

Evidencia de investigación en el sector humanitario

UNA GUÍA PRÁCTICA



AGRADECIMIENTOS

Estamos muy agradecidos con quienes participaron en el evento Evidence Lounge en Ginebra, Suiza, en septiembre de 2017, durante el cual se analizó esta guía, proporcionaron comentarios sobre los primeros borradores de la Guía o ambos: Prisca Benelli (Save the Children), Kate Bingley (Christian Aid), Jonathan Breckon (Alliance for Useful Evidence), Alyoscia D'Onofrio (Comité Internacional de Rescate - IRC), Tarah Friend (Departamento de Desarrollo Internacional del Reino Unido - DfID), Tara Karul (Iniciativa Internacional para la Evaluación de Impacto - 3ie), Paul Knox-Clarke (Red de Aprendizaje Activo sobre Rendición de Cuentas y Resultados de la Acción Humanitaria - ALNAP), Flurina Kuhn (Red de Ciencia e Innovación del Reino Unido), David Loquercio (Norma Humanitaria Esencial en materia de Calidad y Rendición de Cuentas - CHS), Tim Martineau (Escuela de Medicina Tropical de Liverpool - LSTM), Virginia Murray (Public Health England - PHE), Alice Obrecht (ALNAP), Aninia Nadig (Proyecto Esfera), Adina Rom (ETH Zurich), Rodolfo Rossi (Comité Internacional de la Cruz Roja - CICR), Eva Svoboda (Instituto de Desarrollo de Ultramar - ODI y Grupo de Políticas Humanitarias - HPG), Pierluigi Testa (Terre des hommes), Erik Von Elm (Cochrane Suiza) y Karin Wendt (Humanitarian Exchange and Research Centre - HERE).

Además, las siguientes organizaciones contribuyeron con ejemplos (incluidos casos prácticos) y otro contenido: 3ie, ALNAP, Christian Aid, Evidence Aid, Comité Internacional de la Cruz Roja, Comité Internacional de Rescate, Escuela de Medicina Tropical de Liverpool, London School of Hygiene and Tropical Medicine, Instituto de Desarrollo de Ultramar, Public Health England, Save the Children y ShelterBox.



CITA

Este documento se debe citar de la siguiente manera: Blanchet K, Allen C, Breckon J, Davies P, Duclos D, Jansen J, Mthiyane H, Clarke M. (2018) Uso de la evidencia de investigación en el sector humanitario: Una guía práctica. Londres, Reino Unido: Evidence Aid, London School of Hygiene and Tropical Medicine y Nesta (Alliance for Useful Evidence).

AUTORES

Este documento fue redactado por Karl Blanchet^a, Claire Allen^b, Jonathan Breckon^c, Phil Davies^b, Diane Duclos^b, Jeroen Jansen^b, Helen Mthiyane^c y Mike Clarke^b (**a** Health in Humanitarian Crisis Centre, London School of Hygiene and Tropical Medicine; **b** Evidence Aid; **c** Nesta, Alliance for Useful Evidence).

FINANCIACIÓN

Los fondos para este documento fueron proporcionados por la Red de Ciencia e Innovación del Reino Unido (representada por la SIN Suiza, Embajada Británica en Berna).

CORRESPONDENCIA

La correspondencia referente a este documento se debe enviar a Karl Blanchet, Director del Health in Humanitarian Crises Centre, London School of Hygiene & Tropical Medicine, Tavistock Place, Londres, Reino Unido; y Mike Clarke, Director de investigación de Evidence Aid, Centre for Public Health, Queen's University Belfast, ICS Block A, Royal Hospitals, Belfast, Reino Unido.

Karl: Karl.Blanchet@lshtm.ac.uk

Mike: mclarke@qub.ac.uk

EL TRABAJO EN LA GUÍA FUE COORDINADO POR TRES ORGANIZACIONES:

Evidence Aid (www.evidenceaid.org) defiende un enfoque basado en evidencia en el sector humanitario. Su misión es aliviar el sufrimiento y salvar vidas proporcionando la mejor evidencia disponible sobre la eficacia de la acción humanitaria y facilitando su uso. Evidence Aid ha desarrollado una gama de recursos que se encuentran disponibles de manera gratuita en su sitio web. Estos incluyen revisiones sistemáticas de intervenciones para desastres, crisis humanitarias y otras emergencias graves. Evidence Aid organiza eventos para promover el uso de la evidencia en el sector humanitario, incluida una Semana de Evidencia Humanitaria anual que se celebra en noviembre en colaboración con el Centro de Medicina basada en la Evidencia de la Universidad de Oxford, Reino Unido, junto con Evidence Lounges. Los Evidence Lounges reúnen a médicos y miembros de la comunidad de investigación académica para mejorar las colaboraciones y desarrollar maneras prácticas de utilizar la evidencia en el sector humanitario, incluida esta guía.

El **Health in Humanitarian Crisis Centre** (<http://crises.lshtm.ac.uk>) de la London School of Hygiene and Tropical Medicine lleva a cabo investigación primaria y formación sobre salud pública en situaciones de crisis humanitaria, en estrecha colaboración con agencias humanitarias internacionales y centros de investigación en países afectados para abordar los desafíos sanitarios críticos.

En 2018 se lanzó un programa de desarrollo de capacidades e investigación de cuatro años, RECAP, y se ha enfocado en la toma de decisiones y la rendición de cuentas en la respuesta a crisis humanitarias y epidemias. El proyecto RECAP tiene el respaldo de UK Research and Innovation como parte del Global Challenges Research Fund, subvención número ES/P010873/1. El Centro también ofrece cursos para médicos a través de módulos sobre conflicto y salud, cursos cortos y un MOOC gratuito sobre Salud en Situaciones de Crisis Humanitaria.

Nesta (Alliance for Useful Evidence) (www.alliance4usefulevidence.org) es una fundación para la innovación mundial, con base en el Reino Unido. Respaldando nuevas ideas para afrontar los grandes desafíos de nuestra época, desde las presiones del envejecimiento poblacional hasta los servicios públicos insuficientes y un mercado laboral en constante evolución. Ha creado una serie de guías relativas al uso de evidencia para superar este desafío y esta publicación se basa en una de ellas: «Using Research Evidence: A Practice Guide» (Uso de la evidencia de investigación: una guía práctica). Dicha guía fue elaborada por el Equipo de Habilidades de Innovación de Nesta y la Alliance for Useful Evidence; la presente guía para el sector humanitario se desarrolló en consulta con ellos y agradecemos profundamente a Nesta por permitirnos utilizar algunos de sus contenidos.



ÍNDICE

INTRODUCCIÓN 4

¿Quién puede utilizar esta guía?	5
Cómo usar esta guía	5

SECTION A: ¿Qué es la toma de decisiones basada en evidencia y por qué centrarse en la investigación? 6

¿Qué es la «evidencia» y por qué nos centramos en la investigación?	7
El desafío de confiar únicamente en la opinión profesional	9
Mensajes clave de la Sección A	11

SECCIÓN B: ¿Cuándo puede ayudarle la evidencia? 12

Ciclo de Respuesta humanitaria	12
Desarrollo y ejecución de una nueva intervención	14
Adopción de ideas de otras regiones o sectores	18
Creación de una teoría del cambio	19
Mensajes clave de la Sección B	19

SECCIÓN C: ¿Qué tipo de evidencia debe elegir? 21

Parte 1: Diferentes tipos de métodos, diseños y enfoques de investigación	23
Investigación experimental: ¿por qué tanto alboroto?	28
Diseños de investigación experimental	28
Cuestiones prácticas y éticas de los ensayos aleatorios	30
Cuando la aleatorización no es posible o aceptable	30
Diseños cuasi-experimentales	30
Más allá de los experimentos: nuevos enfoques de evaluación	32

¿Mereció la pena? El valor de los análisis de coste-beneficio y de coste-eficacia	32
---	----

Mensajes clave de la Parte 1 de la Sección C	34
--	----

Parte 2: ¿Cómo califica la calidad de la investigación? 35

Revisión por pares: ¿cómo ayuda a los responsables de la toma de decisiones?	35
--	----

Definición de investigación de «alta calidad»	35
---	----

Investigación cualitativa de alta calidad	37
---	----

Evite «elegir a su antojo» usando revisiones sistemáticas	37
---	----

Metaanálisis	39
--------------	----

Diagramas de bosque	39
---------------------	----

Evaluaciones de evidencia rápidas	40
-----------------------------------	----

La importancia de la repetición y la corroboración	41
--	----

Análisis de revisiones: cómo evaluar los conjuntos de evidencia de alta calidad	41
---	----

Mensajes clave de la Parte 2 de la Sección C	42
--	----

SECCIÓN D: ¿Dónde debe buscar la evidencia? 43

Búsqueda de evidencia de investigación	43
--	----

Fuentes de revisiones sistemáticas y otras investigaciones para el sector humanitario	44
---	----

Mensajes clave de la Sección D	45
--------------------------------	----

NOTAS FINALES

Referencias y notas finales	46
-----------------------------	----

Introducción

La evidencia proveniente de la investigación y la evaluación puede ayudarle a comprender qué funciona, dónde, por qué y para quién. También puede indicarle qué no funciona y ayudarle a evitar repetir los fracasos de otros mediante el aprendizaje a partir de las evaluaciones de programas humanitarios infructuosos. La evidencia también puede guiar el diseño de las maneras más eficaces para ejecutar intervenciones específicas.

La evidencia cuestiona lo que podríamos pensar que es sentido común, conocimiento percibido o recibido.

Por ejemplo, por mucho tiempo se creyó que la malnutrición aguda severa requería un tratamiento especializado en centros de hospitalización con productos terapéuticos, a pesar de que este modelo presentaba una serie de desafíos para el tratamiento eficaz tanto para los sistemas de salud como para los pacientes. A mediados de la década de 1990, se desarrollaron alimentos terapéuticos listos para utilizar. En 2000, los proyectos piloto iniciales comenzaron a poner a prueba el enfoque de Gestión Comunitaria de la Malnutrición Aguda (CMAM) durante emergencias humanitarias.¹ Se descubrió que era tan efectivo que recibió el apoyo de agencias de las Naciones Unidas en 2007,² y hoy en día se considera el estándar de cuidado para la gestión de la malnutrición aguda en contextos de emergencia y desarrollo.

Debido a que los recursos financieros y de otros tipos son limitados y a menudo insuficientes en el sector humanitario, no podemos permitirnos malgastar dichos recursos en políticas y programas que no funcionan. Las intervenciones en cualquier sector pueden tener resultados tanto positivos como negativos y consecuencias previstas e imprevistas. Incluso en los casos

en los que invertir dinero para solucionar un problema supuestamente traerá beneficios, necesitamos preguntarnos si ese dinero podría invertirse de manera más eficaz en otro sector, por ejemplo, en una intervención humanitaria diferente. La evidencia de investigación puede ayudarle a tomar decisiones informadas sobre las intervenciones más eficaces que aportar en un contexto de emergencia.

La evidencia es un campo cuestionado, con opiniones discrepantes sobre qué se debería valorar más o considerar más relevante para los responsables de la toma de decisiones.³ No obstante, para el propósito de esta guía práctica, hacemos hincapié en la investigación basada en las nociones científicas de prueba, validez, confiabilidad y que tiene un sesgo reducido. Dicha investigación tiene la ventaja del rigor, la relevancia y la independencia. Nos enfocamos en los estudios de población que apuntan a generar respuestas promedio, que luego necesitan ser contextualizadas, por ejemplo, teniendo en cuenta factores culturales (como el uso de médicos de sexo masculino para examinar y tratar a pacientes de sexo femenino) y asuntos de viabilidad (como la capacidad de realizar una intervención en el período posterior a un desastre, cuando el acceso a las poblaciones o a los recursos puede estar especialmente restringido).

Hemos creado esta guía para ayudarle a aprovechar al máximo la evidencia de investigación cuando se encuentre en una emergencia humanitaria o planificando para una próxima emergencia. Nuestra intención es ayudarle a encontrar y utilizar la evidencia sobre las intervenciones, las medidas y las estrategias que podrían ayudarle a tomar decisiones informadas. Esta guía no se trata de cómo generar más evidencia de investigación. Se trata de utilizar y comprender la evidencia existente y reconocer cuándo hace falta buena evidencia. Debe ayudarle a fomentar su confianza para compilar, asimilar, simplificar e interpretar una base de evidencia sólida de la investigación existente y a pensar cómo podría evaluar sus propios proyectos y encargar investigaciones o evaluaciones.

¿QUIÉN PUEDE UTILIZAR ESTA GUÍA?

Esta guía práctica está dirigida principalmente a los responsables de la toma de decisiones humanitarias y a los médicos que trabajan sobre el terreno o en la sedes de organizaciones donantes, internacionales, nacionales o no gubernamentales. Servirá de ayuda a la hora de tomar decisiones sobre la financiación, la supervisión, la entrega o la evaluación de las intervenciones humanitarias. No está dirigida a evaluadores e investigadores capacitados, sino que intenta fomentar que audiencias más amplias del sector humanitario exijan evidencia de investigación.

CÓMO UTILIZAR ESTA GUÍA

Esta guía está dividida en cuatro secciones principales:

SECCIÓN A ¿Qué es la toma de decisiones basada en evidencia y por qué centrarse en la investigación?

Esta sección analiza a qué nos referimos con la toma de decisiones basada en evidencia y por qué la investigación es un elemento fundamental de esta.

SECCIÓN B ¿Cuándo puede ayudarle la evidencia?

En esta sección se exploran diferentes escenarios en los que el uso de la evidencia puede servirle de ayuda, así como los diferentes tipos de evidencia que puede necesitar en las diferentes etapas de desarrollo o implementación de una nueva intervención o política.

SECCIÓN C ¿Qué tipo de evidencia debe elegir?

Esta sección se centra en los diferentes tipos de evidencia y examina cómo elegir la que mejor se adapte a su situación. También analiza cómo evaluar la calidad de la evidencia.

SECCIÓN D ¿Dónde debe buscar la evidencia?

Esta sección incluye consejos y recursos para ayudarle a encontrar la evidencia que mejor se adapte a su situación.

¿Qué es la toma de decisiones basada en evidencia y por qué centrarse en la investigación?

Esta sección analiza a qué nos referimos con la toma de decisiones basada en evidencia y por qué la investigación es un elemento fundamental de esta.

Para comenzar, seamos claros sobre a qué no nos referimos. No estamos hablando de tomar decisiones siguiendo ciegamente las conclusiones rígidas de la investigación. El criterio profesional y otras fuentes de información, tales como los comentarios de las partes interesadas, siempre serán importantes. Esta guía práctica no se trata de reemplazar el criterio profesional, sino de aumentar el uso de la evidencia en la acción humanitaria.

Una buena manera de comenzar a definir a qué nos referimos con esto se puede tomar prestada de la medicina. Hace más de dos décadas, David Sackett y sus colegas propusieron la siguiente definición que ha perdurado con el paso del tiempo:

*«La medicina basada en evidencia es el uso consciente, explícito y sensato de la mejor evidencia actual en la toma de decisiones sobre el cuidado de pacientes individuales. La práctica de la medicina basada en evidencia significa integrar la experiencia clínica individual con la mejor evidencia clínica externa disponible de la investigación sistemática».*⁴

Este intento por definir la medicina basada en evidencia no fue el primero,⁵ pero ha sido influyente y es igual de relevante para el sector humanitario como para otros sectores. Pone énfasis en cómo la investigación puede complementar el criterio profesional u otras fuentes de información y reconoce la importancia

de la evidencia en asuntos como la viabilidad, la preferencia y la cultura.

En un campo como el del sector humanitario, en el que se requiere más y mejor evidencia,⁶ podríamos asumir que cualquier modelo de buena toma de decisiones debe tener cuidado de no depender únicamente del criterio profesional que no esté avalado por evidencia científica. Más adelante en esta sección, leerá acerca de cómo todos podemos ser «predicablemente irracionales» y, consciente o inconscientemente, cometer errores de criterio importantes. Exploraremos cómo mitigar dichos errores de criterio en las secciones posteriores. Sin embargo, otros modelos de toma de decisiones también han hecho hincapié en la importancia de combinar el conocimiento de la evidencia con el criterio. El sector humanitario es un área sensible en la cual necesitamos tener en cuenta las políticas internacionales y locales y las dinámicas entre los diversos actores involucrados en la prestación de ayuda humanitaria. En ocasiones, esto determinará el acceso a la evidencia y la información, pero también cómo se presta la ayuda humanitaria. No obstante, la importancia de la evidencia perdura y, como se indica en el informe de la ALNAP en 2014, «la incapacidad de producir y utilizar evidencia en la política y en la respuesta hace que la acción humanitaria sea menos eficaz, menos ética y menos responsable».⁷

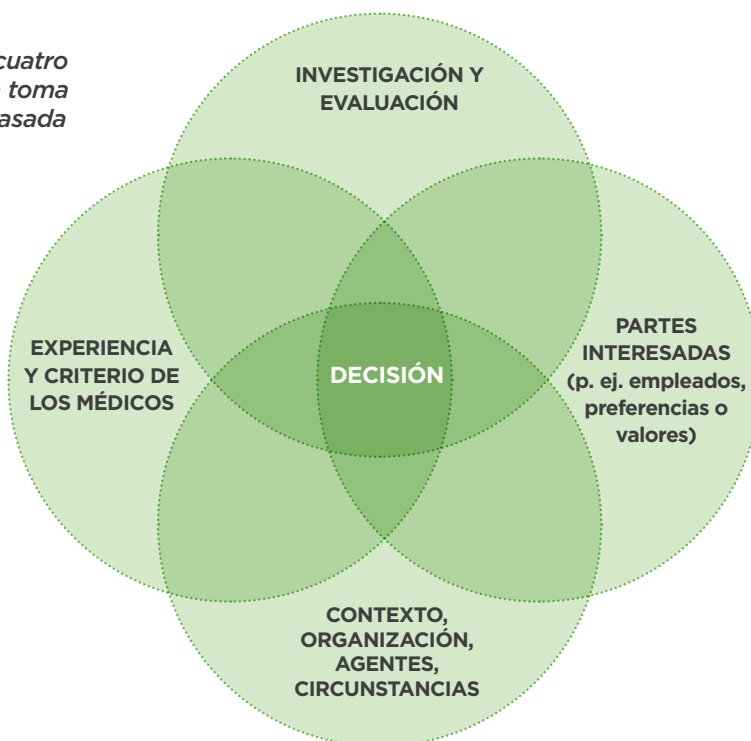
¿QUÉ ES LA «EVIDENCIA» Y POR QUÉ NOS CENTRAMOS EN LA INVESTIGACIÓN?

El Oxford English Dictionary define «evidencia» como «*el conjunto de hechos o la información disponibles que indican si una creencia o proposición es verdadera o válida*»,⁸ y, del mismo modo, en el informe de la ALNAP sobre el estado de la evidencia en el sector humanitario, Paul Knox Clarke y James Darcy la definieron como la «*información que ayuda a corroborar y probar/refutar la verdad sobre una propuesta específica*». ⁷ Nos basamos en estas definiciones porque otras muchas tienden a ser de poca ayuda por ser demasiado inclusivas (en determinadas ocasiones incluyen casi todos los tipos de información) o bien por ser demasiado vagas y abstractas.

La **Figura A.1** muestra los diferentes elementos que deben formar parte de la toma de decisiones basada en evidencia. En esta guía práctica nos centramos en el círculo superior del diagrama: investigación y evaluación.

