



Investigación en Educación Médica

ISSN: 2007-865X

revistainvestedu@gmail.com

Universidad Nacional Autónoma de

México

México

Cobos Aguilar, Héctor
Lectura crítica de investigación en educación médica
Investigación en Educación Médica, vol. 5, núm. 18, abril-junio, 2016, pp. 115-120
Universidad Nacional Autónoma de México
Distrito Federal, México

Disponible en: <http://www.redalyc.org/articulo.oa?id=349745408008>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto



METODOLOGÍA DE INVESTIGACIÓN EN EDUCACIÓN MÉDICA

Lectura crítica de investigación en educación médica



Héctor Cobos Aguilar*

Departamento de Ciencias Clínicas, División Ciencias de la Salud, Universidad de Monterrey, Monterrey, Nuevo León, México

Recibido el 13 de enero de 2016; aceptado el 27 de enero de 2016
Disponible en Internet el 8 de marzo de 2016

PALABRAS CLAVE

Lectura crítica;
Investigación en
educación médica

Resumen

Introducción: La lectura crítica de informes de investigación médica es el sustento de la medicina basada en evidencias. Para facilitar el abordaje de los numerosos artículos publicados se han utilizado diversas guías para cada diseño que han llevado a la formación de CONSORT, STROBE, QUÓRUM y PRISMA.

La investigación médica se encuentra poco desarrollada y aunque existen diversas guías para el abordaje de los artículos publicados, su uso no se encuentra generalizado limitando así su valoración con rigor metodológico.

Se revisan brevemente los diseños de investigación educativa médica más frecuentemente abordados en nuestro medio: validación de instrumentos, intervenciones educativas y encuestas.

En cada diseño se refieren las características más relevantes que se analizan: la postura del autor, el uso de instrumentos válidos y confiables y su proceso de validación, la variable dependiente, las estrategias educativas, los estadísticos no paramétricos utilizados, los resultados y la discusión.

Se presentan sucintamente tres guías relacionadas con los diseños previamente referidos, resumiendo los aspectos metodológicos referidos.

Comentarios: La formación docente en los aspectos teóricos en investigación educativa es insustituible para la crítica de informes de investigación en educación médica. Sin embargo, las guías pueden apoyar inicialmente el análisis crítico de los artículos publicados.

Derechos Reservados © 2016 Universidad Nacional Autónoma de México, Facultad de Medicina. Este es un artículo de acceso abierto distribuido bajo los términos de la Licencia Creative Commons CC BY-NC-ND 4.0.

* Autor para correspondencia. Calle 8 No. 430, Col. Villa Azul, San Nicolás de los Garza, Nuevo León, CP:66420, Tel.: +01(81) 83768504.
Correo electrónico: hector.cobos@udem.edu
La revisión por pares es responsabilidad de la Universidad Nacional Autónoma de México.

KEYWORDS

Critical reading;
Research in medical
education

Critical appraisal in medical education research**Abstract**

Introduction: Critical reading of medical research reports is the essence of evidence-based medicine. In order to simplify the approach to the numerous published articles, various guidelines have been used for each study design, resulting in the creation of CONSORT, STROBE, QUORUM, and PRISMA.

Research in medical education is underdeveloped, and although there are several guidelines focusing on the approach to these articles, they are not widespread, thus limiting their evaluation with methodological rigour.

The most frequently analysed designs used in medical education in our field are briefly reviewed: instrument validation, educational interventions, and surveys.

The most relevant characteristics of each design are referred to, including: the author's position, the use of valid and dependable instruments and their validation process, the dependent variable, educational strategies, non-parametric statistics, results and discussion.

Three guidelines relating to the previously referred designs are also briefly presented, underscoring their methodological aspects.

Comments: Knowledge of the theoretical aspects of educational research is irreplaceable when appraising medical education research reports. However, guidelines may initially support the critical analysis of published articles.

All Rights Reserved © 2016 Universitat Nacional Autònoma de Mèxic, Facultat de Medicina. This is an open access item distributed under the Creative Commons CC License BY-NC-ND 4.0.

Lectura crítica de investigación en educación

La medicina basada en evidencias¹ se fundamenta en la lectura crítica de los informes de investigación médica (LCIIM) publicados, lo que ha refinado su evaluación y permite al lector seleccionar los pertinentes para su conducta médica.

