

Guías SURE para preparación y uso de resúmenes ejecutivos para formulación de políticas

5. Selección y descripción de opciones para abordar el problema

Resumen

La selección y descripción de las opciones para abordar un problema de los sistemas de salud requieren opciones potenciales de identificación, hallazgo de evidencia de los impactos de esas opciones, evaluación de esa evidencia y un resumen de lo que se conoce acerca de los posibles impactos de las opciones viables. En la medida de lo posible, las revisiones sistemáticas se tendrían que utilizar como base para la descripción de los posibles impactos de las opciones incluidas. Es importante evaluar tanto la cantidad de confianza que se puede poner en una evaluación como la confianza que se tiene en los posibles impactos de la opción, basados en la evidencia que se ha revisado. La evidencia de los efectos de los mecanismos de los sistemas de salud y las estrategias de aplicación es a menudo limitada, sobre todo la evidencia en países de bajos y medianos ingresos. Como consecuencia, a menudo puede ser necesario tomar decisiones basadas en evidencia de baja calidad. No obstante, una revisión superficial de las evidencias implica una serie de riesgos, incluyendo descripciones no fiables de las opciones, pérdida de credibilidad, introducción de evaluaciones sesgadas, y exageraciones o subestimaciones del grado de confianza que se puede tener en las estimaciones de los efectos o supuestos de los posibles impactos de las opciones.

Las siguientes preguntas se pueden utilizar para decidir y describir las opciones que va a incluir un resumen ejecutivo para formulación de políticas:

1. ¿Qué opciones se presentan?
2. ¿Qué se sabe acerca de los impactos de las opciones alternativas?
3. ¿Qué confianza podemos tener acerca de los posibles impactos de las opciones alternativas?
4. ¿Cómo se tendría que resumir la información sobre los posibles impactos de las opciones alternativas?

Antecedentes

Una vez que se definió el problema que abordará un resumen ejecutivo para formulación de políticas, el próximo paso es describir las opciones viables para resolver ese problema. Esto implica cuatro pasos:

- Identificación de las posibles opciones
- Hallazgo de evidencia de los impactos de esas opciones
- Evaluación de esa evidencia
- Resumen de lo que se conoce acerca de los posibles impactos de las opciones viables

Temas a considerar

Las siguientes preguntas, que abordan los cuatro pasos anteriores, se pueden utilizar para decidir y describir las opciones que un resumen ejecutivo para formulación de políticas va a incluir:

1. ¿Qué opciones se presentan?
2. ¿Qué se sabe acerca de los impactos de las opciones alternativas?
3. ¿Qué confianza podemos tener acerca de los posibles impactos de las opciones alternativas?
4. ¿Cómo se tendría que resumir la información sobre los posibles impactos de las opciones alternativas?

1. ¿Qué opciones se presentan?

Se puede utilizar una serie de estrategias para identificar opciones potenciales, incluyendo:

- Consideración de los mecanismos alternativos de prestación, financieros y de gobernanza que aborden el problema o sus causas subyacentes
- Uso de marcos desarrollados para hacer frente al problema específico
- Consideración de las intervenciones descritas en las revisiones sistemáticas
- Consideración de las formas en que otras jurisdicciones han abordado el problema
- Consultas con informantes claves
- Lluvia de ideas

Debido a que las causas de los problemas de los sistemas de salud son complejos, las soluciones a estos problemas también pueden ser complejas. Las opciones específicas pueden ser complejas, por ejemplo, la organización de un sistema para reclutar, entrenar, apoyar y retener a los trabajadores comunitarios de salud para proporcionar atención materno-infantil. Además, puede ser necesario integrar diferentes mecanismos de prestación, financieros y de gobernanza para tratar un problema.

Decidir que opciones se tendrían que describir en un resumen ejecutivo para formulación de políticas puede ser relativamente sencillo, por ejemplo, si el resumen ejecutivo se está preparando para describir una solución específica que ya se ha propuesto. Otras veces, decidir que opciones presentar puede requerir varias iteraciones, por ejemplo:

- En primer lugar considerar un amplio rango de opciones y luego restringirlas a las más viables por evaluación de la evidencia para cada una y de su aceptación, o
- En primer lugar considerar los mecanismos potenciales de prestación, financieros y de gobernanza como elementos potenciales de las políticas, luego evaluar la evidencia de esos elementos, y después determinar si y cómo integrar elementos viables en las opciones de política, o
- Considerar primero soluciones posibles y luego que cambios adicionales en los mecanismos de prestación, financieros o de gobernanza pueden ser necesarios, junto con soluciones viables.

Un marco puede ayudar a estructurar la consideración de las posibles soluciones a un problema. A veces el mismo marco que se utiliza para aclarar las causas de un problema se puede utilizar para identificar las posibles soluciones. Por ejemplo, se puede utilizar un marco amplio para los problemas del sistema de salud tanto para aclarar las causas de un problema como para identificar las posibles soluciones.¹ Los marcos más específicos pueden facilitar la

consideración de las posibles soluciones para algunos tipos de problemas. Por ejemplo, un marco para abordar los problemas de los recursos humanos para la salud se puede usar para pensar sistemáticamente en posibles soluciones, así como en las causas de los problemas como la escasez, distribución, rendimiento o costo de los trabajadores de la salud.² Del mismo modo, un marco para la financiación de la atención en salud puede ayudar a pensar de forma sistemática en ambos problemas con la financiación de la atención en salud y las posibles soluciones. Como se señala en la anterior guía SURE³, no hay una forma sencilla de buscar los marcos de este tipo. A menudo, la forma más eficiente de encontrar estos es hablar con personas con experiencia en el área específica. Otras fuentes de marcos como éstos incluyen otros resúmenes ejecutivos sobre el mismo tema o estrechamente relacionados, análisis de políticas y revisiones sistemáticas o reseñas de revisiones sistemáticas.

Las revisiones sistemáticas o las reseñas de las revisiones sistemáticas que describen las opciones para tratar un problema particular pueden estar disponibles. Además de proporcionar un marco, estas podrían proveer un inventario de intervenciones y la evidencia de sus impactos. En el Apéndice 1 se describen estrategias para la búsqueda de revisiones sistemáticas.

También puede ser útil en la identificación de posibles soluciones tener en cuenta las formas en que otras jurisdicciones han abordado problemas similares. Al igual que con los marcos, no hay una forma sencilla de buscar descripciones de las diversas políticas que se han probado en otras jurisdicciones. La manera más eficiente de encontrar estas puede ser hablar con las personas con experiencia en el área específica. En este caso, puede ser particularmente útil contactar con personas de otros países con problemas similares. También puede ser útil consultar con los principales partes interesadas, así como con personas con conocimientos pertinentes para averiguar cuáles son las soluciones que ellos consideran viables, basados en su experiencia, conocimientos y percepciones.

Es probable que la lluvia de ideas o pensamiento creativo sea útil y puede ser fundamental para decidir qué opciones presentar como alternativas viables para solucionar un problema. Esto se puede hacer tanto en forma estructurada, utilizando un marco, como en una manera no estructurada. Es deseable la participación de personas con diferentes perspectivas y personas con un amplio conocimiento del sistema de salud en la lluvia de ideas y puede ser conveniente hacer esto en dos o más fases, por ejemplo: primero generando posibles soluciones e ideas acerca de dónde y cómo identificar otras posibles soluciones consideradas después de tener en cuenta una descripción del problema y sus causas, y luego encontrar y evaluar la evidencia de los posibles efectos de las posibles opciones (o componentes de las opciones), y luego reunirse nuevamente para discutir la viabilidad de las posibles opciones y decidir qué opciones presentar en el resumen ejecutivo para formulación de políticas.

Las opciones que se presentan en un resumen ejecutivo pueden ser opciones mutuamente excluyentes, donde es necesario hacer una elección entre dos o más opciones. En este caso, las opciones tendrían que ser descritas de manera que faciliten la comparación entre ellas y una evaluación bien informada de los pros y los contras de cada opción. Por otra parte, si las opciones que se presentan no son mutuamente excluyentes, un resumen ejecutivo tendría que dejar claro cualquier beneficio potencial de combinar las diferentes opciones, así como los pros y los contras de cada opción.

2. ¿Qué se sabe acerca de los impactos de las opciones alternativas?

Búsqueda de revisiones sistemáticas

El punto de partida ideal para descubrir lo que se conoce acerca de los impactos de las posibles opciones es una revisión sistemática. Hay un número cada vez mayor de revisiones sistemáticas sobre los mecanismos del sistema de salud y las estrategias de implementación. La base de datos del PPD/CCNC es un buen lugar para comenzar la búsqueda de revisiones sistemáticas que abordan los impactos de los sistemas de salud. Otras fuentes en las que se puede buscar revisiones sistemáticas de los impactos de los mecanismos de los sistemas de salud y las estrategias de implementación incluyen la Biblioteca Cochrane y PubMed. En el Apéndice 1, se describen estas bases de datos y sus estrategias de búsqueda.