Tal y como afirman los autores del informe de la Alliance for Useful Evidence «*What Counts as Good Evidence*» (¿Qué se considera una buena evidencia?), *la realización y la publicación de la investigación implican la documentación explícita de la metodología, la revisión por pares y el control externo, dando como resultado rigurosidad y transparencia. Estas características contribuyen al carácter sistemático de la investigación y ayudan a proporcionar un medio para evaluar la fiabilidad de los resultados. También ofrecen la posibilidad de evaluar la validez de una afirmación en comparación con otra.*⁹

Figura A.1: Los cuatro elementos de la toma de decisiones basada en evidencia



Basado en: Barends E, Rousseau DM, Briner RB. (2014) Evidence-based Management: The Basic Principles. Ámsterdam: Center for Evidence-Based Management [www.cebma.org/wp-content/uploads/Evidence-Based-Practice-The-Basic-Principles.pdf]

Esta guía práctica se centra en la investigación, pero hay muchas superposiciones con el campo de la evaluación, de modo que en la Sección C trataremos algunos enfoques para evaluar el impacto y el proceso. También le damos una mayor importancia a la investigación que se ocupa del impacto –si algo ha tenido resultados positivos o negativos– porque las preguntas relacionadas con el impacto son vitales para aquellos que están involucrados en la acción humanitaria. Estos agentes se preocupan por mostrar su «impacto» en las poblaciones, sus «resultados» en términos internacionales o «qué funciona» para los gobiernos y los proveedores locales e internacionales. El lenguaje puede cambiar, pero la idea de su investigación prevalece en el tiempo: comprobar si realmente han conseguido realizar un cambio. Por lo tanto, con esta guía pretendemos ayudarle a decidir cómo dicha investigación puede ayudarle a elegir intervenciones, medidas y estrategias, así como a adoptar las políticas con mayores probabilidades de producir un cambio positivo. A lo largo del documento incluimos ejemplos ilustrativos. Asimismo, en otras colecciones de estudios de caso tiene a su disposición más ejemplos del uso de la evidencia en los sectores humanitarios.¹⁰

Le damos mayor importancia a las investigaciones y las evaluaciones que ya han sido realizadas y se pueden emplear, sin que sea necesario realizar un nuevo estudio. Los responsables de la toma de decisiones tienen un tiempo y unos recursos limitados, de modo que muchos simplemente no

pueden permitirse encargar un estudio de este tipo y esperar hasta que los resultados estén disponibles, lo que puede llevar años. Quien necesita tomar una decisión en este preciso instante, necesita que la evidencia esté disponible ahora, si no ayer, y no dentro de uno o más años. Por consiguiente, los responsables de la toma de decisiones requieren evidencia «lista para usar» y que se pueda combinar con la información del contexto local para fundamentar su elección. Afortunadamente es posible encontrar dicha evidencia y en la Sección D trataremos algunas de las maneras de cómo hacerlo.

La investigación es un proceso participativo con fines didácticos. Busca responder a preguntas, tales como «¿cuál es el tipo de lesión más común después de un terremoto?», «¿cuáles son los efectos en la violencia de género de los diferentes métodos de protección de mujeres y niños?» o «¿cuál es la impermeabilidad de un determinado material cuando es utilizado como refugio?»

La evaluación es un proceso que implica la valoración de los resultados y las observaciones en comparación con los estándares, a fin de tomar decisiones. Las evaluaciones se plantean preguntas, tales como «¿cuál es el tipo de formación en primeros auxilios que debería recibir el personal de primera intervención?», «¿cuál es la mejor manera de proteger a mujeres y niños contra la violencia de género?» o «¿qué material debe emplearse para la construcción de carpas en un asentamiento con lluvias torrenciales?»

Le damos mayor importancia a las investigaciones y las evaluaciones que ya han sido realizadas y se pueden emplear, sin que sea necesario realizar un nuevo estudio.

La investigación no requiere necesariamente una evaluación. Sin embargo, para la realización de una evaluación siempre se requiere una investigación. Una evaluación está relacionada con una intervención que se implementó en su momento, mientras que la investigación es más exhaustiva. Además de incluir evaluaciones, la investigación también puede tratar de responder preguntas conceptuales, como, por ejemplo, a la hora de la planificación de las necesidades que pueden surgir tras un desastre o del desarrollo de una nueva intervención. Sin embargo, independientemente de si pensamos en una investigación o en una evaluación, tenemos que ser conscientes de que la parcialidad puede distorsionar los resultados. En la Sección C nos centramos en algunos de los sesgos que pueden afectar los resultados de un estudio, pero también es importante ser conscientes de los prejuicios de cómo las personas pueden interpretar los resultados de un estudio o que pueden afectar su manera de pensar acerca de la necesidad de una investigación o su impacto. A continuación tratamos algunos de estos sesgos.

EL DESAFÍO DE CONFIAR ÚNICAMENTE EN LA OPINIÓN PROFESIONAL

Una de las razones por las que preferimos una investigación de buena calidad respecto a confiar únicamente en el criterio profesional es el hecho de que prejuicios involucrados en el criterio profesional pueden hacer que éste sea erróneo. Como seres humanos, somos «previsiblemente irracionales»¹¹ y podemos tener sesgos cognitivos que distorsionan nuestra forma de pensar. **Sesgo cognitivo** se refiere a nuestra incapacidad de ser totalmente objetivos, lo que se puede manifestar de diferentes maneras, tales como una distorsión perceptiva, juicios imprecisos e interpretaciones incoherentes o irracionales.

Incluso con las mejores intenciones, los profesionales pueden equivocarse.

Tomemos por ejemplo un tipo de error cognitivo: **sesgo de confirmación**. Se define como la tendencia a centrarse en la evidencia que se ajusta más a lo que creemos, e ignorar o descartar la que no. Incluso los profesionales altamente experimentados pueden caer en esta trampa cognitiva.

Tener en cuenta cómo podemos llegar a conclusiones precipitadas es importante para que seamos precavidos ante los expertos. Pero el sesgo de confirmación también resalta cómo cualquiera –no sólo los expertos– puede ser muy selectivo a la hora de utilizar la evidencia de investigación. Las personas tienden a buscar la evidencia que se ajusta más a sus creencias, ya sea de manera intencional o no. Este hecho puede ser considerado más como una «evidencia basada en política» que como una «política basada en evidencia»: escoger y adaptar la evidencia a las conclusiones que ya hemos fijado en nuestras mentes. Si somos generosos, se podría considerar un error inconsciente, pero en el peor de los casos puede tratarse de una táctica deliberada para apoyar un prejuicio mediante el descubrimiento y la promoción de la evidencia que se ajusta a dicho prejuicio, al mismo tiempo que se ignora todo aquello que no lo hace.

Un fenómeno parecido pero diferente es el del **sesgo de optimismo**, el cual se refiere a «*la tendencia probada de los evaluadores [de proyectos, programas y políticas] a ser demasiado optimistas sobre los parámetros principales del proyecto, incluidos los costes de capital, los costes operativos, la duración del proyecto y la obtención de beneficios*».¹² El sesgo de optimismo puede dar lugar a que se establezcan objetivos para los rendimientos, resultados y efectos que son poco realistas e imposibles de alcanzar. Christoplos afirmó que, tras el tsunami del Océano Índico en 2004, «*los métodos y las directrices se llenaron de optimismo normativo. Se emplearon términos, tales como la "red inconsútil", para describir las situaciones que irían*

surgiendo en caso de que se estuvieran aplicando los métodos adecuados. Se esperaba que las personas, los hogares, las comunidades y las naciones siguieran un proceso continuo, que iba desde el grave sufrimiento humano hasta el desarrollo "normal". A lo largo de los años este optimismo ha ido desapareciendo. La VARD [vinculación entre ayuda de emergencia, rehabilitación y desarrollo] hace más referencia a un misterio que a un camino dorado. Las tensiones políticas intratables y la inseguridad humana se han asociado con la VARD, primero en situaciones posteriores al conflicto y posteriormente cada vez con más frecuencia en los desastres naturales».¹³

El sesgo de optimismo tanto en relación con las intervenciones antiguas como con las nuevas a menudo es ocasionado por la falta de evidencia sobre sus verdaderos efectos. La investigación y la evidencia de las evaluaciones de estas intervenciones, o de otras similares, pueden ayudar a reducir esta incertidumbre.

Tal y como sucede en otros sectores, existen muchos otros sesgos relacionados con la manera de pensar de las personas que pueden tener un efecto negativo en aquellos que trabajan en el sector humanitario. Entre ellos se incluyen:

Prejuicio de retrospectiva: tendencia a ver los acontecimientos pasados como más predecibles de lo que eran antes de que sucedieran.

Aversión a la pérdida: tendencia a preferir no perder antes que ganar.

Efecto de encuadre: extraer diferentes conclusiones sobre la base de una misma información que ha sido presentada de diferentes maneras (p. ej. ¿preferiría que «un 95 % volviera a trabajar» o que «un 5 % no volviera a trabajar»?).

La «heurística de la disponibilidad»:

cuando las personas relacionan el tamaño, la frecuencia o la probabilidad de un problema con la facilidad de recordarlo o imaginarlo .

La «heurística de la representatividad»:

cuando las personas sobrestiman la probabilidad de los acontecimientos vividos.

La «necesidad de coherencia»:

el impulso de establecer patrones y relaciones causales cuando en realidad puede que no existan.

Sesgo metacognitivo:

la creencia de que somos inmunes a los sesgos!

Esto no quiere decir que el criterio profesional siempre esté equivocado. Investigadores como Gary Klein han elogiado el criterio intuitivo de los expertos, como, por ejemplo, en su trabajo sobre la «toma de decisiones naturalista».¹⁴ Los puntos de vista y los instintos de los profesionales pueden ser muy valiosos, pero debemos ser conscientes de sus aspectos negativos. Tal y como afirmó Daniel Kahneman en un artículo conjunto que escribió con el Profesor Klein en *American Psychologist*, *la intuición profesional en ocasiones es maravillosa y en otras es errónea*.¹⁵

ESTUDIO DE CASO

Cómo ha promovido la investigación el desarrollo de la programación de transferencias en efectivo

La programación de transferencias en efectivo (CTP, por sus siglas en inglés) ha sido una de las innovaciones más importantes de los últimos años en el ámbito de la asistencia humanitaria internacional.

En 2016 se pagaron 2800 millones de dólares en asistencia humanitaria mediante bonos y pagos en efectivo, lo que supone un aumento del 40 % desde 2015 y el doble respecto a 2014.¹⁶ El pago en efectivo no solo ha crecido en términos totales, sino también como cuota relativa del total de la asistencia humanitaria internacional, creciendo del 7,8 % en 2015 al 10,3 % en 2016.¹⁷

Dicho crecimiento se ha visto apoyado por un número creciente de evaluaciones, dando lugar a un conjunto de evidencia sobre los efectos de los diferentes programas en los resultados a nivel individual y doméstico. Este conjunto de investigaciones y evaluaciones ha sido fundamental para el desarrollo de la CTP.

ambién sirvió de ayuda para abordar los problemas relativos a las transferencias de efectivo, tales como la corrupción y la inseguridad,¹⁸ y para aumentar la concienciación de que el efectivo genera riesgos distintos, pero no necesariamente mayores, que los de la asistencia en especie.¹⁹ Se ha demostrado que los beneficios de la CTP abarcan múltiples sectores y servirán de ayuda para el empoderamiento de las mujeres, la seguridad alimentaria, la educación y la atención médica. Por lo tanto, la CTP es una herramienta aceptada en casi todas las respuestas de emergencia. Uno de los éxitos clave de la investigación de la CTP es que la evidencia generada ha sido ampliamente utilizada por las instituciones líderes para promover el uso de efectivo en las crisis humanitarias.

Sin embargo, es necesario que se siga investigando para poder ajustar los programas y maximizar sus beneficios.

Mensajes clave de la Sección A

- ✓ No defendemos una toma de decisiones que siga ciegamente las conclusiones de la investigación en detrimento del criterio profesional.
- ✓ Sin embargo, tampoco debe poner la experiencia profesional en un pedestal. En ocasiones, los expertos están terriblemente equivocados, pues no son inmunes al conjunto de sesgos sociales y cognitivos que afectan tanto a los especialistas como a los legos.

- ✓ Los resultados de la investigación son solo un tipo de evidencia, pero presentan la ventaja de ofrecer mayor rigor, relevancia e independencia en comparación con otros tipos de evidencia.
- ✓ Producir nueva evidencia puede resultar costoso y llevar tiempo, pero hay evidencia de buena calidad que está «lista para usarse» (véase Sección D). La investigación y las evaluaciones de las emergencias actuales puede orientar futuras intervenciones.

¿Cuándo puede ayudarle la evidencia?

En esta sección se exploran diferentes escenarios en los que el uso de la evidencia puede servirle de ayuda, así como los diferentes tipos de evidencia que puede necesitar en las diferentes etapas de desarrollo o implementación de una nueva intervención o política.

La evidencia puede mejorar la eficacia de las organizaciones. La evidencia puede reforzar su trabajo: desde campañas más persuasivas hasta la obtención de fondos de subvención, y desde el desarrollo de la capacidad de toma de decisiones de una organización hasta asegurarse de que los programas produzcan los resultados deseados. No importa si usted es una pequeña organización voluntaria o una gran organización internacional. Independientemente de la escala, puede haber investigación existente que se adapte a sus necesidades.

LA EVIDENCIA PUEDE AYUDARLE DE MUCHAS MANERAS, POR EJEMPLO:

- Empezar a hacer algo.
- Dejar de hacer algo.
- Proponer opciones o plantear el caso.
- Identificar prioridades.
- Evaluar necesidades.
- Adaptar los servicios a las necesidades.
- Diseñar e implementar programas más eficaces.
- Organizar recaudaciones de fondos.
- Crear campañas de sensibilización eficaces.
- Reforzar la rendición de cuentas de las partes interesadas.

CICLO DE RESPUESTA HUMANITARIA

La Oficina de Coordinación de Asuntos Humanitarios de las Naciones Unidas (OCAH) desarrolló el ciclo de respuesta humanitaria para mostrar la serie coordinada de medidas necesarias para ayudar a preparar, gestionar y ejecutar la acción humanitaria. Consiste en cinco elementos; cada paso está basado en el anterior y conduce al siguiente.²⁰ La evidencia puede ayudar en la toma de decisiones a medida que avanza en este ciclo.

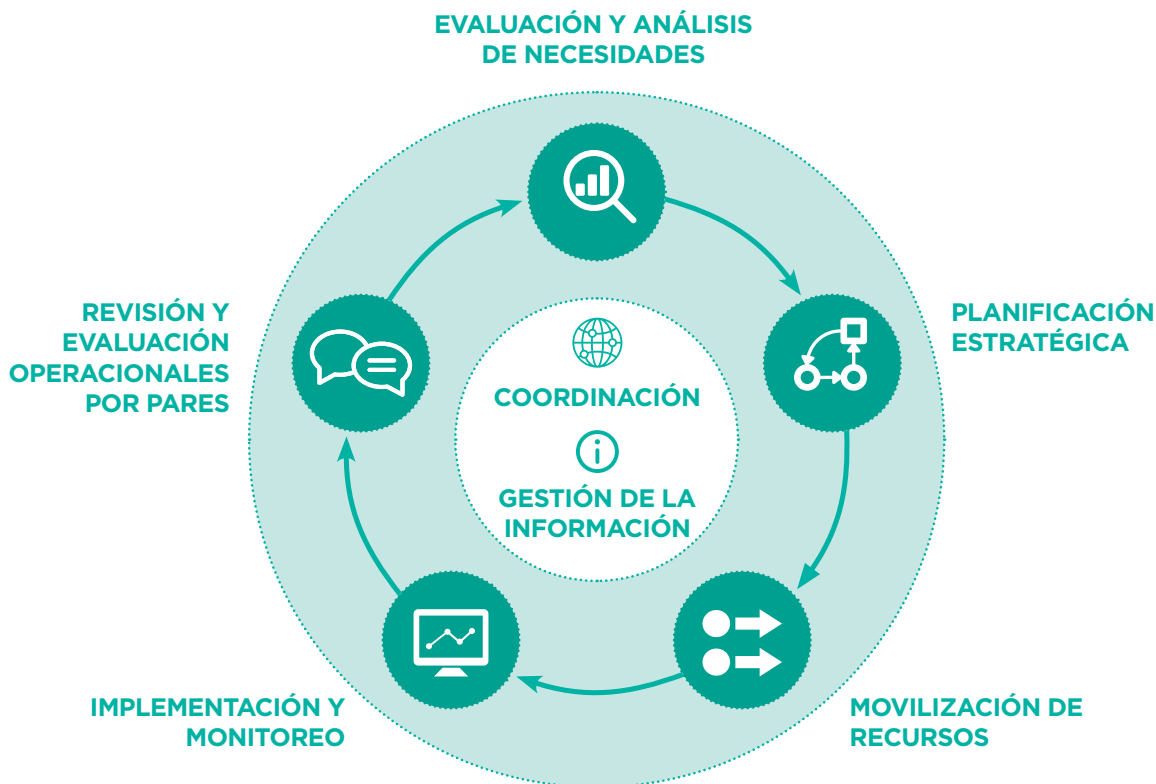
Evaluación y análisis de necesidades:

examinar investigaciones existentes puede ayudar a establecer una base para evaluar las necesidades, así como para comprender la naturaleza, la magnitud y las dinámicas de problemas específicos. Junto con la recopilación de nuevos datos sobre el entorno y las circunstancias a las que se enfrenta (o espera enfrentarse), puede ayudar a identificar oportunidades para responder de manera adecuada. En algunas circunstancias, esto puede resultar especialmente complicado si ha habido pocas oportunidades para realizar investigaciones pertinentes.²¹

Planificación de respuestas estratégicas:

al planificar su respuesta estratégica, debe buscar evidencia sobre lo que se ha probado y evaluado con anterioridad que podría ayudarle a tomar una decisión bien informada. Esto también podría incluir hacer

Figura B.1: Ciclo de respuesta humanitaria



uso de la evidencia sobre cómo comunicar de la mejor manera los posibles riesgos de una situación a aquellos que podrían estar involucrados.²²

Movilización de recursos: la aplicación de una nueva intervención o política requerirá la movilización de recursos. Estos podrían ser financieros para pagar las actividades o el personal necesario para realizarlas. La adopción de un enfoque basado en evidencia para planificar la respuesta estratégica y seleccionar las intervenciones y políticas que se adoptarán debe hacer que estas tengan una mayor probabilidad de ser eficaces y, en consecuencia, más atractivas para los donantes.²³

Implementación y monitoreo: cuando implemente una intervención o una

política, debe pensar en cómo va a monitorear e informar su impacto. Esto le ayudará a demostrar si está logrando marcar una diferencia.

