
Los siete pecados capitales de la evaluación de impacto*

Cesar G. Victora / Iná S. Santos
Universidade Federal de Pelotas, Brasil.

(The seven deadly sins of impact evaluations)

Ideas clave

- La evaluación del impacto de los programas de salud suele ser limitada. Su ausencia permite el mantenimiento de programas inefectivos y el drenaje de recursos.
- El inicio tardío de la evaluación implica generalmente la ausencia de una línea de base para la interpretación posterior de resultados.
- La interpretación del impacto de un programa debe incluir la evaluación del proceso para poder comprender los factores que intervinieron en el resultado.
- La metodología que se utilice en la evaluación de impacto debe permitir establecer si el programa funciona en condiciones reales.
- La inclusión temprana y continuada de los gestores aumenta la probabilidad de que los resultados de la evaluación se tengan en cuenta en la toma de decisiones posterior.
- Los resultados de un programa dependen del contexto social, político, económico y cultural. Un mismo programa puede obtener resultados distintos en contextos diferentes.
- No sólo es importante evaluar el impacto global de un programa, sino hasta qué punto logra alcanzar a los distintos grupos sociales.

*El texto del artículo se basa en la conferencia Ernest Lluch desarrollada por el autor principal durante el XII CONGRESO SESPAS en Barcelona, España, el 20 de junio de 2007. El título del trabajo fue inspirado por la obra del pintor flamenco Hieronymus Bosch (1450-1516), «Los siete pecados capitales», expuesto en el Museo del Prado, en Madrid.

Correspondencia: Cesar G. Victora.
Universidade Federal de Pelotas. Brasil.
Correo electrónico: cvictora@terra.com.br

Introducción

Hay una creciente necesidad de resultados de evaluaciones de impacto de programas y proyectos en salud pública. La definición de nuevas políticas públicas o la reformulación de las ya existentes requieren, cada vez más, que estén basadas en la evidencia de la efectividad de las acciones propuestas. No obstante, a pesar del gran número de programas normalmente implementados en diferentes lugares del mundo, muy pocos han sido o están siendo objeto de evaluaciones de impacto. Las razones de este vacío son varias, e incluyen desde limitaciones económicas (debido a que no se destinan recursos para la evaluación) hasta la falta de una cultura de evaluación entre los gestores y entre las agencias que financian los programas¹. El objetivo de este artículo es señalar, a la luz de la epidemiología y de la experiencia de los autores, siete limitaciones importantes de la evaluación de impacto de programas en salud pública.

No evaluar

El peor de todos los pecados es no evaluar. El hecho de evaluar posee, como mínimo, el efecto pedagógico de ayudar a construir una cultura favorable a la evaluación. Todos los años, se invierten miles de millones de dólares en programas destinados a mejorar la salud, la educación y otros aspectos sociales en los países en vías de desarrollo. De éstos, sólo algunos se benefician de evaluaciones rigurosas, capaces de determinar su impacto real. Eso se aplica tanto a programas financiados por los propios gobiernos de los países en vías de desarrollo como a iniciativas patrocinadas por agencias internacionales. La ausencia de evaluación lleva a que los programas inefectivos continúen el drenaje de los escasos recursos disponibles, los que podrían ser invertidos en iniciativas con un impacto real sobre la salud de la población¹.

Iniciar tardíamente la evaluación

Idealmente, la planificación de la evaluación de los programas se debería realizar simultáneamente a su concepción y diseño. Así, es más fácil identificar indicadores, realizar a tiempo la recogida de los datos de la línea de base y, en caso necesario, seleccionar grupos control para futuras comparaciones.

