

# Evaluación en salud pública: ¿todo vale?

(Evaluation in public health: does anything go?)

«Evaluación» es uno de los términos más utilizados en salud pública. Sin duda, hay un notable —y saludable— interés por conocer el resultado de las intervenciones preventivas y de promoción de la salud, pero no es menos cierto que hay un elevado grado de imprecisión y confusión a la hora de definir qué se pretende con la evaluación en cada caso concreto. En no pocas ocasiones se confunden los indicadores de proceso con los de resultados, o se atribuyen los cambios favorables a una intervención o un programa, sin compararlos con un grupo de control, y sin tratar de descartar las posibles explicaciones alternativas que justifiquen los efectos observados. Los factores que permiten explicar esta situación son múltiples: entre las causas más generales, probablemente hay que señalar una escasa cultura de evaluación, entendida como una parte imprescindible de toda iniciativa que tiene que rendir cuentas (lo que se conoce como *accountability*) ante la sociedad. Como consecuencia, hay una escasa implicación y una limitada formación de los profesionales, en un marco general de ausencia de pautas metodológicas y guías de buena práctica. Además, a diferencia de lo que ocurre en los servicios sanitarios y la gestión sanitaria<sup>1</sup>, donde hay una larga tradición de evaluación y acreditación externas, en muchas ocasiones los programas de salud pública se evalúan por los mismos responsables del programa, lo que supone una menor independencia de criterio y una cierta tentación de autocomplacencia. Finalmente, un elemento adicional que dificulta la posibilidad de realizar una evaluación rigurosa y objetiva guarda relación con la necesidad de dar una respuesta rápida a las preguntas cuando éstas se plantean en los medios de comunicación, porque normalmente éstos no pueden esperar los resultados de los informes finales de evaluación, sobre todo cuando se realizan con mayor rigor, lo que implica tiempos de análisis y elaboración superiores. La reciente ley de tabaquismo o las normativas destinadas a reducir la mortalidad por accidentes de tráfico son ejemplos ilustrativos de ello: independientemente de las evaluaciones formales de su impacto, académicas o administrativas, hemos asistido desde el primer día de su entrada en vigor a noticias, editoriales y valoraciones basados en una diversidad de indicadores, datos y argumentos. Este

interés por la salud pública es positivo, y el hecho de que las valoraciones se apoyen cada vez más en datos objetivos es, asimismo, un buen síntoma. Es legítimo y deseable que la sociedad quiera saber cuál es la utilidad de las medidas de salud pública, especialmente las que afectan nuestros comportamientos cotidianos. Sin embargo, la «evaluación periodística» no puede sustituir la evaluación real, y la profusión de titulares que proclaman (sin matizaciones) el éxito de un programa o una política, y que pueden ser desmentidos por otro titular unos días más tarde, genera una confusión excesiva e innecesaria. Nuestra responsabilidad como profesionales de salud pública es velar por que la evaluación se realice en las condiciones más adecuadas y clarificar en lo posible sus indicaciones, procedimientos y limitaciones, incluidos los necesarios tiempos de espera.

Además de los factores generales descritos, hay algunos condicionantes de carácter metodológico que también contribuyen a dificultar o limitar la posibilidad de realizar de forma adecuada la evaluación cuantitativa de los resultados de las intervenciones en salud pública. Entre éstas, se han señalado principalmente las relacionadas con su naturaleza compleja, ya que suelen estar integradas por diversos componentes<sup>2</sup>, como sucede en el caso de las políticas de salud o los programas dirigidos al cambio de conducta. Por su carácter multicomponente, sus contenidos son difícilmente reproducibles, y su efectividad está en gran parte vinculada al contexto; además, es muy difícil atribuir el efecto a un componente particular del programa. En segundo lugar, la asignación de la intervención impide, en muchas ocasiones, la comparación con un grupo control adecuado, como sucede en los programas o políticas a gran escala, o con programas de tipo social o de otro tipo en los que, por razones éticas, no sea posible dejar de administrar la intervención<sup>3</sup>. Finalmente, en muchas ocasiones el problema estriba en disponer de indicadores de medida de resultados adecuados, bien porque no sean fácilmente visibles o detectables (p. ej., en las intervenciones basadas en actividades culturales o mediáticas, como las obras de teatro, o los recursos educativos de utilización anónima, como los programas o recursos informativos en internet), o porque las consecuencias en salud (p. ej., los cambios en la morbi-