Selección de las revisiones sistemáticas

Habiendo encontrado referencias potencialmente relevantes, es necesario decidir qué revisiones sistemáticas recuperar en texto completo y mirar con más detalle. Tener criterios explícitos de selección puede ayudar a guiar estas evaluaciones, así como a hacer el proceso transparente para los demás. Esos criterios podrían especificar, por ejemplo:

- Requisitos metodológicos mínimos (por ejemplo, sección de métodos, criterios explícitos de selección, descripción de la estrategia de búsqueda que se utilizó)
- Tipos de poblaciones, pacientes o problemas que la revisión necesita abordar
- Tipos de intervenciones, opciones, mecanismos de los sistemas de salud o estrategias de implementación que la revisión necesita abordar

A veces puede ser adecuado contar con un criterio relacionado con el entorno (por ejemplo, atención primaria), sin embargo, considerar solamente revisiones restringidas a países específicos o países de bajos y medianos ingresos pueden excluir evidencia importante (a veces la mejor evidencia disponible). Es probable que sea más informativo incluir revisiones con estudios realizados en distintos contextos, y posteriormente evaluar la aplicabilidad de los resultados a su entorno, utilizando criterios como los de la Tabla 1.

Tabla 1. Evaluaciones sobre la aplicabilidad de los resultados de las revisiones sistemáticas

Las decisiones sobre cómo aplicar los resultados de las revisiones sistemáticas siempre se encuentran ubicadas en contextos específicos. Por lo tanto, es necesario hacer evaluaciones de las posibles diferencias entre donde se hizo la investigación resumida en una revisión sistemática y su contexto. Para intervenciones en sistemas de salud esto incluye consideraciones de las diferencias en:

- elementos estructurales de los sistemas de salud (por ejemplo, que una intervención no podría trabajar de la misma manera)
- realidades y limitaciones del terreno (que podría alterar sustancialmente los beneficios potenciales de la intervención)
- perspectivas e influencias de las partes interesadas del sistema de salud (por ejemplo, que la intervención no se pueda aceptar o recoger de la misma manera)

Además, las condiciones basales (riesgos) pueden producir diferentes efectos absolutos, aunque la efectividad relativa sea la misma. Si bien esto puede no disminuir la confianza en la evidencia, es importante tener esto en mente al aplicar los resultados de los estudios de un contexto a otro.

Evaluación de la cantidad de confianza a tener en una revisión sistemática

Una vez que se encuentra una revisión sistemática relevante, es necesario decidir cuánta confianza tener en la revisión. Esta es una evaluación diferente a la que decide cuánta confianza tener en la evidencia, lo que se aborda en la siguiente sección.

Una serie de listas de verificación está disponible para guiar la evaluación de la confiabilidad de las revisiones sistemáticas. Una lista de verificación adaptada de la utilizada para la preparación de resúmenes SUPPORT se puede ver en el Apéndice 2. Se basa en otras listas similares,^{4,5,6} pero se ha desarrollado con base en la experiencia de aplicación de una lista de verificación ampliamente utilizada^{7,8} para revisiones sistemáticas sobre mecanismos del sistema de salud y estrategias de implementación y la adaptación de la lista de verificación para guiar las evaluaciones sobre hasta qué punto es probable que una revisión proporcione un resumen confiable de la mejor evidencia disponible de los impactos de estas intervenciones complejas.

La lista de verificación se divide en dos partes. La primera evalúa los métodos utilizados para identificar, seleccionar y evaluar críticamente los estudios. La segunda evalúa los métodos utilizados para analizar los resultados de los estudios incluidos. Las evaluaciones resumidas basadas en las preguntas de cada parte determinan si una revisión va a tener limitaciones leves, moderadas o grandes. Una revisión general de las evaluaciones, sobre la base de estas dos evaluaciones resumidas (y otras limitaciones potenciales) ubican luego a las revisiones en tres categorías que pueden guiar el uso de revisiones en resúmenes ejecutivos para formulación de políticas:

- Errores fatales – La revisión tiene limitaciones que son suficientemente importantes como para que sus resultados no sean confiables y no se tendrían que utilizar en un resumen ejecutivo para formulación de políticas (aunque todavía puede ser posible extraer algunos mensajes claves o información útil de la revisión, por ejemplo un marco para identificar las posibles opciones).
- Limitaciones importantes - La revisión tiene limitaciones que son lo suficientemente importantes como para buscar otra revisión sistemática e interpretar los resultados de esta revisión con cautela, si no se puede encontrar una mejor revisión (y potencialmente complementar la información proporcionada por la revisión con búsquedas adicionales o información de los estudios incluidos).
- Confiable - La revisión solo tiene limitaciones de menor importancia y se puede usar como un resumen confiable de la mejor evidencia disponible.

Si no se puede encontrar una revisión sistemática sin limitaciones importantes puede ser necesario buscar estudios individuales ya sea para complementar la información de una revisión o para reemplazar a una revisión sistemática. Si se hace esto, se debe prestar atención a los mismos procesos que se utilizan en una revisión sistemática (como se indica en el Apéndice 2), es decir, en lo posible, se tendría que usar métodos sistemáticos y transparentes (explícitos) para encontrar, seleccionar y evaluar críticamente los estudios, y para sintetizar los resultados de los estudios relevantes. Lo ideal sería que los métodos utilizados para hacer esto se describan en un anexo del resumen ejecutivo para formulación de políticas.

3. ¿Qué confianza podemos tener acerca de los posibles impactos de las opciones alternativas?

En la toma de decisiones de políticas de salud, los formuladores de políticas y las partes interesadas deben compensar los beneficios y desventajas de las estrategias alternativas. Los tomadores de decisiones se verán influenciados no sólo por las mejores estimaciones de las ventajas y desventajas esperadas, sino también por su confianza en estas estimaciones, es decir, la calidad de la evidencia. El sistema GRADE proporciona un marco estructurado y transparente para hacer evaluaciones sobre la calidad de la evidencia.⁹

Utilizando el marco GRADE, se debe hacer valoraciones independientes de la calidad de la evidencia para cada resultado importante. Al igual que los primeros sistemas de clasificación de la calidad de la evidencia, el marco GRADE comienza con el diseño del estudio. Los ensayos aleatorizados proporcionan, en general, mayor evidencia que los estudios observacionales. Por lo tanto, los ensayos aleatorizados sin limitaciones importantes constituyen evidencia de alta calidad. Los estudios observacionales sin fortalezas especiales o limitaciones importantes generalmente proporcionan evidencia de baja calidad. Sin embargo, hay un número de factores que pueden reducir o aumentar nuestra confianza en las estimaciones del efecto.

El marco GRADE considera cinco factores que pueden disminuir la calidad de la evidencia:

1. Limitaciones del estudio
2. Resultados inconsistentes entre los estudios
3. Evidencia indirecta
4. Imprecisión
5. Sesgo de publicación

Y tres que pueden aumentar la calidad de la evidencia:

1. Estimaciones grandes del efecto
2. Gradientes dosis-respuesta
3. Confusión posible que aumentaría la confianza en una estimación

Factores que pueden disminuir la calidad de la evidencia

1. Nuestra confianza en las estimaciones del efecto disminuye si los estudios sufren de limitaciones más importantes que pueden sesgar sus estimaciones del efecto de la intervención. Para los ensayos aleatorizados estas limitaciones incluyen, por ejemplo, falta de ocultamiento de la asignación, falta de cegamiento, especialmente si los resultados son subjetivos y su valoración altamente susceptible al sesgo, grandes pérdidas en el seguimiento, falta de adhesión al análisis con intención de tratar, o fallas en el reporte de los resultados (por lo general aquellos para los que se observó poco o ningún efecto).
2. Las estimaciones del efecto que difieren ampliamente a través de los diferentes estudios para lo que no hay explicaciones convincentes reducen nuestra confianza en conocer cual es el verdadero efecto. La variabilidad puede deberse a las diferencias en las poblaciones (por ejemplo, las intervenciones pueden tener efectos relativos menores pequeños en poblaciones desfavorecidas); a las diferencias en las intervenciones (por ejemplo, efectos menores con incentivos financieros más pequeños), o los resultados (por ejemplo, disminución de los efectos con el tiempo). Cuando la variabilidad existe pero los investigadores no logran identificar una explicación plausible, disminuye la calidad de la evidencia.

3. Los tomadores de decisiones deben considerar tres tipos de evidencia indirecta que pueden disminuir la calidad de las evidencias. El primero ocurre cuando se considera el uso de una de las dos intervenciones, A y B. Aunque las comparaciones randomizadas de A y B pueden no estar disponibles, los ensayos aleatorizados pueden haber comparado A con ninguna intervención y B con ninguna intervención. Tales ensayos permiten comparaciones indirectas de la magnitud del efecto de A y B. Esta evidencia es de menor calidad que las que proporcionarían las comparaciones cabeza a cabeza de A y B. El segundo tipo de evidencia indirecta se produce cuando se utilizan resultados sustitutos, como una medida del proceso que puede o no reflejar exactamente lo que se puede esperar en términos de un resultado importante, como la mortalidad o la morbilidad. El tercer tipo de evidencia indirecta son las diferencias entre la población, la intervención o comparador con la intervención de interés y los incluidos en los estudios pertinentes. Esto es comúnmente referido como "aplicabilidad" (o generalización). En la Tabla I, se describen los criterios para evaluar hasta que punto es probable que los resultados de una revisión sobre los mecanismos de los sistemas de salud sean aplicables en su contexto.
4. Cuando los estudios incluyen relativamente pocas personas y pocos eventos y así tienen amplios intervalos de confianza (o un gran valor de p), se debe tener menos confianza en una estimación.
5. La calidad de la evidencia se reducirá si hay una alta probabilidad de que algunos estudios no se hayan reportado (generalmente los que no muestran ningún efecto). El riesgo de ese "sesgo de publicación" es mayor cuando la evidencia publicada se limita a un pequeño número de ensayos, todos los cuales son patrocinados por personas con un interés personal en los resultados.