Utilizaremos como ejemplo la evaluación de la estrategia de atención integral a las enfermedades prevalentes de la infancia (AIEPI), desarrollada por la Organización Mundial de la Salud y la UNICEF². Esta estrategia se basa en tres pilares: apoyo al sistema de salud, desempeño de los profesionales de la salud, y educación para la salud a la familia y a la comunidad. El estudio de evaluación multicéntrico de la AIEPI incluyó proyectos en cinco países³. En Bangla Desh, la evaluación se inició antes del inicio del programa y mostró un aumento creciente del promedio de consultas por niño sólo unos meses después de la implantación de la AIEPI⁴. Si la evaluación se hubiese iniciado mas tardíamente, se habría perdido la estimación del tiempo de latencia necesario para que el efecto de la intervención sobre el cambio en el comportamiento (que conlleva la búsqueda de servicios de salud), se manifestase. Se habría perdido también la importante observación de que, cuando se implantó el proyecto, los niveles de utilización eran similares en las áreas de intervención y de control.

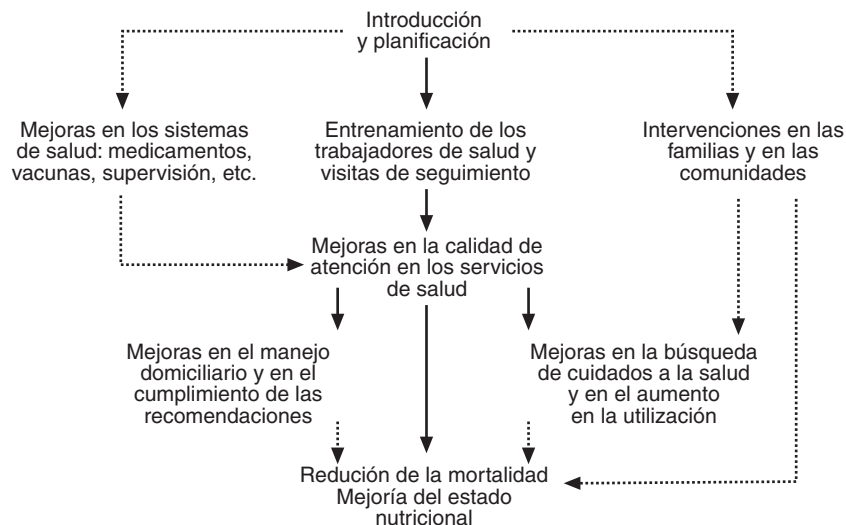
En cambio, el programa de supervivencia y desarrollo infantil acelerado (ACSD, Accelerated Child Survival and

Development) de la UNICEF no comenzó a ser evaluado hasta algunos años después de su implantación. El proyecto, lanzado en 2002 en cerca de 100 distritos de 11 países de África occidental, utiliza las intervenciones más eficaces que se disponen: inmunización de niños y gestantes, distribución de micronutrientes esenciales, promoción de la lactancia materna, suministro de sales de rehidratación oral para el manejo de la diarrea, y mosquiteros tratados para proteger a las gestantes y los niños del paludismo. El primer esfuerzo de evaluación incluyó encuestas de cobertura en diversos distritos 2 años después del inicio de la implementación. La ausencia de una línea de base hace difícil la interpretación de los resultados; por ejemplo, si las diferencias encontradas entre los distritos con y sin ACSD ya estaban presentes antes del inicio de la implementación (datos todavía no publicados; para una descripción del programa, se puede consultar la siguiente página web: http://www.unicef.org/health/index_childsurvival.html).

Evaluar el impacto sin evaluar el proceso

En la figura 1 se muestra de forma esquemática el modelo de impacto de AEDPI y se indican los tres pilares sobre los que se apoya la estrategia de reducción de la mortalidad infantil y la mejora del estado nutricional. Para que el efecto de la intervención pueda verificarse, es necesario documentar mejoras en esos tres elementos. En primer lugar, debe haber mejoras en los sis-

Figura 1. Modelo de impacto de la atención integral a las enfermedades prevalentes de la infancia (AIEPI). Modificado de Bruce et al⁹.



