaron la calidad de las visitas domiciliarias empleando datos de dos operadores del programa *Head Start* en Estados Unidos. Los autores mostraron una asociación significativa entre la calidad de las visitas domiciliarias medida por el HOVRS y la calidad del ambiente del hogar medida por el HOME. Las correlaciones parciales entre los dominios del HOVRS y el puntaje total del HOME, controlando por el operador del programa, se encontraban en un rango de 0,31 a 0,34. Utilizando un análisis con regresiones, los autores exploraron la asociación entre la calidad de las visitas domiciliarias y el vocabulario receptivo del niño medido con el PPVT (conocido como TVIP en su versión para poblaciones de habla hispana) a los 36 meses de edad y encontraron una asociación positiva y significativa (controlando por el operador del servicio). Este coeficiente perdía su significancia una vez que se introducía la calidad del ambiente del hogar en la regresión, lo cual sugiere que el impacto de HOVRS en los resultados de los niños podía estar operando indirectamente a través de las prácticas parentales.

En este estudio, exploramos la asociación entre la calidad de las visitas domiciliarias y los resultados de desarrollo infantil en el contexto de un programa implementado en un país de ingreso medio. Nos concentramos en el servicio de visitas domiciliarias conocido como *Servicio de Acompañamiento a Familias* (Cuna Más, en adelante) que entrega el *Programa Nacional Cuna Más* en Perú. Medimos la calidad de las visitas utilizando la Escala de Calificación de las Visitas Domiciliarias (*Home Visit Rating Scale*), versión A+2 (HOVRS-A+2). Los resultados de desarrollo infantil se miden a través de dos instrumentos: los Cuestionarios de Edades y Etapas (*Ages and Stages Questionnaires*, tercera versión, ASQ-3; Squires et al. 2009) y las Escalas de Bayley

(*Bayley Scales of Infant and Toddler Development*, tercera versión, Bayley-III; Bayley 2006). Este estudio complementa la evaluación de impacto de Cuna Más y provee una mirada profunda a la calidad de la implementación del servicio en una submuestra de familias y de facilitadoras. La evaluación de impacto mostró impactos robustos y significativos sobre el desarrollo infantil en las áreas de desarrollo cognitivo y de comunicación, de una magnitud de entre 0,14 y 0,17 DE (Araujo et al. 2018). La teoría de cambio detrás de los programas de visitas domiciliarias consiste en que las facilitadoras apoyan a los cuidadores en el desarrollo de prácticas parentales que crean un ambiente de aprendizaje favorable para el niño en el hogar, que se traduce en mejores resultados de desarrollo infantil. En línea con esta teoría de cambio, la evaluación de Cuna Más sugiere que los impactos en desarrollo infantil parecen estar mediados por comportamientos parentales como una mayor frecuencia del juego y las actividades de aprendizaje en el hogar, más disciplina positiva y mayor matrícula en el prescolar (MEF-BID 2016). La hipótesis de este estudio es que la *calidad* en la entrega del servicio y, en particular, la calidad de procesos es importante para mejorar las prácticas parentales descritas antes y los resultados sobre el desarrollo infantil. En otras palabras, esperaríamos que los programas de visitas domiciliarias de alta calidad, con mejores interacciones entre facilitadoras, cuidadores y niños, estén más fuertemente asociadas con las mejoras en los cambios de comportamiento de los padres y, consecuentemente, con los resultados de desarrollo infantil. El análisis empírico que presentamos es de naturaleza descriptiva y se concentra en las asociaciones entre variables de desarrollo infantil y las medidas de calidad.

El resto de este estudio se organiza de la siguiente manera. La sección 2 describe brevemente el programa Cuna Más y la sección 3 el diseño del estudio y los datos. La sección 4 presenta un resumen de las estadísticas descriptivas de la muestra de análisis y la sección 5 reporta las asociaciones entre la calidad de las visitas domiciliarias y el desarrollo infantil. Finalmente, la sección 6 concluye.

2. El programa de visitas domiciliarias de Cuna Más

Cuna Más es el proveedor nacional más grande de servicios de desarrollo infantil en Perú. En 2015, alcanzó una cobertura de 93 mil niños de 0 a 36 meses de edad en áreas rurales con sus servicios de visitas domiciliarias. Este es un programa público, completamente subsidiado, que provee visitas domiciliarias de una hora a la semana realizadas por un miembro de la comunidad que ha sido capacitado por el Programa. Aunque el Programa requiere que las facilitadoras cumplan ciertos requisitos mínimos - saber leer y escribir, tener 21 años o más, hablar la lengua local, tener experiencia trabajando con niños y contar con cierto nivel de reconocimiento en la comunidad-, estos pueden ser difíciles de encontrar dado el contexto de pobreza y la ubicación rural y dispersa de las comunidades en las que el Programa opera. Los miembros de los Comités de Gestión nominan una lista de candidatos al Programa y los supervisores locales (Acompañantes Técnicos) seleccionan a las facilitadoras de entre la lista de aplicantes. Cada facilitadora tiene a su cargo a diez familias. Por su parte, por cada 10 facilitadoras hay una supervisora que se espera trabaje de manera cercana ayudándolas a planificar su trabajo y proporcionándoles capacitación continua no

estructurada y mentoría. En la práctica, sin embargo, a los supervisores se les encarga con frecuencia otras tareas administrativas que consumen su tiempo.

Las facilitadoras siguen un currículo estructurado que promueve más y mejores interacciones de calidad entre adultos y niños y aprendizaje a través del juego. Cuando este estudio se realizó, cada visita se organizaba alrededor de tres momentos. El primer momento, *Vida en Familia,* busca ayudar a los cuidadores a identificar oportunidades de aprendizaje y juego en sus rutinas diarias. A menudo se entregan mensajes más generales relacionados a la salud, nutrición y saneamiento. El segundo momento, *Jugando Aprendo,* utiliza actividades y materiales de juego intencionales (por ejemplo, juguetes o rompecabezas de goma que el Programa provee o que son fabricados en casa con el cuidador) para promover el desarrollo del lenguaje, la motricidad fina y las habilidades cognitivas. El tercer momento, *Cuéntame un Cuento,* promueve la adquisición de vocabulario y el desarrollo del lenguaje utilizando libros ilustrados que provee el Programa y ayuda a los cuidadores a aprender a usar estos materiales con bebés y niños muy pequeños. Además de estos momentos, al comienzo de cada visita, la facilitadora y la cuidadora revisan las actividades de la visita previa[1]. De forma similar, al final de cada visita, recapitulan e identifican las actividades específicas que el cuidador y el niño practicarán durante la siguiente semana. Con este fin, todos los materiales que fueron utilizados en la visita permanecen en el hogar y se recogen en la siguiente visita. La facilitadora alterna estos materiales entre las familias con las que trabaja como si se tratara de una 'biblioteca de juguetes'.

El programa Cuna Más está focalizado hacia los más pobres del país, en particular hacia los de los distritos[2] con: (i) una tasa de pobreza de por lo menos 50 por ciento; (ii) una tasa de desnutrición crónica de por lo menos 30 por ciento; (iii) beneficiarios del programa de transferencias monetarias condicionadas *Juntos;* y (iv) donde más del 50 por ciento de la población vive en centros poblados rurales. En estos distritos, Cuna Más interviene exclusivamente en centros poblados rurales que son definidos como aquellos con menos de 2.000 familias o 400 viviendas. Todos los niños de estas comunidades desde su nacimiento hasta los 24 meses son elegibles para inscribirse al Programa y recibir visitas hasta cumplir los 36 meses de edad, que es cuando salen del Programa.

Una evaluación experimental de Cuna Más fue realizada aprovechando su expansión gradual para medir el impacto del Programa en el desarrollo de los niños. La muestra de la evaluación de impacto incluyó 5.859 niños de 1 a 24 meses de edad en línea de base en 360 centros poblados de 180 distritos elegibles para recibir Cuna Más. Estos distritos están repartidos en 12 de los 24 departamentos del país, ubicados en su mayoría en las regiones andina y amazónica. Ciento veinte de estos distritos fueron asignados aleatoriamente al grupo de tratamiento para recibir visitas y 60 distritos fueron asignados al grupo de control, por lo que no recibieron ninguna intervención. Los datos de línea de base fueron recogidos durante la segunda mitad de 2013 y los

[1] Desde el 2017, el Programa empezó a revisar el modelo de intervención y los contenidos para enfatizar los componentes de nutrición sobre los componentes de desarrollo.
[2] Los distritos son el equivalente de las municipalidades en Perú.

datos de seguimiento se recolectaron aproximadamente 24 meses después. DGSE-MIDIS et al. (2015) presenta más detalles sobre la muestra de la evaluación y su diseño y Araujo et al. (2018) reporta los resultados de la misma.

3. Diseño del estudio, muestra y datos

3.1. Muestra del análisis de calidad

La implementación de Cuna Más empezó en agosto de 2013 en todos los distritos de tratamiento en la muestra de la evaluación de impacto. Como se observa en la Figura 1 (Panel A), la recolección de datos de seguimiento se realizó entre junio y diciembre de 2015. En ella, se evaluaron los resultados de desarrollo de los niños utilizando los instrumentos ASQ-3 y Bayley-III. En paralelo, entre agosto y octubre de 2015, se recogieron datos sobre la calidad de las visitas domiciliarias para una submuestra de esos niños. El Programa había operado aproximadamente por dos años cuando se realizó la medición de calidad.

La submuestra en la cual se midió la calidad de las visitas fue seleccionada de la muestra de seguimiento de la evaluación de impacto como se indica a continuación. Primero, identificamos a los niños elegibles que eran aquellos que todavía recibían visitas de Cuna Más al momento de recolección de los datos, es decir, aquellos niños en el grupo de tratamiento y menores de 36 meses de edad, que todavía no habían egresado del Programa. Luego, identificamos a su facilitadora, así como a tres o cuatro niños adicionales visitados por ella, pero que no necesariamente eran parte de la muestra de la evaluación. La submuestra con datos sobre la calidad de las visitas incluyó información de 554 visitas domiciliarias que se hicieron para la misma cantidad de niños, de los cuales 229 eran niños focales (es decir, niños que también eran parte de la muestra de la evaluación). Los participantes vivían en 103 comunidades rurales y fueron visitados por 176 facilitadoras.

3.2. Los datos: medidas, adaptación y capacitación

3.2.1 Medida de la calidad de la visita: HOVRS-A+2

Para cada niño en la muestra, filmamos una visita domiciliaria y calificamos varios aspectos de su calidad y contenido utilizando las versiones adecuadamente traducidas y adaptadas del instrumento Escala de Calificación de las Visitas Domiciliarias versión A+2 *(Home Visit Rating Scale, version A+2*, HOVRS-A+2; Roggman et al. 2014).

Desarrollado a partir del HOVRS original (Roggman et al. 2006), el HOVRS-A+2 es el más reciente en la serie de instrumentos diseñados específicamente para evaluar la calidad de procesos en las visitas domiciliarias. Consta de siete escalas organizadas en dos dominios. El primer dominio, *Prácticas de la Facilitadora*, contiene cuatro escalas que se concentran en las estrategias utilizadas por la facilitadora para realizar efectivamente la visita, es decir, el grado en que la facilitadora está preparada para hacerla, el grado en que observa y responde al cuidador y al niño y demuestra una actitud positiva y cálida hacia ellos, entre otros aspectos. El segundo dominio se refiere al *Involucramiento Familiar*, e incluye tres escalas que buscan medir la naturaleza de las interacciones entre el cuidador y el niño, así como su interés y nivel de participación

durante la visita[3]. Este instrumento es administrado por observación y fue inicialmente desarrollado para ser calificado sobre una filmación de la visita, pero ahora también se administra durante la observación directa de la visita (comunicación personal con L. Roggman en noviembre 16, 2017).

