

Desarrollo y selección de medidas del bienestar infantil

Howard White y Shagun Sabarwal

CENTRO DE INVESTIGACIONES INNOCENTI DE UNICEF

El Centro de Investigaciones Innocenti es la oficina de UNICEF especializada en investigación científica. El objetivo principal del Centro de Investigaciones es mejorar la comprensión internacional de una serie de cuestiones relacionadas con los derechos de la infancia, a fin de facilitar la plena aplicación de la Convención sobre los Derechos del Niño en todo el mundo. El centro tiene el cometido de establecer un marco integral de investigación y conocimiento dentro de la organización para brindar apoyo a los programas y políticas mundiales de UNICEF, y trabaja con los asociados para formular políticas con base empírica en favor de la infancia. Las publicaciones elaboradas por el centro contribuyen al debate global sobre la infancia y los derechos del niño e incluyen una amplia gama de opiniones.

Las opiniones expresadas corresponden a los autores o editores y se publican para estimular un mayor diálogo sobre métodos de análisis de impacto. Esta publicación no refleja necesariamente las políticas o perspectivas de UNICEF.

SINTESIS METODOLOGICAS DEL CENTRO DE INVESTIGACIONES

Las síntesis metodológicas del Centro de Investigaciones de UNICEF pretenden compartir prácticas de investigación, métodos, diseños y recomendaciones de reconocidos investigadores y analistas. La audiencia a la que van dirigidas es principalmente el personal de UNICEF que lleve a cabo, encargue o interprete los resultados de investigación y análisis para la toma de decisiones sobre programas, políticas y actividades de sensibilización.

Esta síntesis metodológica ha seguido un proceso de revisión por pares interna.

El texto no ha sido editado de acuerdo con los estándares de publicación oficiales y UNICEF declina toda responsabilidad por posibles errores.

Se permite la reproducción de cualquier parte de la presente publicación siempre que se incluya referencia a la presente. Si se desea utilizar una parte sustancial o la totalidad de la publicación dirijan su solicitud al Departamento de Comunicación en la dirección de correo electrónico: florence@unicef.org

Para consultas o descargas, pueden encontrar estas síntesis metodológicas en <http://www.unicef-irc.org/KM/IE/>

Recomendamos la siguiente cita para cualquier referencia al presente documento:

White, H., y S. Sabarwal (2014). Desarrollo y selección de medidas del bienestar infantil, *Síntesis metodológicas: evaluación de impacto n.º 11*, Centro de Investigaciones de UNICEF, Florencia.

Agradecimientos: Varios autores han proporcionado orientación en la preparación de esta síntesis. El autor y el Centro de Investigaciones de UNICEF desean agradecer a todos aquellos que han participado en la preparación de la presente publicación, especialmente a:

Por su contribución: Greet Peersman

Por su revisión: Nikola Balvin, Claudia Cappa, Sarah Hague, Hiroyuki Hattori, Priscilla Idele, Julia Krasevec, Rolf Luyendijk, Holly Newby

© Fondo de las Naciones Unidas para la Infancia (UNICEF), septiembre de 2014

Centro de Investigaciones Innocenti de UNICEF

Piazza SS. Annunziata, 12

50122 Florence (Italia)

Tel.: (+39) 055 20 330

Fax: (+39) 055 2033 220

florence@unicef.org

www.unicef-irc.org

1. INDICADORES: BREVE DESCRIPCIÓN

Los indicadores proporcionan una señal a los responsables de adoptar decisiones ya que indican si —y hasta qué punto— una variable de interés (como el uso de los servicios de salud) ha cambiado. Los indicadores pueden utilizarse en todos los niveles del marco de resultados, desde los insumos al impacto, y deben estar relacionados con la [teoría del cambio](#) del programa (véase la Síntesis n.º 2 (La teoría del cambio)). Cualquier programa o política debe contar con un conjunto amplio de indicadores que abarquen todos los niveles de la teoría del cambio. Los indicadores desempeñan una función importante en todas las actividades de [seguimiento](#) y [evaluación](#), incluida la [evaluación de impacto](#) (véase la Síntesis n.º 1 (Sinopsis de la evaluación de impacto)).

Lo más importante en los niveles inferiores de la cadena causal son los indicadores de seguimiento, como los insumos (por ejemplo, el dinero gastado y los kits de inmunización suministrados), las actividades (por ejemplo, los trabajadores de la salud capacitados y los días de inmunización) y los productos (por ejemplo, las clínicas construidas). No obstante, en el caso de los indicadores de resultados y de impacto de nivel superior, el seguimiento nos indica lo que ha sucedido, pero no por qué. Para comprenderlo, debe utilizarse la evaluación de impacto. Esto simplemente significa que, aunque los indicadores pueden vigilar el progreso —como los progresos alcanzados en la consecución de una meta concreta (por ejemplo, la reducción de las tasas de mortalidad de los menores de cinco años)— o proporcionar una señal de advertencia sobre lo que no va bien (por ejemplo, el 90% de los niños en edad escolar de un pueblo no asisten a la escuela), la comprensión de los factores que explican el logro o no de la meta requiere una [investigación](#) o una evaluación más intensiva.

Las diferentes organizaciones y documentos pueden emplear definiciones distintas de un indicador, una medida, un objetivo, una norma de referencia, un estándar o un [índice](#), y algunos de estos términos también se pueden usar de manera intercambiable. De ahí la utilidad de comprobar el significado de los términos utilizados en un contexto específico. Esta síntesis emplea las siguientes definiciones, que se basan en los indicadores fundamentales del Fondo de las Naciones Unidas para la Infancia (UNICEF) para medir sistemas de protección infantil¹:

- Una **medida** es un dato sobre un programa o política; puede ser cuantitativa (número de niños inmunizados en una región en un año) o [cualitativa](#) (en qué medida funciona bien el sistema de inmunización de un país).
- Una **variable** es una medida cuantificable, claramente definida, de una cantidad que puede variar a lo largo del tiempo en una unidad específica de observación, o en el espacio, es decir, en las distintas unidades de observación.
- Un **indicador** es una medida verificable seleccionada por la dirección del programa o política a fin de tomar decisiones al respecto. Por ejemplo, el porcentaje de estudiantes que aprueban una prueba estandarizada.
- Un **objetivo** es el valor de un indicador, o un cambio en el mismo, que se prevé lograr para un momento determinado en el tiempo. Un ejemplo lo constituyen las metas de los Objetivos de Desarrollo del Milenio, como: «Asegurar que, en 2015, los niños y niñas de todo el mundo puedan terminar un ciclo completo de enseñanza primaria»².
- UNICEF emplea los términos **norma de referencia** e **índice** para referirse a lo mismo —un conjunto de indicadores conexos que facilitan comparaciones significativas, precisas y sistemáticas sobre el desempeño. Por ejemplo, el índice de desarrollo humano (IDH) es una estadística compuesta

¹ Fondo de las Naciones Unidas para la Infancia, *Measuring and Monitoring Child Protection Systems: Proposed Regional Core Indicators for East Asia and the Pacific*, Strengthening Child Protection Series No. 3, UNICEF EAPRO, Bangkok, 2012.

² Naciones Unidas, «Objetivo 2: Lograr la enseñanza primaria universal, Meta 2.A: Asegurar que, en 2015, los niños y niñas de todo el mundo puedan terminar un ciclo completo de enseñanza primaria», página web, Objetivos de Desarrollo del Milenio en 2015 y en adelante, Naciones Unidas, <http://www.un.org/es/millenniumgoals/education.shtml>.

sobre la esperanza de vida, la educación y la renta, empleada para clasificar los países en cuatro niveles de desarrollo humano.

- Un **estándar** es un conjunto de normas de referencia, índices o indicadores que brinda información significativa desde el punto de vista del desempeño. En el ejemplo anterior del IDH, los estándares son los cuatro niveles de desarrollo humano: muy alto, alto, medio y bajo. En esta serie de síntesis también se hace referencia a los estándares como [«rúbricas»](#) (véase la Síntesis n.º 4 (Razonamiento evaluativo)).
- Un **hito** representa el logro de un objetivo cuantitativo o cualitativo, por ejemplo, la mejora de las tasas de matriculación como objetivo cuantitativo, o la introducción de la gestión escolar como objetivo cualitativo. El hito puede definir la variable que se utilizará para medir la consecución del hito.

Esta síntesis se centrará principalmente en los indicadores, incluidos los indicadores específicos para los niños. Al diseñar las actividades de seguimiento y evaluación de las intervenciones relacionadas con la infancia, es importante tener en cuenta que varios aspectos del bienestar infantil, como los derechos, la salud y la educación de los niños, requieren indicadores específicos por edad.

Los indicadores pueden ser cuantitativos o cualitativos. Los indicadores cuantitativos son numéricos, por ejemplo, la tasa de inmunización en un país en concreto, el porcentaje de mujeres que asisten a clases prenatales o la tasa de mortalidad de los menores de cinco años. Los indicadores cualitativos usan categorías de clasificación, por ejemplo, si un país ratificó o no la Convención sobre los Derechos del Niño, o si el nivel de ejecución de un programa es «alto», «medio» o «bajo» (cada una de estas categorías está claramente definida).