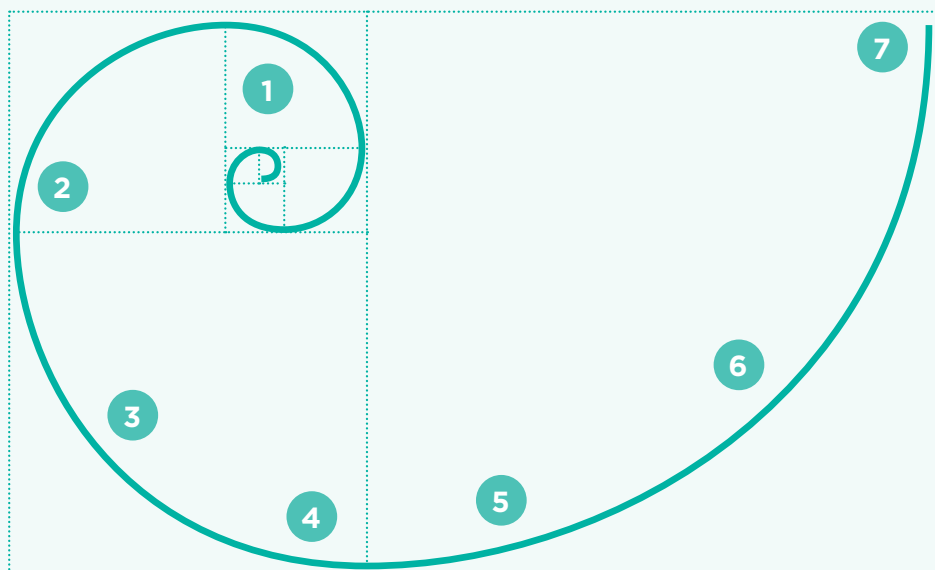
Revisión y evaluación operativa por pares: la revisión operativa por pares puede servir para corregir el rumbo e identificar áreas en las que aplicar medidas correctivas inmediatas. Junto con la evaluación de su propia respuesta estratégica y las comparaciones con la evidencia de intervenciones similares, puede ayudar a determinar si se necesita hacer ajustes o mejoras, en particular para acuerdos de liderazgo, la ejecución de otras fases del ciclo, la coordinación y los mecanismos de rendición de cuentas ante las personas afectadas.

DESARROLLO Y EJECUCIÓN DE UNA NUEVA INTERVENCIÓN

Cuando se desarrolla o ejecuta una nueva intervención o política humanitaria, es aconsejable empezar por considerar el momento adecuado. Diferentes tipos de evidencia serán útiles en distintas etapas de desarrollo o ejecución y debe reflexionar sobre qué tipo de investigación será más apropiada para cada etapa del ciclo de la nueva intervención o política. En los comienzos de una nueva iniciativa, la investigación puede identificar retos emergentes y la magnitud del problema. Puede utilizarse para planificar el desarrollo y la ejecución de una nueva intervención o política.

La espiral a continuación fue elaborada por Nesta para reflejar las diferentes etapas del proceso de innovación y también puede utilizarse para planificar el desarrollo de un nuevo enfoque a un problema social.²⁴ Se necesitarán distintos tipos de evidencia en las diferentes etapas y el proceso deberá ser continuo, de manera que, cuando llegue al final de la espiral, debe empezar de nuevo y encontrarse en un nuevo ciclo. A continuación analizamos las diferentes etapas de esta espiral, mientras que algunos ejemplos específicos de innovaciones exitosas en contextos humanitarios pueden encontrarse en un estudio de 15 proyectos financiados por el Fondo de Innovación Humanitaria.²⁵

Figura B.2: *La espiral de innovación de Nesta*



1 Análisis de oportunidades y retos

2 Producción de ideas

3 Desarrollo y prueba

4 Planteamiento del caso

5 Realización e implementación

6 Crecimiento y dimensionamiento

7 Cambio de sistemas

1. ANÁLISIS DE OPORTUNIDADES Y RETOS

Trabajar con las investigaciones actuales le ayudará a comprender la naturaleza, la magnitud y la dinámica de los problemas que enfrenta y las oportunidades para responder de manera apropiada. (Véase el siguiente caso de estudio.)

CASO DE ESTUDIO



Aplicación de la evidencia sobre los resultados de los sistemas de salud a largo plazo en entornos de crisis prolongada

Con el principio de «no dejar a nadie atrás» integrado en los ODS, existe un acuerdo entre la mayoría de los asociados para el desarrollo que establece que, además de abordar las necesidades de salud humanitarias en entornos de crisis, esto se debe realizar de manera que apoye el desarrollo de un sistema de salud igualitario a largo plazo.

El **ReBUILD Research Programme Consortium** es una asociación internacional de investigación financiada por UK Aid, que trabaja para mejorar el acceso de los pobres a la atención sanitaria eficaz en entornos posconflicto a través de evidencia sólida del fortalecimiento de los sistemas de salud. La investigación de ReBUILD se ha llevado a cabo en entornos posconflicto relativamente estables, si bien todavía frágiles, mediante una perspectiva histórica para comprender los efectos inmediatos, a mediano y a largo plazo en el sistema de salud de las políticas implementadas durante y después del conflicto y la crisis. Los temas incluyen los incentivos y la movilización de los trabajadores sanitarios, el acceso de las comunidades más pobres y más vulnerables a la atención médica, la efectividad de la ayuda y la igualdad de género.

La investigación ha generado una amplia comprensión de la experiencia de las comunidades y los trabajadores sanitarios durante períodos de conflicto y posconflicto, así como recomendaciones para el fortalecimiento de los sistemas de salud posconflicto. Pero la evidencia de ReBUILD también es relevante para los entornos humanitarios, p. ej., los efectos de los enfoques para captar trabajadores

sanitarios en respuestas a emergencias en la distribución de personal a largo plazo, la motivación y las habilidades o el valor de trabajar incluso con agentes débiles durante las emergencias para respaldar la capacidad de coordinación y planificación a largo plazo y los flujos de recursos predecibles.

Por lo tanto, ReBUILD participa también con aquellos que trabajan en el ámbito de la atención médica en entornos de conflicto actual o de crisis prolongada. Los éxitos incluyen informar la toma de decisiones y la implementación en Sierra Leona durante el brote de Ébola y respaldar aspectos de un plan de apoyo de donantes para Siria. Presentar evidencia sobre los resultados a largo plazo a aquellos que trabajan en la interfaz de desarrollo humanitario, donde las necesidades humanitarias son comprensiblemente el principal punto de atención, es desafiante, a menudo con espacio limitado para que la evidencia a largo plazo sea considerada dentro de los objetivos y los enfoques de las organizaciones humanitarias, lo que refleja la amplia brecha que separa los enfoques humanitario y de desarrollo.

ReBUILD colabora con varias organizaciones que tienen en cuenta los enfoques humanitarios basados en evidencia, incluida Evidence Aid, y actualmente trabaja en estrecha colaboración con grupos de trabajo que se centran de manera específica en la interfaz de desarrollo humanitario para la salud, como el **Grupo de Trabajo sobre Entornos Frágiles UHC2030** y el **Grupo de Trabajo Temático Mundial sobre Sistemas de Salud en Estados frágiles y afectados por conflictos**.

2. PRODUCCIÓN DE IDEAS

Después de haber identificado su foco de atención, es hora de buscar las intervenciones y políticas que se han estudiado en el paso. ¿Puede tomar ideas exitosas de otros acerca de qué funcionó anteriormente y evitar las intervenciones que fracasaron?

3. DESARROLLO Y PRUEBA

A medida que se inician las nuevas ideas, será hora de comenzar a pensar en probar y experimentar con diferentes enfoques y en evaluar el impacto de estos.

En la Sección C, analizamos las normas de evidencia que muestran el camino que deben seguir los nuevos programas en términos de proporcionar evidencia del impacto.

4. PLANTEAMIENTO DEL CASO

Tener evidencia del impacto probable que tendrá su intervención le ayudará a plantear el caso a aquellos que podrían financiarlo y aquellos que deberán ejecutarlo en el terreno. Esto lo pondrá en una posición más sólida para persuadirlos de financiar la intervención y apoyar su realización e implementación. (Véase el siguiente caso de estudio.)

CASO DE ESTUDIO



Presentación de evidencia a los líderes de las oficinas nacionales

En junio de 2017, Christine Fernandes, asesora a nivel mundial de Save the Children en materia de nutrición humanitaria para la alimentación de bebés y niños pequeños en emergencias (IYCF-E, por sus siglas en inglés) realizó una presentación sobre la IYCF-E para el director nacional y el equipo de directivos superiores de Save the Children Iraq, quienes se encontraban en proceso de revisión de su estrategia para el período 2017-2018. Hasta ese momento, la nutrición no formaba parte de los sectores de aplicación en Irak. Christine formuló su presentación de 20 minutos como un cuestionario; cada diapositiva comenzaba con una pregunta, como:

¿Cuál cree que es la manera más eficaz de prevenir la muerte de niños menores de cinco años?

Materiales tratados con insecticidas.

Vacuna Hib (meningitis).

Lactancia adecuada.

Alimentación complementaria adecuada.

Suplementos de vitamina A y zinc.

A continuación, procedió a corregir las ideas equivocadas generalizadas con datos concretos. Por ejemplo, citó un artículo de Lancet²⁶ para demostrar que la lactancia adecuada previene el mayor porcentaje de estas muertes en la primera infancia (13 %) y que la alimentación complementaria tiene el segundo mayor beneficio. Solicitar a los asistentes que realicen estimaciones y luego corregir sus ideas también es una manera eficaz de influenciar la opinión de la gente, sobre la base de la evidencia.²⁷ Christine logró convencer a todos los directivos de la Oficina Nacional de que lo correcto era centrarse en apoyar la IYCF-E, lo que llevó a un cambio concreto en la oficina nacional iraquí con la puesta en marcha de un programa de IYCF-E en Bagdad y el apoyo al Ministerio Federal de Salud para la estrategia nacional de nutrición con énfasis en el control de sustitutos de la leche materna y el desarrollo de capacidades a nivel nacional en la inversión de recursos a comienzos de 2018.

5. EJECUCIÓN E IMPLEMENTACIÓN

Una vez implementada una intervención o política, es necesario pensar en una manera de documentar el impacto. Esto sirve para pasar de una posición de retórica en la que dices «confía en mí, el proyecto está funcionando» a una basada en evidencia fiable proveniente de la supervisión y la evaluación. También ayudará con su responsabilidad al demostrar que usted está haciendo una diferencia, que su proyecto tiene una buena relación calidad-precio y que ha puesto su trabajo a disposición para que sea evaluado. (Véase el estudio de caso a continuación).

ESTUDIO DE CASO

Asociación para las evaluaciones de impacto

La Iniciativa Internacional para la Evaluación de Impacto (3ie) procura mejorar la vida de los pobres que viven en países de ingresos bajos y medios, proporcionando y resumiendo evidencia de qué intervenciones de desarrollo funcionan, cuándo, por qué y cuánto cuestan. En 2018, la Ventana Temática de Asistencia Humanitaria de la 3ie está financiando siete evaluaciones de impacto en Chad, la República democrática del Congo, Malí, Níger, Pakistán, Sudán y Uganda. La evaluación de impacto en Pakistán la está realizando la Universidad de Mannheim en Alemania, en colaboración con la Agencia de Cooperación Técnica y Desarrollo (ACTED). La ACTED apoya los programas de asistencia humanitaria en zonas propensas a las catástrofes y la evaluación analiza la eficacia de estas intervenciones para responder a estos peligros y reducir la vulnerabilidad de las comunidades ante los impactos de las emergencias. La colaboración entre la ACTED y la Universidad de Mannheim ha contemplado muchas instancias de fortalecimiento de las capacidades y divulgación exitosas:

Compromiso de múltiples niveles: se han presentado el estudio y la línea base al gobierno pakistaní con el objetivo de informar la estrategia de nutrición integral de Pakistán, aprovechando los conocimientos de distintas partes interesadas que tienen experiencia

trabajando en los ámbitos de nutrición y gestión de los desastres. Las conclusiones también se han difundido a un grupo de trabajo dentro del país que comprende ONG nacionales e internacionales, académicos y especialistas del sector.

Contribución para mejorar la calidad de los datos: se han logrado importantes mejoras en el sistema de Monitoreo y Evaluación (M&E). Ahora se han replicado en otro proyecto que está implementando la ACTED las mejores prácticas sobre recopilación de datos, como controles in situ, sesiones de comentarios de análisis con equipos de campo y retroalimentación periódica sobre las discrepancias en la recopilación de datos, que el equipo de investigación estableció durante la evaluación de impacto.

La ACTED espera incluir estas prácticas en su protocolo regular de M&E una vez que se hayan probado mejor en el campo.

Construir la capacidad para evaluar de la agencia de implementación: el personal de la agencia de implementación fue patrocinado para asistir a un taller sobre la medición del efecto de los programas de reducción de la pobreza y para ayudar a diseñar mejores políticas mediante el uso de las evaluaciones de impacto. Esto les permitirá asimilar mejor las conclusiones de la evaluación, así como proporcionar una base para futuras evaluaciones internas.

6) CRECIMIENTO, ESCALAMIENTO Y DIFUSIÓN

El objetivo principal de muchas innovaciones es replicar y crecer con el fin de generar cambios a gran escala en la práctica diaria y que puedan copiarse en otros entornos. Esto aumentará el alcance para un impacto positivo.

7) CAMBIOS EN LOS SISTEMAS

Lograr cambios a nivel de sistemas es extremadamente complejo. Una buena forma de alcanzar esta escala es contar con diversos estudios independientes que muestren que algo no funciona en las circunstancias específicas de las pruebas iniciales pero es eficaz en muchos lugares.

ADOPCIÓN DE IDEAS DE OTRAS REGIONES O SECTORES

Incluso si encuentra evidencia de que una política o un proyecto han tenido éxito en otro lugar, ¿funcionaría la adopción de esas ideas en su área? Howard White presenta un ejemplo del fracaso de un programa de alimentación complementaria en Bangladesh (el Proyecto de Nutrición Integrada de Bangladesh [BINP]) para lograr los resultados del «*aclamado Proyecto de Nutrición Integrada de Tamil Nadu (TINP) en India*». ²⁸ El programa de Bangladesh tuvo una tasa de participación un 30 % más baja en comparación con el de India y una evaluación comparativa del impacto «*no detectó ningún efecto significativo del programa en el estado nutricional, aunque hubo un efecto positivo en la mayoría de los niños malnutridos*». El profesor White señaló que:

«los destinatarios deben ser los correctos. El programa estaba dirigido a las madres de niños pequeños. Pero a menudo las madres no son quienes toman las decisiones, y rara vez son las únicas que las toman, con respecto a la salud y la nutrición de sus hijos. Para empezar, las mujeres en la parte rural de Bangladesh no van al mercado; son los hombres quienes realizan la compra. Y para las mujeres que viven en hogares comunes, es decir, que viven con su suegra, como es el caso de una minoría

considerable, es la suegra la que ejerce el dominio. De hecho, las tasas de participación en el proyecto son significativamente más bajas en el caso de las mujeres que viven con las suegras en lugares más conservadores del país». ²⁸

También observó que se debe identificar a los niños correctos para que el programa sea eficaz. En las áreas del BINP, los médicos de nutrición de la comunidad que implementaron el programa «*no pudieron identificar correctamente a partir de las gráficas (de crecimiento) qué niños debían ser admitidos al programa*», lo que llevó a una focalización errada de los niños. ²⁸ Estos dos ejemplos de ejecución inapropiada de programas ilustra la importancia de comprender los factores contextualizados a nivel local en el éxito o el fracaso de un programa y la necesidad de utilizar evidencia cualitativa y etnográfica (véase la Sección C), con el fin de conseguir los resultados deseados en diferentes contextos.

A la hora de diseñar nuevos programas que busquen cambiar cómo actúan las personas, como, por ejemplo, respetar el Derecho humanitario internacional o utilizar letrinas en vez de defecar en espacios abiertos, la investigación de los ámbitos de la psicología social, la antropología y el cambio conductual puede ser de ayuda. Esto puede proporcionar ideas basadas en investigaciones sólidas acerca de qué técnica de la gama de técnicas merece la pena adoptar para dar un «empujón» a las personas para que realicen las acciones deseadas.

Hugh Waddington y colegas, ²⁹ por ejemplo, han observado que mejorar el saneamiento haciendo que las personas usen letrinas en vez de defecar abiertamente en lugares públicos requiere más que conocer la teoría microbiana. Los resultados de la evaluación del saneamiento total dirigido por la comunidad (CLTS, por sus siglas en inglés) sugieren que un mayor conocimiento técnico de las letrinas, el uso de subsidios para reducir los costes para los hogares y el cambio de las normas sociales y la acción colectiva a nivel del pueblo son factores clave para provocar el cambio de comportamiento e incrementar la adopción de las letrinas.

CREACIÓN DE UNA TEORÍA DEL CAMBIO

En las primeras etapas del desarrollo o planificación de toda intervención, es importante describir lógicamente lo que hace y por qué importa. Debe ser coherente, claro y convincente. Esta descripción suele denominarse como Teoría del Cambio y tiene por objeto dar una «*explicación clara, concisa y convincente de lo que hace, qué impacto espera alcanzar y cómo cree que lo logrará. Es una base fundamental de cualquier programa y un requisito para una evaluación eficaz*».³⁰

Una Teoría del Cambio es una herramienta útil para especificar más la evidencia que está utilizando de otros y ser más claro sobre cómo está logrando los resultados. Por ejemplo, si desea ejecutar un programa para aumentar la adopción de lavarse las manos en un campo de refugiados, probablemente requiera evidencia que responda a una serie de preguntas. ¿Qué técnicas tienen más probabilidad de generar un cambio eficaz del comportamiento? ¿Los posibles usuarios aceptarán los nuevos lavamanos? ¿Dónde deberían ubicarse los

lavamanos? ¿Cómo estarían compuestas estas instalaciones? ¿Qué información hay que proveer? ¿Qué incentivos hay que introducir para influir en las decisiones domésticas?

Una Teoría del Cambio (ejemplo en la página 20) sirve para especificar sus objetivos y cómo los logrará. Ayuda a evitar que únicamente espere que funcione su enfoque innovador y que sean correctas sus suposiciones. Otro beneficio de aplicar una Teoría del Cambio es que este es el primer paso para diseñar una evaluación eficaz, porque intenta identificar todos los resultados que habrá que medir.

El mensaje más importante es pensar en la adecuación. Es importante que busque una investigación que se adapte a sus necesidades y a la etapa de desarrollo en la que se encuentra. La evidencia que necesitará al planificar un programa nuevo o durante los primeros meses de su implementación será muy distinta de la que necesitará cuando el programa ya esté más establecido. Retomaremos este punto crítico de la adecuación en la Sección C.

Mensajes clave de la Sección B

✓ Hay una amplia gama de situaciones en las que la evidencia puede ser útil. No se limitan a las más obvias relacionadas con la captación de los resultados de las intervenciones y las políticas. Es posible que haya otros beneficios que no ha contemplado, como crear una campaña más persuasiva, obtener fondos o dejar de hacer algo que no está funcionando.