La familia de escalas de HOVRS ha sido ampliamente utilizada en Estados Unidos como una medida de la calidad de las visitas domiciliarias, tanto en el contexto de investigaciones, como en el apoyo a esfuerzos programáticos. Se ha probado que es adaptable a programas con diversos objetivos, a diferentes formatos de visitas y a varias culturas dentro de Estados Unidos, tales como familias latinas hispanohablantes, familias caucásicas rurales y familias afroamericanas.[4]

Adaptación del HOVRS-A+2 y capacitación de la supervisora de su administración[5]. El HOVRS-A+2 fue traducido al español y retraducido al inglés para asegurar su equivalencia lingüística y funcional. La traducción inversa fue revisada por los autores del instrumento, quienes sugirieron modificaciones menores. Los autores capacitaron a la persona que supervisó la codificación de los videos y que fue quien capacitó al equipo de codificadores. La capacitación de la supervisora de la administración del HOVRS-A+2 duró dos días, durante los cuales ella y dos de las investigadoras continuaron calificando videos de visitas domiciliarias hasta obtener una confiabilidad satisfactoria entre observadores.

Capacitación y filmación de los videos. Las visitas domiciliarias fueron filmadas entre agosto y octubre de 2015 por un equipo de 14 camarógrafos capacitados por un profesional durante una semana, que incluyó prácticas de filmación. De las 24 personas que participaron en la capacitación, la mitad tenía experiencia previa en filmación. El equipo de campo fue seleccionado con base en su desempeño durante la capacitación y las prácticas. El contenido de las capacitaciones incluyó especificaciones sobre el uso de una videocámara, la selección del ángulo óptimo para filmar, y otras estrategias para obtener videos de alta calidad considerando las limitaciones del entorno (por ejemplo, filmación en interiores, ruidos de fondo, cámara estática, etc.).

La capacitación también abarcó una sesión de presentación del Programa, la estructura de las visitas domiciliarias y la revisión de un protocolo de filmación diseñado para minimizar la perturbación que esta pudiera crear a la familia y a la visita misma. Los supervisores fueron capacitados para discutir con cada familia y facilitadora el objetivo del ejercicio, previo a comenzar la recolección de la información. No hubo reportes de perturbación o conflictos durante la recolección de los datos.

---

[3] Las escalas del dominio *Prácticas de la Facilitadora* son: (i) Receptividad de la facilitadora hacia la familia (6 ítems); (ii) Relación facilitadora-familia (8 ítems); (iii) Facilitación de la interacción entre cuidador y niño (6 ítems); y (iv) No intrusión/Colaboración de la facilitadora con la familia (5 ítems). Las escalas del dominio *Involucramiento Familiar* son: (i) Interacción entre cuidador y niño durante la visita domiciliaria (7 ítems); (ii) Involucramiento del cuidador durante la visita domiciliaria (6 ítems); y (iii) Involucramiento del niño durante la visita domiciliaria (4 ítems).

[4] Para más información ver: http://healthandwelfare.idaho.gov/Portals/0/Children/HomeVisiting/MIECHV%20Assessment%20Guide%204-9-2013.pdf

[5] La supervisora de la administración del HOVRS-A+2 es la persona que capacitó a los codificadores del HOVRS y los supervisó durante todo el proceso de codificación.

Capacitación de los codificadores. La codificación de los videos se realizó entre noviembre de 2015 y mayo de 2016 por un equipo basado en Quito, Ecuador, con experiencia en codificación de videos para la medición de la calidad de procesos en entornos prescolares y escolares mediante instrumentos de observación. Se capacitó a nueve codificadores y una supervisora (psicóloga) durante seis días. El entrenamiento incluyó entre seis y ocho codificaciones de práctica, hasta obtener una confiabilidad aceptable de forma consistente. Los cinco codificadores con la mayor confiabilidad con respecto al capacitador -todos con estudios de educación superior en psicología, educación o áreas relacionadas- fueron seleccionados para codificar los videos del estudio. La supervisora, una psicóloga, actuó como codificadora de apoyo.

Dos puntajes del HOVRS-A+2 fueron considerados como no confiables si es que (i) el puntaje de una o más escalas difería en más de un punto (criterio de confiabilidad a nivel de escala); o (ii) dentro de una escala dada, el puntaje de dos o más ítems difería en dos o más puntos (criterio de confiabilidad a nivel de ítem)[6]. La confiabilidad entre los codificadores seleccionados y el capacitador fue aceptable: durante la capacitación, la confiabilidad a nivel de ítem osciló entre el 80 y el 100 por ciento en promedio por codificador y se ubicó en más del 90 por ciento para tres de los codificadores. La confiabilidad a nivel de escala osciló entre 61 y 86 por ciento en promedio por codificador, con valores por encima de 80 por ciento para tres codificadores.

Proceso de codificación. Cada video se codificó dos veces, de manera independiente, por dos codificadores seleccionados aleatoriamente en el equipo. Si los puntajes de ambos codificadores no satisfacían los criterios de confiablidad a nivel de ítem o a nivel de escala, se asignaba el video a un tercer codificador -diferente de los dos primeros y también seleccionado aleatoriamente- para una tercera codificación.

Para mantener la confiabilidad durante todo el periodo de codificación, realizamos un ejercicio de calibración mediante el cual el equipo codificaba diariamente un fragmento de video de 20 minutos. Se comparaban los puntajes de los codificadores con los establecidos por la supervisora de la administración de HOVRS-A+2 y se discutían las discrepancias. A la mitad del entrenamiento, modificamos el ejercicio de calibración. En lugar del ejercicio diario, programamos la codificación grupal de una visita entera dos veces por semana pues se observó que codificar fragmentos de 20 minutos generaba menos discrepancias que codificar visitas de una hora[7]. Esta estrategia redujo considerablemente el número de videos que debían ser codificados una tercera vez. El porcentaje de acuerdo entre los codificadores y el capacitador durante este ejercicio fue muy alto, osciló entre el 93 y 98 por ciento.

De los 554 videos en la muestra de calidad, un total de 214 videos (39 por ciento) fueron codificados una tercera vez: 142 videos (66 por ciento) no cumplieron con el criterio de confiabilidad a nivel de escala, 17 (8 por ciento) no cumplieron con el criterio de confiabilidad a nivel de ítem, y 55 videos (26 por ciento) no cumplieron con ninguno de los dos criterios. Después de la tercera codificación, la confiabilidad promedio

[6] Los ítems en el HOVRS-A+2 se puntúan en una escala de 4 puntos, con valores posibles 1, 3, 5 o 7. De esta manera, 1 punto de diferencia al nivel de ítem corresponde, por ejemplo, a la diferencia entre 1 y 3, 3 y 5, o 5 y 7.
[7] Es más probable que ocurran interacciones de naturaleza diferente durante una hora que durante 20 minutos.

incrementó del 65 al 82 por ciento a nivel de escala y del 87 a 94 por ciento a nivel de ítem[8].

El promedio de duración de un video era de cerca de una hora. El equipo codificó no más de tres videos por día en sesiones de cuatro horas, que incluían el ejercicio de calibración. Las interacciones verbales en quechua fueron traducidas al español y subtituladas por un equipo de traductores peruanos antes de la codificación. Esto ocurrió en el 13 por ciento (74) de los videos.

Calificación. Cada *ítem* describe una práctica específica de la facilitadora y su calidad es calificada en una escala de 4 puntos, con puntos de anclaje de 1 'necesita entrenamiento y apoyo', 3 'adecuada', 5 'buena' y 7 'excelente'. El codificador luego asigna a la *escala* un puntaje de 1 a 7 basado en el patrón general de puntajes observado para los ítems[9]. Con el fin de facilitar la interpretación, seguimos las categorías sugeridas en el manual de puntuación: 1 'necesita entrenamiento o apoyo' (si 1 ≤ puntaje promedio de la escala <3), 2 'adecuada' (si 3 ≤ puntaje promedio de la escala <5), 3 'buena' (si 5 ≤ puntaje promedio de la escala <7) y 4 'excelente' (si puntaje promedio de la escala = 7). Los puntajes a nivel de dominio se construyeron como el promedio de los puntajes de las escalas, y el puntaje total como el promedio de los dos dominios.

En este estudio trabajamos con los puntajes de dominio y el puntaje total. Para cada visita, tenemos los puntajes de los dos codificadores con el mayor nivel de confiabilidad a nivel de escala y de ítem. Aunque los codificadores fueron seleccionados aleatoriamente y los criterios estrictos de confiabilidad se mantuvieron durante el proceso, pueden persistir idiosincrasias entre codificadores en la codificación. Por lo tanto, corrimos una regresión con variables dicotómicas de codificador para eliminar el efecto fijo del codificador. Luego promediamos los residuos de estas regresiones para los dos codificadores más confiables y normalizamos el puntaje resultante para que tuviera media cero y desviación estándar de uno.

3.3.3 Resultados de desarrollo infantil: el ASQ-3 y el Bayley-III

Medimos el desarrollo de los niños utilizando la tercera edición de los Cuestionarios de Edades y Etapas (*Ages and Stages Questionnaires*, ASQ-3; Squires et al. 2009) y la tercera edición de las Escalas de Bayley (*Bayley Scales of Infant and Toddler Development*, Bayley-III; Bayley 2006). El ASQ-3 fue administrado a todos los niños en la muestra en dos puntos del tiempo: primero, en línea de base, entre abril y agosto de 2013, cuando ellos tenían entre 1 y 24 meses de edad y antes de que empezaran a recibir visitas domiciliarias; y segundo, en la línea de seguimiento, entre junio y diciembre de 2015, cuando tenían entre 25 y 55 meses de edad y después de haber recibido visitas domiciliarias durante aproximadamente dos años. El Bayley-III fue

[8] Para este ejercicio, la confiabilidad promedio a nivel de escala se construye calculando primero el porcentaje de escalas que cumplen con el criterio de confiabilidad a nivel de escala para cada video y luego calculando el promedio de la muestra. Un procedimiento análogo se utilizó para el promedio de confiabilidad a nivel de ítem.

[9] Los desarrolladores recomendaron que los puntajes a nivel de escala sean asignados a una escala de 1 a 7 con base en un patrón general observado para los ítems. Estos resultados son similares a aquellos que se obtienen de calcular el promedio matemático de todos los ítems de cada escala (correlación entre un rango de 0,92 para *Receptividad de la facilitadora hacia la familia y de* 0,97 para *Involucramiento del niño* durante la visita).

administrado solo en la recolección de datos de seguimiento, entre mediados de julio y de diciembre de 2015 y solo en la submuestra de niños que al momento tenían menos de 42 meses (el instrumento no está diseñado para evaluar a niños mayores).

El ASQ-3 es una prueba de tamizaje. La conforma un conjunto de cuestionarios específicos para cada edad, que deben ser completados por los cuidadores. Cada cuestionario mide la competencia de los niños en cinco dominios: resolución de problemas (cognición), comunicación, motricidad fina, motricidad gruesa y personal-social - con 6 ítems para cada uno de ellos. Por el bajo nivel de educación de los cuidadores en esta muestra, los ítems se administraron o por reporte de la madre durante la entrevista o fueron directamente administrados al niño por el encuestador. Un equipo de psicólogos peruanos adaptó el instrumento al español y al contexto local y determinó cuáles de los ítems serían administrados por reporte del cuidador y cuáles, por administración directa.