temas de salud que garanticen la disponibilidad de vacunas, medicamentos, la supervisión de la calidad de la asistencia, etc.; en segundo lugar, los profesionales deben ser entrenados en la estrategia y recibir visitas de supervisión para mantener los conocimientos adquiridos y, en tercer lugar, se debe educar a las familias y comunidades. De acuerdo con este modelo, las mejoras en los sistemas de salud y el entrenamiento de los profesionales llevarían a un aumento en la calidad de la atención de los servicios de salud y a la consecuente mejora del manejo de casos y de la adhesión a las recomendaciones. Las intervenciones dirigidas a las familias y a las comunidades, a su vez, conllevarían una adecuada búsqueda de atención en salud y el aumento de la utilización de los servicios de salud. Las mejoras en la calidad de la atención, en el manejo domiciliario de las enfermedades y en el aumento de la utilización conducirían al resultado deseado de reducir la mortalidad infantil y mejora del estado nutricional.

Este tipo de modelo conceptual de impacto es esencial para identificar los indicadores de proceso que requieren ser evaluados y los tipos de encuesta necesarios para medirlos, calcular el posible impacto de la estrategia y estimar el tamaño de la muestra de la evaluación. Su elaboración depende de la colaboración entre los gestores –indicando los cambios que esperan obtener con el programa– y los evaluadores. El modelo propone que la evaluación se realice de forma progresiva y se inicie con la medición de los indicadores de oferta, utilización y cobertura, antes de medir el impacto propiamente dicho⁵. De acuerdo con ese modelo (fig. 2), al inicio se debe investigar si las intervenciones y sus estrategias de implementación están bien fundamentadas técnicamente y son apropiadas al contexto

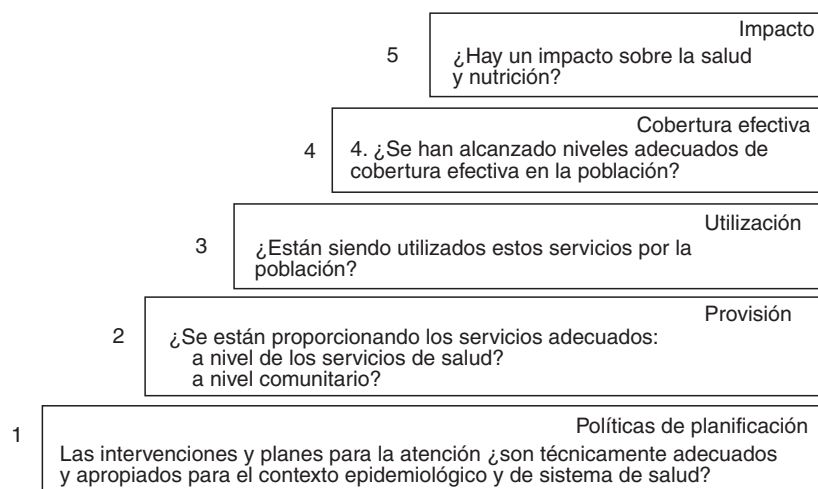
epidemiológico y asistencial del país (evaluación de las políticas y de la planificación basada en resultados). A continuación, si efectivamente se ofrece la atención adecuada en los servicios de salud y en la comunidad (evaluación de oferta). Después, si esta atención es utilizada por la comunidad (evaluación de utilización), de forma que se alcance un nivel adecuado de cobertura efectiva de la población (evaluación de cobertura). Las evaluaciones del impacto sobre la morbilidad y la mortalidad, el estado nutricional u otros indicadores biológicos o del comportamiento únicamente deberían efectuarse si se ha respondido afirmativamente a estas preguntas. La utilización de este procedimiento de evaluación por etapas en la AIEPI, antes descrita, permitió interrumpirla en los países en que las evaluaciones iniciales mostraron que la implementación del programa era insuficiente para alcanzar un impacto detectable en la salud infantil.

La documentación de indicadores de proceso ayuda a entender –en el caso de programas con impacto sobre la salud– qué factores intervinieron en ese impacto; en el caso de los programas sin impacto, la evaluación del proceso permite identificar las medidas que se deben introducir para perfeccionar el programa.