Los criterios SMART³ suelen emplearse con el objeto de especificar qué características se requieren para que un indicador se considere idóneo (véase el cuadro 1).

Cuadro 1. Indicadores SMART

Los criterios SMART son un estándar común empleado para los indicadores. Aunque existen algunas variaciones en cuanto a lo que representa cada letra en el acrónimo SMART, las siguientes consideraciones son importantes.

Específico: el indicador se define con claridad y precisión, por ejemplo, un indicador específico de la «cobertura vacunal» incluiría cuántas dosis de la anatoxina antitetánica debe recibir una embarazada para que se la contabilice como vacunada.

- ¿Es el indicador lo bastante específico para medir el progreso hacia la obtención del resultado?
- ¿Resulta claro qué es exactamente lo que se está midiendo? ¿Se ha especificado el nivel adecuado de desglose? ¿Resulta claro «cuál es el buen camino» (es decir, en qué dirección se espera que evolucione el indicador)?
- ¿Refleja el indicador la esencia del resultado deseado?
- ¿Refleja las diferencias entre las zonas y las categorías de personas?

Medible: el indicador se especifica de tal manera que resulta claro lo que se está midiendo, por ejemplo, «la proporción de escuelas del proyecto que han constituido consejos estudiantiles».

- ¿Se puede establecer un objetivo cuantificado para el indicador?

³ Desarrollado originalmente por Doran, George. T. «There's a S.M.A.R.T. way to write management's goals and objectives», *Management Review* (AMA FORUM), 70 11), 1981, págs. 35-36.

- ¿Es posible verificar los cambios de manera objetiva?
- ¿Refleja una medida fiable y clara de los resultados?
- ¿Es sensible al cambio?

Viable: el indicador debe ser apropiado para la política o programa. Por ejemplo, es poco probable que un indicador que mide únicamente la información proporcionada sobre los riesgos del VIH/sida permita determinar si han cambiado las conductas de riesgo.

- ¿Son los cambios previstos como resultado del programa o política realistas?

Pertinente: el indicador debe ser adecuado para la meta o actividad que se pretende reflejar. Por ejemplo, la mortalidad de los menores de cinco años es un indicador de impacto pertinente si el resultado deseado es una mayor cobertura vacunal, pero los indicadores antropométricos no lo serían tanto.

- ¿Está el indicador asociado de manera plausible con la esfera de influencia del programa o política?
- ¿Refleja el indicador la esencia del resultado deseado?

Rastreadable: debe ser posible recopilar datos para el indicador durante un período establecido de tiempo.

- ¿Es realmente posible disponer de los datos con un esfuerzo y un costo razonables?

Puntos principales

1. Los indicadores proporcionan una señal a los responsables de tomar decisiones ya que indican si —y hasta qué punto— una variable de interés ha cambiado.
2. Los indicadores pueden usarse en todos los puntos de la cadena de resultados: insumos, actividades, productos, resultados e impactos.
3. En cualquier intervención, el sistema de seguimiento y evaluación debe incluir un conjunto equilibrado de indicadores, incluida una variedad de estos en toda la cadena causal.
4. Hay cuatro posibles fuentes de datos para los indicadores: datos administrativos, datos de encuestas o censos existentes, datos de seguimiento del proyecto o datos recopilados en un nuevo estudio. Cada una de estas fuentes tiene sus ventajas e inconvenientes.

2. ¿CÓMO SELECCIONAR O DESARROLLAR INDICADORES?

Los indicadores pueden usarse en todos los puntos de la cadena de resultados: insumos, actividades, productos, resultados e impactos. Se los considera diferentes «tipos» de indicadores⁴:

- insumos: recursos financieros, humanos y materiales empleados en un programa o política. Por ejemplo, los materiales de capacitación elaborados.

⁴ Programa Conjunto de las Naciones Unidas sobre el VIH/SIDA, *Basic Terminology and Frameworks for Monitoring and Evaluation, UNAIDS Monitoring and Evaluation Fundamentals*, ONUSIDA, Ginebra, 2010. Véase http://www.unaids.org/en/media/unaids/contentassets/documents/document/2010/7_1-Basic-Terminology-and-Frameworks-MEF.pdf.

- actividades: acciones emprendidas o procesos mediante los que se movilizan insumos para generar productos determinados. Por ejemplo, las sesiones de asesoramiento que se adhieren a unas normas de calidad.
- productos: efectos inmediatos de las actividades de un programa o política, o los productos o entregables directos de estas actividades. Por ejemplo, el número de vacunas administradas.
- resultados: efectos intermedios de los productos de un programa o política, como por ejemplo un cambio en los niveles de vacunación o los comportamientos clave.
- impactos: efectos de largo plazo y acumulativos a lo largo del tiempo de los programas o políticas sobre lo que se proponen cambiar en última instancia. Por ejemplo, la morbilidad y la mortalidad de los menores de cinco años.

En cualquier intervención, el sistema de seguimiento y evaluación debe incluir un conjunto equilibrado de indicadores, incluida una variedad de estos en toda la cadena causal. El reto consiste en seleccionar indicadores que puedan captar los aspectos importantes y que al mismo tiempo sean realistas en cuanto a qué tipo de datos puede recolectarse⁵. Los indicadores de resultado y de impacto emplean, sobre todo, definiciones estándar aceptadas internacionalmente (es lo preferible), mientras que los indicadores en los niveles inferiores de la cadena causal son más específicos del programa o política para el que se define el sistema de seguimiento y evaluación.

Las evaluaciones deben utilizar los indicadores seleccionados inicialmente por quienes diseñaron la intervención como indicadores del desempeño de un programa o política. Estos, a su vez, deben haberse creado en referencia a la teoría del cambio del programa o la política. A menudo, los indicadores clave del desempeño ya se han definido y recolectado en el nivel de los resultados o del impacto, y constituyen un subconjunto de indicadores compilados en el sistema de información para la gestión de una intervención en particular. También puede incluirse un conjunto clave de indicadores en un «panel de control de indicadores», una herramienta de gestión que ofrece una visión gráfica de la situación actual (y a menudo también las tendencias a lo largo del tiempo) de los indicadores clave del desempeño, a fin de mostrar de «un vistazo» cómo está progresando la intervención. Cuando se carezca de indicadores adecuados, deben crearse ex post sobre la base de la teoría del cambio del programa en consulta con las partes interesadas clave.

En la medida de lo posible, es preferible seleccionar indicadores estándar con definiciones comúnmente aceptadas, que facilitan la comparación de las constataciones entre los estudios. Por ejemplo, los datos antropométricos sobre el peso y la estatura por lo general se expresan como [puntuaciones Z](#) mediante el empleo de normas internacionales elaboradas por la Organización Mundial de la Salud (OMS). Del mismo modo, el diseño de instrumentos de recolección de datos debe basarse en modelos existentes «probados y comprobados» como la [Encuesta Demográfica y de Salud \(EDS\)](#), un programa respaldado por ONUSIDA, o la [Encuesta de Indicadores Múltiples por Conglomerados \(MICS\)](#), una iniciativa impulsada por UNICEF. No obstante, el empleo de indicadores existentes presenta algunas desventajas, sobre todo porque podrían no adaptarse bien al contexto local. Estos inconvenientes pueden mitigarse mediante un enfoque consultivo de la selección o la medición de los indicadores, a fin de garantizar que los indicadores existentes se ajusten de la mejor manera posible al contexto y las preferencias locales.

Quizá en ciertas ocasiones haya que adaptar los indicadores al entorno cultural específico. Por ejemplo, puede usarse una EDS para recopilar datos con ánimo de elaborar indicadores sobre el empoderamiento de las mujeres, pero las preguntas de la [entrevista](#) subyacente (y en consecuencia la definición de los indicadores) variarán según la región. En Asia Meridional, por ejemplo, se interroga a las encuestadas sobre su capacidad para salir de casa, pero las mismas preguntas no son pertinentes en gran parte de África Subsahariana. En esos contextos, se emplean, en cambio, preguntas relativas al control de los recursos por parte de las mujeres y las actitudes hacia la violencia de género. Es importante asegurarse

⁵ Programa de las Naciones Unidas para el Desarrollo, «RBM in UNDP: Selecting indicators», *Signposts of Development*, PNUD, Nueva York. Véase http://toolkit-elections.unteamworks.org/?q=webfm_send/105.

de que los datos recogidos sobre los indicadores corresponden realmente a la información que se busca. Las preguntas y las herramientas de la [encuesta](#) que recopilan datos sobre los indicadores deben someterse a una prueba piloto con un número menor de encuestados en las zonas en las que finalmente se administrará la encuesta.