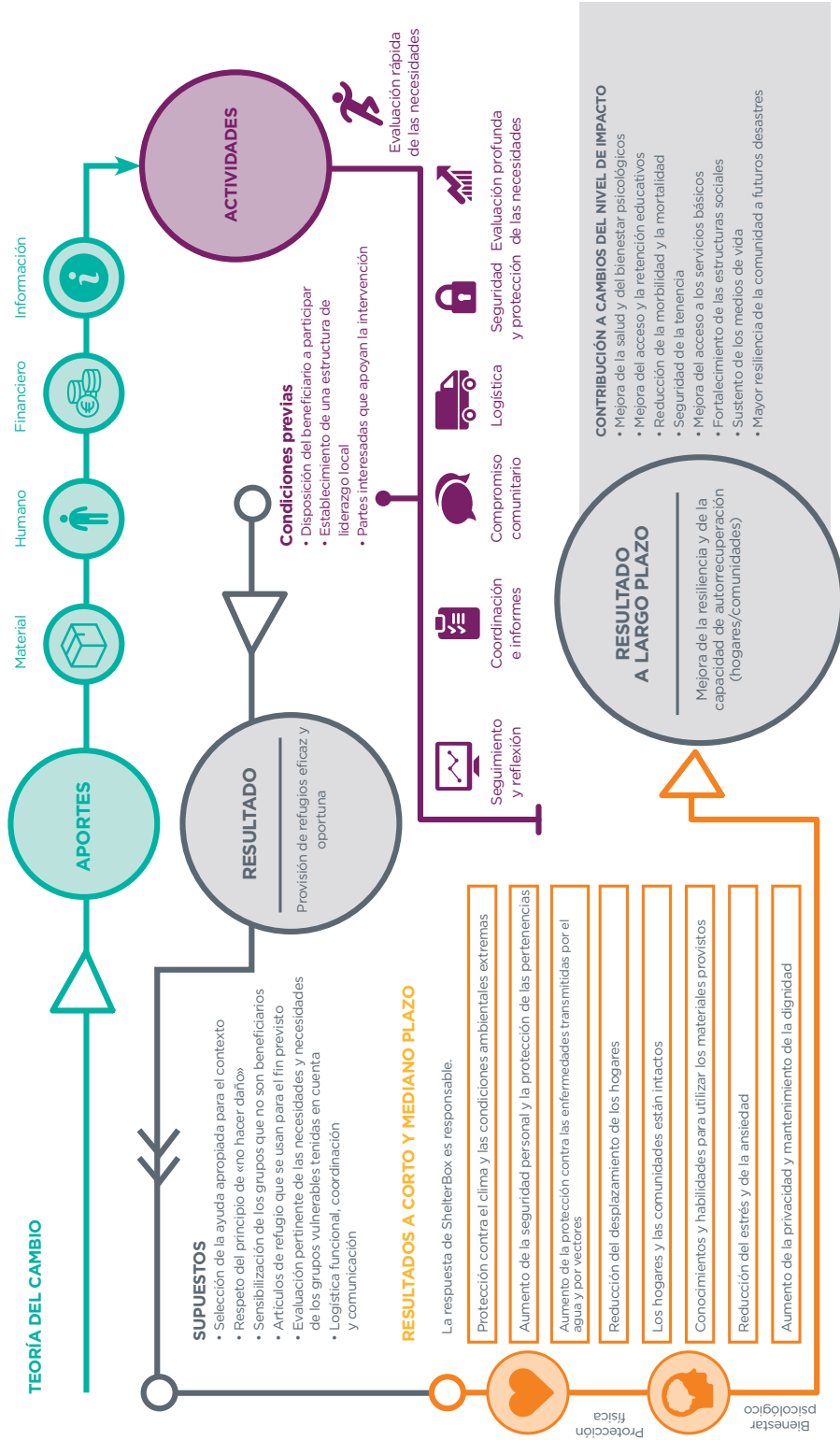
✓ Debe pensar en el momento y la adecuación de la evidencia. ¿En qué parte de la espiral de innovación se encuentra y qué evidencia se adaptaría mejor a sus necesidades? Por ejemplo, en las primeras etapas del desarrollo de un programa innovador, puede ser más valioso aprender de lo que ya han descubierto los demás, analizando las evaluaciones anteriores o investigaciones científicas y sociales más amplias, en vez de emprender una evaluación grande y costosa por su cuenta.

✓ Las etapas iniciales de la innovación son el momento para crear una Teoría del Cambio. Es una herramienta útil para ser más claro sobre la evidencia que está usando y cómo va a obtener los resultados. También le ayudará a diseñar una evaluación eficaz.

✓ Las intervenciones o políticas que están más establecidas pueden montar sus propios experimentos o evaluaciones de impacto. A medida que una innovación crece, también es recomendable considerar reproducciones múltiples para verificar que los beneficios que observó no fueron solo una historia de éxito aislada, sino que pueden darse en otros lugares y contextos.

✓ No importa el tamaño de su organización. Hay muchas investigaciones que podrían serle útiles.

Ejemplo de la Teoría del Cambio en la adaptación de las comunidades



¿Qué tipo de evidencia debería elegir?

Esta sección analiza distintos tipos de evidencia y examina cómo elegir la más apropiada según el problema de interés. También examina cómo juzgar la calidad de la evidencia.


No toda la evidencia es igual.

Un mensaje importante de esta guía práctica es pensar en la utilidad. Muchas personas se meten en problemas por no pensar claramente en qué tipo de diseño de investigación, metodología o métodos de recopilación de datos van a satisfacer sus necesidades realmente. Tienden a ser guiadas por los métodos.³¹ En otras palabras, tendemos a adoptar nuestro enfoque favorito, tales como cuestionarios, ensayos aleatorios o entrevistas, porque nos sentimos más cómodos, en lugar de abrirnos al abanico de métodos y pensar cuáles se adaptan mejor para responder a nuestro desafío y resolver nuestras incertidumbres.

Por ejemplo, la investigación experimental es más adecuada para evaluar el impacto y «lo que funciona» (p. ej., la eficacia), pero otros tipos de investigación pueden aportar otros conocimientos.³² Tal vez no analicen tanto el impacto, pero pueden revelar por qué y cómo funcionan las intervenciones.³³

Asimismo, puede ser importante determinar qué secciones de la población se benefician más de una intervención, lo que a veces queda enmascarado al confiar en el efecto promedio para una población en su conjunto.

ESTUDIO DE CASO



Lecciones aprendidas del uso de la etnografía en la respuesta al Ébola

La epidemia del Ébola en África Occidental reveló la necesidad de fomentar el «compromiso comunitario» a la hora de responder ante los brotes y el papel que desempeñan las redes de científicos sociales para compartir material y comunicar la respuesta. Un ejemplo es la Plataforma antropológica de respuesta al Ébola, cuyos miembros *«trabajan de manera proactiva con organizaciones sanitarias y humanitarias para diseñar, proveer y supervisar intervenciones más reactivas a nivel local y comunicadas a la sociedad e investigación sobre el terreno»*.³⁴ Podemos aprender de cómo los antropólogos usaron la etnografía para incluir las dimensiones sociocultural y política en una respuesta biomédica con el fin de abordar la problemática de las llamadas prácticas de entierro «nocivas». En un contexto en el que los cuerpos transmiten el virus del Ébola, se adoptaron medidas excepcionales por las que solo algunos trabajadores humanitarios estaban autorizados para enterrar los cuerpos de las personas que habían fallecido de Ébola. Los intervinientes del brote del Ébola percibieron las reuniones de

las comunidades locales que acudían a los funerales de sus seres queridos, que habían muerto a causa del virus, como una resistencia tradicional³⁵ a la medicina moderna y como una amenaza para la bioseguridad. Incluir a los antropólogos que tenían amplia experiencia viviendo en los países afectados por la epidemia fue clave para identificar prácticas aceptables y seguras (por ejemplo, mostrar una foto del cuerpo a las familias para responder a los rumores de que los cuerpos desaparecían³⁶), así como para superar las tensiones que surgían de las prácticas de entierro mediante enfoques de colaboración. Una lección clave que surgió fue identificar razonamientos que explicaran las necesidades de las comunidades de enterrar a sus muertos en vez de etiquetar estas prácticas como esencialmente «culturales».³⁷ La respuesta al Ébola demostró *«la necesidad de establecer nuevos mecanismos globales que puedan movilizar rápido a todos los especialistas que pueden aportar información contextual local, médica, epidemiológica y política relevante en situaciones de emergencia sanitaria global»*.³⁸

El tipo de investigación elegido como fuente de evidencia para ayudar en la toma de decisiones debe ajustarse a las necesidades del desafío que se está afrontando.³⁹

La investigación experimental es más adecuada para evaluar el impacto y «lo que funciona».

PARTE 1: DISTINTOS TIPOS DE MÉTODOS DE INVESTIGACIÓN, DISEÑOS Y ENFOQUES

Entonces, ¿cómo sabe qué método de investigación es el mejor para su caso?

Si usted no es investigador, puede ser abrumador enfrentarse a la jerga y la lista interminable de enfoques distintos.⁴⁰ Estos temas se recogen en una nota del Departamento de Desarrollo Internacional del Reino Unido sobre **cómo evaluar la evidencia**⁴¹ y son tan importantes para el sector humanitario como lo son para el desarrollo internacional:

Algunos tipos de investigación (como los diseños experimental y cuasiexperimental) funcionan mejor para identificar la presencia de una relación causal.

Otros tipos de investigación son más apropiados para explicar dichas relaciones causales –véase Más allá de los experimentos: nuevos enfoques de evaluación en la página 32.

Algunos diseños (como los estudios de observación) son más útiles para entender los contextos políticos, sociales y medioambientales.

Algunos pros y contras de varios diseños de investigación se resumen en la **Tabla C.1.**⁴²

Tabla C.1 *Diferentes diseños, métodos y enfoques para investigar evidencias –breve resumen*

TIPOS DE INVESTIGACIÓN Y EVALUACIÓN	¿QUÉ ES?	PROS	CONTRAS
Etnografía	Un informe realizado por alguien con experiencia directa de un tema en particular o afectado por este. El objetivo es recopilar información detallada sobre las experiencias de las personas y sus impresiones.	Fuerte e inmediata; puede ofrecer información detallada de actividades ocultas a gran parte de la población.	Difícil de verificar y puede llevar a la inflación de la prevalencia. Las historias personales conmovedoras pueden dificultar la valoración crítica y las anécdotas personales pueden no ser representativas.
Estudio de caso	Examen detallado e intensivo de uno o varios casos, como una comunidad, una familia, una organización, lugares, un acontecimiento o un individuo.	Fácil y práctico; pueden utilizarlo los médicos y los no especialistas, útil para intervenciones que ya se hayan realizado, puede servir para identificar consecuencias adversas de una intervención, ayuda a describir novedades, genera hipótesis para otros estudios.	No sirve para deducir el resultado de una intervención, el tamaño reducido dificulta generalizarlo a nivel nacional o de población.

TIPOS DE INVESTIGACIÓN Y EVALUACIÓN	¿QUÉ ES?	PROS	CONTRAS
Estudios de control de casos	<p>Compara un grupo que, supongamos, tiene una afección médica con otro grupo que no la tiene y mira hacia atrás para ver qué características difieren entre ambos grupos.</p>	<p>Requiere menos recursos para ejecutarse que los estudios de intervención; es útil cuando los ensayos aleatorios o los estudios de cohorte prospectivos no son prácticos (por ejemplo, estudios de muerte súbita infantil); puede generar hipótesis que se pueden probar en estudios más sólidos.</p>	<p>Es infrecuente en políticas sociales (véase abajo la evaluación «basada en casos» para un enfoque más frecuente en las políticas sociales); más frecuente en epidemiología y salud; proporciona menos evidencia de inferencias causales que el ensayo aleatorio; alto riesgo de muchos sesgos (por ejemplo, sesgo de memoria, sesgo de selección y sesgo del entrevistador).</p>
Evaluación del proceso	<p>Un enfoque respecto a la evaluación que estudia lo que sucede en comparación con las teorías de cambio existentes o los caminos causales identificados durante una evaluación. Puede asociarse con la evaluación realista, el análisis comparativo cualitativo (ACC), el análisis de contribución y el seguimiento de procesos.</p>	<p>Sólido para explicar causas; puede usarse en áreas más desordenadas de intervenciones sociales complejas en las que puede haber muchas causas y en las que el contexto es importante. La evaluación del proceso investigará los mecanismos de cambio, los factores contextuales y la evaluación de la implementación.</p>	<p>Falta de métodos acordados; oportunidades de sesgos; deficiente para estimar las cantidades o el alcance del impacto; evaluación relativamente escasa de las técnicas usadas en contraste con el cuerpo más amplio de literatura que critica los ensayos aleatorios.</p>
Evaluación económica	<p>Las evaluaciones económicas determinan si con una intervención se hace un uso eficiente de los recursos y la compara con otras intervenciones en términos de coste y consecuencias. En el sector humanitario, los métodos más habituales son los análisis de coste-beneficio y de coste-eficacia.</p>	<p>La evaluación económica ofrece una buena idea del nivel de recursos necesario y los gastos por beneficio/efectividad producidos, lo que es un gran indicador para la mayoría de los responsables políticos y se puede usar para la incidencia política.</p>	<p>La evaluación económica implica varias hipótesis y, en ocasiones, medidas indirectas de los costes, que dependen de las opiniones. La medida de efectividad debe facilitarse mediante un ensayo comparativo sólido.</p>

TIPOS DE INVESTIGACIÓN Y EVALUACIÓN	¿QUÉ ES?	PROS	CONTRAS
Transversal	<p>Una muestra representativa de personas encuestadas en un momento determinado. A pesar de que las encuestas, tales como los cuestionarios y las entrevistas estructuradas, se suelen usar en el diseño transversal, no son los únicos medios. Otros métodos incluyen el análisis de contenidos o análisis de estadísticas oficiales.</p>	<p>Los datos cuantitativos se pueden examinar para detectar patrones o asociaciones; son relativamente baratos y éticos; las encuestas se pueden repetir a intervalos, mostrando cambios de tendencias con el tiempo (véanse los estudios de cohortes/ longitudinales a continuación); pueden generar hipótesis que se pueden probar en estudios más sólidos.</p>	<p>Establece la asociación en el mejor de los casos, no la causalidad; es poco posible atribuir algún cambio medido a la intervención o entender qué podría haber ocurrido si no se hubiera realizado la intervención porque cualquier cambio podría deberse a cuestiones más generales, como las condiciones económicas, el clima o campañas en los medios de comunicación, en vez de a la intervención. Otras desventajas son el riesgo de sesgo de memoria, sesgo de conveniencia social, sesgo del investigador; grupos de tamaños desiguales y distribución desigual de factores de confusión.</p>
Estudios de cohorte/ longitudinales	<p>La misma muestra de personas encuestada en diferentes momentos a lo largo del tiempo, a veces desde la infancia hasta la edad avanzada.</p>	<p>La mejor fuente de evidencia sobre la asociación entre la experiencia en la infancia y los resultados en la edad adulta; puede ofrecer un apoyo sólido para algunas intervenciones tempranas; puede usarse para evaluar programas aplicados tras los ensayos aleatorios.</p>	<p>A menudo, los datos afloran demasiado tarde para formular políticas eficaces; los participantes del estudio pueden dejarlo con el tiempo; enfoque costoso al mantenerlo durante décadas.</p>

TIPOS DE INVESTIGACIÓN Y EVALUACIÓN	¿QUÉ ES?	PROS	CONTRAS
<p>Diseño cuasiexperimental</p>	<p>Compara un grupo que, supongamos, tiene una afección médica con otro grupo que no la tiene y mira hacia atrás para ver qué características difieren entre ambos grupos.</p>	<p>Puede ofrecer evidencia razonablemente sólida de la relación entre la intervención y los resultados mensurables; es un método poderoso para explorar el impacto de una intervención cuando no es posible la asignación al azar; se puede aplicar tanto en comunidades grandes como en grupos; no es necesaria la asignación al azar desde el principio (ex-ante), lo que previene algunos de los desafíos y las dificultades de dicha asignación.</p>	<p>La imposibilidad de asegurar la equivalencia de los grupos y de prevenir el cambio con el tiempo puede derivar en conclusiones menos fiables; las técnicas de comparación tienden a necesitar muchos datos tanto en los grupos de intervención como en los de comparación, lo que puede requerir mucho tiempo y cuya recopilación puede ser costosa; es necesario un buen entendimiento de los factores que deben compararse (sin esto, sigue siendo posible que haya diferencias sistemáticas que no se controlan entre los grupos); estos diseños requieren un trabajo analítico complejo y conocimiento especializado.</p>

TIPOS DE INVESTIGACIÓN Y EVALUACIÓN	¿QUÉ ES?	PROS	CONTRAS
<p>Ensayo aleatorio</p>	<p>Un grupo recibe una intervención mientras que otro no la recibe o recibe una intervención diferente; se forman grupos usando un proceso aleatorio, normalmente existe la misma probabilidad de ser asignado a cualquier grupo.</p>	<p>Ofrece los hallazgos más sólidos y fiables, que transmiten la confianza de que cualquier diferencia medida entre los grupos se debe a la intervención; la asignación aleatoria debe prevenir las diferencias sistemáticas entre grupos; mayor confianza en el tamaño del efecto y la relación entre la intervención y el resultado; enfoque reconocido internacionalmente.</p>	<p>Insuficiente en cuanto a tomar en cuenta el contexto (p. ej. entorno cultural, institucional, histórico y económico); difícil de llevar a cabo a nivel de población nacional; al usarlo en un pequeño grupo piloto, puede no ser relevante a nivel de población/nacional (aunque este riesgo lo corren todos los diseños); puede ser difícil manipular las variables para experimentar una política social (p. ej. clase, etnia o lugar de residencia); los errores en la asignación aleatoria pueden invalidar los resultados; puede ser difícil convencer a los responsables de la toma de decisiones de los beneficios de este diseño; posibles cuestiones políticas y éticas relativas a la asignación aleatoria (p. ej. algunos grupos pueden recibir una posible intervención beneficiosa al azar y otros no); puede necesitar más tiempo de gestión y más tiempo de configuración que los cuasiexperimentos.</p>
<p>Revisiones sistemáticas, metaanálisis</p>	<p>Agregación de resultados de estudios admisibles, con criterios de admisibilidad definidos de antemano y metodologías revisadas.</p>	<p>La mejor fuente de garantía de que una intervención funciona (o no); el metaanálisis agrupa los resultados estadísticos; las grandes revisiones tienen un poder estadístico considerable; otros investigadores pueden reproducirlo; puede aplicarse a cualquier tipo de estudio o datos.</p>	<p>Requiere de un número importante de estudios sólidos en un área determinada; la metodología está menos desarrollada para sintetizar datos cualitativos y «literatura gris»; podría producir resultados erróneos si los sesgos de informe selectivo afectaran la investigación disponible para revisión.</p>

INVESTIGACIÓN EXPERIMENTAL – ¿POR QUÉ TANTO ALBOROTO?

La investigación experimental, como los ensayos aleatorios, han recibido mucha atención –puede que demasiada. ¿Por qué han suscitado tanto interés?

Con frecuencia, los médicos y los responsables políticos quieren saber si una intervención ha surtido efecto. ¿Ha habido un vínculo causal entre su nuevo programa y el «impacto», «efecto» y «resultado» final? Quieren ser capaces de atribuir los efectos que han visto a la política o intervención que se está examinando. Los diseños experimentales, a pesar de que a veces son difíciles de aplicar en la práctica, tienen más probabilidades de mostrar esta relación de causa y efecto.

Un estudio experimental podría simplemente examinar y comparar dos grupos de personas antes y después de realizar la intervención en uno de los grupos. Pero si estos grupos no se forman de manera aleatoria, existe el riesgo de que los grupos varíen debido al sesgo de selección y las diferencias en los resultados se deberían entonces a las diferencias en sus características y no a los efectos de la intervención que se está examinando. Por ejemplo, alguien que solicita la intervención puede estar más predispuesto a obtener resultados positivos. Quizás estén más motivados, más sanos, más seguros y, por tanto, sean más susceptibles de obtener mejores resultados independientemente de cualquier intervención. Por otro lado, alguien que desea la intervención puede ser más vulnerable y más susceptible de obtener malos resultados independientemente de cualquier intervención. Estos sesgos distorsionarán los resultados del experimento y pueden implicar una estimación errónea del efecto de la intervención. Asimismo, los sesgos pueden ser impredecibles y si no sabemos la inclinación de dicho sesgo, no podemos simplemente ajustar los resultados del experimento para intentar tenerlos en cuenta.⁴³

En cambio, para reducir el sesgo, podemos usar un proceso de asignación aleatoria: tras identificar a la población del estudio y que hayan aceptado participar, los individuos por comparar entre sí se asignan aleatoriamente a los grupos de estudio que serán comparados, por ejemplo haciendo un sorteo.⁴⁴

DISEÑOS DE INVESTIGACIÓN EXPERIMENTAL

Elegir el diseño experimental apropiado y llevar a cabo el estudio correctamente establece el vínculo causal entre una intervención y un resultado. Cuando se utiliza la asignación aleatoria para asignar a los participantes a la intervención o al grupo de comparación, esto nos ofrece el medio más sólido para determinar si cualquier diferencia de resultado entre los grupos se debe a la intervención o a la casualidad; esta última se minimiza realizando un estudio lo suficientemente grande. La asignación aleatoria minimiza la posibilidad de diferencias sistemáticas entre los grupos, independientemente de si se deben a factores conocidos o desconocidos.