El ASQ-3 ha sido utilizado con frecuencia en evaluaciones de impacto de gran escala (Ángeles et al. 2012; Fernald et al. 2012; Bernal 2015). Al ser una prueba de tamizaje, está diseñado para identificar a los niños en riesgo de retrasos en el desarrollo y por lo tanto es más sensible para la medición del desarrollo en la parte inferior de la distribución de habilidades. Dado que nuestra meta era identificar cambios en los resultados de desarrollo infantil que hubieran resultado de la intervención a lo largo de toda la distribución de habilidades de desarrollo, modificamos la administración y aplicamos tres ítems adicionales, los primeros tres del siguiente cuestionario por edad, sólo para aquellos niños que alcanzaron el puntaje máximo en la escala. Si los ítems están ordenados en orden creciente de dificultad dentro de un mismo cuestionario y entre cuestionarios sucesivos, esta estrategia amplía el techo de la prueba. Adaptaciones similares del ASQ han sido empleadas en otros estudios en países de ingresos medios y bajos (Fernald et al. 2012).

También administramos la prueba Bayley-III a una submuestra de niños a quienes era posible aplicar esta prueba. El Bayley-III es una prueba de diagnóstico, ampliamente reconocida como el 'patrón de oro' para medir el nivel de desarrollo de los niños desde el nacimiento hasta los 42 meses de edad (Frongillo et al. 2014; Fernald et al. 2012; Fernandes et al. 2014). Además, la prueba ha mostrado sensibilidad en la identificación de efectos de programas de visitas domiciliarias similares a Cuna Más en contextos tan diversos como Bangladesh (Nahar et al. 2012) y Colombia (Attanasio et al. 2014).

El Bayley-III evalúa cinco dominios del desarrollo: cognición, lenguaje receptivo y expresivo, motricidad gruesa y fina, socio emocional y comportamiento adaptativo. Debido a restricciones presupuestales y de tiempo, solo administramos las escalas relacionadas más directamente con los objetivos de la intervención, es decir, cognición, lenguaje (expresivo y receptivo) y motricidad fina. Estas escalas se evalúan exclusivamente mediante la observación directa del desempeño del niño en una serie de ítems presentados en orden ascendente de dificultad. Las reglas de base y techo prestablecidas determinan el número de ítems a administrar a cada niño.

El instrumento fue traducido y retraducido para asegurar equivalencia lingüística y funcional. Además, varios ítems necesitaron de adaptación y contextualización. Esto se

llevó a cabo mediante un extenso pilotaje. El proceso fue liderado por una psicóloga peruana, que también estuvo involucrada en la adaptación del ASQ-3 para este estudio (Lazarte 2016 presenta los detalles).

Capacitación y preparación para aplicar las pruebas. El ASQ-3 fue administrado en el hogar de los niños donde también se realizó una encuesta. La línea de base fue realizada por un equipo de 76 encuestadores, que fueron capacitados por 3 entrenadores durante 14 días, de los cuales 5 días fueron de prácticas. En la recolección de los datos de seguimiento, para mejorar la calidad, el equipo de entrevistadores a cargo de la encuesta se redujo a 20 con una capacitación similar (de 13 días) y más administraciones de práctica. Para la muestra donde se midió la calidad de las visitas domiciliarias (una submuestra de la muestra de la evaluación de impacto), un total de 52 entrevistadores recolectaron los datos de la línea de base de ASQ-3 y 18 realizaron el seguimiento.

El Bayley-III requiere ser administrado por profesionales en desarrollo infantil, como psicólogos y educadores, en una sala tranquila que facilite la concentración del niño en la evaluación. Doce personas graduadas como psicólogos fueron capacitadas en el Bayley-III por 5,5 semanas incluyendo entre 15 y 18 prácticas cada uno. De ellos, 8 fueron seleccionados para administrar la prueba y dos fueron asignados a un rol de supervisión. Todos los niños de la muestra fueron evaluados en una sala habilitada para este propósito en su comunidad.

Las evaluaciones de práctica se llevaron a cabo en parejas y continuaron hasta que la coincidencia entre el entrenado y el entrenador fuera >0,9 en cada aplicación. Durante las actividades de campo, 7,55 por ciento de las administraciones de Bayley-III fueron observadas por la capacitadora, quien dio retroalimentación correctiva cuando fue necesario. Las coincidencias entre el calificador y el entrenador durante las pruebas alcanzaron >0,98, asegurando datos de calidad alta. Todas las administraciones de las pruebas se realizaron en presencia del cuidador principal del niño (frecuentemente su madre) después de obtener su consentimiento informado.

Cálculo de un puntaje de desarrollo del niño. Para cada ítem en el ASQ-3, las opciones de respuesta son ‘sí’, ‘a veces’ y ‘todavía no’ y se les asigna 10, 5 o 0 puntos, respectivamente. Para cada escala, el puntaje total crudo es la suma de todos los ítems de la escala. Los ítems sin una respuesta se reemplazan por el promedio de la escala, aunque si en una escala faltan más de dos ítems, entonces no se le puntúa.

En cuanto al Bayley-III, el niño tiene una puntuación de 1 para cada ítem correctamente ejecutado y 0 de lo contrario. Para cada escala, el puntaje crudo es la suma de las respuestas correctas, incluidos los ítems no administrados que preceden a la base. Por construcción, el puntaje crudo del Bayley-III es creciente con la edad.

Dado que el ASQ-3 tiene cuestionarios específicos a la edad, los puntajes crudos deberían ser constantes con la edad. Sin embargo, en la práctica, observamos una tendencia creciente con la edad. Por lo tanto, primero transformamos los puntajes crudos para cada dominio del ASQ-3 o del Bayley-III a puntajes Z, después de ajustar no-paramétricamente por la edad del niño utilizando el método propuesto por Rubio-

Codina et al. (2016a). A diferencia del uso de normas externas construidas a partir de las poblaciones de referencia para cada prueba, este método de estandarización tiene la ventaja de que trata los efectos de edad de manera consistente en las dos pruebas. Por lo tanto, construimos diferentes puntajes agregados. Para el ASQ-3 construimos tres agregaciones: en la primera, los cinco dominios de la prueba (cognición, lenguaje, motricidad fina, motricidad gruesa y personal-social) reciben el mismo peso; en la segunda, empleamos únicamente los dominios de cognición, lenguaje y motricidad fina, cada uno con el mismo peso; y en la tercera, utilizamos solo los dominios de cognición y de lenguaje. Para el Bayley-III, hicimos un ejercicio similar y construimos también tres agregaciones. En la primera, los cuatro dominios (cognición, lenguaje receptivo, lenguaje expresivo y motricidad fina) tienen el mismo peso. En la segunda, primero construimos un puntaje total de lenguaje como el promedio entre los puntajes de lenguaje receptivo y expresivo y luego, en un segundo paso, agregamos la cognición, el lenguaje y los dominios de motricidad fina y cada uno recibió un tercio del peso total. Construimos la segunda agregación de esta manera porque se acerca a la segunda agregación definida para el ASQ-3. Finalmente, la tercera agregación, de forma similar a la definida con ASQ-3, se basa en los dominios de cognición y lenguaje. Todas las agregaciones se volvieron a estandarizar para que tuvieran media cero y una desviación estándar de uno.

3.2.3. Otras medidas

Se realizó una encuesta de hogares tanto en línea de base como en el seguimiento. Esta recogió información socioeconómica básica, así como la historia de la salud de los niños y el conocimiento de la madre sobre el desarrollo infantil, entre otras variables. Además, la encuesta recopiló información de los siguientes índices de calidad del ambiente del hogar: las Escalas de Aceptación y Receptividad del cuestionario de Observación para la Medida del Ambiente (*Home Observation for Measurement of the Environment;* HOME, Caldwell y Bradley 2003), así como el número y variedad de materiales y actividades de juego que el niño usualmente utilizó con el adulto en su hogar durante una semana previa a la encuesta, utilizando los Indicadores de Cuidado Familiar (*Family Care Indicators*, FCI, Frongillo et al. 2003). Los materiales de juego fueron registrados por observación directa y las actividades de juego, por el reporte del cuidador[10]. La encuesta también midió el conocimiento de los padres sobre el desarrollo infantil con el Inventario de Conocimientos sobre el Desarrollo de los Bebés (*Knowledge of Infant Development Inventory,* KIDI, MacPhee 2002). Todos los índices fueron construidos con análisis factorial y normalizados para tener media cero y desviación estándar uno.

[10] El número de variedades de materiales de juego es la suma de los indicadores para: juguetes que hacen o tocan música; juguetes/objetos para apilar/construir; cosas para dibujar/escribir/colorear/pintar; juguetes para moverse; juguetes para el juego dramático; libros ilustrados; libros de dibujo para niños; juguetes para aprender formas y colores. El número de variedades de actividades de juego es la suma de indicadores para: leer libros/mirar libros ilustrados; contar cuentos al niño; cantar canciones con el niño; jugar con el niño con juguetes; pasar tiempo con el niño garabateando/dibujando/coloreando; pasar tiempo con el niño nombrando objetos/contando; y sacar al niño a pasear.

Adicionalmente, en el día de la filmación de la visita, se recolectó otra información como la duración de la relación entre la facilitadora y la familia, el número total de facilitadoras que habían tenido hasta el momento, etc.

4. Muestra y estadísticas descriptivas

Trabajamos con tres muestras diferentes para maximizar el tamaño de la muestra en nuestras estimaciones. Las tres muestras fueron seleccionadas de la muestra que contaba con datos de la calidad de las visitas domiciliarias, como se explica en la Figura 1 (Panel B). La muestra tiene 229 niños focales (aquellos que también eran parte de los datos de la evaluación de impacto) y 176 facilitadoras. Excluimos a los niños focales, facilitadoras y cualquier video de las visitas en todos estos casos: (i) si no teníamos información de identificación de la facilitadora y por lo tanto no podíamos unir esa observación a otras fuentes de datos; (ii) si los niños no tenían resultados de desarrollo en la línea de base y en el seguimiento de la evaluación de impacto; (iii) si la facilitadora no realizó visitas a los niños para los que teníamos los resultados de desarrollo; (iv) si solo se filmó una visita para esa facilitadora. Después de eliminar estos casos, la muestra se redujo a 218 niños que fueron visitados por 155 facilitadoras. Nos referimos a esta muestra como la 'muestra con visita propia' pues, para los niños que se incluyen en ella, las visitas fueron observadas y filmadas.

Para la segunda y tercera muestras, usamos datos administrativos del Programa para identificar a otros niños del grupo de tratamiento de la muestra de la evaluación de impacto que habían recibido visitas de alguna de las 155 facilitadoras entre enero y diciembre de 2015[11], aun cuando sus propias visitas *no hubieran sido filmadas* durante la recolección de datos de medición de la calidad de las visitas. Para estos niños, también tenemos indicadores de desarrollo infantil en línea de base (ASQ-3) y en seguimiento (ASQ-3 o Bayley-III). Específicamente, la segunda muestra incluye 431 niños para los que pudimos identificar a la facilitadora y para los cuales tenemos puntajes ASQ-3 -esto es, los 218 de la 'muestra con visita propia' más 213 niños adicionales que tienen puntajes de ASQ-3. La tercera muestra incluye los 346 niños que pudimos identificar que además fueron medidos con el Bayley-III, esto es 218 de la 'muestra con visita propia', más 123 niños adicionales con puntajes de Bayley-III. Nos referimos a estas muestras como 'muestra ASQ' y 'muestra Bayley', respectivamente.