Suele resultar difícil efectuar mediciones objetivas y exactas de los cambios complejos que tienen lugar en el ámbito del desarrollo. En cambio, se recurre a indicadores indirectos, los cuales se suele considerar que están estrechamente relacionados con los cambios previstos o próximos a ellos. Por ejemplo, la proporción de mujeres parlamentarias puede utilizarse como un indicador indirecto de la participación de las mujeres en los procesos nacionales de toma de decisiones. También pueden emplearse indicadores indirectos cuando es demasiado difícil o costoso medir el indicador deseado, o si lo que se pretende medir tarda mucho tiempo en lograrse.

Una vez seleccionados, los indicadores deben [operativizarse](#) definiendo cómo se recolectarán los datos correspondientes y quién lo hará. Si se emplea una encuesta, deben formularse las preguntas específicas y probarse sobre el terreno.

3. ¿CÓMO RECOLECTAR INDICADORES SOBRE EL BIENESTAR INFANTIL?

Variables de los indicadores

Los indicadores pueden elaborarse a partir de los siguientes tipos de variables:

- Las [variables continuas](#), que pueden asumir cualquier valor, aunque suelen estar dentro de un intervalo; por ejemplo, la edad, los años de educación, la distancia al centro de salud o educativo más cercano, el gasto en artículos específicos o la temperatura corporal.
- Las [variables cualitativas](#), que pueden adoptar solamente ciertos valores predefinidos; por ejemplo el grupo sanguíneo (A, B, AB u O), la región de residencia y la ocupación principal del cabeza de familia.
- Las [variables dicotómicas](#), que tienen solo dos valores posibles, generalmente 0 y 1 (aunque en la recopilación de datos es más habitual 1 y 2). Las variables dicotómicas se pueden utilizar para las características personales (por ejemplo, hombre = 1, mujer = 2), participación en el programa (no participante = 0, participante = 1) o en una variable como la mortalidad (vivo = 0, muerto = 1).
- Las [variables ordinales](#) son valores cualitativos en los que las respuestas se ordenan por categorías, como el nivel de instrucción o el uso de los servicios de salud. Por ejemplo, los valores para la pregunta «¿Cuál es el nivel superior de estudios alcanzado?» podrían ser: 1) Educación primaria incompleta; 2) Educación primaria completa; 3) Educación secundaria incompleta; 4) Educación secundaria completa; y 5) Nivel terciario. O, en el caso de la pregunta «La última vez que su hijo tuvo diarrea, ¿a quién le pidió ayuda o consejo?», podrían ser: 1) Trabajador de la salud; 2) Curandero; 3) Amigo o familiar; y 4) Nadie.
- La [escala de Likert](#) es un diseño empleado en la elaboración de cuestionarios que permite obtener una escala ordinal basada en la autoevaluación del entrevistado. Por ejemplo, la respuesta a la pregunta «¿Cómo lo trató el personal durante su última visita a un centro de salud?» podría ser: 1) Muy bien; 2) Bien; 3) De manera satisfactoria; 4) Mal; o 5) Muy mal.

Fuentes de datos para los indicadores

Existen cuatro posibles fuentes de datos: datos administrativos, datos de encuestas o censos existentes, datos de seguimiento del proyecto o datos recopilados en un nuevo estudio. La tabla 1 muestra las ventajas e inconvenientes de cada una de estas fuentes.

Tabla 1. Ventajas e inconvenientes de las diferentes fuentes de datos

	Ventajas	Inconvenientes
Datos administrativos	<ul style="list-style-type: none"> • Cobertura nacional • Anuales • Bajo costo 	<ul style="list-style-type: none"> • Podrían estar sesgados por la falta o el exceso de presentación de informes • Podrían excluir al sector privado • Podría no ser posible tener un acceso completo a ellos • Los datos suelen recogerse en los centros; por tanto, no existen datos sobre quienes no usan los servicios • Para el cálculo de las tasas y porcentajes solo proporcionan el numerador, no el denominador.
Censo o encuesta existente	<ul style="list-style-type: none"> • Bajo costo, ya que emplean los datos existentes • Suelen tener cobertura nacional y pueden ser representativos de las principales regiones geográficas • Por lo general, las encuestas nacionales contienen datos muy completos 	<ul style="list-style-type: none"> • El tamaño de la <u>muestra</u> de la encuesta en las esferas del programa podría ser insuficiente • Quizá no incluya preguntas relativas a la intervención • El censo tiene un conjunto limitado de datos • Podría ser obsoleto o no existir para el período de tiempo apropiado
Datos de seguimiento del proyecto	<ul style="list-style-type: none"> • Bajo costo 	<ul style="list-style-type: none"> • No suele haber datos del grupo de comparación • Podrían estar sesgados por la falta o el exceso de presentación de informes
Nueva encuesta	<ul style="list-style-type: none"> • Se puede adaptar a las necesidades de datos • Se puede crear una muestra del tamaño requerido 	<ul style="list-style-type: none"> • Tiempo y costo • Podría resultar problemático garantizar la adecuada calidad de los datos

Los datos administrativos por lo general están disponibles, ya que la mayoría de los países disponen tanto de un Sistema de Información sobre la Gestión Educacional como de un Sistema de Información sobre la Gestión Sanitaria que recogen los datos en los centros (escuela, hospital y dispensario) utilizando un censo anual completado por el director del centro. La información de ambos sistemas puede resultar adecuada para la evaluación de impacto; por ejemplo, en Filipinas se llevó a cabo una evaluación de la gestión escolar mediante el Sistema de Información sobre la Educación Básica⁶. El Sistema de Información sobre la Gestión Educacional y el Sistema de Información sobre la Gestión Sanitaria no siempre cuentan con datos [fiables](#) sobre los resultados, pero los datos del Sistema de Información sobre la Gestión Educacional a menudo se pueden combinar con una base de datos distinta integrada por las puntuaciones medias de las pruebas en el ámbito escolar. No obstante, los datos de los centros de salud no pueden emplearse para medir los resultados en materia de salud, ya que no reflejan el estado de salud de aquellos que no utilizan las instalaciones (es decir, los que recurren al autotratamiento o a un curandero, o que solicitan consejo a un farmacéutico).

Los datos a menudo suelen extraerse de varias encuestas diferentes, por ejemplo, una encuesta de hogares, una encuesta de centros (escuelas o dispensarios), una encuesta de proveedores de servicios (docentes o trabajadores de la salud), *así como* una encuesta comunitaria. Si se miden los resultados educativos de los niños, se necesitan datos sobre el entorno doméstico del niño y la educación de los padres. Estos pueden provenir de una encuesta de hogares, pero los datos sobre las instalaciones escolares deben proceder de una encuesta escolar. Deben recopilarse de manera que sea posible vincular los datos de un niño a los datos de la escuela a la que asiste. La vinculación debe efectuarse mediante el empleo de códigos preasignados a la comunidad y los centros, ya que los nombres y la ortografía no suelen seguir la forma estándar.

Las encuestas más apropiadas son las MICS y las EDS. Ambas son encuestas de hogares en lugar de encuestas de centros, por lo que permiten obtener muestras representativas a nivel nacional, e incluyen información detallada sobre la salud maternoinfantil, la nutrición y la educación. Las encuestas sobre la renta y el gasto de los hogares efectuadas bajo los auspicios de los Estudios sobre la Medición de los Niveles de Vida contienen datos sobre el nivel de instrucción, el gasto en salud y educación, así como indicadores [antropométricos](#) para niños menores de cinco años. Algunos países también llevan a cabo encuestas específicas sobre los niños en el ámbito de los hogares, entre las que puede destacarse el proyecto internacional [Niños del Milenio](#), que ofrece datos sobre Etiopía, Andhra Pradesh (India), el Perú y Viet Nam, y el [proyecto Birth to Twenty](#) de Sudáfrica. Ambos efectúan el seguimiento de cohortes de niños desde su nacimiento.

Cómo recopilar indicadores sobre el bienestar infantil

La recopilación de indicadores para niños presenta desafíos particulares.

En primer lugar, existen numerosas cuestiones específicas de los niños, sobre todo en lo relativo a su situación y necesidades únicas. Por ejemplo, las pruebas de los resultados de aprendizaje y las utilizadas para determinar el desarrollo cognitivo y las habilidades sociales son específicas para la edad, y las cuestiones relacionadas con el entorno familiar de acogida son específicas para la edad y para cada niño. Los indicadores deben reflejar con precisión el desarrollo cognitivo y socioemocional durante la infancia.

En segundo lugar, la medición más apropiada para evaluar el bienestar infantil varía según la edad del niño. Los indicadores no solo deben ser específicos para la edad, como se destacó anteriormente, sino que, a fin de vigilar el progreso o medir el impacto, también puede ser necesario ignorar los cambios absolutos en los resultados, ya que podrían no tener sentido en todas las situaciones. Por ejemplo, la alfabetización no es una cuestión pertinente en el caso de un niño de 2 años, pero lo es en el de uno de 10 años.