Usar una metodología inadecuada

Muchas evaluaciones de impacto se realizan mediante estudios de eficacia, que investigan si las intervenciones funcionan cuando se implantan en condiciones ideales. En estos estudios, en general, la intervención se implementa de forma integral, con excelente calidad,

Figura 2. Aproximación en escalera a la evaluación. Modificado de Bruce et al³.



y alcanza una alta cobertura poblacional (incluidos los individuos más pobres) y en poblaciones con gran necesidad de tal intervención. Raramente se describen los factores contextuales (esenciales para decidir sobre la capacidad de generalización de los resultados) presentes en los lugares y servicios donde se realizan las evaluaciones.

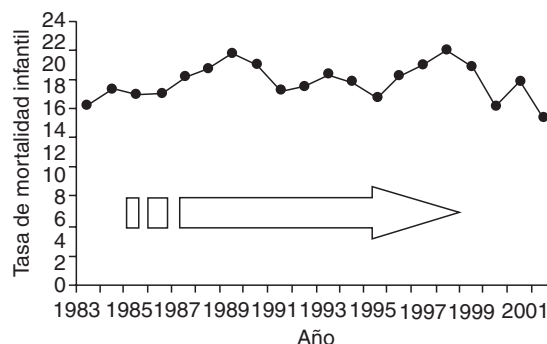
Los estudios de efectividad, por su parte, se planifican para responder si la intervención funciona en condiciones reales. En este caso, las intervenciones se implementan de forma imperfecta, con una calidad intermedia, y alcanzan una cobertura parcial (menor entre los más pobres), en poblaciones con una necesidad de intervención, en general, menos crítica que en los estudios de eficacia. También es posible que estén ausentes ciertos factores contextuales esenciales para el éxito del programa, presentes, en cambio, en las evaluaciones de eficacia.

El modelo de evaluación de Habicht et al⁵ propone un segundo eje (además del eje de los indicadores de oferta, utilización, cobertura e impacto) con tres niveles de inferencia para la evaluación: adecuación, plausibilidad y probabilidad. Las evaluaciones de adecuación responden, por ejemplo, si se alcanzó el objetivo inicial de reducir en un 20% la mortalidad en menores de 5 años, o si las tendencias de los indicadores de impacto iban en la dirección esperada y con la magnitud adecuada. Un programa de reducción de la mortalidad infantil en la ciudad de Pelotas, en el sur de Brasil, iniciado en 2005, por ejemplo, tenía como meta reducir la mortalidad infantil de 20 a 15 por 1.000 nacidos vivos, es decir, menos de 65 muertes anuales en menores de un año. Al final del primer año de intervención, la mortalidad había descendido a 15,5 por 1.000 nacidos vivos, con 66 muertes de menores de un año residentes en el municipio (fig. 3). La evaluación mostró, pues, un éxito parcial.

Los datos de mortalidad por cáncer de cuello uterino en el estado de Río Grande do Sul (Brasil), que muestra la figura 4 para la década de los ochenta, ya señalan una ausencia completa de impacto. Una encuesta de utilización y cobertura del programa, demostró que las mujeres de bajo riesgo repetían muchos exámenes, mientras que las mujeres de alto riesgo no eran atendidas por el programa⁶. Las evaluaciones de adecuación normalmente son muy simples y asequibles a los profesionales con poco entrenamiento en investigación. Principalmente cuando son negativas —es decir, cuando no hubo impacto— pueden ser muy útiles para replantear el programa. En caso de mejora en los indicadores, los métodos descritos más adelante pueden ser útiles para evaluar si el progreso puede ser o no atribuido al programa.

Las evaluaciones de verosimilitud indican si es probable que el impacto observado se deba a la intervención. Estas evaluaciones requieren que se descar-

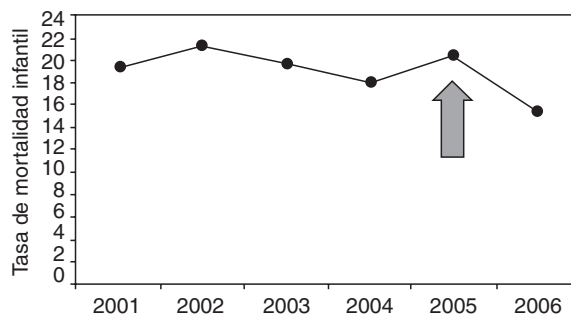
Figura 4. Cáncer de cuello uterino en Río Grande do Sul, Brasil (1980-2000). Modificado de De Quadros et al⁶.