La Tabla 1 muestra estadísticas descriptivas de estas tres muestras de análisis.[12] Las características de los niños y sus familias en línea de base y en el seguimiento se presentan en los paneles A y B, respectivamente. El panel B también muestra información sobre las comunidades donde viven los niños y sus familias. Los niños para los que su visita domiciliaria fue filmada ('muestra con visita propia') tenían en promedio 31,82 meses de edad en la encuesta de seguimiento. Como se esperaba, los demás niños de la muestra de la evaluación de impacto empalmados con la base de calidad son ligeramente mayores dado que estos niños no estaban recibiendo visitas al momento de la filmación (los niños egresan del Programa a los 36 meses). Los niños de la 'muestra ASQ' tenían 36,26 meses de edad y los de la 'muestra Bayley' tenían

[11] Los datos administrativos del Programa no tenían esta información antes del 2015.
[12] Araujo et al. 2018 presenta un conjunto de descriptivas comparables para la muestra completa de la evaluación de impacto de la cual se extrajeron estas submuestras.

33,86 meses de edad, en promedio. En general hay menos niñas que niños en las tres muestras, la proporción de niñas es 47 por ciento, 46 por ciento y 44 por ciento respectivamente. De forma similar, las mamás tienen entre 6 y 7 años de educación, en promedio, en las tres muestras, lo que corresponde a la primaria completa y un año de escuela secundaria. Con respecto a la calidad del ambiente en el hogar, es decir, los índices para el número de variedades materiales y actividades de juego del FCI, vemos que, en promedio, los niños realizaron entre 4 y 5 diferentes actividades (de un total de 7) en la semana anterior a la encuesta de seguimiento y utilizaron entre 2 y 3 tipos de juguetes (de un total de 8). Estos promedios son similares entre las tres muestras. Más de la mitad de los niños en la 'muestra con visita propia' (56 por ciento) tuvo al menos una facilitadora adicional desde que empezó a participar en el Programa y, entre quienes tuvieron más de uno, entre uno y dos adicionales. En promedio, han estado en el Programa por 23 meses.

Utilizando datos administrativos sobre la participación en el Programa, estimamos el número total de visitas que un niño debió haber recibido según la edad que tenía cuando participó por primera vez en Cuna Más. Comparamos esto con el número de las visitas reportadas como las efectivamente entregadas para obtener una estimación de la intensidad de la intervención, es decir, el porcentaje de visitas recibidas dado el máximo posible. Aunque los datos administrativos pueden ser imprecisos e incompletos, una examinación detallada de la distribución de estos números entre las tres muestras sugiere que están concentradas en más de 80 por ciento, lo que significa que la mayoría de los niños recibieron una intervención de una intensidad cercana a la que se pretendía. Los promedios para las tres muestras son similares en magnitud, alrededor de 88-89 por ciento.

Las comunidades donde se realizaron estas visitas son rurales, pequeñas y aisladas. Todas ellas están a más de una hora de distancia de la capital del distrito, ya sea en carro o a pie. Para más de la mitad de ellas caminar es el medio de transporte más común. El acceso a servicios públicos y privados varía bastante: 87 por ciento de las comunidades en la muestra tiene un centro de salud, 96 por ciento tiene un prescolar público, pero solo 11 por ciento tiene un banco. Por su parte, la cobertura de programas sociales emblemáticos -incluyendo la alimentación escolar- es alta en estas comunidades.

La Tabla 2 presenta las características de las facilitadoras a partir de una encuesta realizada y de la lista de chequeo administrada por la supervisora. Estos datos se refieren a las 155 facilitadoras que trabajan con los niños en cada una de las tres muestras y que viven en 92 comunidades.[13] Las facilitadoras tienen, en promedio, 31,44 años de edad y 10,05 años de educación. En promedio trabajan con 9,85 familias y realizan 4,33 visitas por familia por mes. A pesar de que las facilitadoras han estado en el Programa por 17 meses, en promedio, ellas reportan haber trabajado con las familias cuyas visitas fueron evaluadas aproximadamente 14 meses. Esto sucede, en parte, porque las facilitadoras incluyeron la capacitación en su cálculo del tiempo de empleo y también es indicativo de los niveles de rotación de las facilitadoras. Casi

[13] La información de la encuesta del facilitador está solo disponible para 136 de 155 facilitadoras pues esta encuesta no fue administrada a todas las facilitadoras de cada comunidad debido a restricciones presupuestales.

todas las facilitadoras en la muestra (98 por ciento) reportan esta como su ocupación principal, la cual demanda 11,27 horas a la semana.

Veinte y seis por ciento de las facilitadoras reporta haber tenido experiencia previa trabajando con niños en el rango de edad de 0 a 3 años y 27 por ciento reporta haber trabajado con familias antes de su trabajo actual. Entre aquellas que han tenido estas experiencias relevantes previas, el promedio de experiencia es cercano a 2,5 años. El 59 por ciento de facilitadoras en la muestra reportó haber participado en al menos una capacitación en los últimos tres años fuera de las del Programa. Los temas más comunes de las capacitaciones que recibieron fueron nutrición, higiene y salud. Temas más directamente relacionados con el espíritu de la intervención de Cuna Más como el desarrollo infantil, cuidado infantil y aprendizaje temprano son menos frecuentes; a lo sumo una de cada cinco facilitadoras reporta haber tenido alguna capacitación en estos temas.

Como parte de su tiempo en Cuna Más, el 90 por ciento de las facilitadoras ha participado en la capacitación previa ofrecida por el servicio. Además, han recibido en promedio 3,44 sesiones de formación continua en el Programa. Noventa y tres por ciento de las facilitadoras reconoce su interacción con sus supervisoras como parte de su experiencia de capacitación continua. Sin embargo, la frecuencia de esta interacción no es constante. La mayoría de las facilitadoras, 63 por ciento, se reúnen con su supervisora una vez cada dos semanas como lo establecen los lineamientos del Programa. No obstante, 29 por ciento lo hace solo una vez al mes y 7 por ciento lo hace con menor frecuencia (cada dos meses o más).[14] Otra área donde se observa heterogeneidad en las condiciones laborales es la frecuencia en el pago: mientras que el 74 por ciento de las facilitadoras reporta recibir su pago mensual, 17 por ciento reportan recibirlo cada dos meses y el restante 9 por ciento incluso más esporádicamente. Dado que los lideres comunales juegan un rol en facilitar los pagos, estas diferencias no son sorprendentes.

La Tabla 3 resume las estadísticas sobre el desarrollo infantil medido a través del ASQ-3 y el Bayley-III y sobre la calidad de las visitas domiciliarias medidas con el HOVRS-A+2. Los puntajes crudos de línea de base y de línea de seguimiento del ASQ-3 se calcularon sobre un total de 90 puntos y son similares en las tres muestras. Los puntajes crudos del Bayley-III también son similares en magnitud entre las dos muestras para los que están disponibles, aunque ligeramente mayores para la 'muestra Bayley' pues estos niños son mayores el puntaje crudo del Bayley-III es creciente con la edad, por construcción.

En las últimas tres filas de la Tabla 3 reportamos los niveles de calidad de las visitas, esto es el puntaje total del HOVRS-A+2 y los puntajes de los dominios de *Prácticas de las facilitadoras e Involucramiento Familiar*. Para los niños en la 'muestra con visita propia', el puntaje promedio reportado para el HOVRS-A+2 corresponde a la calidad de *su propia visita domiciliaria*, eso es la que fue filmada y codificada (puntaje propio).

[14] No tenemos datos en la capacitación recibida por las supervisoras en la muestra. Por las guías del Programa, sabemos que las supervisoras deben recibir una capacitación antes de prestar el servicio y que ésta tiene una duración de 4 a 5 días y al menos dos talleres de capacitación por año (3 a 5 días cada uno), así como sesiones mensuales de actualización (Josephson et al. 2017).

Para los niños en la 'muestra ASQ' y en la 'muestra Bayley' reportamos *el valor promedio de la calidad de todas las visitas domiciliarias realizadas por su facilitadora* y que fueron filmadas y codificadas. Sin embargo, para los 218 niños para los cuales la visita fue filmada, *el valor promedio de la calidad de la visita hecha por las facilitadoras* excluye su propio puntaje (siguiendo la lógica explicada en la siguiente sección). Los valores son similares entre las muestras y reflejan visitas de rango de calidad 'adecuada' para ambos dominios y para el puntaje total del HOVRS-A+2. Los puntajes del primer dominio *Prácticas del facilitador* son un punto menor a los puntajes en el segundo dominio*, Involucramiento Familiar*. Rubio-Codina et al. (2018) presenta una discusión a profundidad sobre estos puntajes.

5. Metodología empírica y resultados

Esta sección explora la asociación entre la calidad de la visita domiciliaria, la educación de la madre, la calidad del ambiente en el hogar y los resultados de desarrollo infantil. Entender cómo se implementa la visita domiciliaria debería dar información de cuáles son los aspectos críticos de la calidad de los servicios provistos, y de cómo deberían ser fortalecidos para mejorar los impactos en los beneficiarios.

5.1. Correlaciones entre calidad de la visita domiciliaria, calidad del ambiente en el hogar y resultados de desarrollo infantil

Construimos un promedio entre visitas y utilizamos esta variable como proxy de la calidad promedio que cada facilitadora ofrece a los niños a su cargo. Al construir esta media, queremos excluir al niño cuya visita fue observada, por las razones discutidas abajo. Específicamente, si la facilitadora *j* atiende a los niños *i=1,….n* excluimos al niño *n* de la medida de calidad promedio correspondiente a ese niño. En otras palabras, la calidad de la visita para el niño *n* se define como:

$$Q_n = \frac{Q_1 + Q_2 + \ldots + Q_{n-1}}{n-1}$$

Dos razones explican nuestra elección de esta variable como medida de calidad de las visitas domiciliaras para el niño *n*. Primero, si fuéramos a utilizar un solo puntaje de la calidad de la visita domiciliaria para cada facilitadora individual, es probable que la medida sea ruidosa (no representativa) y por lo tanto no capture la asociación entre la calidad de la visita y el desarrollo del niño. Una alternativa seria construir un puntaje promedio de todas las visitas que observamos para esa facilitadora y utilizar este puntaje promedio en las regresiones de todos los niños que ella atiende. Sin embargo, esto puede resultar en un sesgo por variables omitidas. Es plausible que variables no observadas del niño o de la familia -como el temperamento del niño o la habilidad de la madre para reconocer y responder a las necesidades del niño- puedan estar asociadas tanto al nivel de desarrollo infantil *como a* la facilidad con la cual se conduce la visita domiciliaria y, por lo tanto, a su calidad. Las habilidades no observadas de las facilitadoras pueden compaginar mejor con niños o familias con características específicas no observadas y esto podría resultar en visitas de mejor calidad. Asimismo, asumimos que una medida promedio de la calidad de la visita domiciliaria que excluye

al niño cuyo puntaje de desarrollo está siendo asociado con la calidad no tiene este problema.