El abordaje se realiza a través de guías, que facilitan su análisis metodológico y la transformación de *Journal Clubs*². En México, Viniestra la introdujo en especialidades médicas, internado de pregrado y estudiantes de medicina. Paralelamente se investigó el desarrollo de la LCIIM en diferentes grados académicos, la elaboración de instrumentos válidos y confiables para su medición y la evaluación de estrategias educativas para su aprendizaje con resultados adecuados³⁻⁵.

Estos se consolidan con una formación docente paralela, en una visión diferente de la educación médica, con énfasis en la investigación y que pondera la experiencia en estrategias educativas consecuentes para su aprendizaje⁶.

Por lectura crítica (LC) se entiende, una capacidad que, para estar suficientemente desarrollada, supone en el lector una predisposición «casi automática» de hacer consciente, ante la proximidad de una lectura, su propia postura sobre el tema –con grados variables de conocimiento, reflexión y elaboración–, la misma que será confrontada con lo expresado en el texto; el debate constante con el autor le hace posible al lector descubrir los supuestos implícitos, la idea directriz (articuladora) y analizar los puntos fuertes y los débiles de los principales argumentos del escrito; de esta manera puede proponer otros planteamientos que superen los del autor y reafirmar o modificar su postura previa.

La LCIIM tiene algunos ejemplos de investigación breve: la interpretación, con reconocimiento del tipo de estudio, diseño e identificación de características del instrumento y desciframiento de tablas y gráficas. Enjuiciar se relaciona

con distinguir entre los puntos débiles y fuertes de las estrategias y estadísticos utilizados y proponer, son proposiciones de procedimientos que superen a los planteamientos del autor⁷.

El número de publicaciones científicas hace imposible su abordaje por lo que se han utilizado guías de evaluación en investigación empírica^{8,9} que se han consolidado en el análisis de los distintos diseños como CONSORT, TREND, STROBE, QUORUM Y PRISMA¹⁰⁻¹².

Nos hemos referido hasta ahora a la LCIIM, cuya implementación, desafortunadamente, aún no está generalizada en las diferentes instituciones formadoras de médicos. Esta surge de una visión educativa inquisidora, o participativa, contraria a una mirada tradicional, aquiescente, y de la idea de conocimiento que cada una contiene, es decir elaboración *versus* consumo de información. La LCIIM difícilmente se desarrolla en ambientes ajenos a la autonomía del aprendizaje y sin estrategias que promuevan la elaboración del conocimiento¹³.

Definición de las particularidades de la investigación en educación médica, en contraste con el resto de literatura biomédica

La investigación en educación médica (IEM) es escasa en México y su ejercicio –paradójicamente– es una actividad ajena a los entornos educativos^{14,15}. Sus resultados no se utilizan para mejorar los procesos educativos que exploran. En este contexto la IEM se encuentra aún menos expuesta a la lectura crítica, por lo que se pueden observar deficiencias en artículos publicados de intervenciones educativas, por lo menos en indagaciones realizadas analizando sus títulos y resúmenes¹⁶, sin embargo, debiera seguir la evolución que hemos referido brevemente con LCIIM.

Para ello es necesaria la formación en investigación y especialmente de la lectura crítica en los currículos universitarios¹⁷, así como la formación docente en IEM para valorar el rigor metodológico de los artículos publicados y ponderar la pertinencia de su implementación y, a través de preguntas de investigación adecuadamente planteadas, diseñar investigaciones educativas propias que superen los problemas del entorno, fundamentadas con evidencias educativas¹⁸.

La educación basada en evidencias es el medio para mejorar la educación médica en los innumerables aspectos susceptibles de investigar: currículo, estrategias educativas, evaluación, entre otros.

Aplicabilidad de esquemas «convencionales» para la evaluación crítica de la investigación

En relación con lo anterior los esquemas utilizados en la LCIM pueden adaptarse para su uso en la IEM. Los diseños más frecuentemente abordados en nuestro medio se refieren a la construcción y validación de instrumentos, encuestas descriptivas y comparativas, transversales y longitudinales, intervenciones educativas y los metaanálisis, aún menos frecuentes. Algunos autores incluso recomiendan diseños sintéticos educativos¹⁹⁻²¹.