te la influencia de factores externos; por tanto, se requiere un grupo de comparación y de ajuste por factores de confusión. Como la distribución de los grupos no es aleatoria, estos estudios también son conocidos como «casi» experimentales. Por ejemplo, la evaluación del desempeño de los profesionales de salud de 4 municipios en el nordeste de Brasil⁷ mostró que la prescripción adecuada de antibióticos y la administración de la primera dosis de los medicamentos en la propia unidad de salud eran más frecuentes entre los médicos entrenados en la estrategia AIEPI que entre los no entrenados. Los médicos no fueron aleatorizados para recibir o no el entrenamiento, pero sus características personales (sexo, edad, especialidad) fueron tratadas como factores de confusión en el análisis.

Otra evaluación del desempeño sobre la mortalidad infantil entre 1999 y 2002, en 3 estados del nordeste de Brasil, utilizó dos fuentes oficiales de datos: estadísticas vitales de nacimientos y muertes (sistemas SIM y SINASC) y datos de vigilancia poblacional mediante agentes de salud (SIAB). Aunque las diferencias ob-

Figura 3. Programa de reducción de la mortalidad infantil (Pelotas, Brasil). Fuente: Victoria et al (datos no publicados).



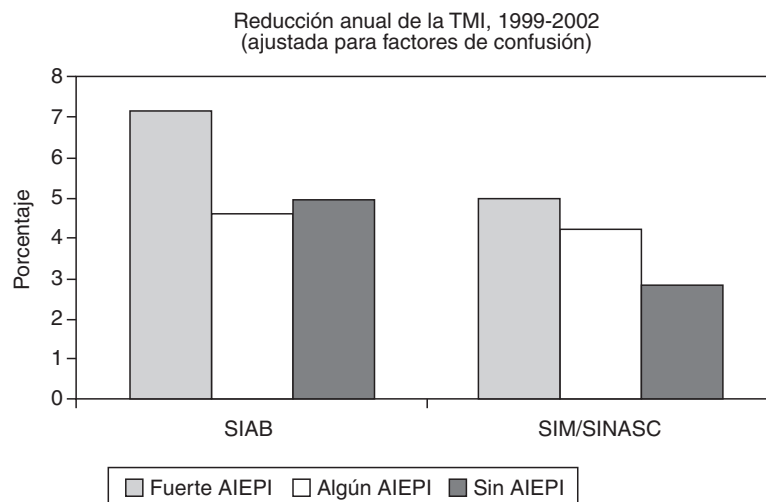
servadas no alcanzaron significación estadística, los resultados fueron similares al usar las dos fuentes de datos: en el análisis ecológico ajustado por las características municipales, las 23 localidades con fuerte implementación de AIEPI revelaron una tendencia a la reducción de la mortalidad en relación con las localidades sin AIEPI (fig. 5). Los datos de SIM/SINASC sugirieron además un efecto del tipo dosis-respuesta, con una reducción moderada de la mortalidad en las localidades con débil implementación de AIEPI, en comparación con los sitios con fuerte implementación o sin AIEPI. Como el número de municipios con AIEPI era muy pequeño (sólo 23 de más de 400), el estudio presentó un poder estadístico bajo para detectar un impacto significativo de AIEPI. Sin embargo, fue tranquilizador observar que los resultados presentaron una tendencia consistente con la mejora de la calidad en la atención, discutida en el párrafo anterior⁸.