La Tabla 4 presenta las correlaciones parciales entre los diferentes resultados de desarrollo infantil -esto es puntajes Z del ASQ-3 y Bayley-III en línea de seguimiento – y (i) las medidas de calidad del ambiente en el hogar (actividades y materiales de juego del FCI, el HOME y el KIDI- todo expresados en puntajes Z) y (ii) los puntajes de los dominios y el total del HOVRS-A+2 para las tres muestras. Presentamos dos medidas diferentes de calidad de las visitas domiciliarias. Para los niños en la 'muestra con visita propia', correlacionamos en las primeras tres columnas del último panel de la Tabla 4, el puntaje de desarrollo del niño *i* con el puntaje de la visita observada y codificada del niño y su facilitadora (puntaje propio); y en las últimas tres filas del panel, el puntaje de desarrollo del niño *i* con el puntaje promedio de las visitas observadas y codificadas para otros niños atendidos por la misma facilitadora del niño *i* (puntaje promedio de las visitas de otros).

Con respecto a la calidad del ambiente en el hogar, observamos que las variedades de materiales de juego están positiva y significativamente correlacionadas con mayores puntajes de desarrollo infantil para todas las muestras y todas las medidas, excepto para el puntaje ASQ-3 en la 'muestra con visita propia'. Las correlaciones para esta medida tienden a ser mayores con los puntajes de Bayley-III que con los puntajes de ASQ-3. El HOME está correlacionado con todas las agregaciones de ASQ en la 'muestra ASQ' y en la 'muestra Bayley' pero solo en la agregación del ASQ que contiene los cinco dominios en la 'muestra con visita propia'. El HOME solo está correlacionado con el Bayley-III en la 'muestra con visita propia'. El KIDI no está correlacionado con ninguno de los resultados de desarrollo infantil excepto, en una oportunidad, con los agregados de cognición y de lenguaje del Bayley en la 'muestra con visita propia'.

Los puntajes para los dominios del HOVRS-A+2 y el total están positiva y significativamente correlacionados con el puntaje de seguimiento del ASQ-3 para todas las muestras, pero no con los puntajes del Bayley-III. Cuando analizamos por separado los dominios del HOVRS-A+2, encontramos que las correlaciones entre el desarrollo infantil y el dominio *Involucramiento Familiar* son más altas que aquellas con el dominio de *Prácticas del Facilitador*.

5.2. Análisis de regresión: asociación de la educación de la madre, ambiente en el hogar y calidad de la visita domiciliaria con los puntajes de desarrollo infantil

Exploramos la asociación entre los puntajes de desarrollo infantil y la educación de la madre, el ambiente del hogar y la calidad de la visita domiciliaria con un análisis de regresión. Para cada una de las tres muestras, se corren regresiones en las cuales la variable dependiente es el desarrollo infantil (los puntajes Z del ASQ-3 y del Bayley-III), controlando por las siguientes variables:

- En la primera especificación (modelo 1), incluimos los mismos controles que se utilizaron en la evaluación de impacto de Cuna Más, reportados en Araujo et al.

2018.[15] También controlamos por la principal variable de interés: la calidad de la visita domiciliaria, medida por el HOVRS-A+2 y expresada en puntaje Z.

- La segunda especificación (modelo 2) añade las medidas de calidad del ambiente en el hogar- esto es las medidas de línea de seguimiento para las actividades y materiales de juego del FCI, el HOME y el KIDI (expresadas en puntajes Z). Incluimos estas variables para analizar el efecto indirecto del HOVRS a través de la calidad de las prácticas parentales.

Los errores estándar (ES) se ajustan a nivel de conglomerado (distrito), dado que es la unidad de aleatorización en la evaluación de impacto. También incluimos controles para las unidades de estratificación utilizadas en el diseño muestral.

Los resultados reportados en las Tablas 5, 6 y 7 (una para cada muestra) difieren entre medidas de desarrollo. El desarrollo infantil en la línea de base no está significativamente asociado con el desarrollo infantil en el seguimiento en ninguno de los modelos. Sorprendentemente, tampoco lo está la educación de la madre. De forma similar, las variedades de actividades de juego no están asociadas con los resultados de desarrollo, mientras que las variedades de materiales de juego se encuentran asociadas con el ASQ-3 y el Bayley-III en las muestras ASQ y Bayley, pero no en la 'muestra con visita propia', que tiene un menor tamaño. El número de visitas que el niño ha recibido no es estadísticamente significativo en ninguna de las especificaciones.

En todas las regresiones que tiene el ASQ-3 como variable dependiente, sin importar la muestra, el HOVRS-A+2 está positiva y significativamente asociado con el desarrollo infantil, lo que sugiere una correlación entre la calidad de la visita domiciliaria y el desarrollo infantil. La magnitud de esta asociación no es pequeña: una mejora en la calidad de la visita domiciliaria de una DE está asociada con un incremento en el puntaje de ASQ-3 de entre 0,149 y 0,257 DE. Sin embargo, el coeficiente del puntaje total del HOVRS-A+2 no es significativo a niveles convencionales en regresiones que utilizan los puntajes Z del Bayley-III como medida de resultado, lo cual es inconsistente con los resultados observados para el ASQ-3.

Realizamos varias pruebas de robustez y todas llevaron a resultados cualitativamente comparables. Específicamente:

- *Ajustar los errores estándar usando la facilitadora como conglomerado* en vez del distrito no modifica los resultados. De forma similar, los resultados cambian mínimamente si no realizamos ningún ajuste de los errores estándar.
- *Estimar todos los modelos utilizando los puntajes del HOVRS-A+2 a nivel de dominio y de ítem* muestra resultados consistentes. Los impactos son significativos únicamente cuando el puntaje ASQ-3 es la variable dependiente, y hay más ítems

[15] Estos corresponden a las características del niño y la familia en la línea de base presentadas en la Tabla 1 así como el puntaje total (todos los dominios) de la línea de base del ASQ-3 y su función polinomica de cuarto grado. También incluyen controles de distrito y centro poblado, es decir el promedio a nivel de centro poblado de todos los controles de niño y hogar de la Tabla 1; el puntaje promedio a nivel de distrito en comprensión de lectura y matemáticas del examen nacional estandarizado Evaluación Censal de Estudiantes; y el promedio a nivel de distrito de los controles de hogar del censo nacional de Perú de 2007.

que están significativamente asociados al desarrollo infantil en el dominio de *Involucramiento Familiar* que en el dominio de *Prácticas de la Facilitadora*, siendo también mayor el coeficiente de la regresión.

- *Eliminar los ítems del HOVRS-A+2 que no están alineados con los objetivos y la estructura de las visitas de Cuna Más* tampoco modifica los resultados principales.
- *Intentar diferentes agregaciones de la calidad de la facilitadora medida por el HOVRS-A+2* no modifica los resultados. Dada la estructura de nuestra muestra, construimos una medida promedio de la calidad de la facilitadora basada en un número de visitas diferente observado para cada facilitadora. Es posible estimar el 'modelo con visita propia' fijando ese número al mínimo común de todas las observaciones de la muestra. Para los niños de la 'muestra con visita propia', en la Tabla 4 hicimos 500 sorteos de nuestros resultados principales, y para cada uno calculamos el promedio del puntaje de calidad para ese niño con un número fijo de dos visitas de la facilitadora (diferente al puntaje propio del niño). Cada sorteo seleccionó aleatoriamente un par diferente de visitas para calcular el puntaje de calidad de visitas recibidas por el niño. Los resultados muestran que el coeficiente en el HOVRS-A+2 y el error estándar del coeficiente son estables entre iteraciones, tanto para el ASQ-3 como para el Bayley-III.
- *Eliminar del análisis las observaciones donde los videos no cumplían los criterios de confiabilidad del HOVRS-A+2* después de codificar dos veces tampoco altera los resultados principales.

6. Conclusiones

En este estudio, investigamos la asociación entre la calidad de la visita domiciliaria y el desarrollo infantil utilizando una muestra de niños y de visitas domiciliarias que son parte de la evaluación de impacto de Cuna Más. La calidad de la visita domiciliaria se mide con el instrumento HOVRS-A+2, utilizado frecuentemente para evaluar la calidad de los servicios de visitas domiciliarias en Estados Unidos. El instrumento fue adaptado para las áreas en las que Cuna Más opera – es decir comunidades rurales y dispersas del Perú. El HOVRS-A+2 fue calificado por codificadores capacitados en grabaciones de video de 554 visitas domiciliarias en igual número de hogares. El desarrollo infantil se evaluó empleando las pruebas ASQ-3 y Bayley-III. El ASQ-3 fue administrado en la

línea de base y en la de seguimiento, para todos los niños en la muestra de evaluación, y el Bayley-III fue administrado a una submuestra de estos niños, únicamente en la línea de seguimiento. Para cada prueba, un puntaje Z de desarrollo fue construido con las escalas cognitiva, lenguaje y de motricidad fina.

En este análisis, trabajamos con tres muestras distintas de niños, dependiendo de la disponibilidad de información sobre su desarrollo y de las grabaciones de visitas propias o de visitas realizadas por sus facilitadoras a otros niños. En general, los resultados son consistentes entre las diferentes muestras. Para las tres muestras, se filmó y calificó los videos correspondientes a las visitas domiciliarias realizadas por un

total de 155 facilitadores en 92 centros poblados rurales. Como se establece en los lineamientos de Cuna Más, las facilitadoras en la muestra estaban a cargo de entre 9 y 10 niños. Las facilitadoras han estado en el Programa por cerca de dos años, la mayoría tienen algo de educación secundaria –algunas con secundaria completa, otras no – y un cuarto de ellas ha trabajado con familias en empleos previos. Los niños de la muestra tenían entre dos y dos años y medio de edad, en promedio, y un poco más de la mitad de ellos ha tenido otra facilitadora desde el inicio de su participación en el Programa (en algunos casos, dos). Los niños de la muestra parecen haber recibido más de 80 por ciento de las visitas que pretendía el Programa, lo cual sugiere un alto cumplimiento, o intensidad del Programa.

Encontramos que los puntajes de la calidad de la visita domiciliaria están significativamente asociados con el puntaje Z del ASQ-3, controlando por las características del niño (incluyendo puntajes de línea de base), la educación de la madre, la calidad del ambiente en el hogar y otros controles reportados en la evaluación de impacto de Cuna Más. Sin embargo, estas asociaciones no son estadísticamente significativas cuando se utiliza el Bayley-III como variable de resultado. Esto es sorprendente e inconsistente, dado que el Bayley-III, al igual que otras pruebas de diagnóstico multidimensional, son consideradas como el estándar de oro para medir el desarrollo infantil (Frongillo et al. 2014; Fernald et al. 2009; Fernandes et al. 2014) y el ASQ-3 ha mostrado tener una menor validez cuando se compara con otras pruebas de diagnóstico y con el Bayley-III (Rubio-Codina et al. 2016a). También es sorprendente la falta de asociación condicional entre la educación de la madre y la calidad del ambiente en el hogar- en particular las actividades y materiales de juego- y los puntajes de desarrollo infantil pues estas correlaciones han sido ampliamente documentadas en la literatura (Rubio-Codina et al. 2016b). Una de las razones puede ser que los tamaños de la muestra son pequeños. Una segunda posibilidad es que, a niveles muy bajos de educación, como es el caso de nuestra muestra, los años de educación de la madre pueden reflejar la falta de acceso y no la falta de motivación de la madre para aprender. Una tercera explicación posible puede ser que, dado que la estrategia de identificación no permite inferencias causales, las variables de calidad de la visita domiciliaria podrían estar capturando *de facto* otras características de la familia y del niño que están asociadas o contribuyen a la calidad de la visita domiciliaria y a un mejor desarrollo infantil.

