

# Iniciativas Disponibles para el Reporte de Resultados en Investigación Biomédica con Diferentes Tipos de Diseño

## Initiatives for Reporting Biomedical Research Results with Different Types of Designs

Carlos Manterola<sup>\*\*</sup>; Tamara Otzen<sup>\*\*\*</sup>; Nicolas Lorenzini<sup>\*\*\*\*</sup>; Andrés Díaz<sup>\*\*\*</sup>;  
Rodrigo Torres-Quevedo<sup>\*\*\*,\*\*\*\*\*</sup> & Nataniel Claros<sup>\*\*\*\*\*</sup>

---

MANTEROLA, C.; OTZEN, T.; LORENZINI, N.; DÍAZ, A.; TORRES-QUEVEDO, R. & CLAROS, N. Iniciativas disponibles para el reporte de resultados en investigación biomédica con diferentes tipos de diseño. *Int. J. Morphol.*, 31(3):945-956, 2013.

**RESUMEN:** La calidad del reporte de los resultados de una investigación no es óptima, razón por la cual, se han desarrollado numerosas iniciativas tendientes a mejorar este aspecto a lo largo de los años. El objetivo de este artículo es mencionar y describir las iniciativas existentes para el reporte de resultados de investigación biomédica en diversos escenarios de investigación clínica y situaciones especiales. Se realizó una búsqueda en las bases de datos THE COCHRANE LIBRARY, MEDLINE, SciELO y Redalyc; y en los buscadores Clinical Evidence, TRIP database, Fistera, Rafabravo, EQUATOR Network, portal de BIREME y Programa HINARI; para obtener las listas de verificación existentes. Los documentos recuperados fueron agrupados de la siguiente forma: relacionados con escenarios de terapia, diagnóstico, pronóstico, evaluaciones económicas y misceláneas. La búsqueda generó un total de 31 documentos. Doce para escenarios de terapia (CONSORT, QUOROM, MOOSE, STRICTA, TREND, MINCIR-Terapia, RedHot, REHBaR, PRISMA, REFLECT, Ottawa y SPIRIT), 5 para diagnóstico (STARD, QUADAS, QAREL, GRRAS y MINCIR-Diagnóstico), 3 para pronóstico (REMARK, MINCIR-Pronóstico y GRIPS), 4 para evaluaciones económicas (NHS-HTA, CHEERS, ISPOR RCT-CEA y NICE-STA); y 7 misceláneos (STROBE, COREQ, GRADE, SQUIRE, STREGA, ORION y MINCIR-EOD). Existen diversas iniciativas y declaraciones. Estas deben ser conocidas y utilizadas por escritores, revisores y editores de revistas biomédicas; de forma tal de incrementar la calidad del reporte de resultados de la investigación biomédica.

**PALABRAS CLAVE:** Reporte de resultados; Listas de verificación; Terapia; Diagnóstico; Pronóstico; Evaluaciones económicas.

---

## INTRODUCCIÓN

La correcta comunicación de los resultados de una investigación se ha convertido en un tema recurrente a medida que se acumula evidencia respecto de lo inadecuado del reporte de resultados dadas las deficiencias observadas en la calidad de la forma en que se realiza.

Desde la concepción y posterior aplicación masiva de la práctica clínica basada en la evidencia, diversas señales han impactado en el quehacer clínico. Uno de ellas dice relación con la forma en que reportamos los resultados de la investigación. Es así como la publicación de la declaración CONSORT (Consolidated Standards of Reporting Trials) en 1996 (Begg *et al.*, 1996), su revisión de 2001 (Moher *et al.*, 2001) y su ulterior extensión en

2004 (Campbell *et al.*, 2004), cuyo objetivo fue mejorar la calidad del informe de los ensayos clínicos (EC); se transformó en un verdadero ícono en esta materia, actuando como motivador de una verdadera avalancha de iniciativas tendientes a mejorar el reporte de resultados, las que en términos generales tienden a definir aquellos aspectos mínimos que deben considerarse en la publicación de un estudio.