⁶ Khattri, Nidhi, *et al.*, «The effects of school-based management in the Philippines: an initial assessment using administrative data», *Journal of Development Effectiveness*, 4 (2), 2012, págs. 277-295.

Por último, existe el problema de la recopilación directa de datos. A menudo, los datos no se pueden recoger directamente si el niño es muy pequeño o si intervienen cuestiones éticas (discutidas más adelante). En ese caso, deben intervenir el o los cuidadores principales. No obstante, esto puede introducir un sesgo, ya que quizá los cuidadores no quieran mostrar una imagen negativa de sí mismos y por tanto sean más propensos a dar las respuestas socialmente deseables (por ejemplo, en relación con la violencia) o deseen proteger al niño u otros miembros de la familia. Una posible forma de contrarrestar este problema es llevar a cabo una verificación independiente siempre que sea posible. Por ejemplo, las estimaciones en materia de inmunización son más fiables si se calculan para los niños provistos de tarjetas de vacunación. Cualquiera que sea el enfoque, los indicadores empleados deben estar formulados y validados correctamente.

Indicadores específicos para niños por sector

Salud

El indicador de salud más intuitivo y sencillo consiste en preguntar a una persona cómo se encuentra ella o sus hijos. La pregunta autoevaluada sobre la salud más habitual es: «En general, ¿diría que su salud es excelente, muy buena, buena, regular o mala?». No obstante, los datos sobre la salud proporcionados por la autoevaluación son muy poco fiables, y está demostrado que las personas pueden interpretar las categorías «excelente», «muy buena», etcétera, de manera muy diferente⁷. Es bien sabido que los estadounidenses afirmaron estar más enfermos que los habitantes del estado indio de Bihar⁸. Por consiguiente, es probable que también exista un sesgo en los informes de los padres sobre la salud de los niños; los más ricos son más propensos a mencionar enfermedades, y más aún en el caso de los niños que en el de las niñas⁹.

Hay dos formas de evitar este sesgo. La primera es ser lo más específico posible acerca de los síntomas. La segunda es utilizar viñetas, también conocidas como viñetas de anclaje, promovidas por la OMS para la Encuesta Mundial de Salud¹⁰. Las viñetas describen brevemente el estado de salud y la incapacidad consiguiente de tres personas; se pide a los encuestados que clasifiquen a las tres en términos de su salud, y que luego se clasifiquen a sí mismos empleando el mismo sistema. Las viñetas son un medio de medición, y su uso culmina en escalas que pueden utilizarse como indicadores.

Los indicadores de mortalidad por lo común se emplean para medir el estado de salud de los niños entre diferentes países, regiones o, en el plano infranacional, entre estados. Entre los indicadores de mortalidad empleados para los niños se incluyen los siguientes: la mortalidad neonatal (la probabilidad de morir durante el primer mes de vida); la mortalidad infantil (la probabilidad de morir antes de cumplir un año); y la mortalidad en la niñez (la probabilidad de morir entre el segundo y el quinto año de vida). Todas las tasas se expresan por cada 1000 nacidos vivos, excepto la mortalidad en la niñez, que se expresa por cada 1000 niños supervivientes a los 12 meses de edad. La mortalidad de los menores de cinco años combina la mortalidad infantil y la mortalidad en la niñez, y proporciona una estimación de la probabilidad de morir antes de cumplir cinco años. La muerte prematura es un indicador claro de la privación de bienestar, que suele describir el grado de desarrollo humano de un país; estos indicadores por lo general no son ambiguos. Existen, sin embargo, dos cuestiones que dificultan la medición de la mortalidad: el tamaño de la muestra y la «censura»; ambas se explican brevemente a continuación.

En el caso de un grupo de niños menores de cinco años, es imposible medir directamente la mortalidad por dos razones. En primer lugar, deben obtenerse datos de los que nacieron durante el período de

⁷ Grol-Prokopczyk, Hannah, *et al.*, «Using Anchoring Vignettes to Assess Group Differences in General Self-Rated Health», *Journal of Health and Social Behavior*, 52 (2), 2011, págs. 246-261.

⁸ Sen, Amartya, «Health: perception versus observation», *British Medical Journal*, 324 (7342), 2002, págs. 860-861.

⁹ *Ibid.*

¹⁰ Organización Mundial de la Salud. Encuesta Mundial de Salud. <http://www.who.int/healthinfo/survey/en/>.

referencia y ya han muerto. Esto se efectúa mediante el módulo «historia de nacimientos», tal como se usa en las MICS y las EDS, aunque en algunos lugares pueden existir reservas culturales respecto a comunicar el nacimiento de niños que ya han muerto (los mortinatos se omiten, ya que la mortalidad se refiere a la proporción de muertes en relación con los nacidos vivos). En segundo lugar, los niños menores de 5 años aún no han cumplido 5 años, y es imposible saber si sobrevivirán hasta entonces. Se trata de un problema conocido como «censura», que puede resolverse por medios estadísticos. Un obstáculo adicional es conseguir los datos correctos sobre la fecha de nacimiento y de muerte, lo que no siempre resulta fácil en un país en desarrollo.

Para medir la mortalidad con la suficiente precisión hacen falta muestras de gran tamaño. Por esta razón, con frecuencia se presentan estimaciones de la mortalidad para intervalos de edad de cinco años en lugar de año por año. Como resultado, la mortalidad generalmente no se utiliza como indicador en las evaluaciones de impacto de los programas o políticas, ya que las exigencias relativas al tamaño de la muestra resultan prohibitivas. En todo estudio específico deben efectuarse cálculos de la potencia estadística, a fin de determinar si el tamaño de la muestra es suficiente.

Además de los indicadores de mortalidad, existen varios indicadores para medir la morbilidad en los niños (aparte de los indicadores de nutrición, tratados más abajo). La morbilidad es más común que la mortalidad. Una forma habitual de medir la morbilidad consiste en preguntar a los cuidadores de los niños si el niño en cuestión ha padecido una enfermedad en particular (la pregunta se hace en relación a una variedad de enfermedades), por lo general dentro de un período de memoria corto. Por ejemplo, tanto las MICS como las EDS preguntan si los niños menores de cinco años en el hogar han tenido diarrea, tos o fiebre durante las últimas dos semanas. Si un niño ha padecido alguna de esas enfermedades, se pregunta a su cuidador el tipo de tratamiento administrado, el tipo de centro de salud visitado, etcétera. Para analizar sus respuestas, es necesario contar con datos sobre la enfermedad del niño para el denominador.

VIH/sida

Los indicadores para los programas sobre el VIH y el sida recogidos mediante encuestas de hogares se refieren a los conocimientos, las actitudes y la conducta sexual. Las preguntas relativas a los conocimientos pueden versar sobre cómo se transmite el VIH, cómo puede prevenirse, las percepciones erróneas acerca del mismo, la transmisión de madre a hijo y a qué servicio hay que acudir para hacerse las pruebas de diagnóstico. Algunas preguntas tratan sobre las actitudes hacia las personas que viven con el VIH. Las preguntas sobre la conducta sexual se refieren a la edad en la primera relación sexual, el número parejas sexuales, tanto recientes como durante toda la vida, el uso del preservativo durante la última relación sexual y la edad de las parejas sexuales recientes. Algunos estudios también incluyen preguntas sobre el uso de alcohol y drogas.

Las intervenciones en materia de prevención, atención y tratamiento del VIH y el sida suelen estar dirigidas a la población en general, sobre todo en los países con una prevalencia alta de VIH. No obstante, en los países con tasas epidémicas bajas o concentradas, el objetivo son principalmente las poblaciones clave con mayor riesgo de exposición al VIH, como los trabajadores sexuales, los hombres que tienen relaciones sexuales con hombres y las personas que se inyectan drogas. La recopilación de datos exactos de las poblaciones clave resulta difícil, ya que dichas conductas se estigmatizan y criminalizan y, por tanto, suelen ocultarse. Algunas experiencias muestran que se puede mejorar la calidad de la recogida de datos si se gana la confianza de estos grupos y se recurre a encuestadores pertenecientes a ellos.

Por último, se pueden recopilar datos sobre el estado serológico de una persona a través de análisis de sangre o frotis practicados durante las encuestas de hogares o en los centros de salud. Dichas pruebas entrañan cuestiones éticas graves, entre ellas la necesidad de establecer protocolos claros para obtener el consentimiento e informar al encuestado del resultado de la prueba, así como garantizar la prestación de asistencia en el caso de que el análisis de VIH arroje un resultado positivo.