No obstante, las asociaciones positivas entre el ASQ-3 y la calidad de la visita domiciliara encontradas en este estudio son consistentes con la teoría en la literatura de los servicios de visitas domiciliarias, que argumenta que la calidad de procesos es clave para alcanzar los resultados deseados en los niños participantes (Paulsell et al. 2010; Wasik y Bryant 2001). En la práctica, sin embargo, los instrumentos para medir la calidad de la visita como el HOVRS-A+2 han sido desarrollados recientemente y solo algunos estudios han reportado asociaciones entre la calidad de procesos de las visitas y el desarrollo infantil. Dos estudios que conocemos que lo utilizan, ambos en el contexto del programa *Early Head Start* de los Estados Unidos, son los de Roggman et al. (2016) y Vogel et al. (2015).

El hecho de que la calidad de procesos esté relacionada con los resultados de desarrollo infantil tiene una implicación de política clara: el énfasis en la capacitación del personal a cargo de la atención de niños y familias. Este personal requiere el dominio de una combinación de habilidades sutiles, como la instrucción explícita, las interacciones sensibles y de calidad, y la retroalimentación receptiva, para lograr los cambios de comportamiento deseados en los beneficiarios. No obstante, frecuentemente el personal de este tipo de programas tiene poca educación y una capacitación precaria y recibe poca o ninguna retroalimentación sobre su desempeño. Si es que la calidad de procesos es realmente el corazón de la entrega del servicio (Wasik y Bryant, 2001), entonces se necesita más investigación para entender qué estrategias de capacitación costo efectivas y cuáles esquemas de mentoría pueden apoyar de mejor manera a este personal en sus responsabilidades diarias. Esto es particularmente importante en programas de visitas domiciliaras como Cuna Más que operan a escala, en contextos de bajos recursos, bajas habilidades y atendiendo a poblaciones muy dispersas.

Referencias

Ángeles, G., Gadsden, P., Galiani, S., Gertler, P., Herrera, A., Kariger, P. y Seira, E. (2011). “Evaluación de Impacto del Programa Estancia Infantiles para Apoyar a Madres Trabajadoras. Informe Final de la Evaluación de Impacto.” Instituto Nacional de Salud Pública. México.

Attanasio O., Fernández C, Fitzsimons E, Grantham-McGregor S, Meghir C, Rubio-Codina M. (2014) “Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial”. *British Medical Journal*, 349:g5785

Araujo, M.C., Dormal, M. Grantham-McGregor, S., Lazarte, F., Rubio-Codina, M., y Schady, N. (2018). “Behavioral change and child development at scale”. Unpublished manuscript.

Bayley N. (2006). Bayley Scales of Infant and Toddler Development—Third Edition: Technical manual. San Antonio, TX: Harcourt Assessment.

Berlinski, S., & Schady, N. (2015). The Early Years: Child Well-being and the Role of Public Policy. Development in the Americas Series.

Bernal, R. (2015). “The Impact of a Vocational Education Program for Childcare Providers on Children’s Well-being”. *Economics of Education Review*, vol. 48(C), pages 165-183.

Boller, K., Vogel, C., Cohen, R., Aikens, N., y Hallgren, K. (2009). “Home Visit Characteristics and Content Form.” Princeton, NJ: Mathematica Policy Research.

Britto, Pia R, Lye. S.L., Proulx K., Yousafzai, A., Matthews, S. Vaivada, T., Perez-Escamilla, R., Rao, N., Ip, P., Fernald, L., MacMillan, H., Hanson, M., Wachs, T., Yao, H., Yoshikawa, H., Cerezo, A., Leckman, J., Bhutta, Z. y the Early Childhood Development Interventions Review Group, for the Lancet Early Childhood Development Series Steering Committee. (2016). Nurturing care: promoting early childhood development, *Lancet* http://dx.doi.org/10.1016/S0140-6736(16)31390-3

Caldwell, B., y Bradley, R. (2003). Home Observation for Measurement of the Environment: Administration manual. Little Rock, AR: University of Arkansas.

DGSE-MIDIS, DGPP-MEF, BID. (2015). “Documento de Línea de Base para la Evaluación de Impacto del Programa Nacional Cuna Más – Servicio de Acompañamiento a Familias”, Lima, febrero de 2015, Documento interno del MIDIS.

Dishion, T. J., Knutson, N., Brauer, L., Gill, A., & Risso, J. (2010). Family Check Up: COACH ratings manual. Unpublished coding manual. Available from the Child and Family Center, 195 West 12th Avenue, Eugene, OR 97401.

Fernald, L., Kariger, P., Hidrobo, M. y Gertler, P. (2012). “Socioeconomic gradients in child development in very young children: evidence from India, Indonesia, Peru, and Senegal.” Proceedings of the National Academy of Science 109 (2): 17273-80.

Fernald, L. C., Kariger, P., Engle, P., y Raikes, A. (2009). Examining early child development in low-income countries. *Washington DC: The World Bank*.

Fernandes M, Stein A, Newton CR, Cheikh-Ismail L, Kihara M, Wulff K, et al. The INTERGROWTH-21st Project Neurodevelopment Package: A Novel Method for the Multi-Dimensional Assessment of Neurodevelopment in Pre-School Age Children. (2014) *PLOS ONE.* 9(11):e113360. doi: 10.1371/journal.pone.0113360 PMID: 25423589

Frongillo EA, Tofail F, Hamadani JD, Warren AM, Mehrin SF. Measures and indicators for assessing impact of interventions integrating nutrition, health, and early childhood development. (2014). *Annals of the New York Academy of Scences,* 1308(1):68–88.

Frongillo EA, Sywulka SM, Kariger P. (2003). UNICEF psychosocial care indicators project. Final report to UNICEF. Ithaca: Division of Nutritional Sciences, Cornell University.

Lazarte, F. (2016). *Informe final de las actividades realizadas para el recojo de información del estudio de Bayley III*. Reporte de consultoría no publicado.

MacPhee, D. (2002). Knowledge of Infant Development Inventory Survey of Child Care Experiences Manual. Mimeo.

MEF-BID, (2016). “Resultados de la evaluación de impacto del Servicio de Acompañamiento a Familias del Programa Nacional Cuna Más”. Reporte no publicado.

Nahar B, Hossain MI, Hamadani JD, Ahmed T, Huda SN, Grantham-McGregor SM, (2012). “Effects of a community-based approach of food and psychosocial stimulation on growth and development of severely malnourished children in Bangladesh: a randomised trial”. *European Journal of Clinical Nutrition*; 66:701-9

Paulsell, D., Avellar, S., Martin, E. S., & Del Grosso, P. (2010). Home visiting evidence of effectiveness review: Executive summary (No. 5254a2ab30e146ce900220dbc4f41900). Mathematica Policy Research.

Roggman, L. A., Cook, G. A., Innocenti, M. S., Jump Norman, V. K., Boyce, L. K., Christiansen, K., y Peterson, C. (2016). “Home Visit Quality Variations in two Early Head Start Programs in Relation to Parenting and Child Vocabulary Outcomes”. *Infant Mental Health Journal.* Vol. 00(0) 1-15.

Roggman, L. A., Cook, G. A., Innocenti, M. S., Jump Norman, V. K., Christiansen, K., Boyce, L. K., Aikens, N., Boller, K., Paulsell, D., & Hallgren, K. (2014). Home Visit Rating Scales—Adapted and Extended (HOVRS-A+) version 2. Adapted from Roggman, L. A., et al (2008). Home Visit Rating Scales. In L. Roggman, L. Boyce, y M.

Innocenti, Developmental Parenting (pp. 209-217). Baltimore: Paul H. Brookes Publishing.

Roggman, L. A., Cook, G. A., Jump, V., Boyce, L. K. & Innocenti, M. S. (2006). Home Visit Rating Scales (HOVRS). Unpublished manuscript. Utah State University, Logan, UT.

Rubio-Codina, M. Araujo, M.C., Attanasio, O., Muñoz, P. Grantham-McGregor, S. (2016a). Concurrent Validity and Feasibility of Short Tests Currently Used to Measure Early Childhood Development in Large Scale Studies. *PlosOne.* Published: August 22, 2016. http://dx.doi.org/10.1371/journal.pone.0160962

Rubio-Codina, M., Attanasio, O., & Grantham-McGregor, S. (2016b). Mediating pathways in the socio-economic gradient of child development: Evidence from children 6–42 months in Bogota. *International journal of behavioral development*, *40*(6), 483-491.

Rubio-Codina, M., Dormal, M. y Araujo, M.C. (2018). "Measuring quality and characterizing Cuna Mas home visits: validation of the HOVRS-A+2 in Peru and of a short checklist for use at-scale". Washington D.C.: Inter-American Development Bank.

Schodt, S., Parr, J., Araujo, M.C., y Rubio-Codina, M. (2015). "Measuring the Quality of Home-Visiting Services. A review of the literature". Inter-American Development Bank, Social Protection and Health Division Technical Note IDB-TN-881, October.

Smith, J. D., Dishion, T. J., Shaw, D. S., y Wilson, M. N. (2013). "Indirect effects of fidelity to the family check-up on changes in parenting and early childhood problem behaviors". *Journal of Consulting and Clinical Psychology*, 81(6), 962.

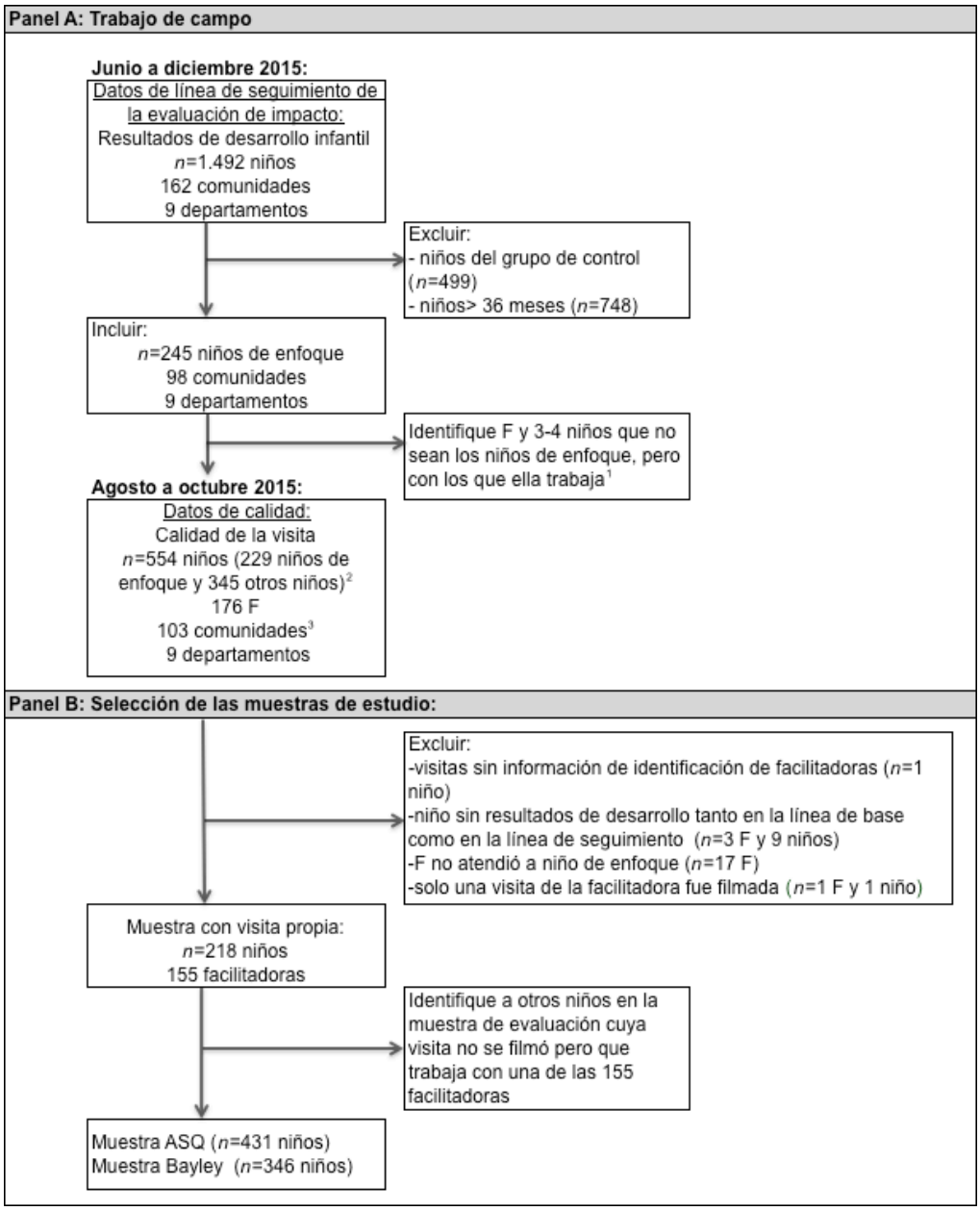
Squires J, Bricker D, Twombly E, Nickel R, Clifford J, Murphy K, et al. (2009). Ages & Stages English Questionnaires, Third Edition (ASQ-3): A Parent-Completed, Child-Monitoring System. Baltimore, MD: Paul H. Brookes Publishing Co.