El objetivo de este artículo es describir las distintas iniciativas existentes para el reporte de resultados de investigación biomédica en diversos escenarios de investigación clínica y situaciones especiales no cubiertas por las anteriores.

\* Departamento de Cirugía, Universidad de La Frontera, Temuco, Chile.

\*\* Centro de Investigación en Ciencias Biomédicas, Universidad Autónoma de Chile, Temuco, Chile.

\*\*\* Programa de Doctorado en Ciencias Médicas, Universidad de La Frontera, Temuco, Chile.

\*\*\*\* Escuela de Psicología, Universidad Autónoma de Chile, Temuco, Chile.

\*\*\*\*\* Programa de Magister en Ciencias Médicas, Universidad de La Frontera, Temuco, Chile.

\*\*\*\*\* Hospital Obrero de La Paz y Universidad de San Andrés, La Paz, Bolivia.

Parcialmente financiado por Proyecto DID-UFRO DI12-0037, Dirección de Investigación de la Universidad de La Frontera.

## MATERIAL Y MÉTODO

Se realizó una búsqueda en las bases de datos THE COCHRANE LIBRARY ([www.thecochranelibrary.com](http://www.thecochranelibrary.com)), MEDLINE (<http://www.ncbi.nlm.nih.gov/pubmed>), SciELO (<http://www.scielo.org/php/index.php>), Redalyc (<http://redalyc.uaemex.mx>); y en los buscadores Clinical Evidence (<http://clinicalevidence.bmj.com/x/index.html>), TRIP database (<http://www.tripdatabase.com>), Fisterra ([www.fisterra.com](http://www.fisterra.com)), Web de información médica-Infodoctor (<http://www.infodoctor.org/rafabravo>), Web de EQUATOR Network (<http://www.equator-network.org>), Portal de evidencias de la Biblioteca Virtual en Salud de BIREME (<http://evidences.bvsalud.org>) y en la página de Acceso a la Investigación en Salud del Programa HINARI (<http://www.who.int/hinari/es>); con el objeto de obtener las diversas iniciativas o listas de chequeo existentes.

La lectura y extracción de datos de los diversos instrumentos analizados fue realizada en forma independiente por cinco investigadores (CM, TO, NL, AD y RT), todos ellos con formación en análisis crítico de la literatura biomédica y epidemiología clínica.

Para cada una de las iniciativas recuperadas, se generó una hoja de recogida con los datos principales de cada iniciativa. Posteriormente, se desarrolló un primer borrador en el que se consideraron los datos obtenidos de cada una de las propuestas. Este, fue discutido en diversas sesiones de trabajo grupal, oportunidades en que las que se resolvieron por consenso aquellas situaciones en las que se generaron diferencias.

Los documentos recuperados fueron agrupados de la siguiente forma: relacionadas con escenarios de terapia o procedimientos terapéuticos, de diagnóstico, de pronóstico, de evaluaciones económicas y misceláneas; y explicados por orden de publicación.

## RESULTADOS

La búsqueda realizada permitió recuperar un total de 31 documentos, los que se describen a continuación agrupados según tipo de escenario clínico.

**Escenarios de terapia.** Para escenarios de terapia o procedimientos terapéuticos, se recuperaron 12 documentos; tres de los cuales se refieren a la utilización de revisiones sistemáticas (RS) y meta-análisis con este tipo de escenarios; y una escala de calidad metodológica (CM) para estudios de terapia.

1. La declaración CONSORT (Consolidated Standards of Reporting Trials), fue desarrollada a mediados de los noventa por el grupo CONSORT, quienes publicaron por primera vez el instrumento generado en 1996 (Begg *et al.*). Con el tiempo se transformó en una herramienta fundamental para el reporte de resultados de EC; siendo utilizada de forma rutinaria por varias revistas biomédicas como pauta de reporte. En 1999, 13 miembros del grupo CONSORT se reunieron para revisar la propuesta original, analizar cada ítem y determinar cuáles eran relevantes. Posteriormente, esta revisión fue compartida con otros comentaristas y todos los cambios fueron discutidos en una reunión de consenso; lo que generó el documento publicado en 2004 (Campbell *et al.*). De este modo, se definió una pauta de chequeo compuesta por 22 ítems que se agrupan en 5 dominios (título y resumen, introducción, métodos, resultados y discusión). No considera asignación de puntaje al aplicar la pauta.