Factores que pueden elevar la calidad de la evidencia

1. Incluso estudios observacionales bien hechos generalmente producen solo evidencia de baja calidad, debido a la gran cantidad de potenciales factores de confusión que no se conocen, o bien, no se miden. Sin embargo, ocasionalmente pueden aportar evidencia de calidad moderada o inclusive alta. Cuanto mayor sea la magnitud del efecto, menos probable es que esto se pueda explicar por factores de confusión y la evidencia es, por lo tanto, más fuerte.
2. La presencia de un gradiente dosis-respuesta puede aumentar nuestra confianza en las estimaciones de los efectos, por ejemplo, si los efectos mayores están asociados con variaciones más intensas de una intervención, como se podría esperar.
3. Cuando se encuentra un efecto, si todos los factores de confusión plausibles disminuirían la magnitud del efecto, esto aumenta la calidad de la evidencia, ya que podemos estar más seguros que el efecto es al menos tan grande como la estimación e inclusive puede ser mayor. Por el contrario, en especial para las preguntas de seguridad, si se encuentra poco o ningún efecto y todos los sesgos plausibles llevarían hacia una sobreestimación del efecto, podemos estar más seguros de que es poco probable que sea un efecto importante.

GRADE proporciona un enfoque claramente articulado e integral para clasificar y resumir la calidad de la evidencia que apoya el tratamiento y las recomendaciones para la prestación de atención sanitaria. A pesar de que siempre serán necesarios los cálculos para cada paso, el enfoque GRADE sistemático y transparente permite el escrutinio y el debate de esos cálculos. Lo ideal sería que las consideraciones de este tipo sean sistemáticas y explícitas. Aunque los detalles de estos cálculos no van a ser probablemente incluidos en un resumen ejecutivo para

formulación de políticas, mantener un "rastros de auditoría" de las evaluaciones que se han hecho puede ser útil en la producción del resumen ejecutivo, puede ayudar a asegurar su confiabilidad y credibilidad, proteger contra la introducción de evaluaciones sesgadas, y puede ser útil disponer de él si esas evaluaciones son posteriormente puestas en tela de juicio o debatidas. En el Apéndice 3 se proporcionan hojas de trabajo para hacer estas evaluaciones.

4. ¿Cómo se debería resumir la información sobre los posibles impactos de las opciones alternativas?

Una "hoja de balance" o "resumen de la tabla de hallazgos" es una manera simple pero poderosa de presentar las ventajas y desventajas de las diferentes opciones.^{10,11} Proveer texto en forma de viñetas, que resumen los mensajes claves cualitativamente junto con un resumen de la tabla de hallazgos, puede mejorar aún más la comprensión.

El objetivo de un resumen de los hallazgos es ayudar a los tomadores de decisiones a desarrollar una comprensión exacta de las consecuencias importantes de las opciones que se comparan. Un resumen de los resultados ayuda a lograr esto de varias maneras. En primer lugar, condensa la información más importante, lo que permite un análisis eficiente. En segundo lugar, centra la atención en los resultados más importantes. Esto aumenta la probabilidad de que los tomadores de decisiones tengan una percepción exacta de lo que se conoce acerca de los impactos de las opciones consideradas y las consecuencias importantes. En tercer lugar, el acto de construir un resumen de los hallazgos es un mecanismo útil para organizar el pensamiento, estructurar el análisis de la evidencia, y centrar el debate. En cuarto lugar, un resumen de los hallazgos puede ayudar a desarrollar evaluaciones más explícitas sobre cuáles son las consecuencias más importantes de las opciones, la evidencia subyacente, y las evaluaciones posteriores sobre el equilibrio entre las ventajas y desventajas relativas de las opciones presentadas en un resumen ejecutivo. Por último, un resumen de los hallazgos puede proporcionar a otros tomadores de decisiones "información en bruto", ayudándolos así a aplicar sus propios juicios acerca del balance entre las consecuencias deseables e indeseables.

Sin embargo, también se debe considerar dos limitaciones importantes cuando se utiliza un resumen de los hallazgos en la toma de decisiones. En primer lugar, cuando hay compensaciones complicadas entre múltiples resultados, las evaluaciones pueden exigir un alto nivel de procesamiento de la información por parte de los formuladores de políticas. En segundo lugar, al sopesar los diferentes resultados, los juicios de valor empleados por los formuladores de políticas podrían quedar implícitos. La elaboración de modelos económicos formales, cuando están disponibles o es posible llevarlos a cabo, puede ayudar a abordar estas limitaciones, haciendo más explícito cualquier supuesto subyacente (incluyendo juicios de valor). Esto permite el uso del análisis de sensibilidad para explorar los efectos tanto de las incertidumbres como de los diversos supuestos en los resultados.

Aunque no es probable que un formato único para un resumen de la tabla de hallazgos sea óptimo para todos los resúmenes ejecutivos para formulación de políticas o incluso para todas las opciones en un solo resumen ejecutivo, la plantilla en la Tabla 2 ilustra la información clave que debe incluir:

1. Un título que indica la comparación que se resume en la tabla

2. Las características de la evidencia, incluyendo los tipos de participantes (pacientes o poblaciones), tipos de contexto (por ejemplo, los países) donde se realizaron los estudios, la intervención y con que se comparó la intervención
3. Los resultados más importantes, incluyendo los beneficios, posibles daños y costos
4. El impacto estimado de la intervención en cada resultado (de preferencia siempre cuantitativamente)
5. La cantidad de información en que se basa la información (número de participantes o de las unidades (por ejemplo: instalaciones) y número de estudios)
6. La calidad de la evidencia para cada resultado (con base en las consideraciones que se acaba de resumir y en el Apéndice 3)

Tabla 2. Plantilla para el resumen de los hallazgos

[Texto] ¹			
Los pacientes o población: [Texto] ²			
Contextos: [Texto]			
Intervención: [Texto]			
Comparación: [Texto]			
Resultados	Impacto	Número de participantes (estudios)	Calidad de la evidencia (GRADE)
[Texto] ³	[Texto] ⁴	[?] ⁵ ([?] Estudios)	. AAA• [Texto] ⁶
[Texto]	[Texto]	[?] ([?] Estudios)	AAA• [Texto]
[Texto]	[Texto]	[?] ([?] Estudios)	AAA• [Texto]
GRADE: Grupo de Trabajo GRADE grados de evidencia (ver Apéndice 3)			

En el Apéndice 3, se pueden encontrar ejemplos de resúmenes de hallazgos. Otros ejemplos se pueden encontrar en los resúmenes SUPPORT (www.support-collaboration.org).

Consideraciones adicionales

Evaluaciones de los impactos sobre la equidad

Los autores del resumen ejecutivo para formulación de políticas deben considerar los posibles impactos sobre la equidad, así como los impactos globales de las opciones, mediante el examen de los hallazgos de una revisión y la consideración de los posibles efectos diferenciales de las intervenciones en poblaciones desfavorecidas. La posibilidad de impactos sobre la equidad se debe considerar en relación con cualquiera de los factores que puede estar probablemente asociado con una desventaja en relación con la opción que se está

considerando, incluyendo situación económica, empleo u ocupación, educación, lugar de residencia, género y etnia.

La siguiente lista de verificación puede ayudar a orientar la consideración de los posibles impactos sobre la equidad:

- ¿Hay razones plausibles para anticipar diferencias en la eficacia relativa de la opción en grupos o contextos desfavorecidos?
- ¿Es probable que haya condiciones basales diferentes entre grupos o contextos de tal manera que la eficacia absoluta de la opción sea diferente, y el problema más o menos importante, en grupos o contextos desfavorecidos?
- ¿Existen consideraciones importantes que se deben hacer durante la implementación de la opción a fin de garantizar que las inequidades se reduzcan, si es posible, y que no aumenten?

Evaluaciones sobre el uso de recursos y costos

En los entornos con recursos limitados, los formuladores de políticas están profundamente preocupados por el uso de recursos, costos y ahorros asociados con las diferentes opciones. Usaremos el término "costos" aquí para referirnos a estas consideraciones de manera general, aunque es importante tener en cuenta que los costos incluyen tanto el uso de recursos (por ejemplo, el tiempo de los trabajadores de la salud) y los valores monetarios (o precios) asociados a los recursos (por ejemplo, los salarios u honorarios).

Los costos y el costo-efectividad de una intervención se deben evaluar en un entorno específico a fin de tener en cuenta las diferencias en los patrones de uso de recursos y los precios. Para ayudar a los tomadores de decisiones a realizar esas evaluaciones, los autores de resúmenes ejecutivos para formulación de políticas deben considerar:

- ¿Cuáles son los costos más importantes, incluyendo los costos de ejecución y la sostenibilidad de la opción?
- ¿Qué información existe sobre los costos, tanto de revisiones sistemáticas como de otras fuentes?¹²
- ¿Hay una incertidumbre importante acerca de los costos a mediano y largo plazo?
- ¿Hay una incertidumbre importante acerca de la aplicabilidad de los costos reportados?

Evaluaciones de la necesidad de seguimiento y evaluación

Debido a que a menudo existe incertidumbre sobre los efectos y el costo-efectividad de las intervenciones, los autores de resúmenes ejecutivos para formulación de políticas deben abordar la necesidad de seguimiento y evaluación de cada opción, incluida la consideración de:

- ¿Es necesario el seguimiento?
- Si el seguimiento es necesario, ¿hay que medirlo?
- ¿Es necesaria una evaluación del impacto?
- Si una evaluación del impacto es necesaria, ¿qué se debe evaluar y cómo?