mortalidad que pueden derivarse de una campaña para promover la alimentación equilibrada) sólo sean detectables a muy largo plazo. Estas limitaciones hacen muy difícil, y en ocasiones desaconsejable, la utilización de diseños experimentales en intervenciones de salud pública, lo que hace necesario tener en cuenta las evidencias de los llamados «diseños evaluativos débiles»<sup>4</sup>. Entre las propuestas que avanzan en esta dirección, cabe situar las listas de recomendación, como el Transparent Reporting of Non-experimental Design studies (TREND), que propuso la sistematización y la estandarización de criterios en la descripción de los estudios no experimentales, con el fin de mejorar la calidad de los artículos científicos y las recomendaciones en las que se pueden basar las recomendaciones en salud pública<sup>5</sup>. Otras propuestas avanzan en recomendaciones metodológicas sobre la adaptación de los procedimientos del ensayo clínico a las particularidades de las intervenciones complejas<sup>2</sup>. En el otro extremo, algunos autores proponen nuevas formas de enfocar la valoración de la efectividad de las intervenciones en salud pública asumiendo que habitualmente no podremos disponer de diseños evaluativos experimentales. Así, Victora et al<sup>4</sup> proponen una valoración «progresiva» en dos etapas. En primer lugar, debemos verificar que se han producido cambios favorables, para lo cual únicamente hace falta disponer de indicadores sensibles, asumiendo una magnitud del efecto suficiente; por ejemplo, las ventas de cigarrillos o la prevalencia de fumadores, en el caso de la evaluación de la ley de tabaquismo; esto es lo que Victora et al denominan «adecuación del efecto». En una segunda fase, si hay indicios de que el cambio se ha producido, debemos tratar de demostrar que éste se ha debido efectivamente a la intervención («plausibilidad del efecto»). Para ello, deberemos reunir todas las evidencias disponibles que sean consistentes con las hipótesis y los mecanismos en los que se basa la intervención, incluidos los indicadores de proceso, y descartar sistemáticamente los sesgos potenciales que puedan ser explicaciones al-

ternativas a los cambios observados, como las tendencias seculares o los factores ajenos al programa que hayan podido influir en los resultados. En cualquier caso, se trata únicamente de una propuesta, pero refleja la necesidad de adaptar nuestros enfoques y métodos a la realidad; no hay duda de que en muchos casos tendremos que tomar decisiones y valorar la efectividad de las intervenciones en ausencia de una evidencia científica completa<sup>6</sup>. La combinación de evidencias parciales, pero consistentes —especialmente si proceden de enfoques distintos y complementarios—, puede ser suficiente para actuar, bajo un cierto «principio de prevención»<sup>7</sup>. Pero no es menos cierto que debemos esforzarnos por ser más activos —y probablemente más imaginativos— en tratar de evaluar lo mejor posible las intervenciones de salud pública: si queremos, podemos.

**Manel Nebot**

*Agència de Salut Pública de Barcelona.  
Barcelona. España.*

#### Bibliografía

1. Guix J. La calidad en salud pública. *Gac Sanit.* 2005;19:325-32.
2. Campbell M, Fitzpatrick R, Haines A, Kinmonth A, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ.* 2000; 321:694-6.
3. Thomson H, Hoskins R, Petticrew M, Ogilvie D, Craig N, Ouinn T, et al. Evaluating the health effects of social interventions. *BMJ.* 2004;328:282-5.
4. Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am J Public Health.* 2004;94:400-5.
5. Des Jarlais DC, Lyles C, Crepaz N, por el TREND group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health.* 2004;94:361-6.
6. Kemm J. The limitations of evidence-based public health. *J Eval Clin Pract.* 2006;12:319-24.
7. Nebot M. Health promotion evaluation and the principle of prevention. *J Epidemiol Community Health.* 2006;60:5-6.