2. La declaración QUOROM (Quality of Reports of Meta-analyses of Randomized Clinical Trials), se gesta en 1996 tras la conformación del grupo QUORUM, integrado por académicos e investigadores de las Universidades de Ottawa, Hamilton y otros centros; y financiado por Abbot Laboratories, Agency for Health Care Policy and Research, Glaxo Wellcome y Merck & Co (Moher *et al.*, 1999). Su objetivo fue crear una herramienta útil para el reporte de resultados de RS basadas en EC. Metodológicamente fue desarrollado a través de un panel de 30 expertos conformado por epidemiólogos clínicos, estadísticos, editores e investigadores; los que identificaron aquellos elementos que deberían ser incluidos en una lista de chequeo. De esta manera, los elementos seleccionados se eligieron por la evidencia científica para evitar resultados sesgados. Posteriormente a través de técnica Delphi modificada, se evaluaron los elementos seleccionados; generándose de este modo, una lista compuesta por 6 dominios (título, resumen, introducción, métodos, resultados y la discusión del informe de un meta-análisis) y 18 ítems; que incluyen un diagrama de flujo. Esta propuesta no consigna asignación de puntaje para la valoración de los estudios.

3. La propuesta MOOSE (Meta-analysis of Observational Studies in Epidemiology), se desarrolló como un proyecto financiado por el Center for Disease Control and Prevention (Stroup *et al.*, 2000). Su objetivo fue desarrollar un instrumento con recomendaciones a los autores, revisores, editores y lectores de metaanálisis de estudios observacionales. Metodológicamente fue desarrollado por un panel de 27 expertos en práctica clínica, ensayos, estadística, epidemiología, ciencias sociales y edición biomédica. Estos, realizaron una revisión de la literatura sobre la realización y presentación de informes de los metaanálisis de estu-

dios observacionales, examinando un total de 32 artículos. Posterior a ello, generaron una lista de chequeo compuesta por 6 dominios (antecedentes, estrategia de búsqueda, métodos, resultados, discusión y conclusión), que incluyen 35 ítems, sin asignación de puntaje para ellos.

4. La recomendación STRICTA (Standards for Reporting Intervention in Controlled Trials of Acupuncture), fue creada en 2001, por un grupo de investigadores en acupuntura reunidos en la Exeter University del Reino Unido (MacPherson *et al.*, 2001). Su objetivo fue generar una pauta de chequeo para estudios de evaluación de intervenciones con acupuntura que utilizaran como diseño el EC. La metodología utilizada en la construcción del instrumento fue una reunión de consenso; lo que dio como resultado una pauta compuesta por 6 ítems, que debe aplicarse en conjunto con la declaración CONSORT, como una extensión de esta. Los ítems que la componen son: fundamentos de la acupuntura, detalles de las punciones, régimen de tratamiento, cointervenciones, experiencia del operador y control (intervenciones). Tal y como lo señala CONSORT, se considera que las recomendaciones STRICTA se encuentran en proceso de mejoramiento, por lo que podrían salir versiones posteriores a la publicada. Esta lista no considera asignación de puntaje.

5. La declaración TREND (Transparent Reporting of Evaluations with Nonrandomized Designs), nació a partir de una iniciativa del Centers for Disease Control and Prevention de Atlanta (Des Jarlais *et al.*, 2004). Su objetivo fue generar una herramienta para el análisis de EC cuando no es posible realizar una asignación aleatoria, considerando que existen estudios distintos a los EC controlados que pueden aportar información de importancia y de calidad. La metodología utilizada en su creación es una reunión de 18 representantes de revistas científicas sobre intervenciones conductuales en VIH; quienes diseñaron la lista de chequeo. Está compuesta por 5 dominios (título y resumen, introducción, métodos, resultados y discusión), que incluyen 21 ítems. No contempla asignación de puntaje para la valoración de los estudios.

