ID: 1657-2513.art.20208

ĬbērAM
Revisión | Review

Fundamentos y Métodos de las Revisiones Sistemáticas

Fundamentals and method of systematic reviews

CC BY NC SA

Silvina **Berra**

ART Volumen 20 #2 julio - diciembre

Revista ARETÉ Fonoaudiología

ISSN-l: 1657-2513 | e-ISSN: 2463-2252

Revista ARETÉ
ISSN-L: 1657-2513 | e-ISSN: 2463-2252 Fonoaudiología

ID: 1657-2513.art.20208

Title: Fundamentals and method of systematic reviews

Título: Fundamentos y métodos de las revisiones sistemáticas

Alt Title / Título alternativo:

[en]: Fundamentos y método de las revisiones sistemáticas

[es]: Fundamentos y método de las revisiones sistemáticas

Author (s) / Autor (es):

Berra

Keywords / Palabras Clave:

[en]: Systematic Reviews, Evidence-Based Practice, Health Research

[es]: revisiones sistemáticas, práctica basada en la evidencia, investigación en salud

Submited: 2020-11-30

Acepted: 2020-12-16

Resumen

Las revisiones sistemáticas (RS) son síntesis cualitativas y, a veces, cuantitativas sobre la evidencia disponible respecto de un tema puntual. Este trabajo contextualiza su origen en el movimiento de la medicina basada en la evidencia y las define como un tipo de investigación orientada a identificar, valorar críticamente y sintetizar conocimiento científico previo. El método de RS se presenta aquí en una serie de fases o pasos: planteo del objetivo, búsqueda y selección de estudios, evaluación de calidad, extracción y análisis de datos, metaanálisis e interpretación de resultados. Los criterios de evaluación de la evidencia científica, un pilar fundamental de este tipo de investigación, son la validez interna, la precisión y la validez externa de los resultados. La calidad de la evidencia científica, junto con la cantidad y la consistencia, son la base sobre las cuales se pueden establecer recomendaciones. El método de RS es útil en la práctica profesional, donde se requiere el manejo de evidencia científica actualizada sobre los asuntos en los que a diario se deciden diagnósticos o intervenciones, así como un tipo de investigación que puede implementarse en grado o posgrado, ya que podría llegar a resultados originales mediante la síntesis de sus antecedentes.

Abstract

Systematic reviews (SR) are qualitative and, sometimes, quantitative syntheses on the available evidence regarding a specific topic. This work contextualizes its origin in the evidence-based medicine movement and defines it as a type of health research aimed to identify, critically appraise and synthesize previous scientific knowledge. The SR method is presented here in a series of steps: objective, search and selection of studies, quality assessment, data extraction and analysis, meta-analysis and interpretation of results. The internal validity, precision and external validity of the results are explained as criteria to assess the quality of the evidence. The quality of the evidence, together with the quantity and consistency, are the basis on which recommendations can be made. The SR method is useful in professional practice, where the management of up-to-date scientific evidence is required to make decisions on diagnoses or interventions daily, as well as a type of research that can be implemented in undergraduate or graduate studies, since it could arrive at original results by synthesizing the previous knowledge.

Citar como:

Berra, S. (2020). Fundamentos y métodos de las revisiones sistemáticas. Areté , 20 (2), 73-82. Obtenido de: https://arete.ibero.edu.co/article/view/art.20208

Silvina **Berra,** PhD MsC BHS

ORCID: *0000-0003-3618-7041*

Source | Filiacion:
Universidad Nacional de Córdoba (UNC)

BIO:
Licenciada en Nutrición (UNC, Argentina), Master en Salud Pública (UPF, España), Doctora en Ciencias de la Salud y la Vida (UPF, España)

City | Ciudad:
Argentina [Arg]

Para su más amplia difusión, esta obra y sus contenidos se distribuyen bajo licencia: **Creative Commons Atribución-NoComercial-CompartirIgual 4.0 International**
Se autoriza cualquier tipo de reproducción y sus diferentes usos, siempre y cuando se preserve la integridad del contenido y se cite la fuente.

Fundamentos y Métodos de las Revisiones Sistemáticas

Fundamentals and method of systematic reviews

Silvina **Berra**

La consideración de las evidencias científicas en la práctica profesional y docente

Cada profesional de la salud se enfrenta diariamente a decisiones frente a las personas que atiende o a las organizaciones en las que lleva a cabo una gestión. En la práctica clínica, luego de establecer el estado de la persona que requiere atención, se decide cuál es la prescripción más adecuada y conveniente para lograr cambios en su salud. Las reflexiones se basan en el conocimiento que cada profesional maneja de su disciplina y, ante más de una alternativa, elige una que supone mejor.

En el campo de la Fonoaudiología, un ejemplo es la elección de la terapia más adecuada y efectiva para una persona con afasia posterior a un accidente cerebrovascular, entre modalidades que pueden variar por su enfoque teórico, la metodología de intervención, la duración, la intensidad y la frecuencia **(Brady, Kelly, Godwin, Enderby y Campbell, 2016)**. Ese supuesto de mayor adecuación o efectividad de una alternativa frente a otra puede haber sido sometido a pruebas científicas mediante estudios experimentales, sus resultados pueden estar publicados y, en tal caso, es posible encontrarlos mediante una búsqueda bibliográfica. De manera similar, en el ámbito de la gestión y política sanitaria, se toman decisiones para establecer programas y políticas sobre los cuales también pueden existir evidencias científicas de sus resultados y sus limitaciones. Por ejemplo, quien debe tomar una decisión acerca del financiamiento o cobertura de las implantaciones bilaterales de implantes activos de oído medio, de tronco cerebral y de conducción ósea se preguntará si son más efectivos, eficientes y seguros en comparación con su inserción unilateral en pacientes con hipoacusia bilateral **(Grau Magaña y Estrada Sabadell, 2020)**.