Las evaluaciones de probabilidad, conducidas a través de ensayos clínicos aleatorizados, priorizan la validez interna: la aleatorización reduce el riesgo de sesgo de selección y de confusión, y el enmascaramiento reduce el riesgo de sesgo de información. Estos estudios son esenciales para determinar la eficacia de nuevas intervenciones, y son particularmente adecuados para verificar el efecto de intervenciones cuya cadena causal entre la intervención y el resultado es corto, como sucede con el efecto biológico de medicamentos, vacunas, suplementos nutricionales, etc. Además, estos estudios generalmente miden la eficacia y la no efectividad, pues la propia presencia de un grupo de evaluadores puede determinar el desempeño del programa, por el efecto Hawthorne⁹.

Una cadena causal de un ensayo comunitario, aleatorizado y de educación nutricional, por ejemplo, requirió el cumplimiento de, por lo menos, 7 pasos, antes de lograr la mejora nutricional en niños menores de 2 años¹⁰: implementación nacional del programa, entrenamiento de los profesionales de salud, aumento del conocimiento de los profesionales de salud en educación nutricional, mejora del desempeño de los profesionales de salud en educación nutricional, aumento del conocimiento en nutrición de las madres atendidas por los profesionales entrenados, mejora de la alimentación de los niños y aumento de la ingestión de calorías. En 6 meses, los niños del grupo intervención que ingresaron en el estudio con 12-17,9 meses de edad ganaron una media de 1,48 kg de peso, mientras que el grupo control (atendido por médicos no entrenados en educación nutricional) ganaron una media de 1,14 kg ($p < 0,05$). En una larga cadena de causalidad como ésta, es necesario un alto desempeño en cada paso del proceso para que un efecto sea detectado. Por ejemplo, sería necesario que el desempeño en el grupo intervención alcanzara el 80% del nivel ideal en cada uno de los 7 pasos de la cadena (el 80% de los profesionales debería estar bien entrenado, el conocimiento de los profesionales entrenados debería haber aumentado en un 80%, el desempeño de los entrenados debería ser un 80% mejor que el de los no entrenados, etc.) para que se pudiese verificar un impacto del 21% —o sea, un 80% elevado a la séptima potencia— sobre el estado nutricional. Difícilmente se alcanza tal nivel de excelencia en condiciones reales, sobre todo en los países en vías de desarrollo, donde se suelen implementar estos programas.

Esto no significa que las respuestas proporciona-

Figura 5. Impacto de la atención integral a las enfermedades prevalentes de la infancia (AIEPI) sobre la mortalidad infantil en 3 provincias de Brasil. TMI: tasa de mortalidad infantil. Modificado de Amaral et al⁸.



das por los ensayos aleatorizados sean erróneas. Por el contrario, unas respuestas ciertas pueden estar respondiendo a preguntas equivocadas. La pregunta que debe responderse no es si el programa funciona en condiciones ideales, sino si funciona en las condiciones reales de cada país.

Pecado: no involucrar a los gestores

El evaluador debe trabajar junto con los gestores y ser sensible a sus preocupaciones^{5,11,12}. La inclusión de los gestores, desde la fase de diseño de la evaluación, garantiza que ésta responda a las preguntas necesarias para la toma de decisiones y contribuya a que las recomendaciones derivadas se implementen realmente. Dentro de un mismo programa, la información necesaria para la toma de decisiones del gestor local es diferente de la exigida por los gestores de nivel estatal o federal. Para el gestor local, los indicadores de oferta (p. ej., ¿la existencia de vitamina A o de suero de rehidratación oral u otro medicamento es suficiente para atender la demanda del programa?) pueden ser los más importantes. Por el contrario, el gestor del nivel estatal podrá estar interesado en conocer las razones por las cuales algunos lugares logran altas coberturas del programa, para poder alcanzar niveles semejantes en todos los municipios. A menos que esas necesidades queden claras desde el inicio de la evaluación, los resultados de ésta pueden ser de poca utilidad.

Una de las experiencias más frustrantes de nuestra carrera de evaluadores fue un estudio sobre la AIEPI en Brasil. A pesar de haber involucrado a los gestores al inicio de la evaluación, hubo un cambio de gobierno y nosotros no dedicamos el tiempo y el esfuerzo necesarios para garantizar la participación del nuevo equipo. Al final de la evaluación, que demostró que el programa era altamente coste-efectivo¹³, éste fue interrumpido por la nueva Dirección Nacional de Salud.