6. La escala MINCIR-terapia (Metodología e Investigación en CIRugía); es una escala creada a partir de un proyecto de investigación financiado por la Universidad de La Frontera (Manterola *et al.*, 2003, 2009). El objetivo fue diseñar y construir una escala válida y confiable para medir CM de estudios referentes a terapia o procedimientos terapéuticos. Metodológicamente fue desarrollado por un panel de expertos. Se aplicó validez de fachada y contenido (en su última versión se determinó validez de criterio y de constructo; y confiabilidad interobservador). Consta de 6 ítems agrupados en 3 dimensiones (diseño del estudio, tamaño de la

muestra estudiada y metodología utilizada), cuya sumatoria genera una escora entre 6 y 36 puntos; siendo el punto de corte 18 puntos el que define el constructo “para estudios de terapia”.

7. La propuesta RedHot (Reporting Data on Homeopathic Treatments), nace como un suplemento de la declaración CONSORT orientado a los estudios homeopáticos. Fue desarrollado de forma colaborativa entre el Departamento de Ciencias de la Salud de la University of York y otros centros (Dean *et al.*, 2007). Su objetivo fue crear un instrumento de consenso para el reporte de tratamientos homeopáticos. La metodología utilizada para su génesis fue una conferencia electrónica internacional de 12 expertos en el campo; posterior a la cual los integrantes de este panel fueron divididos en editores de revistas, profesionales y no profesionales con experiencia en investigación homeopática en EC y RS. Finalmente, se trabajó aplicando técnica Delphi en tres etapas. De este modo, se generó una lista de chequeo compuesta por 8 ítems (justificación, participantes, medicamentos, consultas, médicos, co-intervenciones, intervenciones control y eventos adversos), con sus respectivas descripciones, las que se asocian con algunos ítems de la declaración CONSORT. No se contempla asignación de puntaje para su aplicación.

8. La propuesta REHBaR (Reporting Experiments in Homeopathic Basic Research), nace a partir de una iniciativa de académicos e investigadores de las Universidades de Bologna, Duisburg-Essen y otros centros; y fue cofinanciado por la fundación Karl and Veronica Carstens (Stock-Schröer *et al.*, 2009). Su objetivo fue desarrollar una nómina de criterios que sirvieran de guía a los autores para mejorar la calidad del reporte de resultados de experimentos en la investigación básica en homeopatía. La metodología utilizada para su génesis fue un panel de expertos; y posteriormente, la aplicación de técnica Delphi en 5 etapas. De este modo, se generó una lista compuesta por 4 dominios (título/resumen, introducción material y métodos; y resultados), los que incluyen 23 ítems con sus respectivos descriptores. No considera asignación de puntaje.

9. La declaración PRISMA (Preferred Reporting Items for Systematic Reviews), se desarrolló como una iniciativa de académicos de la Università di Modena e Reggio Emilia, el Cancer Research de Reino Unido y otros centros; y fue financiada por the Canadian Institutes of Health Research y la compañía GlaxoSmithKline de Canadá (Moher *et al.*, 2009). Corresponde a la actualización de la declaración QUORUM, cuyo objetivo fue resolver varios avances conceptuales y prácticos de la ciencia de las RS; entre los cuales cabe señalar que la ejecución de una RS es un proceso repetitivo, que la realización e informe de una investigación

son conceptos diferentes, que el nivel de evidencia no es sinónimo del riesgo de sesgo y que el reporte de los sesgos es fundamental. Metodológicamente fue desarrollado por un panel de expertos, poco después de la reunión del grupo QUORUM; de este modo, un proyecto de la lista PRISMA fue distribuido a los miembros del grupo. Se creó un archivo con comentarios y revisiones de cada entrevistado y la lista de verificación, la que fue revisada en 11 oportunidades por el grupo de trabajo; quienes aprobaron además el diagrama de flujo. La lista de chequeo así construida, quedó compuesta por 7 dominios (título, resumen, introducción, métodos, resultados, discusión y financiamiento), que incluyen 27 ítems, sin asignación de puntaje para ellos.