Nutrición

Los conocimientos y las creencias nutricionales del cuidador principal pueden influir en los comportamientos clave apropiados para garantizar el estado óptimo de nutrición de los niños menores de un año y los niños pequeños. Existen indicadores estándar y preguntas conexas relativos a prácticas fundamentales como la lactancia materna exclusiva o la introducción oportuna de alimentos sólidos, semisólidos y blandos. El seguimiento de estos indicadores estándar suele formar parte del programa evaluado. En ocasiones también pueden vigilarse ciertas prácticas exclusivas de un contexto determinado que influyen en el estado de nutrición de los niños pequeños. Por ejemplo, en Bangladesh existe la creencia generalizada en los beneficios de reducir la ingesta durante el embarazo, lo que puede tener consecuencias negativas para el feto. Para recopilar y analizar datos sobre estas prácticas para las que no se dispone de métodos estándar se ha de decidir cuidadosamente el método adecuado (es decir, cualitativo versus cuantitativo). Además, se deben desarrollar las definiciones de los indicadores, y validar y comprobar sobre el terreno las preguntas asociadas.

Hay una amplia variedad de prácticas nutricionales sobre las que se puede preguntar. La cobertura estándar clave y los indicadores relacionados con la práctica en este campo incluyen las prácticas de alimentación de los niños menores de 1 año y los niños pequeños según lo descrito por la OMS¹¹, el uso doméstico de sal adecuadamente yodada y la cobertura de suplementos de vitamina A para niños de 6 a 59 meses. Dependiendo de las metas de un programa, también puede resultar útil vigilar los indicadores que influyen en el estado nutricional, como la cobertura de agua y saneamiento.

Las encuestas de hogares no siempre son el método recomendado para recopilar datos sobre los indicadores anteriores. Por ejemplo, la suplementación de vitamina A se comunica generalmente a través de fuentes de datos administrativas. Dado el número de factores que influyen en el estado nutricional, cuando se intenta dar seguimiento a los programas de nutrición puede existir un afán de recabar información sobre numerosos indicadores. No obstante, conviene reconocer que la calidad de los datos se resiente a medida que se añaden más preguntas a las encuestas de hogares, y que el número de indicadores incluidos en ellas debe ser equilibrado con respecto a la viabilidad de la recopilación de datos sobre el terreno.

Los indicadores de impacto sobre el estado nutricional más comunes para niños menores de cinco años emplean magnitudes antropométricas basadas en la edad, la altura y el peso: 1) el peso para la estatura puede usarse para determinar *la prevalencia de la emaciación*, que representa cuántos niños son demasiado delgados para su altura; puede cambiar muy rápidamente y a menudo es estacional; 2) la estatura para la edad se puede utilizar para determinar el *retraso del crecimiento*, que representa cuántos niños son demasiado bajos para su edad y es un buen indicador de traumas nutricionales acumulativos durante los períodos más decisivos del crecimiento; y 3) el peso para la edad es una medida compuesta del estado nutricional que comprende elementos tanto de la emaciación como del retraso del crecimiento. Cada una de estas medidas se expresa como la prevalencia de niños que se encuentran por debajo de 2 desviaciones estándar de la mediana de la población de referencia (patrón de crecimiento infantil de la OMS). Se considera que los niños situados por debajo de 3 desviaciones estándar de la mediana de la población de referencia sufren malnutrición grave (es decir, casos de emaciación, retraso en el crecimiento o peso inferior al normal graves).

Las medidas antropométricas por lo general se toman en los niños de hasta cinco años. En el caso de niños mayores pueden emplearse otras medidas, como la circunferencia del brazo o el índice de masa corporal.

¹¹ Organización Mundial de la Salud, *Indicadores para evaluar las prácticas de alimentación del lactante y del niño pequeño, Parte 1, Definiciones*, OMS, Ginebra, 2008. Véase <http://www.who.int/nutrition/publications/infantfeeding/9789241596664/es/>.

Agua, saneamiento e higiene (WASH)

La falta de acceso a fuentes de agua potable e instalaciones de saneamiento mejoradas, y las prácticas de higiene deficientes son factores destacados de la salud de los niños, sobre todo de las enfermedades diarreicas, que causan el 9% de las muertes de niños menores de 5 años y contribuyen en gran medida al retraso en el crecimiento.

Las preguntas normalizadas de las encuestas de hogares sobre agua, saneamiento e higiene de las MICS y las EDS por lo general se refieren a: 1) el uso de diferentes tipos de fuentes de agua potable y de instalaciones de saneamiento (también hacen un seguimiento de la práctica de la defecación al aire libre —la peor forma de eliminación de excrementos, con las mayores consecuencias para la salud pública—); 2) la presencia en las instalaciones de un lugar específico para el lavado de manos que cuente con agua y jabón; y 3) el tratamiento del agua doméstica, si es que se lleva a cabo.

Se ha observado que estas últimas preguntas sobre el tratamiento del agua doméstica, como las relativas a los conocimientos sobre el lavado de manos y las conductas reales al respecto, producen un gran exceso de información. Debe desalentarse su uso en favor de preguntas mucho más sólidas sobre el uso de fuentes de agua potable e instalaciones de saneamiento concretas, y la presencia de un lugar específico para el lavado de manos con jabón, que se registra a través de la observación.

La medición de resultados más común por lo que respecta a las intervenciones de agua y saneamiento es la de la diarrea infantil. No obstante, resulta difícil estimar fracciones atribuibles, y son habituales los factores de confusión, lo que hace que la reducción en la incidencia de diarrea sea un indicador del éxito menos recomendable, a menos que se emplee un diseño de [ensayo controlado aleatorio](#). Un indicador más exacto utilizado en los últimos años es retraso del crecimiento. La malnutrición crónica que causa el retraso en el crecimiento está fuertemente correlacionada con las prácticas de defecación al aire libre y el saneamiento deficiente. Los brotes de cólera proporcionan indicadores idóneos para las intervenciones WASH en los niños. No obstante, son escasas, y a menudo el cólera tiene su origen en los alimentos, no en el agua. El ahorro de tiempo, medido por el tiempo de ida y vuelta (es decir, el tiempo que lleva ir a la fuente, obtener el agua —incluido el tiempo de espera en la cola— y volver) es otro indicador de resultado relevante en el caso de las mujeres y los niños que viven en viviendas sin agua, y debe computarse como un resultado en cualquier evaluación de impacto de las intervenciones de WASH.

Educación

La mayoría de los países cuentan con un Sistema de Información sobre la Gestión Educacional, un censo escolar anual que, por lo general, cada escuela completa y remite al Ministerio o el Departamento de Educación. El Sistema de Información sobre la Gestión Educacional recopila datos básicos sobre las instalaciones escolares y el número de profesores y alumnos. Los datos del Sistema de Información sobre la Gestión Educacional constituyen, por tanto, un ejemplo de datos administrativos que, si son lo bastante amplios, pueden usarse para las evaluaciones de impacto. Los datos del Sistema de Información sobre la Gestión Educacional tienen sus ventajas: 1) se supone que abarcan todas las escuelas, en lugar de una muestra, y, por tanto, permiten el desglose en niveles inferiores (por ejemplo, distrito y escuela); y 2) proporcionan información sobre el entorno escolar y la calidad de la enseñanza impartida.

No obstante, algunos Sistemas de Información sobre la Gestión Educacional excluyen al sector privado, o solo lo incluyen de forma voluntaria, lo que puede ocasionar un sesgo. Otra limitación de los datos de esos sistemas es que no se pueden usar para estimar las tasas de matriculación, ya que: 1) se excluye al sector privado; 2) en algunos países pueden existir sesgos que alienten a exagerar el número de alumnos; y 3) el número de alumnos es el numerador de la tasa de matriculación, pero el Sistema de Información sobre la Gestión Educacional no refleja el denominador.

Una manera de superar algunos de los problemas de estos sistemas de Información es emplear encuestas de hogares. Las encuestas de hogares estándar (por ejemplo, la MICS) recopilan datos sobre si los niños

están en ese momento escolarizados y en qué curso o grado, así como el número de años de educación completados en el caso de quienes han abandonado los estudios. También se recogen datos sobre la asistencia de los niños a la escuela el año anterior, lo que permite el cálculo de las tasas de flujo (es decir, repetición, deserción, permanencia, transición, etc.). Así, las encuestas de hogares proporcionan datos de alta calidad sobre las tasas de asistencia en un determinado año escolar, tanto brutas (el número de estudiantes que asisten a la escuela en varios grados diferentes) como netas (asistencia del grupo de edad oficial a un determinado nivel educativo expresado como porcentaje de la población correspondiente), y posiblemente también información sobre la asistencia durante la semana anterior (es necesario analizar estos datos con cuidado para que la ausencia de la escuela durante los días festivos oficiales no se contabilice como falta de asistencia). Los datos de los hogares también pueden usarse para analizar la transición de una etapa de la educación a otra, la terminación de los estudios, la deserción y la repetición.

Otro problema de los datos del Sistema de Información sobre la Gestión Educacional desde el punto de vista del diseño de numerosos estudios es que se recopilan a escala de las instalaciones. Esto ocasiona dos problemas: en primer lugar, como ocurre con la información recogida en los centros, no se refleja la de quienes no acuden a ellos. En segundo lugar, no está vinculada a los datos de los hogares, a pesar de la importancia del entorno familiar en la determinación de los resultados educativos.