La inclusión precoz y continuada de los gestores es importante porque ellos conocen a fondo el funcionamiento del programa, y su participación ayuda a dismi-

nuir la tensión que generalmente produce una evaluación. La participación activa de los gestores aumenta la probabilidad de que se tomen decisiones efectivas en función de los resultados de la evaluación. De igual manera, es importante que los resultados de la evaluación estén disponibles para los gestores en el tiempo previsto.

Pecado: ignorar el contexto

Es importante resaltar que, aunque el objetivo de muchos programas sea mejorar un indicador de salud específico, éstos se implementan en un contexto más amplio en el que actúan múltiples fuerzas sociales, políticas, económicas y culturales. Diferentes contextos, a su vez, pueden determinar distintos resultados para el mismo programa¹⁴.

Por ejemplo, el desempeño de profesionales de salud entrenados en AIEPI, respecto al manejo adecuado de las enfermedades infantiles, fue notablemente mejor que el de los no entrenados en 3 de los 4 países evaluados (Bangladesh, Brasil y Tanzania). En Uganda, no obstante, el efecto del entrenamiento fue sustancialmente menor¹⁵. Para entrenar rápidamente a todos los profesionales de salud, se optó en este país por hacer «entrenamiento en cascada», con un curso de corta duración—cada profesional entrenado pasaba a entrenar nuevos profesionales, y así sucesivamente—. Esto implicó una reducción en la calidad del entrenamiento, que explica el peor desempeño de los trabajadores entrenados, en relación con lo observado en los otros 3 países. Este aspecto del contexto se refiere a la forma en que el programa fue implementado.

Otro tipo de factor contextual importante se refiere a la realidad epidemiológica de cada país. Por ejemplo, el efecto de AIEPI sobre la mortalidad en menores de 5 años fue constatado en Tanzania, pero no en Perú. En parte, esto se debe al hecho de que la AIEPI prioriza 5 causas de muerte de niños—neumonía, diarrea, paludismo, sarampión y desnutrición—, que representan la gran mayoría de las muertes producidas en África. En Latinoamérica, las causas neonatales son las más importantes y representan hasta la mitad de las muertes, pero la estrategia AIEPI—como fue implementada originalmente—no priorizaba estas enfermedades.

Tabla 1. Los 7 pecados capitales de la evaluación de impacto

1. No evaluar
 2. Iniciar tardíamente la evaluación
 3. Evaluar el impacto sin evaluar el proceso
 4. Usar un método inapropiado
 5. No involucrar a los gestores
 6. Ignorar el contexto
 7. No evaluar las desigualdades
-

Pecado: no evaluar desigualdades

Aún hoy, las intervenciones reconocidamente capaces de mejorar los niveles de salud en el mundo dejan de atender a una alta proporción de la población que las necesita, en especial a las personas más pobres. Los programas que implementan esas intervenciones

son, en general, heterogéneos en calidad y desempeño, no equitativos y efímeros¹⁶. Una evaluación bien delineada precisa medir no sólo el efecto global del programa, sino también investigar en qué medida están siendo alcanzados los diferentes grupos sociales.

La cobertura del programa AIEPI en el nordeste brasileño, por ejemplo, varía inversamente a la tasa de pobreza del municipio –mayor cobertura en los municipios ricos, donde las condiciones de vida ya eran mejores antes de la implementación del programa¹⁷–. De igual forma, la cobertura del examen de prevención del cáncer de cuello uterino en el sur de Brasil es mayor entre las mujeres con bajo riesgo de desarrollar esta enfermedad⁶. Incorporar estudios de desigualdades sociales en la evaluación de programas de salud, en general, requiere poco esfuerzo adicional y puede suponer una importante contribución para la mejora del programa.