10. La declaración REFLECT (Reporting Guidelines for Randomized Controlled Trials in Livestock and Food Safety), fue creada por académicos del ámbito de la medicina veterinaria, de las universidades de Guelph, Iowa State, California y otros centros; y financiado por un Grant de la USDA Food Safety and Response Network (Sargeant *et al.*, 2010). Su objetivo fue generar una guía para mejorar el reporte de EC relacionados con “ganadería y seguridad alimentaria”. La metodología utilizada fue una reunión de consenso en la que se aplicó una modificación de la declaración CONSORT. De este modo, se creó una lista de verificación compuesta por 5 dominios (título y resumen, introducción y antecedentes, método y participantes, resultados y discusión) y 22 ítems que incluyen un diagrama de flujo de los participantes. No hay asignación de puntaje.

11. La declaración de Ottawa sobre la ética del diseño y la conducción de EC con asignación aleatoria por conglomerados; nace como una propuesta de un grupo de investigadores de las Universidades de Western, de Ottawa y otros centros canadienses; y financiado por the Canadian Institutes of Health Research (Weijer *et al.*, 2012). Se realizó una reunión de consenso que se basó en un proyecto de investigación de 5 años; el que incluyó métodos mixtos, incorporando trabajo empírico con informantes clave y análisis ético. Su objetivo es aportar guías para el diseño ético y conducción de este tipo de estudios. Está orientado a investigadores en EC, pero podría resultar de interés para quienes financian investigación, diseñan políticas sanitarias y potenciales participantes en estos estudios. La metodología de esta declaración se sustenta en un proyecto de investigación de 5 años, financiado por la Canadian Institutes of Health Research. Desde el punto de vista metodológico, se desarrolló bajo un enfoque mixto que incluye trabajo empírico y análisis ético. En la parte empírica se realizó una revisión de los EC con asignación aleatoria por conglomerados publicados, una entrevista con investigadores y una evaluación por comités de ética; etapa en la que se identificaron 6 aspectos éticos fundamentales que requerían mayor análisis.

Estos, se analizaron en un encuentro multidisciplinario de expertos, realizada en Ottawa, en 2011; del que emanaron 15 recomendaciones éticas para el diseño y conducción de EC con asignación aleatoria por conglomerados; las que se agrupan en 7 dominios (justificación del diseño, revisión por comité de ética, identificación de participantes, consentimiento informado, controlador de acceso, evaluación de riesgos y beneficios; y protección de los participantes vulnerables). Esta lista de chequeo no contempla asignación de puntaje.

12. La iniciativa SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials), fue desarrollada en el año 2013 por académicos de la Universidad de Toronto (Chan *et al.*, 2013). Su objetivo es mejorar la calidad de los protocolos de EC, al desarrollar una pauta para evaluar el contenido mínimo de un protocolo de EC. Fue realizado consultando a 115 participantes (investigadores de EC, profesionales de la salud, metodólogos, estadísticos, coordinadores de estudios, editores de revistas, representantes de estamentos éticos, financiadores y agencias reguladoras). La guía SPIRIT se diseñó en base a dos RS, un consenso con técnica Delphi; y dos reuniones de consenso cara a cara y un pilotaje. Incluye una lista de 5 dominios (información administrativa, introducción, métodos, ética y difusión; y apéndices) y 33 ítems con detalles de la justificación. No consigna asignación de puntaje para la valoración de los estudios.

**Escenarios de diagnóstico.** Para escenarios de diagnóstico, se recuperaron 5 documentos, 4 listas de verificación y una escala de CM para estudios de diagnóstico.