Aunque la información sobre la escolarización y la asistencia es relevante para entender el estado de la educación de los niños, generalmente también interesa conocer los resultados de aprendizaje. Es posible que existan datos nacionales sobre las calificaciones en los exámenes que estén disponibles en el Sistema de Información sobre la Gestión Educacional o que puedan vincularse a él a nivel escolar, de la clase o individual. Una ventaja de estos datos es que son nacionales, de modo que pueden desglosarse hasta llegar al nivel de los distritos. No obstante, los datos de los sistemas de exámenes nacionales pueden ser inadecuados para su uso en las evaluaciones de impacto, por una serie de razones: 1) no se dispone de datos con los identificadores necesarios para vincularlos a determinadas escuelas; 2) los datos no representan una medida exacta de los resultados de aprendizaje debido a la corrupción o a otros abusos en el sistema de exámenes (por ejemplo, los docentes dan a los estudiantes las respuestas para que su escuela obtenga buenos resultados); y 3) los exámenes nacionales no se llevan a cabo en todos los grados, por lo que podrían no efectuarse en los grados incluidos en la intervención. Si se presenta alguno de estos problemas, entonces será necesario recopilar nuevos datos sobre los resultados de aprendizaje.

Una evaluación de impacto generalmente requerirá datos tanto del hogar como del centro. Los ensayos controlados aleatorios en el ámbito escolar constituyen una excepción, aunque incluso en este caso el análisis que puede hacerse resulta limitado si no se dispone de datos del hogar.

Desarrollo del niño en la primera infancia

Otro aspecto de la medición del bienestar infantil que hasta hace poco había recibido mucha menos atención es la importancia de la asistencia a los centros de educación preescolar junto con otros aspectos del desarrollo del niño en la primera infancia.

El cuestionario de la MICS estándar incluye un módulo exhaustivo sobre el desarrollo del niño en la primera infancia. El módulo recoge información sobre varios aspectos relativos a la primera infancia, como la calidad de los cuidados que reciben los niños, su ambiente doméstico (por ejemplo, la disponibilidad de materiales lúdicos y de lectura), su asistencia a la educación preescolar y el apoyo de los adultos en el aprendizaje. El módulo incluye también un conjunto específico de preguntas destinadas a evaluar el desarrollo socioemocional, físico y cognitivo de los niños (aprendizaje y aritmética/alfabetización).

Protección de la infancia

La MICS proporciona una amplia variedad de indicadores relacionados con la protección de la infancia. Las preguntas del cuestionario estándar abarcan las actitudes respecto al maltrato de la esposa, el trabajo infantil, el matrimonio infantil y la mutilación/ablación genital femenina. La MICS también recopila datos sobre disciplina infantil y niños con discapacidad. En la categoría de protección del niño también se incluyen otras cuestiones importantes relativas a la explotación sexual y la violencia contra los niños. Algunas, como la violencia, son delicadas y, por tanto, suscitan preocupaciones éticas y prácticas por lo que refiere a la recolección de datos. (Otras cuestiones, como el matrimonio, no plantean estos problemas).

Las encuestas estructuradas quizá no sean adecuadas para recopilar datos sobre cuestiones delicadas, ya que su enfoque relativamente formal impide la creación de una relación estrecha y no da lugar a hacer muchos sondeos. Las metodologías cualitativas, como las entrevistas en profundidad, los grupos dirigidos de discusión (véase la Síntesis n.º 12 (Entrevistas)) y otros enfoques innovadores (por ejemplo, el uso de mapas) puede resultar más apropiadas, y suelen utilizarse para recopilar datos sobre cuestiones delicadas, como en el estudio sobre el abuso sexual de niñas en las escuelas de Ghana¹².

Índices de pobreza económica y de activos

Por «pobreza» suele entenderse la pobreza económica, aunque se acepta ampliamente que la pobreza es un concepto multidimensional. La pobreza económica se mide mediante los datos sobre la renta o el gasto; por lo general, la medición de este último se considera más fiable. Los índices de activos se han convertido en los más habituales, ya que son relativamente fáciles de recopilar y desarrollar. Muchas encuestas —como las EDS y el Estudio sobre la Medición de Niveles de Vida— recogen datos sobre los activos, pero no sobre la renta o el gasto.

La pobreza económica se mide a nivel de los hogares, de modo que la [incidencia](#) de la pobreza infantil puede estimarse mediante el cálculo de la proporción de niños que viven en hogares pobres. Los cálculos de la pobreza han de ajustarse para considerar las economías de escala en el consumo (los miembros del hogar comparten algunos bienes, por lo que dos personas que viven juntas gastan menos que dos personas que viven solas) y en función de la composición de los hogares (se emplean escalas de adultos equivalentes para tener en cuenta el hecho de que los niños consumen menos que los adultos y las mujeres menos que los hombres). Los hogares más numerosos son generalmente más pobres que los más reducidos; por consiguiente, la incidencia de la pobreza en los niños será mayor que en los adultos de la población nacional en su conjunto.

Hay algunos gastos específicos dedicados a los niños que pueden analizarse a fines de comparación. Por ejemplo, el gasto diferencial en salud y educación para niños y niñas puede resultar relevante. Pero en la práctica resulta difícil estimar el «gasto total» por niño, de modo que es más lógico hablar de «niños que viven en hogares pobres» que de «niños pobres».

Las estimaciones de la pobreza por lo general se calculan utilizando datos de encuestas de hogares. No obstante, una de sus limitaciones es que no tienen en cuenta a las personas que no viven en un hogar, entre las que se incluyen las poblaciones nómadas, las personas sin techo, los niños de la calle, o las que viven en campamentos de refugiados o en instituciones, como los establecimientos para personas con discapacidad y las prisiones.

Desigualdad

Los indicadores que miden cómo se distribuye la renta o cualquier otro recurso entre las personas o los hogares de una población se denominan «mediciones de la desigualdad vertical» y proporcionan una idea

¹² White, Howard, «Using Household Survey Data to Measure Educational Performance: The Case of Ghana», *Social Indicators Research*, 74 2), 2005, págs. 395-422.

del bienestar relativo. Pueden aplicarse a la distribución de la renta o del gasto, o a otros indicadores de impacto, como la mortalidad, el acceso a los servicios y los activos.

Para medir la distribución de la renta o del gasto es necesario medir los ingresos per cápita ajustados al tamaño del hogar y a su composición para tener en cuenta las economías de escala en el consumo y las necesidades diferentes de consumo de los hombres y las mujeres, y de los adultos y los niños.

La desigualdad puede representarse de manera gráfica en una curva de Lorenz, que describe el porcentaje acumulado de renta en relación con el porcentaje acumulado de población.

Existen varias mediciones, de las cuales la más común es el coeficiente de Gini, una medición de la [dispersión estadística](#) de la distribución de la renta que oscila entre 0 (igualdad perfecta) y 1 (desigualdad perfecta, que ocurriría si una sola persona percibiera todos los ingresos). Las proporciones de renta por quintil de población son mediciones más transparentes. En la mayoría de los casos, UNICEF emplea quintiles de renta o de riqueza para medir la desigualdad de la distribución de la renta, con especial atención a los dos quintiles más bajos¹³. Se puede examinar específicamente el porcentaje de la renta (u otro indicador) que recibe el grupo más pobre (y el grupo más rico) y supervisar cómo cambia con el tiempo.

Las mediciones del acceso a servicios básicos como la educación, la salud y la nutrición suelen presentarse por quintil de renta o de activos.

La desigualdad entre grupos (denominada «desigualdad horizontal»), como la existente entre hombres y mujeres o entre diferentes grupos religiosos, resulta importante para identificar los grupos que sufren exclusión social.

Todas las mediciones anteriores pueden emplearse para medir la desigualdad en un momento determinado o los cambios en la desigualdad a lo largo del tiempo. Mediciones como el coeficiente de Gini y, en cierta medida, el porcentaje de renta y de activos por quintil, necesitan mucho tiempo para registrar cambios sustanciales.

Finalmente, se distingue entre desigualdad de oportunidades y desigualdad de resultados. La igualdad de oportunidades se plasma en la igualdad de acceso a oportunidades adecuadas en materia de educación, salud y empleo, o a créditos y otros requisitos para el autoempleo. La igualdad de resultados se centra en la equidad u otro tipo de indicadores clave de resultados y de impacto, como la renta, la mortalidad y los derechos.

4. CUESTIONES ÉTICAS Y LIMITACIONES PRÁCTICAS

Las cuestiones éticas en torno a las evaluaciones en general y, específicamente, a la recopilación de datos y la elaboración de informes relacionados con los niños, se han abordado en otros escritos (véase la Síntesis n.º 1 (Sinopsis de la evaluación de impacto), la Síntesis n.º 3 (Criterios de evaluación), y la Síntesis n.º 12 (Entrevistas)). Puede obtenerse más información en el recurso en línea sobre la materia [Ethical Research Involving Children](#).