Areté
ISSN-L: 1657-2513

Este tipo de circunstancias motivan el interés por la evaluación de la evidencia científica y llevan a revisar las novedades bibliográficas con cierta frecuencia. Sin embargo, la dinámica de la actividad científica es tal, que la cantidad de información emergente de las investigaciones en algún tema en particular puede ser muy alta y su calidad puede ser heterogénea, por lo que encontrar y evaluar la información puede ser un desafío abrumador para cualquier profesional. Esta es una de las razones por las que emergieron las propuestas metodológicas de evaluación de la evidencia científica, intentando mejorar la capacidad de profesionales de la salud para acceder a esa información, valorar su calidad, cantidad y consistencia, para utilizarla en la toma de decisiones.

En la práctica docente, el método de revisión sistemática merece también atención. Actualmente, las jóvenes generaciones de estudiantes tienen gran facilidad para la búsqueda y el manejo de información a través de tecnologías y recursos digitales. El aprendizaje de los fundamentos y criterios que se utilizan durante una revisión sistemática puede contribuir al aprendizaje de algunas habilidades de gestión de la información y especialmente a un análisis cuidadoso de su calidad. Por otra parte, como se verá, la realización de una RS no solo conlleva la actualización de los conocimientos sino también una mejor comprensión de los métodos de investigación que proporcionan el conocimiento. En efecto, este tipo de investigación implica el manejo y la aplicación de conocimientos de metodología de la investigación en las valoraciones de la calidad de las investigaciones. Asimismo, como resultado de una revisión se pueden identificar vacíos o debilidades en el cuerpo de conocimientos acerca de una cuestión y así promover investigaciones originales y relevantes tanto en el ámbito universitarios como de la práctica profesional o en la fructífera interacción de ambos.

El contexto histórico de desarrollo del método de revisión sistemática

El método de revisión sistemática está ligado a los movimientos de **“evaluación de la evidencia científica”** y el de la **“medicina basada en la evidencia” (evidence based medicine)**, que emergieron con fuerza en las últimas décadas del siglo pasado en el ámbito de la clínica médica, y se extendieron a otros campos profesionales.

Si bien se pueden reconocer antecedentes mucho más antiguos, el movimiento de la medicina basada cobró identidad, estableció métodos y comenzó a globalizarse a partir de un grupo de epidemiólogos y clínicos canadienses, estadounidenses y británicos que plasmaron una propuesta de enfocar la educación y la práctica médica con la consideración de la mejor evidencia científica disponible para la toma de las decisiones clínicas **(Sackett, Rosenberg, Gray, Haynes, y Richardson, W. S., 1996)**. Así, recogieron el pensamiento del médico inglés Archie Cochrane **(1972)**, quien promulgaba la necesidad de distinguir entre aquellas prácticas médicas que eran ciertamente efectivas de aquellas de las que no se conocía su efectividad. Por otra parte, esta propuesta respondía a la necesidad urgente de manejar la extraordinaria producción tecnológica y científica observada en esos tiempos y aprovechaba el desarrollo de los métodos epidemiológicos para producir y evaluar críticamente las pruebas científicas.

Este movimiento cuestionó la tradición de basar las decisiones en el prestigio de expertos, autoridades académicas o científicas, en libros de texto que podrían contener información desactualizada o en la experiencia personal que, aunque posiblemente valiosa, podía tener una heterogeneidad notable entre profesionales. Sin detrimento de la experiencia del profesional o de la preferencia del paciente, el propósito de la medicina basada en la evidencia científica fue la utilización en la práctica clínica, ante cada paciente, del mejor y más actual conocimiento científico disponible, para lo cual se debía realizar una revisión crítica, sistemática y exhaustiva de la información científica.

Cabe mencionar también que, en el contexto de la práctica científica profesional, la palabra “evidencia” se utiliza por su significado en la lengua inglesa, que refiere a ciertos datos para probar o verificar la verdad o falsedad de un supuesto. En cambio, en español, significaría que tales datos no dejan ningún lugar a duda, lo cual es prácticamente imposible cuando se trata de resultados de una investigación. Por eso, la calificación de la evidencia como “científica” le otorga a la expresión una mayor equivalencia semántica con su idea original. Así, la “evidencia científica” es el conocimiento científico sobre una determinada cuestión, reconociendo su carácter de verificable y provisorio a la vez.

La revisión sistemática de la evidencia científica

La revisión sistemática de la evidencia científica (en adelante, RS) es uno de los procedimientos de evaluación del conocimiento científico más difundidos actualmente. Según la definieron Aymerich, Estrada y Jovell **(1999)**, es una metodología que permite conocer de la forma más objetiva posible, por una parte, que un procedimiento determinado -ya sea preventivo, diagnóstico, terapéutico o de rehabilitación- ha demostrado científicamente tener un efecto positivo sobre el curso de la enfermedad a la que se aplica (es decir, su eficacia, efectividad o eficiencia) y, por otra, el riesgo para la persona o para la población que puede tener ese procedimiento (seguridad). En el campo de la evaluación de tecnologías sanitarias, la revisión sistemática es considerada un tipo de investigación en el que la unidad de análisis son investigaciones o artículos científicos originales, es decir, investigaciones publicadas y el método permite buscar, evaluar y sintetizar el conocimiento científico actual sobre un tema provisto por esos artículos, mediante una estrategia planificada y rigurosa. El método de una RS incluye la valoración de la calidad, la cantidad y la consistencia de la información provista por una serie estudios que se dirigen a resolver una cuestión determinada y considera si son suficientes como para recomendar una decisión.

Diferencia entre una revisión sistemática y una revisión narrativa

En la literatura científica hay un gran interés por las revisiones de temas ya que, al reunir información proveniente de numerosas fuentes, tal como en los libros de texto, ahorran el tiempo que el lector necesitaría para hacer una elaboración similar. Es importante distinguir las revisiones sistemáticas entre otros formatos de revisiones.