La introducción de un grupo de control con el cual comparar la intervención minimiza varios sesgos que suelen complicar el proceso de evaluación. Por ejemplo, si introduce un sistema de transferencia de dinero para aumentar las tasas de inmunización y los partos en centros de salud, ¿cómo sabrá si quienes reciban el dinero extra no habrían utilizado los servicios de todas formas? Debe compararlos con un grupo de personas que no reciban el dinero extra.

Como en cualquier diseño de investigación, el poder del ensayo aleatorio también requiere que el estudio se realice correctamente. Esto incluye asegurarse de que la asignación aleatoria de individuos o unidades (p. ej. escuelas, hospitales, vecindarios, etc.) se haga correctamente («fidelidad de asignación») y que únicamente el grupo experimental se haya expuesto a la intervención que se está

examinando («fidelidad al tratamiento»). Un ensayo aleatorio bien realizado también requiere suficientes participantes y poder estadístico para asegurarse de que se equilibren otros factores («de confusión») que podrían haber ocasionado el resultado entre el grupo experimental y los de control. Un bajo índice de abandono («deserción») también es necesario para minimizar el riesgo de sesgo. Por consiguiente, es importante verificar el informe de un ensayo aleatorio para asegurarse de que esos criterios de validez se han cumplido.

Se puede encontrar más información sobre cómo determinar si un ensayo aleatorio se ha realizado e informado correctamente en el documento de orientación sobre cómo realizar ensayos y presentar informes al respecto.⁴⁵

La Iniciativa Internacional para la Evaluación de Impacto (3ie) identificó evaluaciones de intervenciones humanitarias en situaciones de conflicto y reconciliación que usaron ensayos aleatorios y diseños cuasiexperimentales (véase a continuación).⁴⁶

ESTUDIO DE CASO



Establecimiento de la efectividad de las intervenciones humanitarias utilizando métodos experimentales

Las enfermedades transmisibles son de especial preocupación en las poblaciones afectadas por conflictos y desastres que residen en campamentos. En la fase de emergencia aguda, las enfermedades diarreicas pueden representar más del 40 % de las muertes entre los residentes del campamento. Estas enfermedades son consecuencia de una variedad de factores pero, en situaciones de emergencia, el saneamiento y el suministro de agua adecuados (tanto en términos de cantidad como de calidad) son los medios principales para la reducción de la diarrea. Existen claras limitaciones en las tecnologías actuales para el tratamiento de agua y pocos productos son capaces de tratar el agua turbia.

Shannon Doocy y Gilbert Burnham describieron los hallazgos de un estudio de efectividad de doce semanas sobre el tratamiento del agua en el punto de uso con un desinfectante floculante en 400 hogares en campamentos para poblaciones desplazadas en Monrovia, Liberia.⁴⁷

El estudio comparó las tasas de diarrea entre los hogares con tratamiento de

agua con desinfectante floculante y un mejor almacenamiento de agua (grupo de intervención) con los hogares con solamente mejor almacenamiento (grupo de control). La asignación de hogares a los grupos experimental y de control fue aleatoria.

En los hogares de intervención, el tratamiento del agua en el punto de uso con el desinfectante floculante junto con la mejora del almacenamiento redujo la incidencia de diarrea en un 90 % y la prevalencia en un 83 %, en comparación con los hogares de control que contaban únicamente con mejor almacenamiento de agua. En la intervención, los niveles de cloro residual cumplieron o superaron las normas Esfera en un 85 % (IC del 95 %: 83,1–86,8) de las observaciones con una tasa de cumplimiento del 95 %.

Los doctores Doocy y Burnham concluyeron que un desinfectante floculante en el punto de uso redujo significativamente la incidencia y la prevalencia de la enfermedad diarreica en los hogares que residen en campamentos en situaciones de emergencia.

CUESTIONES PRÁCTICAS Y ÉTICAS DE LOS ENSAYOS ALEATORIOS

Tener un grupo de control conduce a una crítica común de los ensayos aleatorios, que es que no son éticos porque solo se da acceso a una intervención experimental, que puede considerarse superior, a una parte de la población. Este «racionamiento» de la intervención se justifica si no sabemos si los beneficios superan a los perjuicios y hacemos el ensayo para averiguarlo. No obstante, si ya hay evidencia contundente de que la intervención será beneficiosa, aún podríamos tener que realizar un ensayo aleatorio para averiguar cuán beneficiosa es y para quién. De ser así, hay maneras de evitar las críticas y los problemas éticos. Por ejemplo, en un ensayo aleatorio con controles en lista de espera, a aquellos asignados al grupo de control se les ofrecerá la intervención ya sea al final del ensayo o cuando su seguimiento personal haya terminado. Otra opción es el diseño «por pasos» o «de tuberías»,⁴⁸ en el que el orden en que los participantes reciben la intervención se determina al azar, lo que permite que se introduzcan gradualmente a lo largo del tiempo, de modo que todos los participantes hayan tenido acceso a la intervención al final del ensayo. Este diseño se utilizó en el estudio Progresas/Oportunidades, uno de los ensayos aleatorios más grandes en el ámbito del desarrollo internacional, que evaluó las transferencias de dinero condicionadas en las poblaciones rurales desfavorecidas en México.⁴⁹ Los ensayos por pasos pueden resultar muy prácticos en un mundo de austeridad, donde no hay suficientes recursos para realizar un despliegue completo de un nuevo programa de una sola vez. Permiten realizar un despliegue gradual utilizando, al mismo tiempo, un ensayo aleatorio para evaluar su efectividad. Sin embargo, los ensayos por pasos son complejos y pueden ser difíciles de implementar y, si la intervención resulta ineficaz o perjudicial, se habrá extendido a todos los participantes del ensayo. En caso de una intervención educativa o de comportamiento o la reestructuración de la prestación de servicios, puede ser difícil e incluso imposible eliminar o «apagar» la intervención después del ensayo.

CUANDO LA ASIGNACIÓN AL AZAR NO ES POSIBLE O ACEPTABLE

En algunas circunstancias, puede que no sea posible o aceptable asignar personas de manera aleatoria a los grupos que se van a comparar. Por ejemplo, podría resultar imposible emplear un ensayo aleatorio para comparar los efectos relativos de diferentes formas de coordinar la respuesta ante una emergencia grave o el impacto de una política nacional destinada a mejorar la inclusión social de los refugiados.⁵⁰ En estos casos, los investigadores podrían usar un diseño cuasi-experimental para investigar la relación entre las intervenciones y los resultados. Kayvan Bozorgmehr y Oliver Razum hicieron esto para comparar el gasto personal en atención médica entre solicitantes de asilo y refugiados en Alemania, que tenían acceso regular o restringido a la atención médica.⁵¹ De manera similar, Rodolfo Rossi y sus colegas utilizaron encuestas antes y después de una campaña de vacunación para investigar su efecto en la cobertura de vacunación de niños pequeños en una zona del Líbano afectada por la crisis en 2015.⁵²

DISEÑOS CUASI-EXPERIMENTALES

Existe una amplia variedad de diseños cuasi-experimentales para comparar los efectos de las intervenciones y estos se usan a menudo cuando los ensayos aleatorios no son factibles. Dichos estudios son similares a los ensayos aleatorios en la medida en que pueden usarse para comparar los resultados de personas o grupos que reciben una intervención o programa y los que no lo reciben. Sin embargo, la recepción o no recepción de la intervención o el programa no se basaría en una asignación aleatoria. En su lugar, se utilizan otros tipos de diseño comparativo, como series de tiempo interrumpido o estudios controlados de antes y después, o análisis, como la discontinuidad de la regresión o el emparejamiento por puntuación de propensión.⁵³ Cabe señalar que los diseños de antes y después sin un grupo de comparación no pueden establecer el impacto de una intervención o programa

porque, en ausencia de una hipótesis de contraste (es decir, una comparación con lo que habría ocurrido sin la intervención), cualquier impacto observado en los resultados puede deberse a factores distintos a la intervención. Los diseños

cuasi-experimentales tienen un mayor riesgo de sesgo que los ensayos aleatorios bien ejecutados, pero aún así podrían permitir realizar inferencias causales sólidas en circunstancias en las que un ensayo aleatorio no es posible o aceptable.⁵⁴

ESTUDIO DE CASO



Determinación de la eficacia de las intervenciones humanitarias utilizando un diseño cuasi-experimental. La contribución de la asistencia alimentaria a las soluciones duraderas en situaciones prolongadas de refugiados

Esta evaluación fue parte de una serie de evaluaciones de impacto encargadas conjuntamente por el Programa Mundial de Alimentos (PMA) y la Oficina del Alto Comisionado de las Naciones Unidas para los Refugiados (ACNUR) en 2011 y 2012 para evaluar el papel de la asistencia alimentaria y su contribución a la autosuficiencia y a las soluciones duraderas para los refugiados y las poblaciones afectadas por ellos.⁵⁵

La evaluación abarcó la prolongada situación de los refugiados en el sureste de Bangladesh, donde aproximadamente 30.000 refugiados rohinyá han recibido asistencia en dos campamentos oficiales durante más de dos décadas. Además de los refugiados registrados, aproximadamente 45.000 rohinyá no registrados residen en sitios improvisados y más de 150.000 residen en comunidades de acogida en el distrito de Cox's Bazar. Para evaluar el impacto de la manera más metodológicamente rigurosa para el contexto, se utilizó un diseño cuasi-experimental. El diseño de la evaluación se centró principalmente en tres grupos poblacionales:

1. Refugiados que viven en dos campamentos de refugiados oficiales;
2. Rohinyás no registrados que residen en dos sitios no oficiales, en los campamentos oficiales o en las comunidades de acogida;
3. Comunidades de acogida.

La comparación cuantitativa clave fue entre los refugiados registrados que recibieron asistencia alimentaria y los Rohinyás no registrados que no la recibieron. La pregunta

principal de la evaluación fue: ¿cuáles son los efectos diferenciales de la asistencia alimentaria a largo plazo en las diferentes poblaciones de refugiados rohinyá y las afectadas por ellos en Bangladesh?

La evaluación concluyó que junto con otras formas de asistencia externa, la asistencia alimentaria era un factor que contribuía a la elección de la actividad económica de los hogares y la adopción de estrategias específicas de afrontamiento. En comparación con sus homólogos no registrados rohinyá, los refugiados registrados participaron en actividades económicas significativamente diferentes, incluido el empleo más cualificado y menos arriesgado con salarios más altos en general. También contaban con un nivel de riqueza significativamente mejor basado en la acumulación de activos. La asistencia alimentaria era un componente integral de sus medios de vida, utilizada principalmente para el consumo y como garantía y transferencia de valor para préstamos e hipotecas. La transferencia de valor de toda la asistencia externa en los campamentos permitió a los refugiados trabajar menos y confiar en esta asistencia externa en tiempos de crisis.

A pesar de estas diferencias, todos los refugiados y los grupos no registrados rohinyá se basaron en actividades económicas para sustentar sus medios de vida. Los rohinyá no registrados emplearon una gama más amplia de mecanismos de afrontamiento, tanto positivos como negativos, y constituyeron una parte importante del mercado laboral de la región.

MÁS ALLÁ DE LOS EXPERIMENTOS: NUEVOS ENFOQUES DE EVALUACIÓN

Cuando pensamos en el impacto, también debemos pensar más allá de la dependencia exclusiva de experimentos como los ensayos aleatorios. Aunque los ensayos y los diseños cuasi-experimentales pueden funcionar bien cuando hay una intervención simple que puede probarse, las intervenciones a menudo no lo son. Las ONG que trabajan en la sociedad civil, por ejemplo, rara vez trabajan solas o tienen la oportunidad de manipular un «tratamiento» experimental claro para un ensayo aleatorio. Los evaluadores están buscando otras formas de hacer las cosas, enfoques que implican ir más allá de una simple causalidad uno-a-uno y que requieren una reflexión cuidadosa sobre su «contribución».

En el complejo mundo de la acción humanitaria, es poco probable que su programa por sí solo sea la condición necesaria o suficiente para el éxito. Podría ser solo un factor entre muchos, parte de un «paquete causal». El éxito del programa depende de qué más esté sucediendo o haya ocurrido a su alrededor. En el pasado, una simple explicación causal de «lo que funciona» pudo haber sido suficiente, pero hoy en día es más común que los investigadores de la evaluación pregunten «¿nuestra intervención marcó una diferencia? ¿Qué habría ocurrido sin ella?»

Hay una variedad de enfoques que se están volviendo más populares entre los evaluadores, como los diseños basados en la teoría⁵⁶ y basados en casos, la teoría de la complejidad, la evaluación realista,⁵⁷ el análisis de contribución,⁵⁸ el seguimiento de procesos⁵⁹ o el análisis comparativo cualitativo.⁶⁰ Una característica común de estos enfoques es el reconocimiento de la complejidad de la causalidad que requiere la identificación de los mecanismos causales en una gama de casos o acontecimientos específicos, así como elementos en la cadena causal que varían. Esto implica «identificar en qué condiciones y en qué combinaciones con otras causas, es probable que el efecto observado de una intervención se produzca, se desbloquee

o se desencadene».⁶¹ Por ejemplo, en el análisis de contribución, la investigación no intenta probar que un factor, como una política específica, «causó» el resultado deseado, sino que pretende explorar la contribución de una política a los resultados observados. Reconoce que los efectos son producidos por varias causas al mismo tiempo, no solo por una.

No todos estos enfoques son nuevos.⁶² De hecho, contar con una buena teoría probablemente siempre ha estado en el corazón de la buena ciencia. Por lo tanto, el uso de un enfoque «basado en la teoría» no es novedoso, pero estas formas de evaluar el impacto han aumentado en popularidad y realmente ayudan a los evaluadores a abordar la causalidad múltiple.

¿MERECIÓ LA PENA? EL VALOR DE LOS ANÁLISIS DE COSTE-BENEFICIO Y DE COSTE-EFICACIA

Después de haber determinado y medido los efectos de una intervención, es importante que el sector humanitario, como cualquier otro sector, sepa si sus programas e intervenciones tienen una buena relación calidad-precio. Cuando los presupuestos son limitados, debemos tomar decisiones financieras difíciles sobre si una cuestión es más valiosa que otra. Los encargados de la formulación de políticas, las entidades que encargan servicios, los organismos que otorgan subvenciones y los donantes de caridad solicitan más y mejores medidas de impacto por cada libra, dólar o euro que gastan. Las técnicas de evaluación económica ayudan a generar esta información sobre el valor.

Existe una gama de técnicas para realizar una evaluación económica. Un análisis de coste-eficacia (ACE) identifica cuáles son las mejores formas de lograr un resultado. Por otra parte, el análisis de coste-beneficio (ACB) nos dice si hay mejores formas de usar los recursos disponibles. El análisis de coste-utilidad (ACU) calcula qué intervención produce la mayor sensación de bienestar subjetivo para las personas afectadas por ella.¹²

Siempre que sea posible, la evaluación económica trata de atribuir valores monetarios a todos los insumos y actividades que intervienen en la producción de los resultados que, a su vez, conducen a los resultados deseados y a sus efectos económicos, sociales y ambientales a más largo plazo. Esto constituye el análisis del coste. Los mismos procedimientos se utilizan para atribuir valores monetarios a todos los resultados y efectos que la política pretende lograr. Cuando hay resultados y efectos negativos, estos se deducen del valor monetario de cualquier resultado / efecto positivo que se haya logrado. Esto constituye el análisis de los beneficios.

Los valores monetarios se derivan lo más posible de donde hay alguna actividad de mercado. El mercado laboral, por ejemplo, proporciona valores monetarios para diferentes tipos y unidades de trabajo. El mercado comercial de la construcción proporciona el valor monetario (el coste) para los bienes que se requieren para construir un hospital o una escuela: ladrillos, acero, vidrio, cemento, etc. Los costes de brindar atención médica se pueden derivar de los costes fijos y variables procedentes de administrar hospitales, centros de salud y clínicas comunitarias, de medicamentos y suministros médicos, entre otros. Esto es lo que significa «monetizar» los costes y los beneficios.

Algunos costes y, sobre todo, beneficios no se negocian en un mercado y, por lo tanto, son más difíciles de monetizar. En tales casos, los economistas estimarán los beneficios (y, a veces, los costes) simulando la actividad del mercado a través de encuestas a las personas sobre la «voluntad de pagar» y la «voluntad de aceptar». La voluntad de pagar (WTP, por sus siglas

en inglés) se ha definido como «la máxima cantidad de dinero a la que una persona está dispuesta a renunciar para recibir un bien».¹² De manera similar, la voluntad de aceptar (WTA, por sus siglas en inglés) es «la mínima cantidad de dinero que habrían de recibir para ser compensados por privarse de un bien o renunciar a él».¹² Los resultados de una encuesta sobre WTP o WTA proporcionan lo que los economistas denominan «preferencias declaradas», es decir, lo que las personas dicen que estarían dispuestas a pagar o aceptar por un bien o servicio. Alternativamente, los economistas intentarán observar cómo se comportarán realmente las personas en una situación de mercado real. Es decir, establecerán una situación en la que se les pide a las personas que paguen diferentes cantidades de dinero por un bien o servicio o se les ofrecerán diferentes cantidades de dinero para que renuncien a un bien o servicio y luego se registra el comportamiento real de las personas. Esto proporciona lo que los economistas llaman «preferencias reveladas».

Además de los procedimientos mencionados anteriormente, la evaluación económica también identifica dónde o sobre quién recaerán los costes y beneficios y en qué lapso de tiempo. El valor económico de los beneficios que se obtendrán en el futuro a mediano y largo plazo se descontará de acuerdo con lo que se conoce como tasa de descuento. Otros ajustes que se realizan a los costes y beneficios incluyen cambios en los precios relativos, diferencias materiales en las opciones fiscales, sesgo de optimismo e imprevistos como inundaciones, sequías y efectos del cambio climático. La evaluación económica es una iniciativa bastante técnica, que generalmente requiere la experiencia de un economista.

Los diseños cuasi-experimentales podrían seguir permitiendo inferencias causales fuertes en circunstancias en las que un ensayo aleatorio no sería posible o aceptable.

Mensajes clave para la Parte 1 de la Sección C

- ✓ No toda la evidencia es igual. Alguna será de mejor calidad y más apropiada a su desafío.
- ✓ El tipo de investigación debe ajustarse a las necesidades del desafío y el método que se utilice debe haberse realizado y analizado de manera adecuada.
- ✓ No se desanime por la larga lista de métodos de investigación, diseños y enfoques. La clave es comprender los supuestos que sustentan estos métodos.
- ✓ Algunos diseños de investigación (como los diseños experimental y cuasi-experimental) funcionan mejor para demostrar la presencia de una relación causal.
- ✓ Otros enfoques de investigación son más apropiados para explicar cómo se producen dichas relaciones causales o para responder a otros tipos de preguntas de investigación.
- ✓ La evaluación y las técnicas basadas en la teoría, tales como el análisis de contribución, son cada vez más populares entre los evaluadores. Estas técnicas pueden ser útiles cuando es difícil hacer un experimento o es imposible atribuir una única política o programa a un solo resultado claro.
- ✓ No se trata solo de si su intervención funcionó o no, sino de si ha sido rentable.
- ✓ Las técnicas como el análisis de coste-beneficio pueden ayudarle a comprender el valor financiero del impacto de una intervención.
- ✓ Podría ser importante establecer qué sectores de la población se benefician más de una intervención. A menudo, esto puede quedar enmascarado al basarse en el efecto promedio que ha tenido en la población en su conjunto.