Conclusión

Se han identificado y discutido siete problemas comunes en las evaluaciones de impacto de los programas de salud. El reconocimiento de estos posibles problemas y la adopción de medidas para prevenirlos, sin duda, contribuiría a mejorar la calidad y la relevancia de las evaluaciones de impacto. Los pecados capitales pueden ser expiados a través de una penitencia adecuada. Sin embargo, los pecados resultantes de una evaluación inadecuada pueden causar daños irreversibles –por lo menos a corto plazo– evaluando positivamente un programa inefectivo, o viceversa, y proporcionando a los gestores sugerencias incorrectas sobre si el programa debe ser expandido, modificado o interrumpido. El trabajo conjunto de gestores e investigadores en el diseño, implementación y análisis de evaluaciones rigurosas es esencial para la priorización de medidas de salud pública.

Bibliografía

- Center for Global Development. Report for the Evaluation Gap Working Group. When will we ever learn? Improving lives through impact evaluation. Washington, DC: 2006.
- Tulloch J. Integrated approach to child health in developing countries. *Lancet*. 1999;354 Suppl 2:16-20.
- Bryce J, Victora CG, Habicht JP, Vaughan JP, Black RE. The multi-country evaluation of the integrated management of childhood illness strategy: lessons for the evaluation of public health interventions. *Am J Public Health*. 2004;94:406-15.
- El Arifeen S, Blum LS, Hoque DM, Chowdhury EK, Khan R, Black RE, et al. Integrated Management of Childhood Illness (IMCI) in Bangladesh: early findings from a cluster-randomised study. *Lancet*. 2004;364:1595-602.
- Habicht JP, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *Int J Epidemiol*. 1999;28:10-8.
- De Quadros CA, Victora CG, Da Costa JS. Coverage and focus of a cervical cancer prevention program in southern Brazil. *Rev Panam Salud Publica*. 2004;16:223-32.
- Amaral J, Gouws E, Bryce J, Leite AJ, Cunha AL, Victora CG. Effect of Integrated Management of Childhood Illness (IMCI) on health worker performance in Northeast-Brazil. *Cad Saude Publica*. 2004;20 Suppl 2:209-19.
- Amaral J, Leite AJ, Cunha AJ, Victora CG. Impact of IMCI health worker training on routinely collected child health indicators in Northeast Brazil. *Health Policy Plan*. 2005;20 Suppl 1:142-8.
- Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am J Public Health*. 2004;94:400-5.
- Santos I, Victora CG, Martines J, Goncalves H, Gigante DP, Valle NJ, et al. Nutrition counseling increases weight gain among Brazilian children. *J Nutr*. 2001;131:2866-73.
- United Nations Children's Fund (UNICEF). A UNICEF Guide for Monitoring and Evaluation: making a difference? New York: UNICEF; 1991.
- Santos IS. Guia Metodológico de Avaliação e Definição de Indicadores: doenças crônicas não transmissíveis e Rede Carmen. Brasília: Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Análise de Situação de Saúde, Série A, Normas e Manuais Técnicos, 2007. Disponible en: <http://www.saude.gov.br/svs>
- Amorim DG, Adam T, Amaral J, Gouws E, Bryce J, Victora CG. The Integrated Management of Childhood Illness improves the efficiency of primary health facilities in Northeast Brazil. *Rev Saude Pública*. En prensa 2007.
- Victora CG, Schellenberg JA, Huicho L, Amaral J, El AS, Pariyo G, et al. Context matters: interpreting impact findings in child survival evaluations. *Health Policy Plan*. 2005;20 Suppl 1:18-31.
- Pariyo GW, Gouws E, Bryce J, Burnham G. Improving facility-based care for sick children in Uganda: training is not enough. *Health Policy Plan*. 2005;20 Suppl 1:58-68.
- Victora CG, Hanson K, Bryce J, Vaughan JP. Achieving universal coverage with health interventions. *Lancet*. 2004;364:1541-8.
- Victora CG, Huicho L, Amaral JJ, Armstrong-Schellenberg J, Manzi F, Mason E, et al. Are health interventions implemented where they are most needed? District uptake of the integrated management of childhood illness strategy in Brazil, Peru and the United Republic of Tanzania. *BWHO*. 2006;84:792-801.