Recomendaciones prácticas para la lectura de investigación en educación médica

Como se refirió antes, una producción elevada de publicaciones condujo a facilitar su lectura con guías o lineamientos para reportes educativos²².

Para facilitar el análisis y crítica de artículos de IEM se presentan tres guías, muy básicas, de los diseños educativos más frecuentemente publicados en nuestro medio, que pueden ser útiles en esta actividad. Son, además, los diseños que el autor ha utilizado. Sin embargo, las guías no sustituyen la formación teórica continua y desafiante del lector en estos aspectos.

En los reportes de IEM es transcendental identificar la postura –generalmente implícita– de los autores. Esta se entiende como «... asumir puntos de vista propios acerca del porqué de una situación problema (educativa) de qué hacer y cómo actuar para atenuar o acrecentar sus efectos y procurar su eliminación o permanencia», es decir, su mirada educativa²³.

Elaboración de instrumentos

Estos artículos²⁴ se refieren a los que se construyen por los autores y aquellos importados de otros ámbitos. Ambos siguen un procedimiento riguroso en su validación²⁵. En los primeros, se indaga la referencia a la validez (medir lo que pretende medir) ya sea intrínseca (teórica, constructo, contenido) y extrínseca (concurrente y predictiva). Se explora asimismo la que se refiere a la validez intrínseca en las mediciones que se pretenden realizar, incluyendo el propósito u objetivo del estudio, que deben estar claramente referidos²⁶. La bibliografía seleccionada desvela la postura educativa. Posteriormente revisamos la validez

de contenido con la inclusión de los aspectos relevantes que conforman los distintos indicadores, ameritando o no un equilibrio. Es crucial evaluar la adecuación gramatical del instrumento, así como la experiencia del autor principal y expertos invitados para su validación. El número de estos puede variar, pero se aconseja contar, por lo menos, con cinco o seis evaluadores. Enseguida hay que revisar el procedimiento de validación, idealmente con la técnica Delphi (paso fundamental)²⁷. A continuación se evalúa el formato de respuestas propuestas, generalmente una escala de Likert, aunque se adecua a la indagación deseada. Posteriormente se buscan las modificaciones sugeridas y realizadas al instrumento y las respuestas consideradas correctas por los expertos, si este fuera el caso, con acuerdo por lo menos de 5/6 o 6/6, equilibradas con las incorrectas y deben equilibrarse. El procedimiento de calificación debe detallarse así como la obtención de los resultados y diferentes estratos con sus intervalos, si fuese necesario. El enmascaramiento de quienes aplican y califican el instrumento debe referirse. Es importante indagar las características de la población o los grupos naturales en la aplicación, y en estos últimos el tamaño muestral pasa a un segundo término, ya que aunque algunos autores recomiendan seleccionar entre 5 y 10 respondedores por reactivo indagado, en general no cuentan con ese número de alumnos. Una vez analizados estos aspectos los autores deben referir la aplicación a un grupo piloto de alumnos semejantes a los que se aplicará finalmente el instrumento. Con ello se intenta su comprensión adecuada, además de la validación por expertos. Es necesario analizar el contexto en el que aplica el instrumento: lugar, horario, tiempo y aceptación para su resolución, si esta influye en la calificación final, etc. Si el instrumento no forma parte del curso regular, proporcionar los resultados individuales o grupales.

De acuerdo al propósito, algunos autores refieren la comparación con otros instrumentos ya validados (validez concurrente) o predicen un resultado (validez predictiva).

Después de la aplicación y calificación se indaga cómo se determinó la consistencia, ya sea con el alfa de Cronbach, en tres o más opciones de respuesta o la fórmula 21 de Kuder-Richardson en dicotómicas²⁸. Es adecuado buscar otros procedimientos como el *test-retest* o la partición del instrumento en mitades.

Cuando evaluamos la adaptación cultural o local de un instrumento importado y validado en su contexto, se revisan los aspectos referidos y se agregan otros como la traducción del idioma original al español y a la inversa, por expertos.