La revisión se considera sistemática y es una investigación en tanto se realice siguiendo un método que tiene, entre sus pilares básicos:

- una búsqueda exhaustiva de artículos científicos, mediante criterios explícitos de selección e inclusión de estudios, en bases de datos bibliográficas pertinentes;
- una valoración pautada de la calidad de cada estudio analizado; y

ART 20 (2) pág. 73-82
Areté
ISSN-L: 1657-2513

➔ la presentación de resultados que incluye la descripción de las características metodológicas de los estudios originales, la valoración de su calidad y los datos extraídos de sus resultado

Las revisiones no sistemáticas suelen denominarse "narrativas". En comparación con una revisión sistemática, las revisiones narrativas suelen no tener un planteo claro de una pregunta estructurada que permita establecer unos criterios explícitos de inclusión de estudios, no presenta los términos y abasto de la estrategia de búsqueda bibliográfica o la búsqueda no es exhaustiva, no realiza una evaluación pautada de la calidad de los trabajos incluidos y la presentación de resultados no muestra las características metodológicas de los estudios. Si bien podrían llegar a conclusiones similares a las de una revisión sistemática, tienen menos garantías de validez por cuanto podrían no haber encontrado e incluido todos los estudios que podrían responder la pregunta o podría dar igual importancia a estudios de diferente calidad.

A modo de ejemplo se presenta una tabla extraída y adaptada del trabajo de Di Paoli, Berra y Paolini **(2020)** que compara revisiones sistemáticas que evalúan propiedades psicométricas de instrumentos de valoración de la calidad de vida relacionada con la voz o resultados percibidos por pacientes en el área de la salud vocal, donde se puede observar que algunas no explicitan el periodo de búsqueda o las bases de datos consultadas **(Zraick y Atcherson, 2012; Zraick y Risner, 2008)** por lo que no se puede valorar la exhaustividad de la búsqueda. Por otra parte, mientras algunas siguen un método explícito de selección y de evaluación de la calidad de los trabajos incluidos **(Branski y otros, 2010; Franic, Bramlett y Bothe, 2005; Speyer, y otros, 2019)**, otras toman de otras fuentes y reúnen información previamante producida **(Zraick y Risner, 2008)**. Con esto se puede decir que las revisiones de Zraick y Atcherson **(2012)** y de Zraick y Risner **(2008)**, aunque su contenido puede ser de sumo interés, no cumplen los criterios metodológicos de una revisión sistemática, mientras que la de Speyer y cols. **(2019)** los cumple acabadamente.

Características de revisiones sistemáticas que evalúan propiedades psicométricas de los instrumentos de calidad de vida relacionada con la voz o resultados percibidos por pacientes en el área de la salud vocal.

Autoría	Franic y cols.	Zraick y Risner	Branski y cols.	Zraick y Atcherson	Speyer y cols.
Año de publicación	2005	2008	2010	2012	2019
Período de búsqueda	No especifica, aparente hasta 2003	No especifica	Entre 1966 y 2007	No especifica	Hasta noviembre del 2016
Bases de datos consultadas	PsychInfo, Medline, Int. Pharmacy Abstracts, Soc. Sci. Citations Index, Current Contents, Health Star.	No especifica	Medline, Health and Psych. Instruments y Cum. Index to Nursing & Allied Health	Libros de texto o capítulos de libros, actas de congresos, revisiones publicadas y búsquedas académicas y generales en internet	Embase y PubMed
Estrategia de búsqueda	Sí	No	Sí	No	Sí, con detalle en tablas suplementarias
Criterios de inclusión de estudios	Muy precisos	No	Poco precisos	Poco precisos	Muy precisos
Método de evaluación de calidad de los estudios	Criterios de McHorney and Tarlov's, de Davis and Pathak's, de EPEC's y de Smith y cols.	No evalúa. Incluye resultados de Franic y cols. (2005)	Estándares de Scientific Advisory Committe of the Medical Outcomes Trust y FDA	FORCAST, Flesch Reading Ease (FRE) y FOG	Taxonomía COSMIN
Métodos de los estudios	No los expone	No los expone	No los expone	No los expone	Sí, en tabla suplementaria

El método de revisión sistemática

El método de revisión sistemática de la evidencia científica incluye las siguientes fases:

1. Planteo del objetivo y de los criterios de elegibilidad de estudios
2. Búsqueda y selección de los estudios y/o documentos relevantes
3. Evaluación de la calidad de los estudios
4. Elección de variables, extracción de datos y presentación de resultados
5. Análisis de heterogeneidad y metaanálisis (si es adecuado)
6. Interpretación de los resultados y recomendaciones

En un protocolo o artículo de investigación que presente una revisión sistemática, el planteo de la pregunta y el objetivo formarán parte de la introducción del trabajo, mientras que la sección de métodos podría organizarse indicando: a. diseño (método de revisión sistemática), b. criterios de elegibilidad, c. estrategias de búsqueda, d. evaluación de la calidad de los estudios, e. variables y extracción de datos, y f. análisis de datos y recomendaciones.

A continuación se describe cada una de las fases en sus aspectos principales, dejando para una audiencia interesada en más detalles las referencias de las guías de la Colaboración Cochrane **(Higgins y Green, 2011)** o el manual para el desarrollo de guías de práctica clínica, que incluyen revisiones sistemáticas, como las de los sistemas nacionales de salud de España **(Grupo de trabajo para la actualización del**

ART 20 (2) pág. 73-82
Areté
ISSN-L: 1657-2513

Manual de Elaboración de GPC, 2016) o Gran Bretaña **(Scottish Intercollegiate Guidelines Group, 2011).**

1. Planteo del objetivo y de los criterios de elegibilidad de estudios

Se recomienda que el diseño de una revisión sistemática se inicie avanzando, en esta primera en tres aspectos que se irán elaborando iterativamente hasta tener claridad del alcance y los límites de la revisión sistemática: a) el planteo de la pregunta de investigación o de revisión, b) el planteo del objetivo de la revisión sistemática, y c) la definición de criterios de elegibilidad de estudios y/o documentos

Se recomienda que el diseño de una revisión sistemática se inicie con la pregunta que debe responderse, de forma similar a la pregunta de cualquier otra investigación. En los inicios del movimiento de la evaluación de la evidencia científica, en el área clínica, se recomendó que la pregunta incluya a las personas que son objeto de estudio, las intervenciones o prácticas que se comparan y los indicadores de resultados o términos sobre los que se pretende evaluar sus efectos, tales como indicadores clínicos de diferentes niveles **(Richardson, Wilson, Nishikawa y Hayward, 1995)**, lo que se recuerda con la clave "PICO" que refiere a las palabras inglesas: **Patient, Intervention-Comparison, Outcome.**