Conclusión

La evidencia de los efectos de los mecanismos y estrategias de implementación de los sistemas de salud es a menudo limitada, sobre todo la evidencia en países de bajos y medianos ingresos. Como consecuencia, a menudo puede ser necesario que las decisiones se hagan con base en evidencia de baja calidad. No obstante, una revisión superficial de la evidencia implica una serie de riesgos, incluyendo descripciones no confiables de las opciones, pérdida de credibilidad, introducción de evaluaciones sesgadas y exageraciones o subestimaciones del grado de confianza que se puede tener en las estimaciones o supuestos sobre los impactos de las opciones. En consecuencia, aunque es probable que un proceso más sistemático y transparente lleve más tiempo que una revisión superficial de la evidencia; también es probable que esta inversión adicional esté garantizada con respecto a un resumen ejecutivo para formulación de políticas específico y que mejore a través del tiempo en la medida en que las decisiones de política sanitaria estén bien informadas por la mejor evidencia de investigación disponible.

La forma en que se seleccionaron y describieron las opciones incluidas en un resumen ejecutivo se debe evaluar de manera que se pueda aprender las lecciones para futuros resúmenes ejecutivos y se pueda mejorar esta guía. En el Apéndice 4, se adjunta un formulario para evaluar esta guía.

Recursos

Referencias

¹ Lavis JN, et al. Enhancing the retrieval of systematic reviews that can inform health system management and policymaking.

² Chopra M, Munro S, Lavis JN, Vist G, Bennett S. Effects of policy options for human resources for health: an analysis of systematic reviews. *Lancet* 2008; 371:668-74.

³ SURE guides for preparing and using policy briefs: 4. Clarifying the problem.

⁴ West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, et al. Systems to rate the strength of scientific evidence [Evidence report/technology assessment no 47]. 2002 [www.thecre.com/pdf/ahrq-system-strength.pdf]. Rockville (MD): Agency for Healthcare Research and Quality; AHRQ Publication No 02-E016.

⁵ Proposed Evaluation Tools for COMPUS.: Assessment, November 29, 2005. 2005, Canadian Coordinating Office for Health Technology: Ottawa. www.cadth.ca/media/pdf/COMPUS_Evaluation_Methodology_draft_e.pdf

⁶ Oxman AD, Schünemann HJ, Fretheim A. Improving the Use of research evidence in guideline development: 8. Synthesis and presentation of evidence. *Health Res Policy Syst* 2006; 4:20.

⁷ Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA, Streiner DL. Agreement among reviewers of review articles. *Journal of Clinical Epidemiology* 1991; 44:91-98.

⁸ Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology* 1991; 11:1271-1278.

⁹ Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ, and the GRADE Working Group. What is 'quality of evidence' and why is it important to clinicians? *BMJ* 2008; 336:995-8.

¹⁰ Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Liberati A, Vist GE, Schunemann HJ, GRADE Working Group: Incorporating considerations of resources use into grading recommendations. *BMJ* 2008, 336:1170-3.

¹¹ Eddy DM: Comparing benefits and harms: the balance sheet. JAMA 1990, 263:2493. 2498, 2501.

¹² Oxman AD, Fretheim A, Lavis JN, Lewin S. SUPPORT Tools for evidence-informed health Policymaking (STP) 12: Finding and using research evidence about resource use and costs. Health Research Policy and Systems 2009, 7(Suppl 1):S12.

Apéndice 1. Estrategias para la búsqueda de revisiones sistemáticas

Con frecuencia, tendrá sentido buscar en estas tres fuentes de revisiones sistemáticas en el orden en que se enumeran a continuación, utilizando búsquedas posteriores para complementar las búsquedas anteriores mediante búsquedas más específicas para llenar las brechas (por ejemplo, enfocada en opciones para las que no se encontraron revisiones sistemáticas) o para encontrar revisiones más recientes (si sólo se encontraron revisiones obsoletas). A continuación se resumen las estrategias adicionales que se pueden usar para llenar las brechas o encontrar revisiones más recientes seguidas de una breve descripción de las ventajas, desventajas y estrategias de búsqueda para las bases de datos PPD/CCNC, Biblioteca Cochrane y PubMed.

Base de datos PPD/CCNC <http://www.researchtopolicy.ca/search/reviews.aspx>

Ventajas:

- Enfocada exclusivamente en mecanismos de los sistemas de salud
- Contiene información bibliográfica con enlaces de fácil uso cuando estos están disponibles para las revisiones Cochrane, otras revisiones de impactos, y revisiones que abordan otros tipos de preguntas (por ejemplo, revisiones de estudios cualitativos), así como una visión general de las revisiones sistemáticas y resúmenes ejecutivos para formulación de políticas
- Indexada y se puede buscar usando un marco para mecanismos de los sistemas de salud

Desventajas:

- Puede estar atrasada con respecto a la Biblioteca Cochrane y PubMed y puede no incluir todas las revisiones sistemáticas sobre mecanismos de los sistemas de salud de PubMed, debido a las limitaciones de búsqueda en PubMed
- No es fácil realizar búsquedas que no utilicen términos en el marco

Búsqueda en PPD/CCNC

- Se puede buscar haciendo clic en los tipos de mecanismos del sistema de salud de interés
- Menos de 1.000 registros actualmente, por lo tanto la búsqueda es relativamente fácil, aunque actualmente se limita a la utilización de términos de indexación

La Biblioteca Cochrane www.thecochranelibrary.com

Ventajas

- Contiene el texto completo de las revisiones Cochrane en la Base de Datos Cochrane de Revisiones Sistemáticas (CDSR)
- Además contiene resúmenes estructurados de otras revisiones de efectos (Base de datos de Resúmenes de Revisiones de Efectos (DARE) y Base de Datos de Evaluación de la Tecnología en Salud (HTA))
- Ahora se actualiza mensualmente

Desventajas

- Puede estar atrasada con respecto a las bases de datos PubMed, DARE y HTA y puede no incluir todas las revisiones sobre mecanismos de los sistemas de salud que se encuentran en PubMed

Búsqueda en la Biblioteca Cochrane Plus

- Se puede buscar mediante palabras de texto, términos indexados y lógica booleana
- En la actualidad hay menos de 10.000 registros en las bases de datos CDSR, DARE y HTA, por lo tanto la búsqueda es todavía relativamente fácil
- Se puede buscar en línea en las versiones más actualizadas de las bases de datos DARE y HTA (<http://www.crd.york.ac.uk/crdweb/>)

PubMed www.pubmed.gov

Ventajas

- Contiene información bibliográfica, resúmenes y términos indexados
- Está más actualizada y posiblemente es más amplia que PPD/CCNC y la Biblioteca Cochrane Plus con respecto a revisiones sistemáticas que se han publicado en revistas
- Es posible realizar búsquedas más sofisticadas
- Se actualiza diariamente

Desventajas

- Tiene más de 19 millones de registros de los cuales sólo una pequeña proporción son revisiones sistemáticas sobre mecanismos de los sistemas de salud y por lo tanto resulta difícil hacer las búsquedas - puede requerir algún tipo de capacitación y práctica
- No incluye literatura gris ni todas las revistas y por lo tanto puede no tener algunas revisiones sistemáticas que están incluidas en las bases de datos de PPD/CCNC, DARE y HTA

Búsqueda en PubMed

- Se puede buscar usando palabras de texto, términos indexados y lógica booleana
- Además, se puede buscar "artículos relacionados" a los artículos que ya se tienen identificados mediante un sofisticado algoritmo
- Existe la posibilidad de utilizar un "delimitador" (hedge) o un filtro de revisiones sistemáticas (<http://www.ncbi.nlm.nih.gov/sites/pubmedutils/clinical>)

Estrategias adicionales de búsqueda

Resúmenes SUPPORT (www.support-collaboration.org) Los resúmenes SUPPORT son resúmenes concisos de revisiones sistemáticas sobre los efectos de las intervenciones de los sistemas de salud (así como las intervenciones en salud materno infantil) para países de bajos y medianos ingresos. Estos incluyen el resumen de las tablas de hallazgos preparadas utilizando el marco GRADE y las evaluaciones de la relevancia de los hallazgos para países de bajos y medianos ingresos.

Google scholar ([scholar.google.com / schhp? Hl = es & tab = es](http://scholar.google.com/schhp?hl=es&tab=es)) se puede utilizar para búsquedas de referencias y, a veces puede ser útil para identificar revisiones publicadas y revisiones en la literatura gris mediante el uso de características avanzadas y términos precisos (por ejemplo, poner términos como "revisión sistemática" entre comillas de modo que se busque la frase exacta).

Conctacto personal con personas que tienen experiencias en sistemas de salud y/o revisiones sistemáticas puede ser útil, particularmente para encontrar revisiones sistemáticas no publicadas, más recientes y difíciles de encontrar de otro modo.

Otras bases de datos tales como la ISI Web of Science (para búsqueda de referencias) y EMBASE puede requerir suscripción y asistencia de un bibliotecario investigador. Una lista de una colección de base de datos, sitios webs y revistas relevantes para países de bajos y medianos ingresos se puede encontrar en el sitio web de Norwegian Cochrane Effective Practice and Organisation of Care (EPOC) Satellite (<http://www.epocoslo.cochrane.org/en/newPage2.html>).