Las pruebas estadísticas utilizadas se valoran también. En educación generalmente se utilizan variables cualitativas por lo que los estadísticos no paramétricos son los más utilizados y dependen de las variables estudiadas²⁹.

Se espera que los resultados puedan ser claros y de acuerdo al objetivo del estudio. La discusión debe señalar las limitaciones del estudio, así como sus fortalezas y debilidades. Una muestra, si no el instrumento completo, debe anexarse para su análisis.

Es frecuente que muchos artículos publicados omitan algunos de los pasos referidos lo que traduce una validez y consistencia insuficientes que afectan los resultados y la aplicabilidad de ese instrumento particular (tabla 1).

Tabla 1 Guía breve para la lectura crítica de un artículo de elaboración de instrumentos

1. *Introducción (postura)*
2. *Objetivo*
3. *Validez*
 - a) Teórica
 - b) Contenido
 - c) Predictiva
 - d) Concurrente
4. *Técnica Delphi*
5. *Calificación*
6. *Prueba piloto*
7. *Población*
8. *Enmascaramiento en calificación*
9. *Consistencia*
 - a) Alfa de Cronbach
 - b) Kuder-Richardson
 - c) Test-retest
10. *Resultados en cuadros y/o gráficas*
11. *Comentarios*
Derivados de resultados
 - a) Fortalezas
 - b) Debilidades
 - c) Aplicabilidad

Tabla 2 Guía breve para la lectura crítica de un artículo de intervención educativa

1. *Introducción (postura)*
2. *Objetivo*
3. *Definición operacional de variables*
 - a) Independiente (maniobra educativa)
 - b) Dependiente (resultado)
 - c) Confusoras (profesor, alumno, contexto)
4. *Conformación de los grupos*
 - a) Aleatorización
 - b) Quasiexperimento
5. *Instrumentos*
 - a) Validez referida
 - b) Consistencia referida
6. *Medición inicial y final (mínimo)*
 - a) Se define el contexto y condiciones de medición
7. *Estadísticos pertinentes*
 - a) Paramétricos
 - b) No paramétricos
 Se fundamenta adecuadamente su uso
8. *Resultados en cuadros y/o gráficas*
 - a) Claros, con valores de p o IC
9. *Comentarios*
 - a) Derivados de los resultados
 - b) Refiere fortalezas y debilidades
 - c) Aplicabilidad

Intervenciones educativas

Son los artículos que evalúan estrategias educativas y en ellos se deben buscar ciertas características que les confieren validez³⁰. La validez se refiere a la confiabilidad de las estrategias y traduce la relevancia y trascendencia del estudio: desarrollar habilidades complejas o memorizar conceptos a corto plazo en alumnos, por ejemplo. Se analiza la operacionalización de la variable dependiente (aprendizaje, desarrollo, etc.), la escala de medición y el instrumento propuesto (válido y confiable) y los criterios para su calificación y ceguera.

La conformación de los grupos (dos o más) es imprescindible, por aleatorización *versus* grupos naturales. Sus características se detallan para identificar posibles sesgos (calificaciones, género, escuela de procedencia, medio socioeconómico, p.ej.).

En los grupos se exploran las intervenciones educativas (variable independiente) a contrastar: duración, elaboración de tareas, experiencia docente, impacto en la calificación del alumno, entre otras, y las variables contextuales (ambiente educativo, aula, auditorio, permiso para asistencia, etc.). Se analizan pérdidas de alumnos y sus motivos, si ocurren.

Los estadísticos no paramétricos se fundamentan: entre dos grupos, inicio y término de las estrategias (U de Mann-Whitney) o más términos (Kruskal-Wallis). Cambios intragrupal, inicio y término (Wilcoxon) o correlación de variables (S de Spearman).

Los resultados se expresan claramente en cuadros, gráficas y/o figuras. La discusión se relaciona con los resultados

obtenidos, sin serendipia. Se refieren las limitaciones del estudio, debilidades y fortalezas (tabla 2).

Encuestas

También se valora la postura en la introducción. El objetivo se expresa claramente (con frecuencia es confuso). Se revisa la variable dependiente, bien operacionalizada, el instrumento (válido y confiable), su calificación y contexto de aplicación, referido antes. Además se indaga quién, cómo y cuándo se aplica, así como quién y cómo lo califica (enmascaramiento).