Considerando esos términos con cierta amplitud, se describen estos componentes como los principales para integrar la pregunta para una revisión sistemática, que se deberán seleccionar en función del interés particular de quien realice la revisión:

- **Personas o población:** en estudios clínicos se deberá definir con precisión el diagnóstico o situación clínica de interés, mientras que en revisiones de salud pública se definirá una población delimitada geopolíticamente y/o un problema de salud-enfermedad sobre el que se requiere información.
- **Intervenciones o exposiciones:** según el ámbito del estudio, se podrá evaluar desde una intervención clínica con dianas muy precisas, que pueden ser tratamientos farmacológicos o quirúrgicos, procedimientos o instrumentos diagnósticos, otros tipos de prácticas terapéuticas, actividades o programas preventivos o políticas de salud con objetivos que podrán ser más o menos variados para modificar indicadores del problema de salud de interés. Tradicionalmente, evaluar una intervención conlleva comparar al menos dos alternativas viables en la misma situación clínica o población. De ellas, típicamente en el campo clínico, una es nueva y se evalúa bajo el supuesto de que aportará algún beneficio respecto de las que se aplican hasta el momento.
- **Resultados:** medida de resultado o de efecto de las intervenciones sobre la salud de las personas, que pueden ser indicadores bioquímicos, clínicos, sanitarios, etc.

Los elementos hasta aquí mencionados coinciden con el objeto de estudio y las variables u objetos de conocimiento de las preguntas de investigación. A diferencia del problema de investigación, en estos casos no se incluye una delimitación del contexto espacial y temporal porque inicialmente interesa buscar exhaustivamente los conocimientos científicos sobre el tema, con independencia del contexto en el que se generaron. Sin embargo, si se considera que el problema de salud, el tipo de intervención o sus efectos pueden tener una significativa influencia del contexto sociocultural, es posible que se sea conveniente agregar como elemento importante de la revisión un término de delimitación geográfica, política, cultural o contextual. De hecho, a partir de la clave PICO, ampliamente utilizada, fueron proponiéndose adaptaciones para contemplar las necesidades de búsqueda en otras áreas y metodologías. En este sentido, puede ser necesario ampliar el objeto de estudio a una población o cierto grupo específico de usuarios de un servicio, enfocar la evaluación a políticas o programas o a ciertos ámbitos particulares de los servicios sanitarios, a buscar no solo indicadores de resultados sino también de estructura, de organización o de proceso **(Wildridge y Bell 2002)**, indicar el tipo de investigación si se requiere enfocar a investigación cuantitativa o cualitativa, o más precisamente descriptores de los tipos de diseños o perspectivas que podrían ser útiles para responder la pregunta de interés **(Cooke, Smith y Booth, 2012)**.

El método de revisión sistemática puede dirigirse también a la caracterización y comparación de instrumentos diagnósticos o de medición para una misma variable o indicador de salud. En estos casos, igualmente se deberá identificar cuál es el concepto que miden, o bien la situación clínica, sanitaria o de otro tipo en que se sitúa el interés por la revisión y cuáles son las características o variables sobre las que se desea tener conocimiento científico. Por ejemplo, como se ha mostrado antes, se puede dirigir una revisión sistemática a la caracterización de instrumentos para medir calidad de vida relacionada con la voz de las personas y, sobre esos instrumentos, conocer su origen y desarrollo, los conceptos subyacentes a la medición, su validez, su fiabilidad y otras posibles propiedades métricas **(Di Paoli, Berra y Paolini, 2020)**.

A modo de recomendación final se advierte que, como en toda estrategia de búsqueda, agregar campos de descriptores puede mejorar la especificidad, pero consecuentemente aumenta el riesgo de no incluir artículos igualmente relevantes **(Methley, Campbell, Chew-Graham, McNally y Cheraghi-Sohi, 2014)**, o al menos, no mejoran la exhaustividad de una búsqueda con los descriptores básicos inicialmente presentados **(Kloda, Boruff, y Cavalcante, 2020)**.

Junto al desarrollo de la pregunta de investigación se consolidando el objetivo de la revisión sistemática. Si bien la pregunta puede no expresarse como tal sino como una inquietud que se presente en la introducción del trabajo de investigación, el objetivo seguramente deberá plantearse explícitamente. Los objetivos de una revisión sistemática en el área clínica son evaluar o comparar resultados de diferentes alterativas terapéuticas o diagnósticas, en términos de eficacia, efectividad, eficiencia y/o seguridad. Sin embargo, como se ha dicho, una revisión sistemática con otras inquietudes o en otras áreas puede tener el objetivo de buscar, evaluar y sintetizar información científica referida a un asunto particular, señalando clara y explícitamente cuáles son las características del asunto que son objeto de conocimiento. El objetivo se compone de los mismos elementos que forman parte de la pregunta y, además, se deberá indicar el periodo de tiempo que abarcará la revisión de la literatura científica.

Definir los criterios de elegibilidad de los artículos y estudios es una tarea crucial en este momento inicial. Tales criterios se utilizarán para el diseño de la estrategia de búsqueda bibliográfica y para la selección definitiva de los artículos y estudios que vayan a ser incluidos y evaluados en la revisión. Es decir, se definen los criterios de inclusión de estudios para ser revisados sistemáticamente. Algunos criterios pueden aparecer como consecuencia directa de los componentes de la pregunta de revisión, como las características de las personas que deben ser sujeto de las investigaciones, pero otras pueden no ser tan obvias, como establecer qué tipos de estudios son los más adecuados para resolver la pregunta de investigación planteada **(Sackett y Wennberg, 1997)**. Por ejemplo, si se trata de la evaluación de una intervención, el diseño más conveniente será un ensayo clínico aleatorizado;

mientras que para estudiar los efectos de una exposición con la que no se puede experimentar, el diseño adecuado podría ser un estudio observacional de cohortes. Para definir criterios de elegibilidad, se pude considerar y establecer explícitamente:

- Características de las personas incluidas en los estudios; en caso de diagnósticos, especificar criterios e incluso las técnicas e instrumentos de determinación, si es necesario.
- Variables de interés (intervenciones o exposiciones y resultados) sobre las que debe haber información suficiente.
- Tipos de estudios o diseños de investigación pertinentes.
- Límite temporal de la búsqueda (años que abarca).
- Bases de datos bibliográficas y otras fuentes.
- Lengua en las que el equipo de investigación podrá revisar y evaluar las publicaciones.
- Tipos y estado de las publicaciones.