1. La propuesta STARD (Standard for Reporting of Diagnostic Accuracy), fue creada en 2001 por el grupo Screening Test Methods Working Group de la colaboración Cochrane (Bossuyt *et al.*, 2003). Su objetivo fue generar un estándar para el informe de estudios de precisión diagnóstica. La metodología utilizada en la construcción del instrumento fue primero la elaboración de una lista de 75 puntos posibles. Posterior a esto, se realizaron búsquedas bibliográficas en las bases de datos: MEDLINE, EMBASE, BIOSIS; y en la base de datos metodológica de la Colaboración Cochrane, hasta junio de 2000. Se examinaron las referencias bibliográficas de los artículos seleccionados; posteriormente se contactaron a expertos del campo de la investigación diagnóstica, para finalmente ampliar la lista inicial. Finalmente, utilizaron un panel de expertos compuesto por investigadores, editores, especialistas en metodología y organizaciones profesionales, quienes evaluaron todos los puntos inicialmente considerados, para elaborar la lista final que quedó compuesta por 5 dominios (título/resumen/palabras clave; introducción, métodos, resultados y discu-

sión), un diagrama de flujo y un total de 25 ítems con formato de respuesta abierta de cumple o no cumple. Esta herramienta asigna puntuación y además informa sobre el método de selección de los pacientes, el orden de realización de las pruebas y el número de pacientes que deben evaluarse mediante la prueba índice y la prueba de referencia.

2. La herramienta QUADAS (Quality Assessment Diagnostic Accuracy Studies), se desarrolló como un proyecto colaborativo entre el Centre for Reviews and Dissemination (University of York) y la Academic Medical Centre (Universidad de Amsterdam); financiado por el programa Health Technology Assessment y publicado en su primera versión en 2003 (Whiting *et al.*, 2003). Una versión modificada, fue utilizada por la Colaboración Cochrane en RS de precisión de pruebas diagnósticas. Su objetivo fue generar una herramienta para la valoración de la calidad de estudios primarios de precisión diagnóstica incluidos en una RS. Metodológicamente fue desarrollado por un panel de expertos que diseñaron una lista de ítems que recogieron de la literatura médica y consideraron relevantes. Posteriormente, utilizando el método Delphi con 4 etapas seleccionaron 14 ítems, cada uno de los cuales se puntuaba como “sí”, “no” o “dudoso”; de modo tal que la opción “sí” indicaba siempre una buena respuesta. Posteriormente y con aportaciones de la Cochrane, se desarrolló la segunda versión (QUADAS-2); una herramienta mejorada, con un nuevo diseño, que se basa tanto en la experiencia con la función original como en la nueva evidencia sobre las fuentes de sesgo y variación de los estudios de exactitud diagnóstica. Esta, se ejecutó en 4 fases: definición del alcance de ésta, revisión de la evidencia, reunión de consenso de “cara a cara”, y perfeccionamiento del instrumento generado a través de un pilotaje. QUADAS-2, publicada en 2011 (Whiting *et al.*, 2011); se se aplica en 4 fases: resumen de la pregunta de la revisión, adaptación al estudio que se analiza, construcción de un diagrama de flujo para el estudio primario; valoración del riesgo de sesgo y aplicabilidad. Esta herramienta permite una adecuada apreciación de los sesgos y de la aplicabilidad de los estudios de exactitud diagnóstica.

3. La herramienta QAREL (Quality Appraisal tool for Studies of diagnostic reliability), fue desarrollada como una iniciativa de académicos e investigadores del School of Biomedical and Health Sciences, el Sydney School of Public Health y otros centros (Lucas *et al.*, 2010). Su objetivo fue desarrollar una herramienta de evaluación de calidad específica de la confiabilidad de estudios de prueba diagnósticas; la que puede ser utilizada también en RS de prueba diagnósticas. Metodológicamente fue desarrollada aplicando estrategias cualitativas con un panel de expertos y reuniones a través de videoconferencia. Diseñaron una lista inicial de 42 de ítems que obtuvieron de la literatura, la que incluyó la valoración de herramientas relacionadas y pre-