Es importante para el sector humanitario saber si sus programas e intervenciones son rentables.

PARTE 2: ¿CÓMO CALIFICA LA CALIDAD DE LA INVESTIGACIÓN?

Otra forma de ayudarle a elegir qué tipo de investigación necesita es hacer preguntas como: ¿en qué investigación puedo confiar? ¿Qué evidencia es lo suficientemente buena para cubrir mis necesidades? ¿Podría ser un artículo en una revista científica de alto impacto? ¿Qué pasa con las evaluaciones internas realizadas por mi propia organización? ¿Deberían también tener un lugar en la mesa de evidencia?

Ser publicado en una revista de investigación arbitrada es una forma de ayudarle a sentirse seguro acerca de una investigación, pero no es una garantía de alta calidad. En un famoso artículo, John Ioannidis de la Universidad de Stanford en los Estados Unidos causó revuelo al argumentar que «la mayoría de los hallazgos publicados probablemente son falsos». Examinó los artículos más citados (aquellos con más de 1000 citas) en algunas de las revistas médicas más reconocidas del mundo, en gran parte de *The Lancet*, *el New England Journal of Medicine* y *el Journal of the American Medical Association*.⁶³ De aquellos que afirmaban ser eficaces, cuyos resultados se habían probado en estudios posteriores, se encontró que el 41 % estaba equivocado o que el impacto era mucho menor del que había sugerido el estudio original.

¿CÓMO AYUDA LA REVISIÓN POR PARES A LOS RESPONSABLES DE LA TOMA DE DECISIONES?

También debemos tener en cuenta que la revisión por pares, la cual es una piedra angular de las revistas académicas, está lejos de ser perfecta. Puede haber sesgos inconscientes, como el «gregarismo», en el que el comportamiento de los revisores está influenciado por el comportamiento de sus homólogos. Y todo el sistema de publicaciones científicas puede verse distorsionado por el sesgo de publicación: los resultados

positivos tienen más posibilidades de ser publicados,⁶⁴ mientras que los datos negativos se ocultan. Como ejemplo, una encuesta realizada por investigadores de la Universidad de Stanford encontró que la mayoría de los «estudios nulos» de las ciencias sociales nunca se publican: solo el 20 % había aparecido en una revista y el 65 % ni siquiera se había redactado.⁶⁵ Esto podría tener graves consecuencias para los responsables de la toma de decisiones que revisan la evidencia. Si nunca se ven los estudios negativos sobre una nueva intervención, se podría concluir erróneamente que todo está bien y que es eficaz.

Sin embargo, a pesar de estos problemas, la revisión por pares sigue siendo, para la mayoría, la «regla de oro» para decidir qué se incluye en la literatura científica: proporcionando una comparación con el mal trabajo.⁶⁶ Utilizar un artículo de una revista o un informe que haya sido revisado de manera independiente por otros expertos es una forma de ayudarle a tener más confianza en que la investigación es fiable. A pesar de que pueden pasar años antes de que un estudio sea publicado (y esto puede ser demasiado tiempo si tiene que tomar una decisión rápida), la revisión por pares aporta algunos controles de calidad a la literatura para ayudarle a sentirse seguro.

DEFINICIÓN DE INVESTIGACIÓN DE «ALTA CALIDAD»

La revisión por pares puede darnos un mínimo de comodidad. Pero, ¿qué debe hacer si va a incluir evidencia que no ha sido verificada por otros expertos? ¿Cómo decide si esto le proporciona una base lo suficientemente sólida para sus decisiones?

Claramente, sería un error ignorar evidencia de investigación importante solo porque no se ha publicado en una revista arbitrada. O perder la gran cantidad de evaluaciones de donantes o ONG que nunca se publican

en los medios académicos, sino que están disponibles solo en la «literatura gris». Existe una amplia variedad de definiciones de «literatura gris», pero generalmente se refiere a documentos que no están publicados o que se han publicado sin revisión por pares. También puede referirse a investigaciones que aún están en marcha o que están siendo preparadas para su publicación. Los informes gubernamentales, las declaraciones y los sumarios sobre políticas y las actas de conferencias también son tipos de «literatura gris». Este tipo de literatura es importante porque puede contener evidencia de resultados negativos e intervenciones infructuosas, lo cual es importante para el equilibrio de la toma de decisiones basada en la evidencia. Se puede buscar esta literatura utilizando bases de datos electrónicas, como Open Grey (www.opengrey.eu), actas de conferencias y registros de adquisiciones de los patrocinadores de investigación. Los sitios web de organizaciones que tienen interés o experiencia en un cierto tema son otra fuente de «literatura gris». Se puede contactar con estas organizaciones para identificar a los investigadores y a los responsables de la toma de decisiones que tienen experiencia en un asunto o tema particular.

Un buen comienzo al tratar de evaluar la calidad de la evidencia es definirla. Sin embargo, uno de los problemas es que palabras como «calidad», «normas», «solidez», «sesgo» y «fuerza» se usan a menudo como si fueran intercambiables y sin definir claramente lo que significan. Esto genera muchos malentendidos. Por ejemplo, en algunas guías,⁶⁷ la «calidad» de investigación significa usar diseños y métodos particulares, como un ensayo

aleatorio. Este énfasis en minimizar el sesgo como un medio para garantizar la calidad surge de algunos de los enfoques clínicos y de salud formales para evaluar la calidad de la evidencia, como el sistema GRADE⁶⁸ o la Escala de Métodos Científicos de Maryland⁶⁹. Estos enfoques de evaluación de la calidad para las evaluaciones experimentales generalmente se basan en la validez interna de los estudios, la calidad de los informes y la validez externa.

La calidad se refiere a qué tan bien se han realizado, informado y analizado los estudios,⁷⁰ así como a la integridad de los investigadores en cuanto a no distorsionar o falsificar los datos.⁷¹ Algunas personas también vinculan la calidad con la relevancia del estudio para la política y la práctica.⁷²

Al intentar responder una pregunta causal, debe considerar si el diseño de investigación utilizado para un estudio es apropiado para determinar la causalidad y si el diseño se implementó correctamente en el estudio. Las evaluaciones de impacto de alta calidad responderán a las preguntas de atribución: demostrarán que la intervención causó los resultados. Esto requiere una comparación o un grupo de control lo más parecido posible al grupo de intervención en todos los aspectos excepto en la intervención como tal. Si esto es cierto y el estudio ha sido bien realizado, puede estar más seguro de que, por ejemplo, los efectos en la prevención de la violencia, la reducción del estrés familiar o el regreso más rápido al trabajo se deben a la intervención. También es importante considerar si los efectos encontrados en el estudio se reproducirán en otros lugares. Esto impulsa la demanda de métodos mixtos de investigación y evaluación y también puede requerir información de investigación cualitativa.

Claramente, sería un error ignorar evidencia de investigación importante solo porque no se ha publicado en una revista arbitrada.

INVESTIGACIÓN CUALITATIVA DE ALTA CALIDAD

Muchos han destacado el valor de la investigación cualitativa de buena calidad⁷³ para los encargados de la formulación de políticas. La evidencia cualitativa puede ser tan científicamente creíble como la cuantitativa y es probable que una combinación de ambas ayude a los responsables de la toma de decisiones a combinar la evidencia de los efectos de una intervención con el conocimiento para decidir sobre la aplicabilidad de esa evidencia a su situación. La clave es que la evidencia de investigación que utilice para enfrentar cualquier desafío sea adecuada para su propósito. Debe haber utilizado el método apropiado, haber recopilado y analizado sus datos (ya sean cualitativos o cuantitativos) utilizando métodos bien definidos y replicables e informar los hallazgos de una manera transparente que minimice la parcialidad.

Al igual que con la investigación cuantitativa, existen normas y directrices para evaluar la calidad de la investigación y la evaluación cualitativa.⁷⁴ Mientras que la calidad de los estudios experimentales se basa generalmente en la validez interna, la adecuación de los informes y la validez externa, la calidad de la investigación cualitativa por lo general se construye entorno a las dimensiones de contribución, defensa, rigor y credibilidad. La contribución se refiere a si el estudio promueve un conocimiento o una comprensión más amplios sobre una política, una práctica, una teoría o un campo sustantivo particular. La defensa se refiere a si el estudio proporciona una estrategia de investigación adecuada para abordar las preguntas de evaluación planteadas. El rigor se refiere a la recopilación, el análisis y la interpretación sistemáticos y transparentes de datos cualitativos. La credibilidad se refiere a qué tan bien fundados y plausibles son los argumentos sobre la evidencia generada.

EVITE «ELEGIR A SU ANTOJO» USANDO REVISIONES SISTEMÁTICAS

Los responsables de la toma de decisiones deben pensar en la calidad de todo el conjunto de evidencia, no solo en piezas de evidencia aisladas. Deben utilizar colecciones agregadas

de investigación y la síntesis precisa de la información existente se ha descrito como «*la oferta individual más importante por parte de los académicos para el proceso político*».⁷⁵ Gran parte de la reflexión sobre la calidad expuesta anteriormente se ha centrado en estudios aislados de la «investigación primaria»; en otras palabras, estudios individuales, como experimentos, encuestas o una serie de entrevistas. Pero, es importante considerar las «síntesis de investigación», el resumen y la combinación de una serie de estudios individuales. Esta idea no es nueva y el concepto de tomar decisiones sobre la base de la evidencia acumulada se ha enfatizado durante siglos.⁷⁶

Nuestra discusión anterior sobre los sesgos cognitivos mostró cómo las personas pueden caer inconscientemente en la trampa de buscar evidencia que se ajuste a sus creencias. Por lo tanto, debemos tener cuidado al recopilar la investigación existente y estar dispuestos a incluir la investigación, incluso si no nos gustan sus conclusiones. Una forma de evitar «elegir a su antojo» es utilizar lo que se conoce como revisiones sistemáticas.⁷⁷ Estas tienen el objetivo de ser exhaustivas y encontrar la mayor investigación posible que sea relevante para responder a la pregunta que ayudará con el desafío al que nos enfrentamos. Usan métodos explícitos para identificar lo que se puede decir de manera confiable sobre la base de los estudios disponibles y evaluar el diseño de la investigación y los métodos de estos estudios para determinar su calidad.

Algunas de las características clave⁷⁸ de una revisión sistemática son:

Planteamiento claro de la pregunta de investigación, de los objetivos y de los criterios de elegibilidad para los estudios.

Metodología explícita y reproducible para minimizar la parcialidad.

Búsqueda sistemática para identificar tantos estudios como sea posible que cumplan con los criterios de elegibilidad.

Evaluación formal de la validez de los hallazgos de los estudios incluidos.

Presentación sistemática y síntesis de las características y los hallazgos de los estudios incluidos.

ESTUDIO DE CASO



Utilización de revisiones sistemáticas para formular orientaciones para el personal

El Comité Internacional de Rescate (IRC, por sus siglas en inglés) deposita una gran confianza en la evidencia a la hora de elaborar sus herramientas y documentos de orientación sobre los programas para el personal de campo. Esto lo llevan a cabo mediante la realización de revisiones de la evidencia de muchas fuentes de alta calidad en torno a intervenciones o enfoques específicos y utilizan lo aprendido sobre los efectos, contextos, poblaciones y condiciones para informar si es necesario adaptar tales intervenciones a contextos específicos y cómo hacerlo.

Esto no está exento de desafíos, ya que la mayoría de los estudios disponibles no brindan información crítica, como la fidelidad de la implementación, los mecanismos causales y la sensibilidad de género.

La labor de toda la agencia del IRC para garantizar que la evidencia esté fácilmente a disposición de su personal se realiza mediante el desarrollo del marco de resultados y evidencia (oef.rescue.org). En esta plataforma electrónica accesible al público, el IRC ha definido los resultados y subresultados en los que el IRC se centra en su trabajo, las teorías de cambio generales (o caminos) a través de las cuales se pueden lograr esos resultados y los indicadores para medirlos.

Para cada resultado y subresultado, el IRC ha resumido la evidencia cuantitativa mejor disponible sobre la eficacia de las intervenciones que buscan cambiar / mejorar el resultado o subresultado pertinente, con un enfoque principal en la evidencia de las revisiones sistemáticas. Para los temas en los que las revisiones sistemáticas aún no existen, el IRC ha identificado y resumido las evaluaciones de impacto individuales.

Los enfoques sistemáticos también tienen el valor de ser explícitos acerca de cómo buscaron estudios de investigación y decidieron sobre su elegibilidad. Por lo tanto, al menos en teoría,⁷⁹ otras personas podrían replicar la revisión sistemática.

En 2018, Christl Donnelly y sus colegas sugirieron cuatro principios para ayudar a los investigadores, los encargados de la formulación de políticas y otros a encargar, realizar, compartir, evaluar y utilizar síntesis de evidencia, incluidas las revisiones sistemáticas.⁸⁰ Propusieron que la síntesis de evidencia debería ser inclusiva, rigurosa, transparente, accesible y elaborada en estas cuatro áreas de la siguiente manera:

Inclusiva

Involucra a los encargados de la formulación de políticas y es relevante y útil para ellos.

Toma en consideración muchos tipos y fuentes de evidencia.

Utiliza una serie de capacidades y personas.

Rigurosa

Utiliza el conjunto de evidencia más completo posible.

Reconoce y minimiza la parcialidad.

Se revisa de manera independiente como parte de un proceso de control de calidad.

Transparente

Describe claramente la pregunta de investigación, los métodos, las fuentes de evidencia y el proceso de control de calidad.

Comunica las complejidades y los puntos de controversia.

Reconoce las suposiciones, limitaciones e incertidumbres, incluida cualquier brecha en la evidencia.

Declara los intereses personales, políticos y organizativos y gestiona los conflictos.

Accesible

Está redactada en lenguaje sencillo.

Está disponible en un marco de tiempo adecuado.

Está disponible en línea de forma gratuita.

METAANÁLISIS

El metaanálisis es una parte importante de las revisiones sistemáticas donde todos o algunos de los estudios incluidos son lo suficientemente similares («homogéneos») en términos de población, intervención, comparadores y resultados para que sus conclusiones estadísticas se agrupen y se agreguen. Esto da lugar a una estimación acumulativa del efecto que suele tener un riesgo de sesgos más bajo que la estimación del efecto para los estudios individuales aislados y mayor precisión. El metaanálisis suele representarse con un diagrama de bosque⁸¹, como el que se muestra en la **figura C.1**, tomado de una revisión sistemática reciente del impacto de los programas de apoyo para las poblaciones afectadas por emergencias humanitarias. El metaanálisis es quizás mejor conocido por combinar los resultados de ensayos aleatorios, pero también se puede realizar con datos de otros tipos de estudios, como los que han usado diseños de control de casos, transversales o de cohortes.

DIAGRAMAS DE BOSQUE

La **figura C.1** presenta las conclusiones de 28 evaluaciones de impacto en las que se compararon personas en emergencias humanitarias que habían recibido programas de salud mental y apoyo psicosocial con las que no habían recibido este tipo de intervención.⁸² En este diagrama de bosque, los datos presentan los efectos de los programas en la gravedad de los trastornos por estrés postraumático

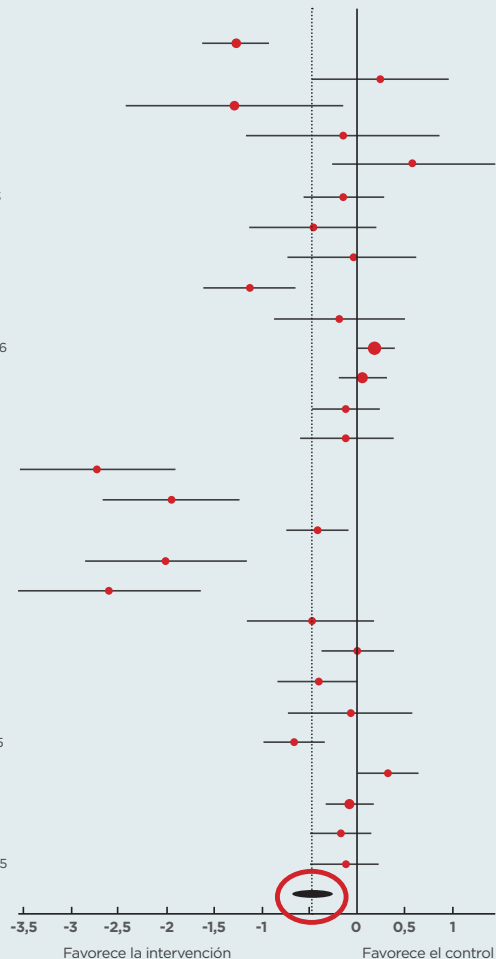
Figura C.1: Diagrama de bosque del efecto de los programas de salud mental y apoyo psicosocial en el trastorno por estrés postraumático⁸²

Medida: continua: d(Hedges g)

Heterogeneidad: $Q=206$; $df=27$; $I^2=86,9\%$; $\tau^2=0,29$

Modelo de efectos aleatorios: $-0,463$ ($-0,689$, $-0,237$)

Berger (2009)	-1,269 (-1,607 -0,932) W:5,1
Catani (2009)	0,251 (-0,47 0,971) W:1,1
Chen (2014)	-1,277 (-2,41 -0,144) W:0,5
Chen (2014)	-0,138 (-1,152 0,875) W:0,6
Cluver (2015)	0,59 (-0,257 1,436) W:0,8
Dybdahl (2001)	-0,137 (-0,558 0,284) W:3,3
Ertl (2011)	-0,457 (-1,12 0,206) W:1,3
Ertl (2011)	-0,035 (-0,7 0,629) W:1,3
Gordon (2008)	-1,116 (-1,595 -0,637) W:2,5
Jordans (2010)	-0,18 (-0,858 0,498) W:1,3
Khamis (2004)	0,205 (0,006 0,404) W:14,6
Khamis (2004)	0,066 (-0,178 0,311) W:9,7
Lange-Neilsen (2012)	-0,11 (-0,462 0,243) W:4,6
Layne (2008)	-0,109 (-0,599 0,38) W:2,4
McMullen (2013)	-2,727 (-3,539 1,915) W:0,9
O'Callaghan (2013)	-1,944 (-2,66 -1,228) W:1,1
O'Callaghan (2014)	-0,405 (-0,72 -0,09) W:5,8
O'Callaghan (2015)	-1,994 (-2,848 1,141) W:0,8
O'Callaghan (2015)	-2,595 (-3,558 1,633) W:0,6
Pityratstian (2015)	-0,473 (-1,137 0,19) W:1,3
Qouta (2012)	0,01 (-0,367 0,387) W:4,1
Qouta (2012)	-0,401 (-0,813 0,011) W:3,4
Schauer (2008)	-0,062 (-0,7 0,576) W:1,4
Tol (2008)	-0,656 (-0,98 -0,332) W:5,5
Tol (2012)	0,329 (0,006 0,652) W:5,6
Tol (2012)	-0,074 (-0,329 0,18) W:9
Tol (2014)	-0,166 (-0,479 0,148) W:5,9
Tol (2014)	-0,102 (-0,426 -0,222) W:5,5
TOTAL:	-0,463 (-0,689 -0,237)



(TEPT) como una variable continua. Para cada estudio en este diagrama de bosque, el punto rojo representa el efecto promedio del tratamiento de la intervención y las líneas paralelas a cada lado de los puntos rojos representan el intervalo de confianza para tal estudio. La línea negra vertical sólida que corre desde el 0 en el eje horizontal indica que no hay ninguna diferencia entre usar o no usar el programa, y los resultados de todos los estudios se agrupan para proporcionar la estimación general de los efectos de los programas. Esta nueva estadística de resumen es el diamante negro (en el círculo rojo) en la parte inferior del diagrama de bosque. Esto representa la *estimación acumulativa del efecto* de agrupar y agregar los tamaños de los efectos promedio y las variaciones de todas las 28 evaluaciones de impacto incluidas en la revisión. Esto nos permite concluir que, en promedio, los programas de apoyo psicosocial y salud mental tienen un pequeño efecto positivo en el TEPT en comparación con no utilizar estas intervenciones.