2. Búsqueda y selección de los estudios y/o documentos relevantes

En esta fase se diseña y ejecuta una búsqueda bibliográfica en la que los términos de la estrategia se reunirán en conjuntos en función de los componentes de la pregunta de revisión, los tipos de estudio que se han considerado adecuados y otras claves como puede ser el tipo de artículos (por ejemplo, originales y otras revisiones sistemáticas, pero no editoriales o cartas), las lenguas viables para las personas que hacen la revisión y el límite temporal que se pondrá a la búsqueda. La búsqueda produce una cantidad de referencias de cada base de datos, que conviene registrar en términos cuantitativos.

La identificación de los estudios se puede realizar en dos momentos, a los que les denominamos preselección y selección de estudios. Con la información que proporcionan las bases de datos bibliográficas se debe hacer una preselección, mediante la constatación en títulos y resúmenes de los criterios de inclusión y exclusión. En esta acción se desechan rápidamente una buena cantidad de publicaciones que evidentemente no cumplen los criterios de inclusión definidos. Ante la menor duda, es conveniente consultar el texto completo.

Cuando se consultan varias bases de datos bibliográficas, que es lo correcto, se producen resultados duplicados. Para evitar la consecuente duplicación de tarea que se puede llegar a producir como consecuencia, es conveniente que los trabajos preseleccionados se vayan listando en una planilla de datos manteniendo los campos de autores, títulos, resumen, datos bibliográficos y base de datos en la que se obtuvo.

Aquellos que parezcan pertinentes y los aún dudosos, deben localizarse en sus formatos de documentos completos para confirmar en el resto de la publicación el cumplimiento de los criterios de selección establecidos. A medida que se va constatan los criterios de inclusión de los artículos se registrará que es un artículo seleccionado, o en su defecto, excluido. EN este momento también puede ser importante identificar diferentes publicaciones de un mismo estudio, casos para los cuales se debe haber definido si la unidad de análisis son estudios o son artículos. Los registros seleccionados constituirán la base de datos principal de la revisión sistemática.

De este paso de la revisión sistemática se pueden realzar diagramas que muestren desde la cantidad de referencias obtenidas en cada base de datos consultadas, pasando por los diferentes momentos y criterios de exclusión hasta el número de artículos finalmente incluidos en la revisión.

3. Evaluación de la calidad de los estudios

Como se ha mencionado, la evaluación de la evidencia científica se hace en términos de su la calidad, cantidad y consistencia. En este momento se realiza la evaluación de la calidad de los estudios, que se ha señalado como uno de los aspectos característicos y fundamentales de una revisión sistemática. En esta fase, dos personas deben evaluar cada estudio de forma independiente, es decir, que cada uno realice y registre sus valoraciones sin consultar con su par evaluador. Al finalizar cada evaluación, podrán consensuar sus opiniones y, si es necesario, pedir la opinión de una tercera persona. Si es posible, otra persona debe enmascarar los estudios de forma que se oculten los datos de autores y de la revista, para que un supuesto de prestigio no influya en quien evalúa.

Existen instrumentos para esta tarea, conocidos como guías de lectura crítica, listados de verificación o **checklists**. Estas guías el examen sistemático de todos los aspectos metodológicos que se relacionan con los criterios de calidad de la evidencia científica comentados antes. Existen numerosos tipos de ***checklists*** entre los que se pueden elegir los más adecuados según las preferencias de quienes realicen la revisión. Estas herramientas, han sido confeccionadas por instituciones que se dedican a la evaluación de tecnologías sanitarias o a la confección de guías de práctica clínica, como ***S*cottish Intercollegiate Guidelines Network** (SIGN), Osteba (Servicio de Evaluación de Tecnologías Sanitarias del País Vasco) o el **Centre for Reviews and Dissemination** (CRD) de la Universitat de York. Asimismo, existen éstos y otros recursos facilitados por iniciativas de formación de profesionales de la salud como **Evidence Based Medicine Working Group** (Universidad McMaster, Canadá) y el Programa CASP (**Critical Appraisal Skills Programme**).

Dentro de ellos, hay instrumentos específicos para cada tipo de estudio (ensayos clínicos aleatorizados, cohortes, casos y controles, etc.). La mayoría de estas herramientas están diseñados como planillas en las que se registran puntajes y valoraciones que cada revisor hace sobre un determinado estudio. Así, cada evaluador completa un listado de verificación y luego las dos evaluaciones se comparan y se discuten aquellos aspectos en los que hay discordancias entre revisores, hasta llegar a una calificación final única de cada estudio.

La evaluación de la evidencia científica tiene el propósito de decidir si la **calidad, cantidad y consistencia** de la información provista por una serie estudios sobre una cuestión determinada son suficientes como para tomar una decisión. La fuerza de la evidencia científica para hacer una recomendación depende de las observaciones que se hagan sobre esas tres características.

La **calidad** de la evidencia científica depende fundamentalmente de la rigurosidad con la que se ha obtenido el conocimiento científico. Por ello, la evaluación de la calidad requiere la revisión cuidadosa de todos los aspectos metodológicos de cada uno de los estudios que se tengan en cuenta. Los aspectos a considerar son, en este orden, la validez interna, la magnitud y precisión de los resultados y la validez externa o aplicación de las conclusiones del estudio.