Recuperación de copias de texto completo de revisiones sistemáticas

[Necesitamos discutir y acordar sobre un sistema para que SURE / EVIPNet apoyen el acceso a copias de texto completo y potencialmente a otras publicaciones.]

Apéndice 2. Lista de verificación SURE para realizar evaluaciones sobre cuanta confianza depositar en una revisión sistemática¹³

Revisión:
Evaluada por:
Fecha:

Sección A: Métodos usados para identificar, incluir y evaluar críticamente los estudios

<p>A.1 ¿Se informaron los criterios usados para decidir qué estudios incluir en la revisión?</p> <p>Los autores especificaron:</p> <ul style="list-style-type: none"> <input type="radio"/> Tipos de estudios <input type="radio"/> Participantes <input type="radio"/> Intervención (es) <input type="radio"/> Resultado (s) <p><i>Guía de codificación – verifique las respuestas anteriores</i> <i>SI: Las cuatro deben ser sí</i></p>	<ul style="list-style-type: none"> <input type="radio"/> Sí <input type="radio"/> No sé/parcialmente <input type="radio"/> No
<p><i>Comentarios (anote limitaciones o incertidumbres importantes)</i></p>	
<p>A.2 ¿La búsqueda de la evidencia fue razonablemente completa?</p> <p>Se hizo lo siguiente:</p> <ul style="list-style-type: none"> <input type="radio"/> Se evitó el sesgo de idioma (ninguna restricción de inclusión basada en el idioma) <input type="radio"/> Ninguna restricción de inclusión basada en el estado de publicación <input type="radio"/> Búsqueda en bases de datos relevante (Medline + Biblioteca Cochrane) <input type="radio"/> Se verificaron las listas de referencias de los artículos incluidos <input type="radio"/> Se contactó a los autores/expertos <p><i>Guía de codificación – verifique las respuestas anteriores</i> <i>SI: Las cuatro deben ser sí</i> <i>PARCIALMENTE: Bases de datos relevantes y listas de referencias no deben estar marcadas</i></p>	<ul style="list-style-type: none"> <input type="radio"/> Sí <input type="radio"/> No sé/parcialmente <input type="radio"/> No
<p><i>Comentarios (anote limitaciones o incertidumbres importantes)</i></p>	
<p>A.3 ¿La revisión está razonablemente actualizada?</p> <p><i>¿Las búsquedas son suficientemente recientes como para que sea improbable encontrar investigación más reciente o que cambie los resultados de la revisión?</i></p>	<ul style="list-style-type: none"> <input type="radio"/> Sí <input type="radio"/> No sé/parcialmente <input type="radio"/> No
<p><i>Comentarios (note limitaciones o incertidumbres importantes)</i></p>	

<p>A.4 ¿Se evitó el sesgo en la selección de los artículos?</p> <p>¿Los autores especificaron:</p> <ul style="list-style-type: none"> <input type="radio"/> Criterios de selección explícitos <input type="radio"/> Tamizaje independiente del texto completo por al menos 2 revisores <input type="radio"/> Se proporcionó lista de estudios incluidos <input type="radio"/> Se proporcionó lista de estudios excluidos <p><i>Guía de verificación - verifique lo anterior</i> <i>SI: Los cuatro debe ser sí</i></p>	<ul style="list-style-type: none"> <input type="radio"/> Sí <input type="radio"/> No sé/parcialmente <input type="radio"/> No
<p><i>Comentarios (anote limitaciones o incertidumbres importantes)</i></p>	
<p>A.5 ¿Los autores usaron criterios apropiados para evaluar el riesgo de sesgo al analizar los estudios que están incluidos?*</p> <ul style="list-style-type: none"> <input type="radio"/> Se reportaron los criterios usados para evaluar el riesgo de sesgo <input type="radio"/> Se reportó una tabla o resumen de la evaluación de cada estudio incluido para cada criterio <input type="radio"/> Se usaron criterios sensibles que se enfocan en el riesgo de sesgo (y ninguna otra cualidad de los estudios como la precisión o aplicabilidad) <p><i>Guía de verificación - verifique lo anterior</i> <i>SI: Los cuatro deben ser sí</i></p>	<ul style="list-style-type: none"> <input type="radio"/> Sí <input type="radio"/> No sé/parcialmente <input type="radio"/> No
<p><i>Comentarios (anote limitaciones o incertidumbres importantes)</i></p>	
<p>A.6 En general – ¿cómo calificaría los métodos usados para identificar, incluir y evaluar críticamente los estudios?</p> <p><i>La calificación A de resumen de evaluación se relaciona con las 5 preguntas anteriores.</i></p> <p><i>Si la opción “No” o “Parcial” se usa en alguna de las preguntas anteriores, es probable que la revisión tenga limitaciones importantes.</i></p> <p><i>Ejemplos de errores fatales pueden incluir no reportar los criterios explícitos de selección o no evaluar el riesgo de sesgo en los estudios incluidos.</i></p>	<ul style="list-style-type: none"> <input type="radio"/> Errores fatales (limitaciones que son suficientemente importantes como para que los resultados no sean confiables y no se deban usar en el resumen ejecutivo para formulación de políticas) <input type="radio"/> Limitaciones importantes (limitaciones que son suficientemente importantes como para justificar buscar otra revisión sistemática e interpretar los resultados de esta revisión con precaución si no se puede encontrar una revisión mejor) <input type="radio"/> Confiables (solo limitaciones menores)
<p><i>Comentarios (anote cualquier error fatal o limitación importante).</i></p>	

Sección B: Métodos usados para analizar los hallazgos

<p>B.1 ¿Las características y resultados de los estudios incluidos se reportados confiablemente?</p> <p>Hubo:</p> <ul style="list-style-type: none"> <input type="radio"/> Extracción independiente de datos por al menos 2 revisores <input type="radio"/> Una tabla o resumen de las características de los participantes, intervenciones y resultados de los estudios incluidos <input type="radio"/> Una tabla o resumen de los resultados de los estudios incluidos. <p><i>Guía de codificación –verifique las respuestas anteriores</i> <i>SI: Los tres deben ser sí</i></p>	<ul style="list-style-type: none"> <input type="radio"/> Sí <input type="radio"/> No sé/parcialmente <input type="radio"/> No <input type="radio"/> No aplicable (por ej. no hay estudios incluidos)
<p><i>Comentarios (anote limitaciones o incertidumbres importantes)</i></p>	
<p>B.2 ¿Se reportaron los métodos usados por los autores de la revisión para analizar los hallazgos de estudios incluidos?</p>	<ul style="list-style-type: none"> <input type="radio"/> Sí <input type="radio"/> Parcialmente <input type="radio"/> No <input type="radio"/> No aplicable (por ej. no hay estudios o no hay datos)
<p><i>Comentarios (anote limitaciones o incertidumbres importantes)</i></p>	
<p>B.3 ¿La revisión describió la magnitud de la heterogeneidad?</p> <ul style="list-style-type: none"> <input type="radio"/> ¿La revisión se aseguró de que los estudios incluidos fueran suficientemente similares de modo que tuviese sentido combinarlos, dividir sensiblemente los estudios incluidos en grupos homogéneos o concluir sensiblemente que no tenía sentido combinar o agrupar los estudios incluidos? <input type="radio"/> ¿La revisión discutió hasta que punto había diferencias importantes en los resultados de los estudios incluidos? <input type="radio"/> Si se hizo un meta-análisis, ¿se reportó I^2, prueba de heterogeneidad de chi cuadrado u otra estadística apropiada? 	<ul style="list-style-type: none"> <input type="radio"/> Sí <input type="radio"/> No sé/parcialmente <input type="radio"/> No <input type="radio"/> No aplicable (por ej. no hay estudios o no hay datos)
<p><i>Comentarios (anote limitaciones o incertidumbres importantes)</i></p>	

<p>B.4 ¿Los hallazgos de los estudios relevantes se combinaron (o no se combinaron) apropiadamente en relación a la pregunta primaria que la revisión enfoca y los datos disponibles?</p> <p>¿Cómo se hizo el análisis de datos?</p> <ul style="list-style-type: none"> <input type="radio"/> Solo descriptivo <input type="radio"/> Recuento de votos basado en la dirección del efecto <input type="radio"/> Recuento de votos basado en la significancia estadística <input type="radio"/> Descripción del rango de los tamaños del efecto <input type="radio"/> Meta-análisis <input type="radio"/> Meta-regresión <input type="radio"/> Otros: especificar <input type="radio"/> No aplicable (por ej. no hay estudios o no hay datos) <p>¿Como se ponderaron los estudios en el análisis?</p> <ul style="list-style-type: none"> <input type="radio"/> Pesos iguales (esto se hace cuando se usa el recuento de votos) <input type="radio"/> Por calidad o diseño del estudio (esto se hace raramente) <input type="radio"/> Inversa de la varianza (esto es lo que hace generalmente en un meta-análisis) <input type="radio"/> Número de participantes <input type="radio"/> Otros, especificar: <input type="radio"/> No claro <input type="radio"/> No aplicable (por ej. no hay estudios o no hay datos) <p>¿La revisión abordó los errores de unidad de análisis?</p> <ul style="list-style-type: none"> <input type="radio"/> Si – consideró el agrupamiento (clustering) en el análisis (por ej. usó el coeficiente de correlación intra-grupo) <input type="radio"/> No, pero mencionó el problema de los errores de la unidad de análisis <input type="radio"/> No se mencionó el tema <input type="radio"/> No aplicable – no hay ensayos agrupados o estudios incluidos <p><i>Guía de codificación – verifique las respuestas anteriores</i></p> <p><i>Si es narrativa O recuento de votos (cuando los análisis cuantitativos hubiesen sido posibles) O tabla, gráfico o meta-análisis inapropiado O no se abordaron los errores de unidad de análisis (y tendrían que haber sido), la respuesta es probablemente NO.</i></p> <p><i>Si la tabla, gráfico o meta-análisis eran apropiados Y los pesos eran apropiados Y se consideró la magnitud de la heterogeneidad, la respuesta es probablemente SI.</i></p> <p><i>Si no hay estudios/no hay datos: NO APLICABLE</i></p> <p><i>Si no hay seguridad: NO SE/PARCIALMENTE</i></p>	<ul style="list-style-type: none"> <input type="radio"/> Sí <input type="radio"/> No sé/parcialmente <input type="radio"/> No <input type="radio"/> No aplicable (por ej. no hay estudio o no hay datos)
<p><i>Comentarios (anote limitaciones o incertidumbre importantes)</i></p>	