viamente publicadas como STARD y QUADAS. La lista de chequeo así desarrollada quedó compuesta por 7 dominios (espectro de los sujetos en estudio, espectro de los examinadores, enmascaramiento del examinador, intervalo de tiempo entre las mediciones, aplicación e interpretación de la prueba, orden del examen y análisis estadístico de los datos). Esta lista se aplica en base a preguntas que tienen 3 alternativas de respuesta “sí” (buena calidad del estudio), “no” (mala calidad), “no claro”; además, algunos artículos incluyen la opción “no aplicable”. Es importante señalar que los ítems clasificados como “no” pueden invalidar el estudio y poner en peligro los resultados de éste independiente del número de ítems calificados como “sí”.

4. La propuesta GRRAS (Guidelines for Reporting Reliability and Agreement Studies), se generó por iniciativa de un grupo de académicos e investigadores de las universidades de Charité-Universitätsmedizin, Herlev y otros centros (Kottner *et al.*, 2011). El objetivo fue desarrollar una herramienta que permitiera cubrir la información referente a confiabilidad y estudios de acuerdo en mediciones en una amplia gama de disciplinas, especialmente en asistencia sanitaria; y el objetivo específico fue establecer los elementos que deben ser abordados cuando se reportan este tipo de estudios. Metodológicamente se desarrolló a partir del trabajo de un grupo internacional conformado por 8 investigadores con experiencia en el desarrollo de instrumentos y la evaluación de confiabilidad y grado de acuerdo de observaciones, o en RS de estudios de confiabilidad; quienes elaboraron las directrices para la presentación de resultados en este tipo de investigaciones. La lista de chequeo quedó compuesta por 6 dominios (título y resumen, introducción, métodos, resultados, discusión y material auxiliar), que incluyen 15 ítems. No se contempla asignación de puntaje a los dominios ni ítems.

5. La escala MINCIR-diagnóstico; (Metodología e Investigación en CIRugía); es una escala que fue creada a partir de un proyecto de investigación financiado por la Universidad de La Frontera (Burgos *et al.*, 2011). El objetivo fue diseñar y construir una escala válida y confiable para medir CM de estudios referentes a estudios de pruebas diagnósticas. Metodológicamente fue desarrollado por un panel de expertos. Se determinó validez de contenido de criterio y de constructo; y confiabilidad interobservador. Está compuesta por 8 ítems agrupados en 3 dimensiones (diseño del estudio, tamaño de la muestra estudiada y metodología utilizada), cuya sumatoria genera una score entre 8 y 45 puntos; siendo el punto de corte 20 puntos el que define el constructo “CM para estudios de pruebas diagnósticas”.

**Escenarios de pronóstico.** Para escenarios de pronóstico, se recuperaron 3 documentos, dos listas de verificación y una escala de CM para estudios de pronóstico.

1. La recomendación REMARK (REporting recommendations for tumor MARKer prognostic studies), fue creada en 2003 por 9 miembros de un comité de investigadores pertenecientes al US National Cancer Institute (Bethesda), Cancer Research UK Medical Statistics Group y otros centros; y financiado por OSI Pharmaceuticals, Inc., Boulder (McShane *et al.*, 2005). Tiene por objetivo generar un conjunto de directrices relativas a las recomendaciones para la publicación de estudios sobre marcadores tumorales para valorar modelos pronósticos relacionados a marcadores tumorales. La metodología utilizada en la construcción del instrumento fue en primera instancia constituir un grupo heterogéneo de clínicos, especialistas en estadística y científicos de laboratorio; para luego realizar una revisión bibliográfica sensible sobre la evidencia empírica relativa a las deficiencias en la publicación de este tipo de materias y los métodos de análisis utilizados. Revisaron además otras directrices previamente generadas (CONSORT, QUOROM y STARD), posterior a lo cual tres miembros del equipo escribieron un borrador inicial, el que remitieron al resto de los integrantes. Realizaron sesiones por videoconferencia y comentarios por correo electrónico. El instrumento final quedó finalmente compuesto por 4 dominios (introducción, método, resultados y discusión) y 20 ítems con formato de respuesta abierta “cumple o no cumple”. Contempla una puntuación al aplicar el instrumento, cuyo máximo es de 20 puntos. A pesar que REMARK fue elaborado para estudios de pronóstico sobre marcadores tumorales, la mayor parte de las recomendaciones pueden ser utilizadas para estudios de pronóstico en general.