Se explora si se aplicará a un grupo o más, transversal o longitudinal, así como el tamaño muestral de los grupos y sus características sociodemográficas, de acuerdo al propósito del mismo, pues se identifican más fácilmente las variables confusoras.

Los estadísticos paramétricos o no paramétricos se relacionan con la escala de medición de la variable dependiente y los segundos se han referido en el diseño previo.

Los resultados, los cuadros y gráficas deben elaborarse con claridad. La discusión se fundamenta en los resultados y refiere las limitaciones, fortalezas y debilidades del estudio (tabla 3).

Revisiones sistemáticas y metaanálisis

Cada vez más se utilizan estos diseños y también se han construido diferentes índices cualitativos para su evaluación³¹. Se publican poco en nuestro país, pero sustentan de manera

Tabla 3 Guía breve para la lectura crítica de un artículo de encuesta educativa

1. *Introducción (postura)*
2. *Objetivo*
3. *Variable dependiente*
 - a) Definición operacional
 - b) Escala de medición
4. *Instrumentos*
 - a) Validez referida
 - b) Consistencia referida
5. *Población*
 - a) Grupos naturales
 - b) Selección y tamaño muestral
6. *Aplicación*
 - a) Contexto
 - b) Control de sesgo (entrevistador/alumno)
7. *Medición*
 - a) Calificación ciega
8. *Estadísticos pertinentes*
 - a) Paramétricos
 - b) No paramétricos
 - c) Se fundamenta su uso
9. *Resultados en cuadros y/o gráficas*
 - a) Claros, bien construidos
10. *Comentarios*
 - a) Derivados de los resultados
 - b) Refiere fortalezas y debilidades
 - c) Aplicabilidad

más adecuada la educación basada en evidencias³². No se abordan en este reporte.

Conclusiones

La educación predominante en nuestro país no utiliza a la LCIM o la de IEM como herramientas de aprendizaje aunque no difieren metodológicamente.

Es posible utilizar guías para evaluar las publicaciones en IEM pero esta actividad no sustituye la formación teórica, permanente, desafiante y continua del lector.

Sin profesores que critiquen y realicen IEM de calidad no podemos esperar alumnos, futuros docentes, que enfrenten sus problemas educativos con educación basada en evidencias.

Responsabilidades éticas

Protección de personas y animales. El autor declara que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad de los datos. El autor declara que en este artículo no aparecen datos de pacientes.

Derecho a la privacidad y consentimiento informado. El autor declara que en este artículo no aparecen datos de pacientes.

Financiación

Ninguna.

Autoría/colaboradores

HCA elaboró y revisó el manuscrito.

Conflicto de intereses

El autor declara no tener ningún conflicto de intereses.