Existen también cuestiones éticas y prácticas relacionadas con la recopilación y el uso de datos sobre los indicadores, entre ellos:

- datos inexactos o no disponibles
- corrupción de los datos (errores en los datos ocurridos durante la redacción, la lectura, el almacenamiento, la transmisión o el procesamiento de datos informatizados que introducen cambios no intencionados en los datos originales)

¹³ Stewart, Frances, «Approaches towards Inequality and Inequity: Concepts, measures and policies», *Office of Research Discussion Paper* No. 2013-01, Centro de Investigaciones de UNICEF, Florencia, 2013. Véase http://www.unicef-irc.org/publications/pdf/stewart%20inequality_inequity_layout_fin.pdf.

- manipulación de los datos (distorsión intencional de la definición del indicador o de los datos para alcanzar los estándares de desempeño)
- medir lo que se puede medir con facilidad en lugar de lo importante (algunos resultados e impactos, en especial los relacionados con el desarrollo y los derechos del niño, son más difíciles de medir, y a veces resulta más costoso hacerlo. Pero no debe constituir un motivo para que los investigadores recurran a medir lo que puede medirse en lugar de lo que debe medirse.

Todas estas cuestiones pueden conducir a que se extraigan falsas conclusiones sobre el programa o política (o sus componentes).

5. ¿CÓMO UTILIZAR LOS INDICADORES EN LA EVALUACIÓN DE IMPACTO?

El propósito de los indicadores correctamente seleccionados en la gestión basada en los resultados es vigilar de manera rutinaria las condiciones clave de la implementación de un programa o política, así como los resultados específicos derivados de la intervención (por lo general los productos directos, y a menudo también los resultados), con el fin de fundamentar las decisiones sobre el programa o la política. Como se señaló antes, una evaluación de impacto puede recurrir a los datos existentes complementados por otras fuentes. Idealmente, la medición del desempeño en curso y la evaluación en profundidad deben diseñarse de modo que funcionen como un sistema integrado¹⁴.

A veces se cree que para la evaluación de impacto solo se requieren datos sobre los indicadores de resultados o de impacto. No obstante, para el análisis de la cadena causal también se necesitan datos sobre las variables intermedias, como los cambios en los conocimientos y la conducta. Además, se ha observado que, en la cadena de resultados, suelen pasarse por alto los procesos específicos (y su calidad) mediante los cuales las actividades del programa convierten los insumos (los diferentes tipos de recursos) en resultados (productos, resultados e impactos) —un descuido denominado el «vacío intermedio»—. Entre esos procesos figuran las disposiciones institucionales (incluidas las gubernamentales) y los procedimientos que determinan la toma de decisiones, y el mecanismo mediante el cual se implementan —o no— esas decisiones (denominado «gobernanza»). Si bien ya se han elaborado los indicadores para evaluar los insumos, los productos y los resultados en la mayoría de los sectores, incluida la protección de la infancia, las herramientas disponibles para medir los indicadores de gobernanza aún se encuentran en una etapa incipiente¹⁵.

Un reto importante de la evaluación es que muchas de las llamadas «mediciones», incluso si se efectúan con cuidado y precisión, son únicamente indicadores de variables relevantes, y por tanto deben complementarse con datos triangulados. En otras palabras, los indicadores solo indican; no explican. La constatación de que se ha producido un cambio no explica por qué ha ocurrido. Los indicadores constituyen solo una parte del análisis lógico y sustantivo necesario para entender el contexto, el programa o política y sus efectos, incluido el impacto. La Síntesis n.º 1 (Sinopsis de la evaluación de impacto) proporciona más información sobre la combinación de métodos en las etapas de diseño y análisis.

¹⁴ Barrados, María, y J. S. Blain, «Improving Program Results Through the Use of Predictive Operational Performance Indicators: A Canadian Case Study», *American Journal of Evaluation*, 34 (1), 2013, págs. 45-56.

¹⁵ Fondo de las Naciones Unidas para la Infancia, «Measuring and Monitoring Child Protection Systems: Proposed Regional Core Indicators for East Asia and the Pacific», *Strengthening Child Protection Series* n.º 3, UNICEF EAPRO, Bangkok, 2012.

6. EJEMPLOS DE INDICADORES EMPLEADOS EN LOS ESTUDIOS DE UNICEF

Las evaluaciones de impacto de UNICEF utilizan una serie de indicadores, algunos de los cuales son habituales en las intervenciones para el desarrollo —como la renta per cápita y la escolarización— y otros son específicos de las intervenciones —como las preguntas sobre el abuso de sustancias entre los adolescentes— (tabla 2). Todos ellos son indicadores de resultados y de impacto que han demostrado ser apropiados para evaluar las intervenciones. Como se mencionó anteriormente, es mejor emplear los indicadores existentes en lugar de diseñar otros nuevos: el empleo en Pakistán de las escalas de Bayley sobre el desarrollo de lactantes y niños en sus primeros pasos es un buen ejemplo de esta práctica.

Tabla 2. Indicadores empleados en los estudios de UNICEF seleccionados

Estudio	Indicadores
«Innovative Features in Conditional Cash Transfers: An impact evaluation of Chile Solidario on households and children», <i>Innocenti Working Paper</i> No. 2012-03, Centro de Investigación de UNICEF Innocenti, Florencia, 2012.	<p>Monetario: la renta del hogar per cápita (miles de USD)</p> <p>Trabajo: a) empleo del cabeza de familia (%); b) empleo en el sector informal del cabeza de familia</p> <p>Educación: a) el nivel superior de estudios alcanzado de los miembros de la familia; b) escolarización de los niños en edad escolar</p> <p>Salud: usuarios del sistema de salud pública</p>
<i>The Pakistan Early Child Development Scale Up Trial: Outcomes on Child Development, Growth and Health</i> , Yousafzai, A.K., et al., PEDS Trial Outcome Data Report, redactado para UNICEF Pakistan, Karachi, 2012.	<p>Desarrollo en la primera infancia: puntuaciones del desarrollo cognitivo, motor, socioemocional y del lenguaje evaluadas mediante las escalas de Bayley sobre el desarrollo de lactantes y niños en sus primeros pasos (tercera edición)</p> <p>Nutrición: crecimiento indexado mediante puntuaciones Z de longitud/estatura para la edad, peso para la edad y peso para la estatura</p>
<i>The South African Child Support Grant Impact Assessment: Evidence from a survey of children, adolescents and their households</i> , UNICEF Sudáfrica, Pretoria, 2012.	<p>Educación: a) nivel de estudios alcanzado; b) puntuaciones en pruebas de habilidades matemáticas, lectura y vocabulario; c) promedio de días que no asistió a la escuela</p> <p>Trabajo infantil: adolescentes que trabajan fuera del hogar (días de trabajo en el mes anterior)</p> <p>Salud: probabilidad de enfermedad (medida durante los 15 días anteriores a la encuesta)</p> <p>Conductas sexuales de riesgo entre los adolescentes: a) actividad sexual («no ha tenido relaciones sexuales»); b) número de parejas sexuales, embarazo («no ha estado embarazada»)</p> <p>Uso de sustancias: a) alcohol («nunca ha bebido alcohol»); b) drogas («nunca ha consumido drogas»)</p>

7. LECTURAS Y ENLACES CLAVE

Lecturas

Barrados, Maria, y J. S. Blain, «Improving Program Results Through the Use of Predictive Operational Performance Indicators: A Canadian Case Study», *American Journal of Evaluation*, 34 1), 2013, págs. 45-56.

Child Protection Monitoring and Evaluation Reference Group, *Ethical Principles, Dilemmas and Risks in Collecting Data on Violence against Children: A review of available literature*, UNICEF, Sección de Estadística y Supervisión/División de Políticas y Estrategia, Nueva York, 2012. Véase http://www.childinfo.org/files/Childprotection_EPDRCLitReview_final_lowres.pdf. (Consultado el 1 de enero de 2014).

Fondo de las Naciones Unidas para la Infancia, «Measuring and Monitoring Child Protection Systems: Proposed Core Indicators for East Asia and the Pacific», *Strengthening Child Protection Series* No. 3, UNICEF EAPRO, Bangkok, 2012. Véase http://www.unicef.org/eapro/Measuring_and_monitoring.pdf.

Instituto de Estadística de la UNESCO, *Children Out of School: Measuring Exclusion from Primary Education*, Instituto de Estadística de la UNESCO, Montreal, 2005. Véase <http://www.uis.unesco.org/Library/Documents/oosc05-en.pdf>. (Consultado el 2 de enero de 2014).

Kusek, Jody Zall, y Ray Rist, *Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners*, Banco Mundial, Washington D. C., 2004. Específicamente el capítulo 3, «Selecting Key Performance Indicators to Monitor Outcomes», págs. 65-79. Véase http://www-wds.worldbank.org/external/default/WDSPContentServer/WDSP/IB/2004/08/27/000160016_20040827154900/Rendered/PDF/296720PAPER0100steps.pdf.