EVALUACIONES RÁPIDAS DE EVIDENCIA

La preparación de las revisiones sistemáticas y de los metaanálisis puede tomar mucho tiempo. Esto significa que si no está disponible una revisión sistemática actualizada, las personas que necesitan tomar una decisión urgente podrían tener que realizar sus propias búsquedas de la evidencia pertinente y luego evaluar y sintetizar esto más rápido que como ocurriría en una revisión sistemática formal. Afortunadamente, la ya enorme cantidad de revisiones sistemáticas sigue creciendo rápidamente y describimos cómo encontrarlas, así como diversas recopilaciones que están disponibles en la sección D. Sin embargo, si no puede encontrar lo que busca entre las revisiones sistemáticas existentes o las revisiones que encuentra son obsoletas, podría tener que pensar en encargar una «revisión sistemática

reducida», como una evaluación rápida de evidencia.⁸³ Estas revisiones rápidas normalmente toman de 1 a 3 meses y están programadas para satisfacer las necesidades de los encargados de la formulación de políticas y los médicos que no pueden esperar una revisión sistemática completa. Utilizan la misma estructura básica y las etapas de una revisión sistemática completa, pero no son tan profundas, exhaustivas o integrales. Tomarán más atajos en la búsqueda, la evaluación crítica, la extracción y recopilación de datos y el análisis estadístico de los estudios incluidos. Las conclusiones también se presentan de forma más corta y menos detallada que en una revisión sistemática completa y podrían no ser de más de 25 páginas, con un resumen ejecutivo de tres páginas y un documento informativo de una página para los responsables de la toma de decisiones.

Las limitaciones de las evaluaciones rápidas de evidencia consisten en que no son tan integrales o exhaustivas como las revisiones sistemáticas y tienen una mayor probabilidad de ser objeto de sesgos que una revisión sistemática completa. Por consiguiente, se necesita mayor precaución al basar una decisión en las pruebas de una evaluación rápida de evidencia que en una revisión sistemática completa. A pesar de estas limitaciones, los encargados de la formulación de políticas y los ejecutores de programas las encargan y utilizan a menudo, especialmente cuando el tiempo es esencial y no se dispone de revisiones sistemáticas.

Un tipo particular de revisión rápida, denominado evaluación rápida de necesidades de investigación, también se puede utilizar para identificar rápidamente las brechas en la evidencia. El equipo de apoyo rápido para la salud pública del Reino Unido para brotes de enfermedades incluye un plan para realizar estas evaluaciones con Evidence Aid, para identificar las incertidumbres que podría abordar la investigación en las etapas tempranas de una emergencia humanitaria relacionada con un brote de enfermedad.

LA IMPORTANCIA DE LA REPETICIÓN Y LA CORROBORACIÓN

Reflexionar sobre el concepto de brechas en la evidencia, nos conduce a una de las otras cuestiones que hay que considerar al evaluar la calidad de un resumen de estudios de investigación: el número de estudios que es necesario incluir para sentirse seguro de que el conjunto de evidencia es lo suficientemente sólido. ¿Sería suficiente solo un par de estudios si son realmente buenos? ¿O se necesitarían docenas o incluso cientos?

La realidad es que no existe un número mágico de estudios. Sin embargo, no podemos ignorar el hecho de que el tamaño del conjunto de evidencia es importante: existe fuerza en los números y debemos contar con repetición y corroboración. Incluso los estudios que han recibido muchos elogios necesitan repetirse. Amgen, una empresa farmacéutica californiana, intentó replicar 53 estudios históricos de cáncer. El trabajo, publicado en *Nature*,⁸⁴ confirmó los hallazgos de tan solo seis (el 11 %) de estos importantes estudios. Esto es sumamente preocupante, ya que los estudios han influenciado a las empresas farmacéuticas y los tratamientos contra

el cáncer mundialmente y la incapacidad de reproducirlos podría indicar que sus conclusiones no son confiables. Por otra parte, no queremos ver que los estudios se repitan una y otra vez mucho después de que la base empírica sea sólida.⁸⁵ Los estudios innecesarios representan un desperdicio de la investigación⁸⁶ y, en el contexto de los ensayos aleatorios, podrían ser poco éticos si algunos participantes siguen siendo aleatorizados a intervenciones que distan de ser óptimas.

ANÁLISIS DE REVISIONES: CÓMO EVALUAR LOS CONJUNTOS DE EVIDENCIA DE ALTA CALIDAD

Debemos recordar que los métodos formales para reunir las investigaciones, por ejemplo en revisiones sistemáticas, no siempre conducen a una respuesta de alta calidad. Sigue siendo necesario evaluar la calidad del diseño y aplicación de la revisión de la evidencia y la calidad de cualquier síntesis solo puede ser tan buena como la calidad de los estudios sobre los que se basa. Incluso el análisis mejor realizado no puede convertir una investigación de baja calidad en una respuesta de alta calidad.

También existen herramientas formales de evaluación para valorar la calidad de las revisiones sistemáticas, como AMSTAR.⁸⁷ Asimismo, se puede utilizar una de las listas de control originales para evaluar la calidad de las revisiones sistemáticas,⁷⁸ que formulan una serie de preguntas:

¿La pregunta está claramente definida?

¿La búsqueda de estudios pertinentes es exhaustiva?

¿Los criterios de inclusión son adecuados?

¿Se aborda adecuadamente validez de los estudios incluidos?

¿Se ha obtenido la información faltante de los investigadores originales?

¿Qué tan susceptibles son los resultados a cambios en la manera de realizar la revisión?

¿Se han interpretado los análisis por subgrupos cuidadosamente?

Una nota del Departamento de Desarrollo Internacional del Reino Unido sobre cómo evaluar la solidez de la evidencia enumera cuatro puntos por considerar al examinar los conjuntos de evidencia:⁴¹

La calidad (técnica) de los estudios que constituyen el conjunto de evidencia (o el grado en que se abordó el riesgo de sesgo).

El tamaño del conjunto de evidencia.

El contexto en el que se sitúa la evidencia.

La coherencia de los resultados producidos por los estudios que constituyen el conjunto de evidencia.

¿Las conclusiones surgen de la evidencia que se analiza?

¿Las recomendaciones están vinculadas a la solidez de la evidencia?

¿Son válidos los juicios sobre las preferencias (valores)?

¿Se confunde «pruebas de ausencia de efecto» con «no hay pruebas del efecto»?

La evaluación de todos estos criterios siempre será algo bastante subjetivo y se necesita tener en cuenta el contexto de la cuestión política. Es difícil dar consejos generales sobre lo que constituiría el

conjunto de evidencia correcto para una política. También estamos comenzando a ver la combinación de las conclusiones de múltiples revisiones sistemáticas en «análisis de revisiones». Por supuesto, esto no puede funcionar si no hay un conjunto de revisiones sistemáticas existentes, y en algunos sectores hay pocas. Sin embargo, estos análisis de revisiones pueden proporcionar una evaluación y un resumen de evidencia de múltiples revisiones sistemáticas de la misma intervención; o proporcionar una reseña de la evidencia de una recopilación de revisiones de diferentes intervenciones en el mismo ámbito temático.⁸⁸

Mensajes clave para la Parte 2 de la Sección C

✓ Para encontrar evidencia en la que pueda confiar, busque investigación revisada por pares. Pero tenga en cuenta que la revisión por pares está lejos de ser perfecta. Podría haber sesgos inconscientes, como el «gregarismo», sesgo de publicación hacia resultados positivos o incluso la falsificación o distorsión deliberada de datos.

✓ Al analizar las preguntas de impacto y «lo que funciona», utilice los marcos y normas formales de evidencia, como los utilizados por Nesta y otros.

✓ Un estudio nunca es suficiente. Evite tomar decisiones basadas en estudios individuales y busque repeticiones múltiples. Hay fuerza en los números y se

necesita repetición y corroboración. Incluso los estudios que han recibido muchos elogios necesitan repetirse, pero es necesario evitar realizar los estudios una y otra vez mucho después de que la base empírica sea suficientemente sólida.

✓ Utilice revisiones sistemáticas, cuyo propósito es ser exhaustivas y analizar la calidad de los estudios, por lo general sobre la base de los métodos y el diseño de la investigación.

✓ Si no tiene tiempo de realizar una revisión sistemática o no puede encontrar una ya completada que responda a su pregunta, podría realizar o encargar una evaluación rápida de evidencia.⁸⁹

Incluso el análisis mejor realizado no puede convertir una investigación de baja calidad en una respuesta de alta calidad.

¿Dónde debería buscar evidencia?

Esta sección analiza distintos tipos de evidencia y examina cómo elegir la más apropiada según el problema de interés. También analiza cómo evaluar la calidad de la evidencia.

BÚSQUEDA DE RESULTADOS DE INVESTIGACIONES

Con el rápido crecimiento del número de personas y organizaciones que trabajan en revisiones sistemáticas, la literatura ahora cuenta con una cantidad cada vez mayor de estas revisiones, por lo que es posible que encuentre una que satisfaga sus necesidades. Una estimación reciente es que hay más de 200.000 revisiones sistemáticas en todas las áreas temáticas. Por supuesto, solo una pequeña parte de ellas sería relevante para el sector humanitario, pero (como se señala a continuación) hay conjuntos de revisiones disponibles gratuitamente en colecciones en línea, como las comisariadas por Evidence Aid. También hay grandes organizaciones internacionales, como Cochrane (www.cochrane.org) y Campbell Collaboration (www.campbellcollaboration.org), dedicadas a la preparación y el mantenimiento de revisiones sistemáticas, que las publican en línea.

Sin embargo, si necesita realizar sus propias búsquedas de estudios sobre los efectos de la acción humanitaria, el punto de partida es reconocer que no siempre son fáciles de encontrar y es posible que necesite la ayuda de un bibliotecario o un especialista en información que pueda asesorarlo y quizá le ayude en el diseño y la ejecución de la búsqueda. La evidencia está dispersa en miles de informes repartidos en miles

de revistas, libros y sitios web, muchos de los cuales no son gratuitos o fáciles de consultar. Al considerar dónde buscar, es importante tener en cuenta dónde es probable que haya sido publicada e indexada la evidencia que le interesa y qué tan exhaustivo desea ser. Por ejemplo, si se publicaron en una revista científica, los artículos podrían estar disponibles a través de una de las miles de bases de datos bibliográficas electrónicas. Estas incluyen, por ejemplo, PubMed para la atención médica, LILACS para artículos de especial relevancia para Sudamérica o ERIC para literatura educativa. También puede ser necesario identificar literatura gris, como informes gubernamentales o de ONG e investigaciones presentadas en conferencias, lo que requerirá búsquedas en bases de datos electrónicas, como Open Grey (www.opengrey.eu), y en los sitios web de las organizaciones que quizás hayan llevado a cabo investigaciones en la materia.

Al planificar una búsqueda, puede ser útil dividir los elementos de búsqueda en los tipos de intervención o política que le interesan, la población o los entornos pertinentes, las mediciones de resultados que serían más útiles para su toma de decisiones y los tipos de estudio que desea encontrar. Luego puede decidir qué área o áreas son las más importantes y útiles para identificar el material que más le interesa.

Debe tener en cuenta los tipos de fuentes por consultar, que pueden incluir:

Bases de datos bibliográficas (p. ej., PubMed, LILACS, ERIC).

Sitios web de revistas y conferencias (p. ej., aquellos centrados en desastres o acción humanitaria).

Registros de investigaciones (tales como a través del portal de ensayos de la OMS, www.researchregistry.com y PROSPERO para revisiones sistemáticas).

Recopilaciones en línea de evidencia de investigaciones (véase a continuación, p. ej., ALNAP, 3ie).

Sitios web de organizaciones relevantes (p. ej., para informes gubernamentales y de ONG).

Referencias en artículos.

Correspondencia con investigadores y evaluadores.

Al elegir los términos por incluir en su búsqueda, debe considerar:

Sinónimos de diferentes épocas y lugares.

Otras palabras y oraciones relacionadas con lo que le interesa.

Palabras más amplias.

Palabras más específicas o enfocadas.

Términos de indexación o palabras clave asignadas a los artículos por los autores originales o la base de datos bibliográfica.

Y, por último, debe decidir si aplicará alguna restricción según el idioma o el período en que se realizó o publicó la investigación.

FUENTES DE REVISIONES SISTEMÁTICAS Y OTRAS INVESTIGACIONES PARA EL SECTOR HUMANITARIO

Debido a estos retos al buscar evidencia, es posible que sea más eficaz y fácil utilizar repositorios confiables de investigaciones y revisiones sistemáticas, en lugar de depender de búsquedas al azar en internet. Afortunadamente, varias organizaciones trabajan actualmente para recopilar evidencia pertinente en fuentes en línea. Estas incluyen las siguientes:

La Red de Aprendizaje Activo sobre Rendición de Cuentas y Resultados de la Acción Humanitaria (ALNAP, por sus siglas en inglés) es organizada por el Instituto de Desarrollo de Ultramar (ODI, por sus siglas en inglés) en Londres, Reino Unido, con una página web que contiene más de 15.000 recursos, incluidas las publicaciones de la ALNAP relativas a la evidencia y a los detalles de miles de evaluaciones.

Luego del tsunami en el Océano Índico en 2004, se sentaron las bases de Evidence Aid en Cochrane, cuando se publicaron gratuitamente en línea los textos completos de varias docenas de revisiones de Cochrane relacionadas con la respuesta a los desastres. Se estableció como organización benéfica en 2015 y ya ha recopilado cientos de revisiones sistemáticas accesibles gratuitamente desde su página web (www.evidenceaid.org), incluidos, entre otros, conjuntos de Colecciones Especiales de revisiones relacionadas con vendavales, terremotos, Ébola, el estado de salud de los refugiados y solicitantes de asilo, el trastorno de estrés postraumático y la prevención de la malnutrición aguda en situaciones de emergencia y crisis humanitarias.

La Iniciativa Internacional para la Evaluación de Impacto (3ie) se creó en 2008 y actualmente ofrece cuatro bases de datos para consulta en línea (www.3ieimpact.org). Dos de ellas, la Base de Datos de Revisiones Sistemáticas de 3ie y la

Base de Datos de Evaluaciones de Impacto, catalogan la evidencia de la eficacia de las intervenciones en el sector humanitario. También incluyen revisiones sistemáticas y evaluaciones de impacto sobre el contexto general del desarrollo internacional, muchas de las cuales son relevantes para las intervenciones en situaciones de emergencia.

Estamos llevando una lista más completa y actualizada de este tipo de recursos en línea, en: www.evidenceaid.org/online-collections-of-research-for-the-humanitarian-sector

La lista proporciona una gran variedad de recursos de investigación en línea, varios de ellos gratuitos y de fácil acceso. Deberían ser útiles para cualquier encargado de la formulación de políticas, ONG o profesional de primera línea en el sector humanitario, y proporcionan un acceso fácil a evidencia fiable y de alta calidad sobre la eficacia de las intervenciones. Si desea sugerir recursos adicionales para esta lista, póngase en contacto con Evidence Aid:

info@evidenceaid.org

CONCLUSIÓN

Como conclusión de esta guía sobre el uso de la evidencia en el sector humanitario, le recomendamos aprovechar los resúmenes de investigaciones accesibles, procesables y disponibles de manera gratuita, como las revisiones sistemáticas contenidas en las páginas web que hemos enumerado. Esto le ayudará a moverse rápido hacia las fuentes de evidencia que podrían ayudar a informar su política y su práctica.

Mensajes clave para la Sección D

✓ La evidencia de investigación pertinente para el sector humanitario está dispersa en decenas de miles de informes distribuidos en miles de revistas, libros y páginas web.

✓ Buscar esta evidencia puede requerir ayuda de un especialista en información y diseñar la búsqueda requiere considerar con cautela lo que se está buscando, las fuentes de búsqueda, los términos para usar en la búsqueda y cualquier restricción relacionada con el idioma o el período.

✓ Las revisiones sistemáticas proporcionan resúmenes y síntesis de la evidencia de investigación.

✓ Hay disponibles muchos repositorios de investigación en línea, lo que mejora el acceso a las revisiones sistemáticas y a otros tipos de evidencia de investigación.

Podría resultarle más eficaz y fácil usar repositorios de investigación confiables y revisiones sistemáticas.

Notas finales

- ¹ Collins S, Sadler K. (2002) Outpatient care for severely malnourished children in emergency relief programmes: a retrospective cohort study. *Lancet* 360: 1824-30.
- ² OMS, Programa Mundial de Alimentos, Comité Permanente de Nutrición del Sistema de las Naciones Unidas y UNICEF. (2007) Community-based management of severe acute malnutrition: a joint statement by the WHO, World Food Programme, UN System Standing Committee on Nutrition and UNICEF. Nueva York: UNICEF.
- ³ Price AI, Djulbegovic B. (2017) What does evidence mean? Most languages translate “evidence” into “proof”. *Journal of Evaluation in Clinical Practice* 23(5): 971-3.
- ⁴ Sackett D, et al. (1996) Evidence based medicine: what it is and what it isn't. *British Medical Journal* 312: 71. Véase www.bmj.com/content/312/7023/71.
- ⁵ La definición de medicina basada en evidencia surgió en un momento en el que la mayoría de las decisiones médicas se basaban en la experiencia, la autoridad y la eminencia. La práctica médica no se basaba en la mejor evidencia científica disponible. Algunos comentaristas e investigadores sostienen que la política social se encuentra en el mismo lugar que la medicina hace 20 o 30 años, es decir, que la autoridad domina la toma de decisiones, en lugar de la evidencia de investigaciones.
- ⁶ Allen C, et al. (2016) Evidence Aid. *Oxford Public Health* August: 51-54; Blanchet K, et al. (2017) Evidence on public health interventions in humanitarian crises. *Lancet* 390: 2287-96; y Christoplos I, et al. (2017) Strengthening the quality of evidence in humanitarian evaluations. ALNAP Method Note. Londres: ALNAP/ODI (véase www.alnap.org/system/files/content/resource/files/main/alnap-eha-method-note-5-2017.pdf).
- ⁷ Knox Clarke P, Darcy J. (2014) Insufficient evidence? The quality and use of evidence in humanitarian action. Estudio ALNAP. Londres: ALNAP/ODI. Véase www.alnap.org/system/files/content/resource/files/main/alnap-study-evidence.pdf.
- ⁸ www.oxforddictionaries.com/definition/english/evidence.
- ⁹ Nutley S, et al. (2013) What Counts as Good Evidence? Londres, Reino Unido: Alliance for Useful Evidence.
- ¹⁰ Hallam A, Bonino F. (2013) Using Evaluation for a Change: Insights from humanitarian practitioners. Estudio ALNAP. Londres: ALNAP/ODI. Véase www.alnap.org/system/files/content/resource/files/main/alnap-study-using-evaluation-for-a-change.pdf.
- ¹¹ Ariely D. (2009) Predictably Irrational: The Hidden Forces that Shape Our Decisions. Londres, Reino Unido: HarperCollins.
- ¹² HM Treasury. (2011) The Green Book: Appraisal and Evaluation in Central Government. Londres, Reino Unido: HM Treasury.
- ¹³ Christoplos I. (2006) Links between Relief, Rehabilitation and Development in the Tsunami Response: A Synthesis of Initial Findings. Estocolmo, Suecia: Swedish International Development Cooperation Agency.
- ¹⁴ Zsombok CE, Klein G, (editores). (2014) Naturalistic decision making. Nueva York, EE.UU.: Psychology Press.
- ¹⁵ Kahneman D, et al. (2009) Conditions for intuitive expertise: a failure to disagree. *American Psychology* 64: 515-26. Véase www.ncbi.nlm.nih.gov/pubmed/19739881.
- ¹⁶ Estas estimaciones provienen de Cash Learning Partnership (CaLP). (2018) The State of the World's Cash Report. Oxford: CaLP; Doing Cash Differently: Report of the High Level Panel on Cash Transfers. (2015); y Development Initiatives. (2017) Global Humanitarian Assistance Report.
- ¹⁷ ODI, Development Initiatives. (2016) Counting Cash: Tracking Humanitarian Expenditure on Cash-Based Programming.
- ¹⁸ Oxfam. (2006) Good Practice Review 11: Cash Transfer Programming in Emergencies. Oxford: Oxfam.
- ¹⁹ Jackson R (Save the Children UK), Kukrety N (Oxfam GB) (2012) Institutionalising cash transfer programming. Véase <https://odihpn.org/magazine/institutionalising-cash-transfer-programming/>.