Vogel, C. A., Caronongan, P., Thomas, J., Bandel, E., Xue, Y., Henke, J., Aikens, N., Boller, K. & Murphy, L. (2015). Toddlers in Early Head Start: A Portrait of 2-Year-Olds, Their Families, and the Programs Serving Them. Volume I: Age 2 Report. OPRE Report 2015-10. Mathematica Policy Research. February 2015. Available at: https://www.mathematica-mpr.com/our-publications-and-findings/publications/volume-i-toddlers-in-early-head-start-a-portrait-of-2yearolds-their-families-and-the-programs

Wasik, B. H., y Bryant, D. M. (2001). *Home visiting: Procedures for helping families*, *2nd ed*. Thousand Oaks, CA: Sage.

Tablas y Figuras

Figura 1: Diagrama de flujo de los participantes en el estudio



Notas. F=Facilitadora

[1]Para seis facilitadoras (3% de la muestra) solo tuvimos una visita filmada y para diez facilitadoras (6%) solo tenemos dos visitas filmadas.

[2] Incluye el 229 de los 245 niños de enfoque con información completa de HOVRS y la lista de verificación.

[3] Incluye 97 de las 98 comunidades que se identificaron y 6 comunidades nuevas no incluidas en la muestra de evaluación donde viven otros niños diferentes a los niños de enfoque.

Tabla 1: Características de los niños, las familias y la comunidad, por muestra

	N	Media/ Proporción	DE	N	Media/ Proporción	DE	N	Media/ Proporción	DE
	Muestra con visita propia			Muestra ASQ			Muestra Bayley		
Panel A: Información en línea de base									
Niña	218	0,47		431	0,46		346	0,44	
Edad de la madre	218	27,50	7,71	431	28,49	7,68	346	27,97	7,57
Educación de la madre (años)	207	6,83	3,77	405	6,65	3,74	327	6,61	3,71
Madre es indígena	218	0,01		431	0,02		346	0,02	
Hogar con ambos padres	218	0,89		431	0,86		346	0,88	
Tamaño del hogar	218	5,34	2,04	431	5,46	2,11	346	5,47	2,04
Número de habitaciones en la vivienda	218	2,42	1,26	431	2,52	1,31	346	2,48	1,31
Número de activos (0-11)	218	2,86	2,23	431	2,71	2,15	346	2,68	2,17
Hogar con suministro de agua entubada	218	0,64		431	0,67		346	0,65	
Hogar conectado al sistema de alcantarillado	218	0,28		431	0,30		346	0,27	
Material principal del piso no es tierra	218	0,22		431	0,22		346	0,21	
Panel B: Información en línea de seguimiento									
Características de los niños y sus familias									
Edad (meses)	218	31,82	2,85	431	36,26	6,37	346	33,86	4,09
Tiempo en el Programa (meses)[1]	217	23,18	5,99	-	-	-	-	-	-
Al menos otra facilitadora desde la inscripción en el Programa [1]	218	0,56		-	-	-	-	-	-
Número de facilitadoras desde la inscripción en el Programa (si es que más de una)[1]	120	1,84	0,96	-	-	-	-	-	-
Proporción de visitas recibidas	218	0,89		431	0,88		346	0,88	
FCI, número de variedades de actividades de juego	218	4,47	2,09	431	4,35	2,14	346	4,36	2,16
FCI, número de variedades de materiales de juego	215	2,32	1,71	428	2,24	1,67	343	2,21	1,63
HOME, puntaje total (codificado a la inversa)	218	9,99	1,82	431	9,87	1,90	346	9,89	1,87
KIDI	215	0,59	0,14	421	0,59	0,14	340	0,59	0,14
Características de la comunidad (N=92)[2]									
Tiempo hasta la capital del distrito a pie (horas)[3]	46	1,79	3,57						
Tiempo hasta la capital del distrito en carro (horas)[3]	43	1,02	0,84						
Centro de salud = 1	92	0,87	0,34						
Preescolar del Ministerio de Educación = 1	92	0,96	0,21						
Banco o caja municipal/rural =1	92	0,11	0,31						
Programa de desayuno escolar "Qali Warma" = 1	92	0,97	0,18						

Notas: [1]Información de la lista de chequeo administrada a los supervisores. El resto de variables proviene de la encuesta a la comunidad y al hogar

[2] Las características de la comunidad son iguales para las tres muestras

[3] Solo para las comunidades que lo reportan como medio principal de transporte

Tabla 2: Características de las facilitadoras (N=155), todas las muestras

	N	Media/ Proporción	DE
Edad (años)	136	31,44	8,09
Años de educación	136	10,05	2,52
Tiempo en el programa (meses)[1]	155	16,79	9,69
Tiempo trabajando con la familia (meses)[1]	155	14,45	9,52
Número de familias que visita (por mes)	136	9,85	2,89
Número de visitas por familia (por mes)	136	4,33	2,06
Proporción para los que ser facilitador es su ocupación principal (semana pasada)	136	0,98	
Número de horas que trabaja como facilitadora (por semana)	136	11,27	5,12
Proporción con experiencia con niños entre 0 y 3 años	136	0,26	
Años de experiencia con niños entre 0 y 3 años	35	2,38	3,08
Proporción con experiencia trabajando con familias	136	0,27	
Años de experiencia trabajando con familias	37	2,52	3,39
Proporción que ha sido capacitada en los últimos 3 años	136	0,59	
Proporción que ha sido capacitada en:			
Cuidado de niños	136	0,19	
Higiene	136	0,49	
Salud	136	0,41	
Nutrición	136	0,42	
Desarrollo infantil	136	0,21	
Aprendizaje temprano	136	0,14	
Proporción que fue capacitada al entrar al Programa	136	0,90	
Número de capacitaciones recibidas (adicional a la capacitación cuando ingresó al Programa)	136	3,44	2,53
Proporción capacitada por las supervisoras durante las visitas	136	0,93	
Frecuencia de capacitación por supervisor:			
Cada dos semanas	126	0,63	
Mensual	126	0,29	
Cada dos meses/otro	126	0,07	
Frecuencia del pago			
Mensual	136	0,74	
Cada dos meses	136	0,17	
Trimestral/otro	136	0,09	
Puntaje KIDI (Proporción de respuestas correctas)	136	0,75	0,09

Notas: [1]Información de la lista de chequeo administrada a los supervisores. El resto de variables proviene de la encuesta a hogares y a las facilitadoras.

Tabla 3: Puntajes crudos de ASQ, Bayley y HOVRS-A+2, por muestra

	N	Media/ Proporción	DE	N	Media/ Proporción	DE	N	Media/ Proporción	DE
	Muestra con visita propia			Muestra ASQ			Muestra Bayley		
Puntajes crudos ASQ-III									
Línea de base									
ASQ total (todos los dominios)	218	45,29	11,48	431	45,27	11,38	346	45,60	11,18
ASQ, cognición, lenguaje, y motricidad fina	218	47,45	13,34	431	46,42	13,32	346	47,65	12,81
ASQ, cognición y lenguaje	218	46,67	13,86	431	46,38	14,21	346	47,70	13,49
Seguimiento									
ASQ total (todos los dominios)	218	43,54	9,10	431	43,76	9,68	346	43,42	9,58
ASQ, cognición, lenguaje, y motricidad fina	218	40,47	8,67	431	40,49	9,69	346	40,46	9,36
ASQ, cognición y lenguaje	218	40,30	10,31	431	40,15	11,01	346	40,25	11,04
Bayley-III puntajes crudos									
Bayley total (todos los dominios)	218	42,34	3,16	-	-	-	346	43,49	3,51
Bayley, cognición, lenguaje, y motricidad fina	218	46,38	3,20	-	-	-	346	47,46	3,62
Bayley, cognición y lenguaje	218	48,27	3,25	-	-	-	346	49,24	3,54
HOVRS-A+2 puntajes crudos[1]									
HOVRS, Dominio de Prácticas del Facilitador	218	3,43	0,52	431	3,44	0,42	346	3,43	0,43
HOVRS, Dominio de Involucramiento Familiar	218	4,47	0,73	431	4,50	0,56	346	4,49	0,56
HOVRS, todos los dominios	218	3,88	0,56	431	3,90	0,45	346	3,88	0,46

Notas: [1] En la muestra con visita propia, el puntaje de la visita del niño; en las muestras ASQ y Bayley, el puntaje promedio de todas las visitas del facilitador del niño, sin incluir el puntaje del propio niño.