La **validez interna** es el grado en que los resultados se deben al efecto estudiado y no a la presencia de algún error sistemático (sesgos) en la ejecución del estudio. La validez interna se

ART 20 (2) pág. 73-82
Areté
ISSN-L: 1657-2513

asume en la medida que se reduzca al mínimo la posibilidad de que los resultados estén afectados por sesgos. Una adecuada validez interna implica un bajo riesgo de error sistemático o sesgo. Existen múltiples tipos de errores sistemáticos que, en general, se agrupan en sesgos de selección y sesgos de información. El riesgo de incurrir en sesgos varía según el tipo de estudio que se realice ya que hay diseños y técnicas que pueden reducirlo mucho más que otros.

Cuando se valoran los **resultados** de un estudio se tiene en cuenta la magnitud y la precisión de los mismos. La **magnitud** de los resultados es la distancia o diferencia obtenida entre los resultados de los diferentes grupos comparados en la investigación. Según el tipo de estudio, las medidas de la magnitud de los resultados pueden ser diferentes, tales como diferencias en las frecuencias observadas y esperadas del evento de interés que pueden ser cuantificadas, por ejemplo, como riesgo relativo u otras. La **precisión** de un estudio es el grado en que su resultado está libre de error aleatorio, es decir, un concepto que juzga la diferencia entre la estimación de una característica en una muestra y el parámetro o valor real de tal característica en la población. El error aleatorio es inherente a toda investigación que se realiza con una muestra de la población. Como resultado del estudio de una muestra se obtiene una estimación puntual, que probablemente será diferente al resultado si repetimos la investigación en otra muestra, aunque ésta sea de la misma población. Por eso interesa calcular un intervalo de confianza, cuyos valores mínimo y máximo indican el rango en que se encontraría el valor en toda la población. Se dice que un resultado es más preciso cuanto menor es la amplitud del intervalo de confianza, la cual disminuye a medida que aumenta el tamaño de la muestra.

La **validez externa** es el grado en que los resultados son correctos para sujetos de otras poblaciones. Se refiere a la potencialidad de generalizar o extrapolar los resultados obtenidos en una muestra a una población (por ejemplo, de pacientes) que puede tener características diferentes a la población de estudio. El término **generalizar** se suele utilizar cuando se hace una inferencia desde la población de estudio a la población diana. El término **extrapolar** se utiliza cuando se quiere aplicar los resultados de un estudio a una población diferente de la que fue la población de estudio. Desde la perspectiva de un lector interesado en aplicar los resultados de un estudio a su propia población de interés (comunidad, pacientes, etc.), durante la evaluación de la validez externa deberá analizar críticamente las similitudes y diferencias entre los sujetos estudiados y los que recibirían la intervención, mediante la comparación de las características de ambas poblaciones.

- La validez interna de un estudio es un requisito indispensable antes de considerar la validez externa. Si existen dudas sobre los resultados del estudio, es infructuoso reflexionar acerca de su extrapolación a otras poblaciones.
- Para un lector crítico evaluando la validez interna y externa de una investigación es importante tener clara la definición de población de estudio. La validez interna puede estar limitada por algún tipo de diferencias entre la población de estudio y la muestra, por errores en la selección de la muestra; la validez externa estará amenazada por las diferencias entre la población de estudio y la población diana (generalización) u otra población de interés a la que quiera aplicar los resultados (extrapolación).

Finalmente se valora la **cantidad** y la **consistencia** de las evidencias científicas acerca de la cuestión, que da mayor seguridad acerca de los resultados. Cuantos más estudios se hayan realizado para intentar responder la pregunta de investigación, mayor confianza tendremos en sus resultados. La consistencia refiere a que la relación comprobada en un estudio se observa de manera similar en otros estudios, es decir, que los resultados no son contradictorios entre unos y otros.

En síntesis, la calificación de la calidad del estudio es una valoración global que se emite tras haber evaluado validez interna, resultados y validez externa. Por ejemplo, el SIGN propone utilizar la siguiente notación

- "++" cuando se cumplen todos o la mayoría de los criterios; y en los puntos en que no se han cumplido, se considera muy poco probable que dicho incumplimiento pueda afectar a las conclusiones del estudio o revisión
- "+" cuando se cumplen algunos de los criterios; y se considera poco probable que los criterios que no se han cumplido o que no se describen adecuadamente puedan afectar a las conclusiones del estudio o revisión.
- "-" cuando se cumplen sólo unos pocos criterios o ninguno de ellos, por lo que se considera probable o muy probable que esto afecte a las conclusiones

Algunas revisiones sistemáticas excluyen de los pasos siguientes a aquellos estudios que fueron calificados negativamente, ya que, al tener un alto riesgo de sesgos, sus resultados pierden interés.

4. Elección de variables, extracción de datos y presentación de resultados

Todas las características de los estudios que interese describir o comparar, exponiéndolas en tablas o en textos, evaluar, otorgando niveles de calidad o sintetizar como resultados de la revisión sistemática son equivalentes a variables de cualquier otro tipo de investigación. En esta sección del protocolo se deberán listar, definir e indicar cómo se valorará cada una de las características de los estudios.

Durante la ejecución del método de revisión sistemática, esos datos se extraen de los artículos o documentos seleccionados y puede hacerse en formularios independientes para cada estudio o en planillas en las que se recoja toda la información de todos los estudios. Algunos checklist de evaluación incluyen ya algunos datos mínimos que se recomienda extraer y, adicionalmente, se pueden adaptar para que incluyan todo el registro necesario de las variables de interés. En cualquier caso, estos instrumentos de recolección de datos deben estar suficientemente desarrollados antes de comenzar la etapa de extracción, para evitar que cada persona participante de la revisión se maneje con criterios diferentes.

La extracción debe ser cuidadosa y exhaustiva a fin de tener reunida de forma homogénea toda la información necesaria de los estudios cuando sea el momento de tabular datos y analizarlos. La tarea de extracción de los datos se hace también por las dos personas que evaluaron cada estudio. Dado que estos datos luego formarán parte de las tablas de presentación de resultados, por lo tanto, puede ser conveniente hacer una prueba piloto con uno o dos estudios para acordar el estilo y el detalle de registro de datos, a fin de que luego sea fácilmente trasladable a las tablas de resultados. Aún cuando se han establecido previamente los criterios para la extracción de los datos, ambas personas pueden hacerlo de manera diferente y la comparación de ambos trabajos permitirá una mejor recolección de los datos de cada artículo.