- ²⁰ Véase www.humanitarianresponse.info/en/programme-cycle/space.
- ²¹ La discusión de los desafíos de la toma de decisiones tras las circunstancias excepcionales de una emergencia grave por radiación está disponible en Carr Z, et al. (2016) Using the GRADE approach to support the development of recommendations for public health interventions in radiation emergencies. *Radiation Protection Dosimetry* 171: 144-55; y Ohtsuru A, et al. (2015) Nuclear disasters and health: lessons learned, challenges, and proposals. *Lancet* 386: 489-97.
- ²² Bradley DT, et al. (2014) The effectiveness of disaster risk communication: a systematic review of intervention studies. *PLOS Currents Disasters* August 22; Edition 1.
- ²³ Una discusión sobre la importancia de prestar atención a la forma en que los donantes recopilan, utilizan y comparten evidencia e información está disponible en Obrecht A. (2017) Using Evidence to Allocate Humanitarian Resources: Challenges and Opportunities. Documento de trabajo de la ALNAP. Londres: ALNAP/ODI. Véase <https://reliefweb.int/sites/reliefweb.int/files/resources/alnap-ear-resource-allocation-2017.pdf>.
- ²⁴ Nesta. (2013) Understand how innovation works. Vídeo disponible en www.nesta.org.uk/resources/understand-how-innovation-works.
- ²⁵ Obrecht A, Warner AT. (2016) More than just luck: Innovation in humanitarian action. Estudio HIF/ALNAP. Londres: ALNAP/ODI.
- ²⁶ Jones G, et al. (2003) How many child deaths can we prevent this year? *Lancet* 362: 65-71. Véase [https://linkinghub.elsevier.com/retrieve/pii/S0140-6736\(03\)13811-1](https://linkinghub.elsevier.com/retrieve/pii/S0140-6736(03)13811-1).
- ²⁷ Nyhan B, Reifler J. (2010) When corrections fail: The persistence of political misperceptions. *Political Behavior* 32: 303-30.
- ²⁸ White H. (2009) Theory-based impact evaluation: principles and practice, working paper 3. Delhi, India: International Initiative for Impact Evaluation, página 4.
- ²⁹ Waddington H, et al. (2009) Water, Sanitation and hygiene interventions to combat childhood diarrhoea in developing countries: a systematic review. Delhi, India: International Initiative for Impact Evaluation.
- ³⁰ Nesta y TSIP. (2014) Guidance for Developing a Theory of Change for Your Programme. Véase www.nesta.org.uk/sites/default/files/theory_of_change_guidance_for_applicants_.pdf.
- ³¹ Stern E. (2015) Impact Evaluation: A Design Guide for Commissioners and Managers of International Development Evaluations in the Voluntary and Community Sector. Londres, Reino Unido: Big Lottery Fund, Bond, Comic Relief and the Department for International Development.
- ³² Gerdin M, et al. (2014) Optimal evidence in difficult settings: improving health interventions and decision making in disasters. *PLoS Medicine* 11(4): e1001632.
- ³³ Mulgan G. (2015) The six Ws: a formula for what works. Londres: Nesta. Véase www.nesta.org.uk/blog/six-ws-formula-what-works.
- ³⁴ www.ebola-anthropology.net/about-the-network.
- ³⁵ Faye SL. (2015) L'“exceptionnalité” d'Ebola et les “réticences” populaires en Guinée-Conakry. Réflexions à partir d'une approche d'anthropologie symétrique. *Anthropologie & Santé*. Véase <https://journals.openedition.org/anthropologiesante/1796>.
- ³⁶ <http://pubman.mpdl.mpg.de/pubman/item/escidoc:2096578/component/escidoc:2103624/AAA-Ebola-Report-1.pdf>.
- ³⁷ Fairhead J. (2016) Understanding Social Resistance to the Ebola Response in the Forest Region of the Republic of Guinea: An Anthropological Perspective. *African Studies Review* 59: 7-31. doi:10.1017/asr.2016.87.
- ³⁸ Abramowitz S, et al. (2015) Social science intelligence in the global Ebola response. *Lancet* 385: 330.
- ³⁹ Petticrew M, et al. (2003) Evidence, hierarchies and typologies: horses for courses. *Journal of Epidemiology and Community Health* 57: 527-9.
- ⁴⁰ Una lista oficial y exhaustiva de los marcos y métodos de investigación de las ciencias sociales está disponible en Luff R, et al. (2015) Review of the Typology of Research Methods within the Social Sciences. Londres, Reino Unido: ESRC/National Centre for Research Methods. Véase <http://eprints.ncrm.ac.uk/3721>.

- ⁴¹ Department for International Development. (2014) How to Note: Assessing the Strength of Evidence. Véase www.gov.uk/government/uploads/system/uploads/attachment_data/file/291982/HTN-strength-evidence-march2014.pdf.
- ⁴² Adaptado de HM Treasury, DECC y DEFRA. (2012) Quality in policy impact evaluation: understanding the effects of policy from other influences. Londres, Reino Unido: HM Treasury/DEFRA/DECC; Frost S, et al. (2006) The Evidence Guide; Using Research and Evaluation in Social Care and Allied Professions. Londres, Reino Unido: Barnardo's; Petticrew M, Roberts H. (2003) Evidence, hierarchies and typologies: horses for courses. *Journal of Epidemiology and Community Health*. 57: 527-9; y Stern E. (2015) Impact Evaluation; A Design Guide for Commissioners and Managers of International Development Evaluations in the Voluntary and Community Sector. Londres, Reino Unido: Big Lottery Fund, Bond, Comic Relief and the Department for International Development, Tabla 2, página 18.
- ⁴³ Odgaard-Jensen J, et al. (2011) Randomisation to protect against selection bias in health care trials. *Cochrane Database of Systematic Reviews* (4):MR000012.
- ⁴⁴ White H. (2013) An introduction to the use of randomised control trials to evaluate development interventions. *Journal of Development Effectiveness* 5(1): 30-49.
- ⁴⁵ Véase por ejemplo: Schulz KF, et al. (2010) CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *PLoS Medicine* 7(3): e1000251; Higgins JP, et al. (2011) CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 343:d5928.
- ⁴⁶ Puri J, et al. (2015) What methods may be used in impact evaluations of humanitarian assistance? documento de trabajo 22, Delhi, India: International Initiative for Impact Evaluation. Véase www.3ieimpact.org/media/filer_public/2014/12/08/wp_22_humanitarian_methods_working_paper_top.pdf.
- ⁴⁷ Doocy S, Burnham G. (2006). Point-of-use water treatment and diarrhoea reduction in the emergency context: an effectiveness trial in Liberia. *Tropical Medicine and International Health* 11: 1542-52.
- ⁴⁸ White H, et al. (2014) Randomised Controlled Trials (RCTs). *Methodological Briefs*. Evaluación de impacto n°. 7 Florencia, Italia: Unicef. Véase www.unicef-irc.org/publications/pdf/brief_7_randomised_controlled_trials_eng.pdf.
- ⁴⁹ Gertler P. (2000) Final Report: The Impact of PROGRESA on Health. Washington, DC, EE.UU.: International Food Policy Research Institute (IFPRI). Véase www.ifpri.org/publication/impact-progres-a-health.
- ⁵⁰ Hainmueller J, et al. (2017) Catalyst or crown: does naturalization promote the long-term social integration of immigrants? *American Political Science Review* 111(2): 256-76. doi:10.1017/S0003055416000745.
- ⁵¹ Bozorgmehr K, et al. (2015) Effect of restricting access to health care on health expenditures among asylum-seekers and refugees: A quasi-experimental study in Germany, 1994–2013. *PLoS ONE* 10: e0131483.
- ⁵² Rossi R, et al. (2016) Vaccination coverage cluster surveys in Middle Dreib - Akkar, Lebanon: comparison of vaccination coverage in children aged 12-59 months pre- and post-vaccination campaign. *PLoS ONE* 11(12): e0168145.
- ⁵³ Alexander J, Bonino F (2015). A discussion on designs, approaches and examples, ALNAP Discussion Series Improving the quality of EHA evidence, Method Note 4, January 2015. Véase www.alnap.org/system/files/content/resource/files/main/alnap-eha-method-note-addressing-causation-jan2015.pdf.
- ⁵⁴ Waddington H, et al. (2017) Quasi-experimental study designs series paper 6: risk of bias assessment. *Journal of Clinical Epidemiology* 89: 43-52; y Bärnighausen T, et al. (2017) Quasi-experimental study designs series—paper 7: assessing the assumptions. *Journal of Clinical Epidemiology* 89: 53-66.
- ⁵⁵ Nielsen NS, et al. (2013). The Contribution of Food Assistance to Durable Solutions in Protracted Refugee Situations; its impact and role in Bangladesh: A Mixed Method Impact Evaluation, Volume I-Evaluation Report, Geneva, Switzerland: World Food Program/UNHCR.

- ⁵⁶ Carter R. (2012) Helpdesk Research Report: Theory-based evaluation approach. Universidad de Birmingham, Reino Unido: Birmingham: Governance and Social Development Resource Centre. Véase www.gsdrc.org/docs/open/hdq872.pdf.
- ⁵⁷ Westthorp G. (2014) Realist Evaluation: An Introduction. Londres, Reino Unido: Methods Lab, Overseas Development Institute. Véase www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9138.pdf.
- ⁵⁸ Mayne J. (2012) Contribution Analysis: Coming of Age? *Evaluation* 18(3): 270-80.
- ⁵⁹ Para análisis de red y seguimiento de procesos, véase: Befani B, et al. (2014) Process tracing and contribution analysis: a combined approach to generative causal inference for impact evaluation. *IDS Bulletin* 45(6): 17-36.
- ⁶⁰ Baptist C, et al. (2015) Coffey How To Guide: Qualitative Comparative Analysis – A Rigorous Qualitative Method for Assessing Impact. Véase www.coffey.com/assets/Ingenuity/Qualitative-Comparative-Analysis-June-2015.pdf.
- ⁶¹ Stern E, et al. (2012) Broadening the range of designs and methods for impact evaluations, Londres, Reino Unido: Department for International Development.
- ⁶² Por ejemplo, véase la guía de Carol Weiss: Weiss C. (1998) *Evaluation: methods for studying programs and policies*. Nueva York, EE.UU.: Prentice Hall.
- ⁶³ Ioannidis JA. (2005) Contraindications and Initially Stronger Effects in Highly Cited Clinical Research. *Journal of the American Medical Association*. 294(2): 218-28.
- ⁶⁴ Para una revisión sistemática del sesgo de publicación, véase Hopewell S, et al. (2009) Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database of Systematic Reviews* (1): MR000006. doi: 10.1002/14651858.MR000006.pub3.
- ⁶⁵ Franco A, et al. (2014) Social science. Publication bias in the social sciences: unlocking the file drawer. *Science* 345: 1502-5.
- ⁶⁶ Sense about Science. (2006) I don't know what to believe: Making sense of science stories. Véase www.senseaboutscience.org/resources.php/16/i-dont-know-what-to-believe.
- ⁶⁷ Véase por ejemplo, el suplemento de la guía Magenta: HM Treasury, DECC y DEFRA (2012) *Quality in policy impact evaluation; understanding the effects of policy from other influences*. La guía muestra cómo los diseños de investigación de mayor calidad pueden ayudar a superar el desafío de atribuir los resultados medidos a la política en cuestión (en lugar de otras influencias), mientras que los diseños de menor calidad reducen la confianza de si fue la política la que logró esos resultados (página 5).
- ⁶⁸ <http://gradeworkinggroup.org>.
- ⁶⁹ Farrington RP, et al. The Maryland Scientific Methods Scale, en: Farrington DP, et al. *Evidence-based crime prevention*. Londres, Reino Unido: Routledge, 2002 capítulo 2.
- ⁷⁰ Para un modelo de enfoques diferentes sobre la calidad que incluye las cuatro dimensiones de (1) calidad metodológica, (2) calidad de informes, (3) adecuación y (4) pertinencia para la política y la práctica; véase Boaz A, Ashby D. (2003) Fit for purpose? Assessing research quality for evidence based policy and practice. Documento de trabajo 11. Londres, Reino Unido: ESRC UK National Centre for Evidence Based Policy and Practice.
- ⁷¹ Callaway E. (2011) Report finds massive fraud at Dutch universities. *Nature* 479(15). Véase www.nature.com/news/2011/111101/full/479015a.html.
- ⁷² Boaz A, et al. (2003) Fit for purpose? Assessing research quality for evidence based policy and practice. Documento de trabajo 11. Londres, Reino Unido: ESRC UK National Centre for Evidence Based Policy and Practice.
- ⁷³ La investigación cualitativa explora e intenta entender las creencias, experiencias, actitudes, comportamiento e interacciones de las personas. Genera datos no numéricos que podrían recolectarse mediante, por ejemplo, entrevistas profundas, grupos de enfoque, análisis documental y observación de los participantes.

- ⁷⁴ Spencer L, et al. (2002) Quality in Qualitative Evaluation: A framework for assessing research evidence. Londres, Reino Unido: Cabinet Office; CASP. (2018) CASP Checklist for Qualitative Research. Oxford, Reino Unido: Critical Appraisal Skills Programme; National Institute for Health and Clinical Excellence. (2012) The guidelines manual: Appendix H: Methodology checklist: qualitative studies. Londres, Reino Unido: National Institute of Health and Clinical Excellence; y O'Brien BC, et al. (2014) Standards for reporting qualitative research: a synthesis of recommendations. *Academic Medicine* 89: 1245-51.
-
- ⁷⁵ Whitty CJM. (2015) What makes an academic paper useful for health policy? *BMC Medicine* 13: 301.
-
- ⁷⁶ Véase, por ejemplo, la discusión sobre la importancia de usar evidencia existente a la hora de diseñar nuevos estudios en Clarke M. (2004) Doing new research? Don't forget the old: nobody should do a trial without reviewing what is known. *PLoS Medicine* 1: 100-2; y la historia de las revisiones sistemáticas en Chalmers I, et al. (2002) A brief history of research synthesis. *Evaluation and the Health Professions* 25: 12-37; y Clarke M. (2016) History of evidence synthesis to assess treatment effects: personal reflections on something that is very much alive. *JLL Bulletin: Commentaries on the history of treatment evaluation. Journal of the Royal Society of Medicine* 109: 154-63.
-
- ⁷⁷ Allen C. (2014) A resource for those preparing for and responding to natural disasters, humanitarian crises, and major health care emergencies. *Journal of Evidence-Based Medicine* 7: 234-7; y Gurevitch J, et al. (2018) Meta-analysis and the science of research synthesis. *Nature* 555: 175-82.
-
- ⁷⁸ Oxman AD. (1994) Checklists for review articles. *BMJ* 309: 648-51.
-
- ⁷⁹ Para una crítica de la aplicación de las síntesis de investigación a la política, véase Pawson R. (2001) Evidence Based Policy: In Search of a Method. Documento de trabajo 3. Londres, Reino Unido: ESRC UK National Centre for Evidence Based Policy and Practice.
-
- ⁸⁰ Donnelly CA, et al. (2018) Four principles for synthesizing evidence. *Nature* 558: 361-4.
-
- ⁸¹ Lewis S, Clarke M. (2001) Forest plots: trying to see the wood and the trees. *BMJ* 322:1479-80.
-
- ⁸² Bangpan M, et al. (2017) The impact of mental health and psychosocial support programmes for populations affected by humanitarian emergencies. Oxford, Reino Unido: Oxfam.
-
- ⁸³ Ganann R, et al. (2010) Expediting systematic reviews: methods and implications of rapid reviews. *Implementation science* 5: 56; y Tricco AC, et al. (2015) A scoping review of rapid review methods. *BMC Medicine* 13: 224.
-
- ⁸⁴ Begley CG, et al. (2010) Drug development: Raise standards for preclinical cancer research. *Nature* 483: 531-3.
-
- ⁸⁵ Una revisión de investigaciones que habían incluido uno o más metaanálisis acumulativos encontró muchos ejemplos que demuestran que se habrían visto resultados estables (beneficiosos, nocivos y neutros) si se hubiera hecho un metaanálisis de la investigación existente antes de comenzar un nuevo ensayo aleatorio, lo que habría conducido a la implementación más rápida de intervenciones eficaces: Clarke M, et al. (2014) Accumulating research: a systematic account of how cumulative meta-analyses would have provided knowledge, improved health, reduced harm and saved resources. *PLoS ONE* 9(7): e102670.
-
- ⁸⁶ Chalmers I, et al. (2014) How to increase value and reduce waste when research priorities are set. *Lancet* 383: 156-65.
-
- ⁸⁷ Shea BJ, et al. (2007) Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology* 7:10; y Shea BJ, et al. (2017) AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of health care interventions, or both. *BMJ* 358: j4008.
-
- ⁸⁸ Smith V, et al. (2011) Methodology in conducting a systematic review of systematic reviews of health care interventions. *BMC Medical Research Methodology* 11: 15.
-
- ⁸⁹ Véase, por ejemplo: Brennan RJ, et al. (2005) ERapid health assessment in Aceh Jaya district, Indonesia, following the December 26 tsunami. *Emergency Medicine Australasia* 17: 341e350; y Beebe J. (2014) Rapid Qualitative Inquiry: a Field Guide to Team-based Assessment, segunda edición. Lanham, Maryland, EE.UU.: Rowman & Littlefield.
-

Karl Blanchet, Director del Centro de Crisis Sanitarias y Humanitarias, Escuela de Higiene y Medicina Tropical de Londres
Karl.Blanchet@lshtm.ac.uk | crises.lshtm.ac.uk

Mike Clarke, Director de investigación de Evidence Aid, Centro de Salud Pública, Universidad de la Reina de Belfast
mclarke@qub.ac.uk | www.qub.ac.uk