<p>B.5 ¿La revisión examinó hasta qué punto factores específicos podían explicar las diferencias en los resultados de los estudios incluidos?</p> <p><input type="radio"/> ¿Los factores que los autores de la revisión consideraron como probables factores explicativos fueron descritos claramente?</p> <p><input type="radio"/> ¿Se usó un método sensible para explorar hasta qué punto los factores claves explicaban la heterogeneidad?</p> <ul style="list-style-type: none"> <input type="radio"/> Descriptivo/textual <input type="radio"/> Grafico <input type="radio"/> Meta-regresión <input type="radio"/> Otro 	<ul style="list-style-type: none"> <input type="radio"/> Sí <input type="radio"/> No sé/parcialmente <input type="radio"/> No <input type="radio"/> No aplicable (por ej. Demasiado pocos estudios, no hay diferencias importantes en los resultados de los estudios incluidos, o los estudios incluidos fueron tan disímiles que no tendría sentido explorar la heterogeneidad de los resultados)
<p><i>Comentarios (anote limitaciones o incertidumbres importantes)</i></p>	
<p>B.6 En general - ¿cómo clasificaría los métodos usados para analizar los hallazgos relacionados a la pregunta primaria abordada en la revisión?</p> <p><i>La calificación B de la evaluación resumida se relaciona con las 5 preguntas con relación al análisis.</i></p> <p><i>Si se usa la opción “No” o “Parcial” para cualquiera de la 5 preguntas precedentes, es probable que la revisión tenga limitaciones importantes.</i></p> <p><i>Ejemplos de errores fatales pueden incluir no reportar características críticas de los estudios incluidos o no reportar los resultados de los estudios incluidos.</i></p>	<ul style="list-style-type: none"> <input type="radio"/> Errores fatales (limitaciones que son suficientemente importantes como para que los resultados de la revisión no sean confiables y no se deban usar en el resumen ejecutivo para formulación de políticas) <input type="radio"/> Limitaciones importantes (limitaciones que son suficientemente importantes que sería mejor buscar otra revisión sistemática o interpretar los resultados de la revisión con precaución si no se puede encontrar una mejor revisión) <input type="radio"/> Confiable (solo limitaciones menores)
<p><i>Use comentarios para especificar si es relevante, indicar incertidumbre o se necesita discutir</i></p>	

Sección C: Evaluación general de la calidad de la revisión

C.1 ¿Hay algún otro aspecto de la revisión no mencionado antes que puede llevarlo a cuestionar los resultados?	<input type="radio"/> Preocupaciones metodológicas adicionales <input type="radio"/> Solidez <input type="radio"/> Interpretación <input type="radio"/> Conflictos de interés (de los autores de la revisión o para los estudios incluidos) <input type="radio"/> Otros <input type="radio"/> No se identificaron otros temas de calidad
C.2 Basado en las evaluaciones previas ¿como usted clasificaría la calidad de la revisión? <input type="radio"/> Errores fatales (excluir); brevemente (y amablemente) enuncie las razones para excluir la revisión completando la siguiente oración: <i>Esta revisión no fue incluida en este resumen ejecutivo para formulación de políticas por las siguientes razones:</i> <i>Comentarios (brevemente resume cualquier mensaje clave o información útil que se pueda extraer de la revisión para los formuladores de políticas o gerentes):</i> <input type="radio"/> Limitaciones importantes ; brevemente (y amablemente) enuncie las limitaciones más importantes editando la siguiente oración, si es necesario, y especifique cuales son las limitaciones importantes: <i>Esta revisión tiene limitaciones importantes:</i> <input type="radio"/> Confiable ; brevemente anote cualquier comentario que se debería destacar con relación a la confiabilidad de esta revisión editando la siguiente oración, si es necesario: <i>Esta es una revisión sistemática de buena calidad solo con limitaciones menores.</i>	

NOTAS

* **Riesgo de sesgo** es la magnitud hasta la cual el sesgo puede ser responsable de los hallazgos de un estudio.

Sesgo es un error sistemático o desviación de la verdad en los resultados o inferencias. En los estudios de los efectos de la atención en salud, los principales tipos de sesgo surgen de diferencias sistemáticas en los grupos que son comparados (sesgo de selección), la atención que se provee, o la exposición a otros factores aparte de la intervención de interés (sesgo de ejecución), abandonos o exclusiones de personas ingresadas al estudio (sesgo de atrición) o como se evaluaron los resultados (sesgo de detección). Las revisiones de los estudios también pueden ser particularmente afectadas por el sesgo de reporte, en el que un subgrupo sesgado de todos los datos relevantes está disponible.

Algunas veces las evaluaciones del riesgo de sesgo son también llamadas evaluaciones de la **validez** o **calidad** de un estudio.

Validez es la magnitud hasta la cual un resultado (de una medición o estudio) es probablemente cierto.

Calidad es una vaga noción de la fuerza o validez de un estudio, indicando frecuentemente la magnitud del control sobre el sesgo.

Apéndice 3. Hojas de trabajo SURE para preparar un resumen de hallazgos usando GRADE

Estas hojas de trabajo son para 1) identificar los resultados más importantes, 2) evaluar la calidad de la evidencia para cada uno de los resultados y preparar un resumen de hallazgos para un resumen ejecutivo para formulación de políticas usando el enfoque GRADE – basado en los resultados de una revisión sistemática.

Instrucciones

1. Identifique la principal comparación de relevancia para el resumen ejecutivo para formulación de políticas. Prepare más de un resumen de hallazgos si una revisión contiene más de una comparación que sea relevante al resumen ejecutivo.

2. Seleccione los resultados más importantes para esta comparación

Sugerencias

- a) Genere una lista de resultados relevantes (ver **hoja de trabajo 1**)
 - Liste los resultados que los autores de la revisión sistemática identificaron como resultados primarios
 - Agregue otros resultados para los cuales se reportan datos
 - Agregue cualquier otro resultado que no fue reportando en la revisión pero que puede ser importante para alguien que toma una decisión – desde la perspectiva de aquellos que serán afectados por la decisión. Asegúrese de considerar los beneficios potenciales, efectos adversos y uso de recursos (costos).
 - Determine que resultados son suficientemente importantes como para ser incluidos en la tabla SoF (hoja de trabajo 1)
- b) De esta lista elija los resultados que piensa son más importantes y se deben incluir en la tabla SoF; transfíralos a una tabla SoF vacía (ver **hoja de trabajo 2**).
 - Incluya los resultados que son críticos para una decisión aún si la revisión no provee ninguna evidencia.

3. Evalúe la calidad de la evidencia para cada resultados usando el enfoque GRADE

Sugerencias

- Complete la **hoja de trabajo 2** para determinar la calidad de la evidencia para el resultado
- Consulte los **criterios para evaluar la calidad de la evidencia**
- Complete la columna Calidad de la Evidencia en la **hoja de trabajo 3**.

4. Resuma los hallazgos para la resultante (cuantitativamente si es posible), en una forma que sea entendible para los formuladores de políticas y otras partes interesadas.

5. Complete la tabla SoF (**hoja de trabajo 3**).

6. Prepare viñetas para resumir la información en el resumen de la tabla de hallazgos en lenguaje sencillo. Sea consistente en como traduce los hallazgos en afirmaciones cualitativas (**hoja de trabajo 4**).

Hoja de trabajo 1: Evaluación de la importancia relativa de los resultados y decisión sobre cuales incluir en la tabla Resumen de Hallazgos

Revisión:
Evaluada por:
Fecha:

Clasifique la importancia relativa de cada resultado en una escala de 9 puntos yendo de 1 (no importante) a 9 (crítico).

1 – 3 no importante y no incluido en la tabla SoF

4 – 6 importante pero no crítico para tomar una decisión (la inclusión en la tabla SoF puede depender de cuantos resultados importantes hay)

7 – 9 crítico para tomar una decisión y definitivamente se debe incluir en la tabla SoF t

Resultado	Iniciales de las personas que evalúan la importancia relativa de los resultados				Consenso
	Importancia relativa (1-9)				
1.					
2.					
3.					
4.					
5.					
6.					
7.					
8.					
9.					
10.					
11.					
12.					
13.					
14.					
15.					
16.					
17.					