Las tablas constituyen el formato visual más práctico para sintetizar la información que contienen los estudios y que interesarán al lector de la revisión. Habitualmente, cada fila de la tabla corresponderá a un estudio evaluado y en las columnas de las tablas se presentan

las variables o características que se comparan. Sin embargo, según la cantidad de estudios y de variables que se quieran incorporar, este esquema puede invertirse.

Las revisiones que abarcan numerosas variables pueden tener diferentes tablas de resultados, lo que se debe señalar muy claramente en el título de la tabla. En este punto cabe recordar que uno de los criterios que define a una revisión sistemática es la presentación de resultados que incluye la **descripción de las características metodológicas de los estudios originales, la valoración de su calidad y los datos extraídos de sus resultados**. De ahí que es recomendable que los resultados se presenten en al menos dos tipos de tablas:

- Tablas de características de los estudios incluidos, incluyendo por columnas: la referencia del estudio (primer autor y año); diseño del estudio; ámbito; población, muestra o características de los participantes (diagnóstico) y valoración de su calidad metodológica;
- Tablas de resultados por estudios incluidos, incluyendo por columnas: la referencia del estudio (primer autor y año): exposiciones o intervenciones y control; resultados; calificación de la calidad del estudio; observaciones adicionales.

Si se diseñara una única tabla con una síntesis de los resultados más importantes, para una presentación breve, se debería incluir la identificación del estudio, los elementos principales de la pregunta de revisión y calidad de los estudios.

Ejemplo de tabla con los componentes mínimos que podrían exponerse si la pregunta de investigación responde al formato PICO.

Estudio	Población	Intervención / comparación	Resultado	Observaciones	Calidad de la evidencia
Autor, año					
Autor, año					
Autor, año					
Autor, año					

La tabulación de la información y su presentación en el cuerpo del informe constituyen una síntesis cualitativa de la información que proporcionan los artículos. A continuación, puede considerarse la realización de una síntesis cuantitativa, como es un metaanálisis. Si esto no es factible, según lo que se verá a continuación, el siguiente paso será el de la interpretación de los resultados a partir de su síntesis cualitativa, para establecer recomendaciones.

5. Análisis de heterogeneidad y metaanálisis

Cuando los estudios que forman una revisión son muy similares a punto tal que sus protocolos de investigación han sido prácticamente iguales en cuanto a la muestra y la medición de las variables de interés de la revisión, sus resultados podrían reunirse para obtener una medida sintética y más precisa, como si se hubiera realizado un único estudio con todas las personas que fueron incluidas en todos los estudios. Para esto se debe comprobar, en primer lugar, la homogeneidad de los estudios, es decir, que en sus principales elementos sean coincidentes. Esto significa que los criterios de inclusión de pacientes, las características de las intervenciones que se comparan y los métodos del estudio son similares, así como que los resultados son consistentes. Si bien esto se puede examinar a simple vista en la tabla de resultados, existen pruebas estadísticas de homogeneidad. Si los estudios son heterogéneos no debe hacerse un metaanálisis.

Las técnicas para realizar el metaanálisis son relativamente sencillas, apoyadas en softwares específicos. Los resultados de un metaanálisis proporcionan una medida global como síntesis cuantitativa de los resultados de todos los estudios incluidos. El resultado global tendrá mayor precisión (por ejemplo, un intervalo de confianza más pequeño) por la ganancia que representa en términos de tamaño muestral, la conjunción de todos los estudios.

6. Interpretación de los resultados y recomendaciones

Como se mencionó, una revisión sistemática tiene el propósito de presentar una síntesis de la evidencia científica existente para responder a una pregunta concreta, con una la valoración de la calidad, la cantidad y la consistencia de esa información disponible. En el escenario típico de la revisión sistemática, que es el ámbito clínico, esta síntesis debe ser un recurso útil para quien deba tomar una decisión ante una situación clínica o sanitaria determinada. Entonces, la interpretación de los resultados presentados debe orientarse a los conceptos antes desarrollados de calidad de la calidad de la evidencia científica, considerando validez interna, magnitud y precisión de los resultados y validez externa o aplicación de los estudios incluidos en la revisión. Concretamente, interesa destacar **la cantidad de estudios de alta calidad que son consistentes en un resultado preciso (Guyatt y otros, 2008)**.

De esta síntesis pueden emerger conclusiones y recomendaciones tales como: "La síntesis de los datos existentes ha sido limitada por las dispares medidas de resultado utilizadas entre los estudios. Solo dos de los siete estudios publicados hasta la fecha, con un total de 35 participantes, informan mejoras estadísticamente significativas en el grupo de terapia de integración auditiva y solo para dos medidas de resultado. Por lo tanto, no hay suficiente evidencia de que la terapia de integración auditiva u otras terapias de sonido sean efectivas como tratamientos para los trastornos del espectro autista." Las recomendaciones que emanan de las revisiones científicas realizadas rigurosamente facilitan a los profesionales cierta confianza en las decisiones que deben tomar frente a sus pacientes y a la gestión de los recursos.

El método de revisión sistemática permite detectar fácilmente las limitaciones de los estudios, con lo que vale la pena dedicar una sección de los resultados a señalar cuáles son las debilidades más frecuentes en las investigaciones valoradas y, en la discusión, proponer mejoras metodológicas para avanzar en el conocimiento del tema. Asimismo, se deben identificar y hacer explícitas limitaciones de la propia revisión sistemática.