Lippman L. H., *et al.*, «Positive indicators of child well-being: A conceptual framework, measures and methodological issues», *Innocenti Working Paper* 2009-21, Centro de Investigaciones Innocenti de UNICEF, Florencia, 2009. Véase http://www.unicef-irc.org/publications/pdf/iwp_2009_21.pdf. (Consultado el 1 de enero de 2014).

McCormick, Marie C., «Issues in Measuring Child Health», *Ambulatory Pediatrics*, 8 2), 2008, págs. 77-84.

Perrin, Burt, «Effective Use and Misuse of Performance Measurement», *American Journal of Evaluation*, 19 3), 1998, págs. 367-379.

Stewart, Frances, «Approaches towards Inequality and Inequity: Concepts, measures and policies», *Office of Research Discussion Paper* No. 2013-01, Centro de Investigaciones de UNICEF, Florencia, 2013.

White, Howard, *et al.*, «Comparative Perspectives on Child Poverty: A Review of Poverty Measures», *Working Paper* No. 1, Young Lives, 2002. Véase <http://www.younglives.org.uk/publications/WP/comparative-perspectives-on-child-poverty-a-review-of-poverty-measures>. (Consultado el 28 de diciembre de 2013).

Enlaces

Anchoring Vignettes, sitio web, <http://gking.harvard.edu/vign>.

Banco Mundial, «About LSMS», página web sobre el Estudio sobre la Medición de Niveles de Vida, Banco Mundial, <http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTLSMS/0,,contentMDK:21478196~menuPK:3359066~pagePK:64168445~piPK:64168309~theSitePK:3358997,00.html>.

CE-DAT, «Complex Emergency Database», página web con datos sobre una serie de indicadores para los países en situación de emergencia compleja, Centro de Investigación sobre la Epidemiología de los Desastres, <http://cedat.be/>.

Childinfo, «Multiple Indicator Cluster Surveys (MICS)», página web sobre informes de los países preparados mediante el empleo de datos de la Encuesta de Indicadores Múltiples por Conglomerados, UNICEF, <http://www.childinfo.org/mics.html>.

The DHS Program, página de inicio del sitio web sobre datos e informes de la Encuesta de Demografía y Salud, <http://www.measuredhs.com/>.

Fondo de las Naciones Unidas para la Infancia, «Encuestas de Indicadores Múltiples por Conglomerados (MICS)», página web, UNICEF, http://www.unicef.org/spanish/statistics/index_24302.html.

GLOSARIO

<u>Antropometría</u>	<i>Medición del tamaño y las proporciones del cuerpo humano, por ejemplo las puntuaciones del peso para la estatura.</i>
<u>Dispersión estadística</u>	<i>Variabilidad o diferencial de una variable o una distribución de probabilidad. Algunos ejemplos comunes de mediciones de la dispersión estadística son la varianza, la desviación estándar y la amplitud intercuartílica. Términos relacionados: variabilidad o variación estadística. (Definición de la Universidad de Princeton)</i>
<u>Encuesta</u>	<i>Herramienta de investigación que incluye al menos una pregunta —que puede ser abierta o cerrada— y que emplea un método oral o escrito para formular esas preguntas. El objetivo de una encuesta es obtener información específica sobre un grupo específico o una muestra representativa de un grupo en particular.</i>
<u>Encuesta de Indicadores Múltiples por Conglomerados (MICS)</u>	<i>Iniciativa de encuesta de hogares a escala internacional desarrollada por UNICEF para corregir las carencias de datos necesarios para vigilar la situación de los niños y las mujeres de todo el mundo.</i>
<u>Encuesta Demográfica y de Salud (EDS)</u>	<i>Encuestas de hogares representativas a escala nacional que proporcionan datos sobre una amplia variedad de indicadores de seguimiento y evaluación de impacto en los ámbitos de la población, la salud y la nutrición. (Definición de EDS).</i>
<u>Ensayos controlados aleatorios</u>	<i>Diseño de investigación o evaluación con dos o más grupos seleccionados de forma aleatoria (un grupo experimental y un grupo de control) en los que el investigador controla o introduce una intervención (por ejemplo un nuevo programa o política) y mide su impacto en la variable dependiente al menos dos veces (medición anterior y posterior al ensayo). En concreto, los ensayos controlados aleatorios —que tienen su origen en el contexto clínico y se conocen como la «regla de oro» de la investigación médica y sanitaria— se suelen utilizar para responder a las preguntas de investigación de la evaluación, que tratan de evaluar la eficacia de las intervenciones de un programa o política en entornos de desarrollo.</i>
<u>Entrevista</u>	<i>Método de recolección de datos en el que el entrevistador hace preguntas al entrevistado. Esta metodología es habitual en la investigación cualitativa.</i>
<u>Ética/Investigación ética</u>	<i>Investigación que respeta las directrices sobre lo que es ético, moral y responsable en los contextos de investigación (por ejemplo, no plagiar el trabajo de otros, no presentar datos cuestionables, no destruir u ocultar las fuentes, etc.). La investigación ética es reflexiva y considera explícitamente su impacto tanto en los participantes como en la comunidad en general a lo largo de todo el ciclo de investigación.</i>

<u>Evaluación</u>	<i>Apreciación sistemática y objetiva para determinar la pertinencia, la idoneidad, la eficacia, la eficiencia, el impacto y la sostenibilidad de las iniciativas para el desarrollo, sobre la base de criterios y normas de referencia acordados entre los asociados y partes interesadas clave. Implica un proceso riguroso, sistemático y objetivo en el diseño, análisis e interpretación de la información para responder a preguntas específicas. Proporciona una valoración de lo que funciona y por qué, pone de relieve los resultados previstos e imprevistos, y aporta lecciones estratégicas para orientar a los encargados de adoptar decisiones e informar a las partes interesadas.</i>
<u>Evaluación de impacto</u>	<i>Evaluación que proporciona información sobre los impactos que produce una intervención. Puede realizarse una evaluación de impacto de un programa o política o del trabajo preliminar, como la creación de capacidad, la promoción de políticas y el apoyo a la creación de un entorno propicio. Supone examinar no solo las metas y objetivos, sino también los impactos no previstos.</i>
<u>Fiabilidad</u>	<i>Los datos se miden y recopilan sistemáticamente según las definiciones y metodologías estándar; los resultados son los mismos cuando se repiten las mediciones.</i>
<u>Incidencia</u>	<i>Número de casos nuevos o recién diagnosticados de una enfermedad que ocurren durante un período determinado. La tasa de incidencia es el número de casos nuevos de una enfermedad dividido por el número de personas en riesgo de contraer la enfermedad.</i>
<u>Índice</u>	<i>Conjunto de indicadores relacionados que permiten efectuar comparaciones significativas, precisas y sistemáticas en lo relativo al desempeño. Por ejemplo, el índice de desarrollo humano (IDH). Término relacionado: norma de referencia.</i>
<u>Investigación</u>	<i>Enfoque sistemático para la recopilación y análisis de los datos y la información, con el fin de generar nuevos conocimientos y responder una pregunta específica o probar una hipótesis. La investigación añade nuevos conocimientos y se documenta en formatos que permiten el examen y la duplicación.</i>
<u>Mediana</u>	<i>Medida de la tendencia central. El valor medio de una distribución cuyas puntuaciones están ordenadas de mayor a menor.</i>
<u>Muestra</u>	<i>Población investigada en un estudio en particular. Por lo general, se intenta seleccionar una «muestra de la población» que se considere representativa de los grupos de personas a quienes se generalizan o transfieren los resultados.</i>
<u>Operativización</u>	<i>Proceso de definición de conceptos en términos de procedimientos simples y observables. Una definición operativa establece un concepto o una variable de manera que pueda medirse y observarse. Por ejemplo, la devoción religiosa se puede definir operativamente en términos de la frecuencia de asistencia a la iglesia. Otra forma habitual de operativización es el empleo de constructos y escalas.</i>
<u>Puntuación Z</u>	<i>Puntuación estándar, expresada en términos de desviaciones estándar de la media.</i>

<u>Rúbrica</u>	<i>Herramienta de evaluación que utiliza un conjunto de criterios y una guía de calificación/puntuación predeterminada por los evaluadores. Las rúbricas pueden emplearse para evaluar presentaciones, proyectos, carteras, etc.</i>
<u>Seguimiento</u>	<i>Recolección y análisis de datos a medida que avanza el proyecto para asegurar la adecuación de la evaluación, su diseño y la protección de los participantes.</i>
<u>Teoría del cambio</u>	<i>Explica cómo se entiende que las actividades produzcan una serie de resultados que contribuyen a lograr los impactos finales previstos. Puede elaborarse para cualquier nivel de intervención: un acontecimiento, un proyecto, un programa, una política, una estrategia o una organización.</i>
<u>Validez</u>	<i>Medida en que un estudio refleja o valora con precisión el concepto específico que el investigador pretende medir. Cabe la posibilidad de que un método sea fiable, al medir de manera coherente el mismo concepto, pero no válido. Véase: fiabilidad</i>