Tabla 4: Correlaciones de Pearson entre el ASQ y el Bayley con la calidad del ambiente en el hogar y la calidad de la visita domiciliaría, por muestra

	Calidad del ambiente en el hogar/ Conocimiento de los padres				Calidad de la visita domiciliaria (HOVRS)					
	FCI actividades de juego	FCI materiales de juego	HOME	KIDI	PF[1]	IF[1]	Total[1]	PF[2]	IF[2]	Total[2]
Panel A: Muestra con visita propia (N=218)										
ASQ total (todos los dominios)	0,05	0,16**	0,20***	0,07	0,09	0,22***	0,18***	0,11*	0,20***	0,17**
ASQ, cognición, lenguaje y motricidad fina	-0,03	0,05	0,08	0,04	0,12*	0,20***	0,18***	0,10	0,17**	0,14**
ASQ, cognición y lenguaje	-0,02	0,02	0,08	-0,01	0,17**	0,24***	0,22***	0,13*	0,19***	0,17**
Bayley total (todos los dominios)	0,10	0,17**	0,11*	0,09	0,00	0,16**	0,09	0,05	0,10	0,08
Bayley, cognition, language, and fine motor skills	0,10	0,14**	0,11	0,10	0,00	0,16**	0,09	0,06	0,11	0,09
Bayley, cognition and language skills	0,09	0,17**	0,14**	0,12*	0,00	0,16**	0,10	0,07	0,10	0,09
Panel B: Muestra ASQ (N=431)										
ASQ total (todos los dominios)	0,10*	0,22***	0,20***	0,09	-	-	-	0,12**	0,21***	0,17***
ASQ, cognición, lenguaje y motricidad fina	0,03	0,09*	0,13**	0,05	-	-	-	0,13**	0,20***	0,18***
ASQ, cognición y lenguaje	0,04	0,06	0,10*	0,01	-	-	-	0,16***	0,22***	0,21***
Panel C: Muestra Bayley (N=346)										
ASQ total (todos los dominios)	0,10*	0,22***	0,20***	0,09	-	-	-	0,12**	0,21***	0,17***
ASQ, cognición, lenguaje y motricidad fina	0,03	0,09*	0,13**	0,05	-	-	-	0,13**	0,20***	0,18***
ASQ, cognición y lenguaje	0,04	0,06	0,10*	0,01	-	-	-	0,16***	0,22***	0,21***
Bayley total (todos los dominios)	0,10*	0,15***	0,05	0,03	-	-	-	0,03	0,07	0,05
Bayley, cognición, lenguaje y motricidad fina	0,10*	0,14***	0,05	0,03	-	-	-	0,03	0,07	0,05
Bayley, cognición y lenguaje	0,06	0,14***	0,06	0,03	-	-	-	0,05	0,08	0,07

Notas: Puntajes en el seguimiento, PF=Prácticas del facilitador; IF=Involucramiento Familiar, Coeficientes significativos al * p<0,10, ** p<0,05, *** p<0,01
[1]Puntaje de la visita al niño, neto del efecto de codificador.
[2]Puntaje promedio de todas las visitas del facilitador del niño, sin incluir el puntaje del niño propio y neto del efecto del codificador.

Tabla 5: Asociación entre el Desarrollo infantil y la calidad de la visita domiciliaria, 'muestra con visita propia' (N=218)

	ASQ total (todos los dominios)		ASQ, cognición, lenguaje, y motricidad fina		ASQ, cognición y lenguaje		Bayley total (todos los dominios)		Bayley, cognición, lenguaje, y motricidad fina		Bayley, cognición y lenguaje	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Niña	0,212	0,190	0,271*	0,279	0,163	0,145	0,191	0,161	0,122	0,098	0,190	0,153
	(0,152)	(0,165)	(0,161)	(0,168)	(0,141)	(0,152)	(0,152)	(0,156)	(0,148)	(0,150)	(0,159)	(0,159)
Educación de la madre (años)	0,027	0,020	-0,001	-0,005	-0,026	-0,026	0,004	-0,003	0,013	0,007	0,002	-0,011
	(0,027)	(0,026)	(0,038)	(0,035)	(0,035)	(0,033)	(0,039)	(0,041)	(0,039)	(0,041)	(0,040)	(0,042)
Puntaje total ASQ en línea de base (todos los dominios)	-0,083	-0,073	-0,085	-0,065	-0,129	-0,121	-0,122	-0,119	-0,110	-0,112	-0,161	-0,146
	(0,140)	(0,147)	(0,171)	(0,173)	(0,164)	(0,170)	(0,157)	(0,168)	(0,153)	(0,161)	(0,167)	(0,172)
Número de visitas recibidas	-0,005	-0,005	-0,002	-0,003	-0,001	-0,001	0,002	0,002	0,001	0,001	0,004	0,004
	(0,004)	(0,004)	(0,004)	(0,004)	(0,004)	(0,004)	(0,004)	(0,003)	(0,003)	(0,003)	(0,004)	(0,003)
HOVRS total puntaje Z	0,198**	0,191**	0,196*	0,175	0,214*	0,206*	0,066	0,058	0,066	0,054	0,016	0,007
	(0,090)	(0,089)	(0,115)	(0,111)	(0,123)	(0,119)	(0,067)	(0,068)	(0,071)	(0,072)	(0,077)	(0,071)
Número de actividades de juego FCI puntaje Z		-0,015		-0,055		-0,035		-0,007		-0,006		-0,021
		(0,076)		(0,080)		(0,087)		(0,095)		(0,093)		(0,103)
Número de materiales de juego FCI puntaje Z		0,074		-0,052		0,024		0,054		0,020		0,108
		(0,110)		(0,118)		(0,107)		(0,085)		(0,081)		(0,093)
HOME puntaje z (codificado a la inversa)		0,118		-0,005		0,025		0,070		0,078		0,114
		(0,088)		(0,085)		(0,091)		(0,089)		(0,086)		(0,090)
KIDI puntaje z		0,037		0,078		0,012		0,038		0,042		0,097
		(0,077)		(0,081)		(0,095)		(0,064)		(0,065)		(0,068)

Notas: Los resultados en la línea de seguimiento se expresan en puntajes Z. Otros controles incluyen las características en la línea de base del niño y su familia presentadas en la Tabla 1, el puntaje total ASQ-3 (todos los dominios) de la línea de base y su función polinomica de cuarto grado, el promedio al nivel de distrito de todos los controles del niño y del hogar en la línea de base de la Tabla 1, el puntaje promedio de a nivel de distrito en comprensión de lectura y matemáticas de la "Evaluación Censal de Estudiantes", y el promedio a nivel de distrito de los controles a nivel de hogar del Censo Nacional de Perú 2007. Coeficientes significativos al * p<0,10, ** p<0,05, *** p<0,01.

Tabla 6: Asociación entre el Desarrollo infantil y la calidad de la visita domiciliaria, 'muestra ASQ' (N=431)

	ASQ total (todos los dominios)		ASQ, cognición, lenguaje, y motricidad fina		ASQ, cognición y lenguaje	
	(1)	(2)	(1)	(2)	(1)	(2)
Niña	0,126	0,097	0,200**	0,175*	0,153*	0,124
	(0,085)	(0,090)	(0,094)	(0,095)	(0,091)	(0,090)
Educación de la madre (años)	0,027	0,017	0,024	0,018	0,014	0,008
	(0,020)	(0,019)	(0,021)	(0,021)	(0,022)	(0,021)
Puntaje total ASQ en línea de base (todos los dominios)	0,053	0,066	-0,037	-0,019	-0,117	-0,097
	(0,106)	(0,108)	(0,112)	(0,112)	(0,111)	(0,111)
Número de visitas recibidas	-0,001	-0,001	-0,000	-0,000	-0,002	-0,002
	(0,003)	(0,003)	(0,003)	(0,003)	(0,003)	(0,003)
HOVRS total puntaje Z	0,149**	0,152***	0,157**	0,160**	0,187***	0,189**
	(0,056)	(0,056)	(0,065)	(0,065)	(0,069)	(0,071)
Número de actividades de juego FCI puntaje Z		0,077		-0,011		-0,000
		(0,053)		(0,052)		(0,053)
Número de materiales de juego FCI puntaje Z		0,148**		0,093*		0,095**
		(0,064)		(0,055)		(0,047)
HOME puntaje z (codificado a la inversa)		0,077		0,043		0,067
		(0,074)		(0,060)		(0,058)
KIDI puntaje z		-0,007		0,009		0,009
		(0,045)		(0,041)		(0,051)

Notas: Los resultados en la línea de seguimiento se expresan en puntajes Z. Otros controles incluyen las características en la línea de base del niño y su familia presentadas en la Tabla 1, el puntaje total ASQ-3 (todos los dominios) de la línea de base y su función polinomica de cuarto grado, el promedio al nivel de distrito de todos los controles del niño y del hogar en la línea de base de la Tabla 1, el puntaje promedio de a nivel de distrito en comprensión de lectura y matemáticas de la "Evaluación Censal de Estudiantes", y el promedio a nivel de distrito de los controles a nivel de hogar del Censo Nacional de Perú 2007. Coeficientes significativos al * p<0,10, ** p<0,05, *** p<0,01.

Tabla 7: Asociación entre el Desarrollo infantil y la calidad de la visita domiciliaria, 'muestra Bayley' (N=346)

	ASQ total (todos los dominios)		ASQ, cognición, lenguaje y motricidad fina		ASQ, cognición y lenguaje		Bayley total (todos los dominios)		Bayley, cognición, lenguaje y motricidad fina		Bayley, cognición y lenguaje	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Niña	0,051	-0,004	0,140	0,107	0,070	0,030	0,115	0,087	0,061	0,039	0,108	0,074
	(0,095)	(0,102)	(0,107)	(0,112)	(0,107)	(0,111)	(0,116)	(0,116)	(0,115)	(0,116)	(0,119)	(0,117)
Educación de la madre (años)	0,020	0,010	0,018	0,013	0,010	0,005	0,012	0,010	0,015	0,014	0,010	0,006
	(0,024)	(0,023)	(0,026)	(0,025)	(0,026)	(0,025)	(0,024)	(0,025)	(0,025)	(0,026)	(0,026)	(0,027)
Puntaje total ASQ en línea de base (todos los dominios)	0,072	0,098	0,004	0,028	-0,053	-0,035	0,056	0,072	0,047	0,057	0,037	0,064
	(0,131)	(0,133)	(0,150)	(0,150)	(0,144)	(0,142)	(0,111)	(0,118)	(0,105)	(0,110)	(0,116)	(0,120)
Número de visitas recibidas	-0,001	-0,001	-0,000	0,000	-0,002	-0,001	0,002	0,002	0,001	0,001	0,004	0,004
	(0,003)	(0,003)	(0,003)	(0,003)	(0,004)	(0,004)	(0,003)	(0,003)	(0,003)	(0,003)	(0,003)	(0,003)
HOVRS total puntaje Z	0,207***	0,217***	0,224***	0,226***	0,254***	0,257***	0,036	0,041	0,041	0,046	0,003	0,005
	(0,067)	(0,072)	(0,078)	(0,079)	(0,084)	(0,085)	(0,062)	(0,065)	(0,066)	(0,069)	(0,068)	(0,069)
Número de actividades de juego FCI puntaje Z		0,045		-0,032		0,012		-0,022		-0,009		-0,055
		(0,056)		(0,055)		(0,055)		(0,064)		(0,066)		(0,063)
Número de materiales de juego FCI puntaje Z		0,200***		0,104		0,125**		0,120**		0,102*		0,140**
		(0,068)		(0,068)		(0,057)		(0,057)		(0,053)		(0,061)
HOME puntaje z (codificado a la inversa)		0,110		0,045		0,042		-0,008		-0,008		0,011
		(0,070)		(0,067)		(0,072)		(0,066)		(0,063)		(0,070)
KIDI puntaje z		0,021		0,016		0,007		-0,017		-0,019		0,006
		(0,054)		(0,050)		(0,053)		(0,064)		(0,060)		(0,061)

Notas: Los resultados en la línea de seguimiento se expresan en puntajes Z. Otros controles incluyen las características en la línea de base del niño y su familia presentadas en la Tabla 1, el puntaje total ASQ-3 (todos los dominios) de la línea de base y su función polinomica de cuarto grado, el promedio al nivel de distrito de todos los controles del niño y del hogar en la línea de base de la Tabla 1, el puntaje promedio de a nivel de distrito en comprensión de lectura y matemáticas de la "Evaluación Censal de Estudiantes", y el promedio a nivel de distrito de los controles a nivel de hogar del Censo Nacional de Perú 2007. Coeficientes significativos al * p<0,10, ** p<0,05, *** p<0,01.