Hoja de trabajo 2: Evaluación de la calidad de la evidencia entre estudios para un resultado (Ver criterios y definiciones)

Comparación _____

Evaluación de la calidad

Nro. de estudios	Diseño	Riesgo de sesgo	Inconsistencia	Indirectos	Imprecisión	Otros*	Calidad
Resultado:							
Resultado:							
Resultado:							
EJEMPLO: El uso de trabajadores de la salud empíricos comparado con los servicios usuales de atención de la salud							
Resultado: Captación de inmunidad en niños							
4	Ensayos randomizados (4)	Riesgo de sesgo serio (-0.5)	Inconsistencia importante (-0.5)	Sin resultados indirectos graves	Sin imprecisión grave	Ninguno	Moderado (3)

*Otras consideraciones para bajar de categoría incluyen el sesgo de publicación. Otras consideraciones para subir de categoría incluyen una fuerte asociación con confundidores no plausibles, una relación dosis respuesta, y si todos los confundidores plausibles o sesgos disminuirían el tamaño del efecto si hay evidencia de un efecto o lo aumentarían si hay evidencia de ausencia de efecto dañino (seguridad).

NOTAS – explicaciones de las evaluaciones que ha hecho, por ej. Se bajó la categoría de la evidencia de alta a moderada debido al riesgo de sesgo que estaba al borde de ser grave (debido al seguimiento incompleto y a la ausencia de cegamiento en algunos de los ensayos) y a la inconsistencia de los resultados entre estudios que estaba al borde de ser importantes (yendo de no concluyente a un aumento relativo de 36%)

1.

Nro. de estudios	Diseño	Riesgo de sesgo	Inconsistencia	Indirectos	Imprecisión	Otros*	Calidad
Resultado:							
Resultado:							
Resultado:							
EJEMPLO: El uso de trabajadores de la salud empíricos comparado con los servicios usuales de atención de la salud							
Resultado: Captación de inmunidad en niños							
4	Ensayos randomizados (4)	Riesgo de sesgo serio (-0.5)	Inconsistencia importante (-0.5)	Sin resultados indirectos graves	Sin imprecisión grave	Ninguno	Moderado (3)

*Otras consideraciones para bajar de categoría incluyen el sesgo de publicación. Otras consideraciones para bajar de categoría incluye una fuerte asociación con confundidores no plausibles, una relación dosis respuesta, y todos los confundidores plausibles o sesgos disminuirían el tamaño del efecto si hay evidencia de un efecto o lo aumentarían si hay evidencia de ausencia de efecto dañino (seguridad).

NOTAS – explicaciones de las evaluaciones que ha hecho, por ej. Se bajó la categoría de la evidencia de alta a moderada debido al riesgo de sesgo que estaba al borde de ser grave (debido al seguimiento incompleto y a la ausencia de cegamiento en algunos de los ensayos) y la inconsistencia de los resultados entre estudios que estaban al borde de ser importantes (yendo de no concluyente a un aumento relativo de 36%)

Nro. de estudios	Diseño	Riesgo de sesgo	Inconsistencia	Indirectos	Imprecisión	Otros*	Calidad
Resultado:							
Resultado:							
Resultado:							
EJEMPLO: El uso de trabajadores de la salud empíricos comparado con los servicios usuales de atención de la salud							
Resultado: Captación de inmunidad en niños							
4	Ensayos randomizados (4)	Riesgo de sesgo serio (-0.5)	Inconsistencia importante (-0.5)	Sin resultados indirectos graves	Sin imprecisión grave	Ninguno	Moderado (3)

*Otras consideraciones para bajar de categoría incluyen el sesgo de publicación. Otras consideraciones para bajar de categoría incluye una fuerte asociación con confundidores no plausibles, una relación dosis respuesta, y todos los confundidores plausibles o sesgos disminuirían el tamaño del efecto si hay evidencia de un efecto o lo aumentarían si hay evidencia de ausencia de efecto dañino (seguridad).

NOTAS – explicaciones de las evaluaciones que ha hecho, por ej. Se bajó la categoría de la evidencia de alta a moderada debido al riesgo de sesgo que estaba al borde de ser grave (debido a al seguimiento incompleto y a la ausencia de cegamiento en algunos de los ensayos) y la inconsistencia de los resultados entre estudios que estaban al borde de ser importantes (yendo de no concluyente a un aumento relativo de 36%)

Criterios de evaluación de calidad

Calidad de la evidencia	Diseño del estudio	Disminuir si *	Aumentar si *
Alta (4)	Ensayo aleatorizado	Limitaciones del estudio -1 Grave -2 Muy grave	Asociación fuerte +1 Fuerte, sin confundidores plausibles
Moderada (3)			
Baja (2)	Estudio observacional	Inconsistencia -1 Grave -2 Muy grave Resultados Indirectos -1 Grave -2 Muy grave Imprecisión -1 Grave -2 Muy grave Sesgo de publicación -1 Probable -2 Muy probable	+2 Muy fuerte, sin mayores amenazas a la validez Dosis respuesta +1 Evidencia de un gradiente Todos los confundidores plausibles +1 Todos los confundidores plausibles o sesgo reducirían el tamaño del efecto si hay evidencia de un efecto, o lo aumentarían si hay evidencia de ausencia de efecto dañino (seguridad)
Muy baja (1)			

* 1 = mueva hacia arriba o abajo un grado (por ejemplo de alta a intermedia)

2 = mueva hacia arriba o abajo dos grados (por ejemplo de alta a baja)

0.5= dudoso (borderline)

Calidad de evidencia entre estudios para el resultado

- ⊕⊕⊕⊕ Alta = Estamos confiados en que el efecto verdadero está cerca de la estimación del efecto.
- ⊕⊕⊕• Moderada = Es probable que el efecto verdadero esté cerca de la estimación del efecto pero hay una posibilidad de que sea sustancialmente diferente.
- ⊕⊕•• Baja = El efecto verdadero puede ser sustancialmente diferente de la estimación del efecto.
- ⊕••• Muy baja = Cualquier estimación del efecto es muy incierta.

Hoja de trabajo 3: Resumen de la Tabla de hallazgos

Pacientes o población:

Contextos:

Intervención:

Comparación:

Resultados	Impacto	Número de participantes (estudios)	Calidad de la evidencia (GRADE)

GRADE: Grados de evidencia del Grupo de Trabajo GRADE

⊕⊕⊕⊕ Alta = Estamos confiados en que el efecto verdadero está cerca de la estimación del efecto.

⊕⊕⊕⊕• Moderada = Es probable que el efecto verdadero esté cerca de la estimación del efecto pero hay una posibilidad de que sea sustancialmente diferente.

⊕⊕⊕••• Baja = El efecto verdadero puede ser sustancialmente diferente de la estimación del efecto.

⊕⊕•••• Muy baja = Cualquier estimación del efecto es muy incierta.

Ver ejemplo.

Resumen de hallazgos – ejemplos

Ejemplo 1: Uso de trabajadores de la salud empíricos (comunidad) como una ayuda en la atención habitual

Pacientes o población: Madres o niños menores de cinco años

Contextos: Mixto (países de ingresos altos para inmunización, mixto para amamantamiento, países de bajos ingresos para morbilidad y mortalidad en niños)

Intervención: Trabajadores de la salud empíricos (TSEs) (miembros de la comunidad que no son profesionales de la salud y han recibido alguna capacitación para promover la salud o proveer algunos servicios de atención en salud)

Comparación: Cuidado usual (variado entre estudios)

Resultados	Impacto		Número de estudios	Calidad de la evidencia (GRADE)*
	Relativo	Absoluto (Ilustrativo)		
Mortalidad en niños menores de cinco	25% reducción relativa	De 5 a 4 por 100	3	⊕⊕○○ Baja
Mortalidad neonatal	24% reducción relativa	De 4 a 3 por 100	4	⊕⊕○○ Baja
Morbilidad en niños menores de cinco (por ej. fiebre, diarrea)	14% reducción relativa	De 50 a 43 por 100	7	⊕⊕○○ Baja
Búsqueda de atención para niños menores de cinco	33% aumento relativo	De 20 a 27 por 100	3	⊕⊕○○ Baja
Inmunizaciones infantiles completas	22% aumento relativo	De 50 a 61 por 100	4	⊕⊕⊕⊕ Moderada
Iniciación de amamantamiento	36% aumento relativo	De 50 a 68 por 100	12	⊕⊕⊕⊕ Moderada
Amamantamiento exclusivo	178% aumento relativo	De 20 a 36 por 100	10	⊕⊕⊕⊕ Moderada

* Grados de evidencia del Grupo de Trabajo GRADE

⊕⊕⊕⊕ Alta = Estamos confiados en que el efecto verdadero está cerca de la estimación del efecto.

⊕⊕⊕⊕• Moderada = Es probable que el efecto verdadero esté cerca de la estimación del efecto pero hay una posibilidad de que sea sustancialmente diferente.

⊕⊕••• Baja = El efecto verdadero puede ser sustancialmente diferente de la estimación del efecto.