ART 20 (2) pág. 73-82
Areté
ISSN-L: 1657-2513

Calidad y limitaciones de las revisiones sistemáticas

Como para otros tipos de estudios y publicaciones, existen herramientas para guiar la confección y la lectura crítica de revisiones sistemáticas. Las instituciones dedicadas a la elaboración de revisiones sistemáticas, como ejemplo, merece ser citada la Colaboración Cochrane, disponen de extensos manuales para sus grupos de trabajo. La declaración PRISMA establece las pautas para la publicación de revisiones en revistas. La herramienta AMSTAR se dirige a evaluar la metodología de realización de una revisión sistemática y también el programa CASP ofrece una herramienta de lectura crítica de revisiones sistemáticas.

Referencias

Aymerich, M., Estrada, M.D., Jovell, A.J. (1999). Revisión sistemática de la evidencia científica. En: ***Jovell AJ, Aymerich M. Evidencia científica y toma de decisiones en sanidad.*** Monografies Mèdiques de l'Acadèmia de Ciències Mèdiques de Catalunya i de Balears. Barcelona.

Brady, M.C., Kelly, H., Godwin, J., Enderby, P., Campbell, P. (2016). ***Speech and language therapy for aphasia following stroke.*** Cochrane Database of Systematic Reviews. Issue 6. Art. No.: CD000425. DOI: *10.1002/14651858.CD000425.pub4.*

Branski, R. C., Cukier-Blaj, S., Pusic, A., Cano, S. J., Klassen, A., Mener, D., . . . Kraus, D. H. (2010). Measuring quality of life in dysphonic patients: a systematic review of content development in patient-reported outcomes measures. ***Journal of Voice***, 193-8.

Cochrane, A.L. (1972). E***ffectiveness and efficiency: Random reflections on health services***. Nuffield Trust.

Cooke, A., Smith, D., Booth, A. (2012). Beyond PICO: the SPIDER tool for qualitative evidence synthesis. ***Qualitative health research,*** 22(10), 1435–1443. *https://doi.org/10.1177/1049732312452938*

Di Paoli, F., Berra, S., Paolini, G. (2020) ***Revisión Sistemática de las Características de los Instrumentos de Autopercepción de la Voz disponibles en Argentina y en el contexto internaciona***l. Trabajo Final de Investigación para la Licenciatura en Fonoaudiología. Escuela de Fonoaudiología (FCM, UNC), Córdoba, Argentina.

Franic, D. M., Bramlett, R. E., Bothe, A. C. (2005). Psychometric evaluation of disease specific quality of life instruments in voice disorders. ***J Voice,*** 300-15.

Grau Magaña, M., Estrada Sabadell, M.D. (2020). E***ficacia, efectividad, eficiencia y seguridad de la implantación bilateral de implantes auditivos activos: implante activo de oído medio, implante de tronco cerebral e implante de conducción ósea***. Barcelona: Agència de Qualitat i Avaluació Sanitàries de Catalunya. Departament de Salut. Generalitat de Catalunya (Informes de Evaluación de Tecnologías Sanitarias).

Grupo de trabajo para la actualización del Manual de Elaboración de GPC (2016). ***Elaboración de Guías de Práctica Clínica en el Sistema Nacional de Salud.*** Actualización del Manual Metodológico [Internet]. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad; Zaragoza: Instituto Aragonés de Ciencias de la Salud (IACS); [Febrero 2018]. Disponible en: *http://portal.guiasalud.es/emanuales/elaboracion_2/?capitulo*

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schünemann, H. J., & GRADE Working Group (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. ***BMJ (Clinical research ed.)***, 336(7650), 924–926. *https://doi.org/10.1136/bmj.39489.470347.AD*

Higgins, J.P.T., Green, S. (editors). (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. ***The Cochrane Collaboration,*** 2011. Disponible en: *www.cochrane-handbook.org*.

Kloda, L. A., Boruff, J. T., Cavalcante, A. S. (2020). A comparison of patient, intervention, comparison, outcome (PICO) to a new, alternative clinical question framework for search skills, search results, and self-efficacy: a randomized controlled trial. Journal of the Medical Library Association : ***JMLA***, 108(2), 185–194. *https://doi.org/10.5195/jmla.2020.739*

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. ***BMC health services research***, 14, 579. *https://doi.org/10.1186/s12913-014-0579-0*

Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. (1995). The well-built clinical question: a key to evidence-based decisions. ***ACP journal club***, 123(3), A12–A13.

Sackett, D. L., & Wennberg, J. E. (1997). Choosing the best research design for each question. ***BMJ (Clinical research ed.)***, 315(7123), 1636. *https://doi.org/10.1136/bmj.315.7123.1636*

Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. ***BMJ (Clinical research ed.)***, 312(7023), 71–72. *https://doi.org/10.1136/bmj.312.7023.71*

Scottish Intercollegiate Guidelines Group. SIGN 50: ***A guideline developers' handbook***. Edinburgh: SIGN Publication No.50; 2011. Edición revisada en 2011. Disponible en: *https://www.sign.ac.uk/assets/sign50_2011.pdf*

Sinha, Y., Silove, N., Williams, K., Hayen, A. (2004). ***Auditory integration training and other sound therapies for autism spectrum disorders***. Cochrane Database of Systematic Reviews. Issue 1. Art. No.: CD003681. DOI: *10.1002/14651858.CD003681.pub2.*

Speyer, R., Kim, J. H., Doma, K., Chen, Y. W., Denman, D., Phyland, D., Parsons, L., Cordier, R. (2019). Measurement properties of self-report questionnaires on health-related quality of life and functional health status in dysphonia: a systematic review using the COSMIN taxonomy. ***Quality of life research***, 28(2), 283–296. *https://doi.org/10.1007/s11136-018-2001-6*

Wildridge, V., Bell, L. (2002). How CLIP became ECLIPSE: a mnemonic to assist in searching for health policy/ management information. ***Health information and libraries journal,*** 19(2), 113–115. *https://doi.org/10.1046/j.1471-1842.2002.00378.x*

Zraick, R. I., & Atcherson, S. R. (2012). Readability of Patient-Reported Outcome Questionnaires for Use With Persons With ***Dysphonia. Journal of Voice***, 635-641.

Zraick, R. I., & Risner, B. Y. (2008). Assessment of quality of life in persons with voice disorders. ***Current Opinion in Otolaryngology & Head and Neck Surgery,*** 188-93