Referencias

1. Espino SS, Figueroa DR, Baptista GH, Ramírez CJA. Medicina basada en evidencias: Introducción. *Perinatol Reprod Hum.* 2010;24(3):194–201.
2. Kitching AD, Guyatt GH. The Critical Appraisal Club. *CMAJ.* 1987;136:819–21.
3. Viniegra VL. La experiencia reflexiva y la educación. *Rev Invest Clin.* 2008;60(2):133–56.
4. Cobos AH, Pérez CP, Vázquez AMZ, Garcialuna ME. Lectura crítica de investigación en el internado médico de pregrado. SIIC En Internet www.siicsalud.com, Sección Expertos Invitados, septiembre 2014 [consultado 19 Feb 2016]. Disponible en: www.siicsalud.com/dato/experto.php/137883.
5. Cobos AH, Pérez CP, Ramírez MM, Tapia OJA. Lectura crítica de investigación en estudiantes de medicina. Efecto a un año. *Rev Med Inst Mex Seguro Soc.* 2006;44 Supl. 3:85–91.
6. Vázquez MCA, Insfrán Sánchez MD, Cobos AH. Papel de la experiencia docente en estrategias educativas promotoras de la participación. *Rev Med IMSS.* 2003;41(1):23–9.
7. Viniegra VL. La investigación en la educación, papel de la teoría y la experiencia. Segunda edición. México, D.F.: Ed. Leonardo Viniegra; 2000. p. 87–155.
8. Alderson D. How to critically appraise a research paper. *Surgery.* 2012;30(9):477–80.
9. Cronin P, Rawson JV, Heilbrun ME, Lee JM, Kelly AM, Sanelli PC, et al. How to critically appraise the clinical literature. *Acad Radiol.* 2014;21:1117–28.
10. Moraga CJ VR. Pautas de chequeo Parte I: CONSORT y TREND. *Rev Chil Cir.* 2015;67(2):225–32.
11. Elm EV, Altman DG, Egger M, Pocock ST, Gøtzsche PC, Vandenbroucke JP. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007;370:1453–7.
12. Moraga CJ VR. Pautas de chequeo, Parte II QUOROM y PRISMA. *Rev Chil Cir.* 2015;67(3):325–30.
13. Viniegra VL. La educación y la crítica del conocimiento. Una perspectiva transdisciplinaria para entender la subjetividad y sus posibilidades cognitivas. México, D.F.: Ed. Hospital Infantil de México; 2015. p. 175–212.
14. Hamui SL. Los artículos de educación médica. Desafíos actuales para publicar. *Rev Med Inst Mex Seguro Soc.* 2012;50(3):233–4.
15. Hernández PF. Investigación educativa. *Arch Med Urgen Méx.* 2013;5(1):3–4.
16. Cook DA, Beckman TJ, Bordage GA. Systematic review of titles and abstracts of experimental studies in medical education: many informative elements missing. *Med Educ.* 2007;41:1074–81.
17. Laidlaw A, Aiton J, Struthers J, Guild S. Developing research skills in medical students: AMEE Guide No. 69. *Med Teach.* 2012;34:754–71.
18. Mc Kimm J, Morris C. Approaches to medical education research. *Br J Hosp Med.* 2010;71(7):404–8.

19. Lynch DC, Whitley TW, Willis SE. A rationale for using synthetic designs in medical education research. *Adv Health Sci Educ.* 2000;5:3-103.
20. Mc Gaghie WC, Bordage G, Crandall S, Pangaro L. Method research design. Review criteria. *Acad Med.* 2011;76(9):929-30.
21. Curry LA, Nembhard IM, Bradley EH. Qualitative and mixed methods provide unique contributions to outcomes research. *Circulation.* 2009;119:1442-52.
22. Kelly AM. Evaluating and writing education papers compared with noneducation papers. *Acad Radiol.* 2012;19:1100-9.
23. Viniestra VL. La crítica: aptitud olvidada por la educación. *Vislumbres en el lado opaco de la medicina.* México, D.F: IMSS; 2000.
24. Cobos AH, Pérez CP, de la Garza QH, Ochoa CCE. Proceso de validación de un instrumento para medir la lectura crítica de informes de investigación médica. *Inv Ed Med.* 2015;4(16):200-6.
25. Roncero C. La validación de instrumentos psicométricos: un asunto capital en la salud mental. *Salud Mental.* 2015;38(4):235-6.
26. Cook DA, Bordage G, Schmidt HG. Description, justification and clarification: a framework for classifying the purposes of research in medical education. *Med Educ.* 2008;42:128-33.
27. García VM, Suárez MM. El método Delphi para la consulta a expertos en la investigación científica. *Rev. Cub. Salud Pública.* 2013;39(2):253-67.
28. Da Silva FC, Gonçalves E, Arancibia BAV, Bento GG, Castro TLS, Hernández SSS, et al. Estimaciones en las investigaciones en salud: el uso del coeficiente alfa. *Rev Peru Med Exp Salud Publica.* 2015;32(1):169-78.
29. Siegel S, Castellan NJ. *Estadística no paramétrica. Aplicada a las ciencias de la conducta.* 4.^a ed. Mexico, D.F: Ed. Trillas; 2012.
30. Cook DA, Beckman TJ, Bordage G. Quality of reporting of experimental studies in medical education: a systematic review. *Med Educ.* 2007;41:737-45.
31. Gómez Ortega OR, Amaya Rey MC. ICRESAI-IMeCI: instrumentos para elegir y evaluar artículos científicos para la investigación y la práctica basada en evidencia. *Aquichan.* 2013;13(3):407-20.
32. Liberati A, Altman DA, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2000;9:e1-34.