⊕•••• Muy baja = Cualquier estimación del efecto es muy incierta

Ejemplo 2. Substitución de enfermeras por médicos en atención primaria

Pacientes o población: Todos los pacientes presentes en atención primaria

Contextos: Primariamente Canada, EEUU y el RU

Intervención: Substitución de enfermeras por médicos (enfermera dirigiendo atención primaria)

Comparación: Atención rutinaria proveida por médicos (médico dirigiendo atención primaria)

Resultados	Impacto	Número de Estudios	Calidad de la evidencia (GRADE)*
Resultados de pacientes	Enfermeras y médicos puede conducir a resultados de salud similares para los pacientes.	4	⊕⊕•• Baja
Calidad de la atención	La extensión hasta la cual la atención proveida por las enfermeras fue más o menos apropiada que la atención proveida por médicos no fue reportada.	0	
Satisfacción del paciente	En promedio los pacientes probablemente están más satisfechos con la atención proveida por las enfermeras pero algunos prefieren la atención proveida por las enfermeras y algunos prefieren la atención proveida por los médicos.	3	⊕⊕⊕• Moderada
Costos directos	Los costos de salarios más bajos de las enfermeras pueden estar compensados por un uso aumentado de recursos o menor productividad de modo que puede haber poca si existe alguna diferencia en el costo de la atención proveida por enfermeras comparado con el costo de la atención proveida por los médicos. Debido a que la diferencia en el salario entre enfermeras y médicos puede variar de un lugar a otro a través del tiempo, el ahorro neto, si existe, es probable que dependa del contexto.	2	⊕••• Muy baja

* Grados de evidencia del Grupo de Trabajo GRADE

⊕⊕⊕⊕ Alta = Estamos confiados en el efecto verdadero está cerca de la estimación del efecto.

⊕⊕⊕• Moderada = Es probable que el efecto verdadero esté cerca de la estimación del efecto pero hay una posibilidad de que sea sustancialmente diferente.

⊕⊕•• Baja = El efecto verdadero puede ser sustancialmente diferente de la estimación del efecto.

⊕••• Muy baja = Cualquier estimación del efecto es muy incierta

Limitaciones: La revisión en la cual está basada este resumen de hallazgos tenía algunas importantes limitaciones. La búsqueda solo va hasta el 2002 y solo se incluyeron los artículos escritos en inglés u holandés. Algunos análisis fueron potencialmente erróneos.

Ejemplo 3.

RESUMEN DE HALLAZGOS PARA LA COMPARACION PRINCIPAL (Explicación)

Reuniones educativas con o sin otras intervenciones comparadas con ninguna intervención				
Paciente o población: Profesionales de atención en salud				
Contexto: atención primaria y secundaria				
Intervención: Reuniones educativas con o sin otras intervenciones*				
Comparación: Ninguna intervención				
Resultados	Mejoría absoluta ajustada (diferencia de riesgo) ¹ Mediana (Rango de intercuartil	Número de estudios	Calidad de la evidencia (GRADE)	Comentarios
Conformidad con la práctica deseada	Mediana 6% (1,8 a 15,9)	30	+++0 Moderada ^H	El efecto parece ser mayor cuando hay una asistencia mayor a las reuniones educativas y con reuniones educativas mixtas interactivas y didácticas. Las reuniones educativas no parecen ser efectivas para resultados menos serios
Resultados de pacientes	Mediana 3,0% (0,1% a 4,0%)	5	+++0 Moderada ^I	
* El efecto de las reuniones educativas solamente sobre la práctica profesional fue la misma que en las intervenciones multifacéticas que incluyeron reuniones educativas.				
^H Las diferencias de riesgo post-intervención está ajustadas a las diferencias pre-intervención entre los grupos de comparación				
^I Hemos reducido la evidencia de alta a moderada debido a la inconsistencia en los resultados que no se pudo explicar totalmente				

Grados de evidencia del Grupo de Trabajo GRADE

Calidad alta (++++): Es muy improbable que la investigación futura cambie nuestra confianza en la estimación del efecto

Calidad moderada (+++0): Es probable que la investigación futura tenga un impacto importante sobre nuestra confianza en la estimación del efecto y puede cambiarla

Calidad baja (++00): Es muy probable que la futura investigación tenga un impacto importante sobre nuestra confianza en la estimación del efecto y es probable que la cambie.

Calidad muy baja (0000): Estamos muy inseguros de la estimación

Hoja de trabajo 4: Mensajes claves en lenguaje sencillo

Prepare un número pequeño de viñetas (por ej. 1 a 3) que resuman los contenidos de la Tabla Resumen de hallazgos como mensajes claves. Use lenguaje consistente, como el siguiente (adaptado de sugerencias para resúmenes en lenguaje sencillo de Cochrane).

	Diferencia importante	Diferencia pequeña (puede no ser importante)	Poca o ninguna diferencia
Evidencia de Alta Calidad	Mejora/ disminuye/previene/conduce a [resultado]	Mejora levemente/ disminuye levemente/conduce levemente a menos (más) [resultado]	Resulta en poca o ninguna diferencia en [resultado]
Evidencia de calidad moderada	Probablemente mejora/ disminuye/previene/conduce a [resultado]	Probablemente mejora levemente/ disminuye levemente/ conduce levemente a menos (más) [resultado]	Probablemente conduce a poca o ninguna diferencia en [resultado]
Evidencia de baja calidad	Puede mejorar/disminuir/prevenir/conducir a [resultado]	Puede aumentar levemente/disminuir levemente/conducir levemente a menos (más) [resultado]	Puede conducir a poca o ninguna diferencia en [resultado]
Evidencia de muy baja calidad	Es incierto si la [intervención] mejora, disminuye, previene o conduce a [resultado] porque la calidad de la evidencia es muy baja		
No hay datos o no hay estudios	El [resultado] no se midió o no se reportó, o no se encontraron estudios que evaluaran el impacto de la [intervención] sobre el [resultado]		

Resumen en lenguaje sencillo de los hallazgos más importantes

∅

∅

∅

Ejemplos

Uso de trabajadores de la salud empíricos como un agregado a la atención usual (Ejemplo 1):

- ∅ Probablemente aumenta la cobertura de inmunización y el amamantamiento
- ∅ Puede aumentar el comportamiento de búsqueda de atención para niños menores de cinco y reducir la morbilidad y mortalidad en niños menores de cinco y neonatos

Sustitución de enfermeras por médicos en atención primaria (Ejemplo 2):

- ∅ Enfermeras y médicos pueden conducir a resultados de salud similares en los pacientes.
- ∅ Es incierto si hay alguna diferencia en el costo de la atención proveída por enfermeras comparada con el costo de atención proveída por médicos.

Reuniones educativas para profesionales de la salud (Ejemplo 3):

- ∅ Probablemente mejora la conformidad con la práctica deseada y los resultados de los pacientes

Apéndice 4. Evaluación de la guía SURE para decisión y descripción de opciones

El equipo que prepare el resumen ejecutivo para formulación de políticas debe completar la siguiente evaluación y enviarla a Andy Oxman oxman@online.no.

1. ¿Cuán útil fue la guía para decidir sobre las opciones que se deben presentar en el resumen ejecutivo para formulación de políticas?

Muy inútil	Moderadamente inútil	Levemente inútil	Neutral	Levemente útil	Moderadamente útil	Muy útil
1	2	3	4	5	6	7

¿Cómo se podría mejorar la guía con respecto a este punto? _____

2. ¿Cuán útil fue la guía para encontrar y evaluar evidencia de los impactos de las opciones?

Muy inútil	Moderadamente inútil	Levemente inútil	Neutral	Levemente útil	Moderadamente útil	Muy útil
1	2	3	4	5	6	7

¿Cómo se podría mejorar la guía con respecto a este punto? _____

3. ¿Cuán útil fue la guía para realizar evaluaciones de cuán confiados se puede estar en los probables impactos de las opciones?

Muy inútil	Moderadamente inútil	Levemente inútil	Neutral	Levemente útil	Moderadamente útil	Muy útil
1	2	3	4	5	6	7

¿Cómo se podría mejorar la guía con respecto a este punto? _____

4. ¿Cuán útil fue la guía para decidir como resumir y presentar información sobre los posibles impactos de las opciones?

Muy inútil	Moderadamente inútil	Levemente inútil	Neutral	Levemente útil	Moderadamente útil	Muy útil
1	2	3	4	5	6	7

¿Cómo se podría mejorar la guía con respecto a este punto? _____

5. ¿Las preguntas a considerar fueron útiles?

Muy inútil	Moderadamente inútil	Levemente inútil	Neutral	Levemente útil	Moderadamente útil	Muy útil
1	2	3	4	5	6	7

¿Cómo se podría mejorar las preguntas a considerar? _____

6. ¿Los antecedentes fueron útiles?

Muy inútil	Moderadamente inútil	Levemente inútil	Neutral	Levemente útil	Moderadamente útil	Muy útil
1	2	3	4	5	6	7

¿Cómo se podría mejorar los antecedentes? _____

7. ¿Qué recursos se tendrían que incluir en esta guía? _____

8. ¿Cuán útiles fueron los apéndices?

Muy inútil	Moderadamente inútil	Levemente inútil	Neutral	Levemente útil	Moderadamente útil	Muy útil
1	2	3	4	5	6	7

¿Cómo se podría mejorar los apéndices? _____

9. ¿Cuán útil fue la estructura de la guía?

Muy inútil	Moderadamente inútil	Levemente inútil	Neutral	Levemente útil	Moderadamente útil	Muy útil
1	2	3	4	5	6	7

¿Cómo se podría mejorar la estructura de la guía? _____

10. En general, ¿cuán útil fue la guía para decidir y describir opciones?

Muy inútil	Moderadamente inútil	Levemente inútil	Neutral	Levemente útil	Moderadamente útil	Muy útil
1	2	3	4	5	6	7

¿De qué otras maneras se podría mejorar la guía? _____