2. La escala MINCIR-pronóstico; (Metodología e Investigación en CIRugía); fue creada a partir de un proyecto de investigación financiado por la Universidad de La Frontera (Losada *et al.*, 2005). El objetivo fue diseñar y construir una escala válida y confiable para medir CM de estudios referentes a pronóstico. Metodológicamente fue desarrollado por un panel de expertos. Se determinó validez de contenido de criterio y de constructo; y confiabilidad interobservador. Está compuesta por 25 ítems agrupados en 6 dimensiones (variable de desenlace, variable de exposición, seguimiento, análisis y conclusiones, diseño y variables de confusión; y tamaño de la muestra), cuya sumatoria genera una score entre 6 y 36 puntos, sin definir el constructo “CM para estudios de pronóstico” a partir de algún punto de corte.

3. La declaración GRIPS (Genetic RIsK Prediction Studies), referente al fortalecimiento del reporte de los estudios de predicción de riesgo genético fue publicada en 2011 y en diferentes revistas (Janssens *et al.*, 2011). La importancia de esta declaración es que existe un número creciente de marcadores genéticos que podrían ser usados en futuros es-

tudios de predicción de riesgo de patologías y no existían guías de cómo se debe reportar la investigación que se genera en esta área. Su objetivo fue mejorar la transparencia y calidad del reporte de estudios de predicción de riesgo genético. Para realizar esta declaración se constituyó un grupo de trabajo multidisciplinario (investigadores, epidemiólogos, genetistas, metodólogos, estadísticos y editores de revistas), quienes discutieron en base a un borrador, elaborado a partir de las guías de reporte ya existentes (STREGA, REMARK, STARD) y se enfatizó en los tópicos pertinentes a los estudios de predicción de riesgo. El resultado fue una lista de chequeo compuesta por 25 ítems, los que se agrupan en 6 dominios (título y resumen, introducción, métodos, resultados, discusión y otros). Para cada uno de estos ítems se describe el tipo de información, así como el contenido mínimo que se debe reportar.

#### Escenarios de evaluaciones económicas

Para escenarios de evaluaciones económicas, se recuperaron 4 documentos.

1. Recomendaciones de la NHS-HTA (National Health Service y Health Technology Assessment), fue desarrollada por investigadores de la University of York y financiada por la NHS Research and Development HTA Programme en Reino Unido (Drummond *et al.*, 2005). Su objetivo fue desarrollar recomendaciones para incrementar la generalización de evaluaciones económicas, las que representan un componente importante de las HTAs. La metodología utilizada fue recolectar la información existente en base a la cual se propuso un listado de recomendaciones. Estas recomendaciones se componen de dos iniciativas: Las recomendaciones para el reporte de resultados de evaluaciones económicas de EC, compuesta por 8 ítems que representan diferentes elementos de un estudio (centros donde se desarrolla el estudio, enrolamiento de pacientes, alternativas de tratamiento, reporte de costes y beneficios de cada perspectiva en evaluación, fuentes utilizadas y sus costos, valores de preferencia del estado sanitario, análisis de variabilidad y otros aspectos analíticos); una lista de verificación para la evaluación de la generalización de los estudios basados en EC (compuesta por 10 preguntas); y otra lista de chequeo para la evaluación de la generalización de los estudios de modelación (compuesta por 7 preguntas). Ninguno de cuyos instrumentos considera un sistema de puntuación.

2. La declaración CHEERS (The Consolidate Health Economic Evaluation Reporting Standards), es un reporte del grupo ISPOR (International Society for Pharmacoeconomics and Outcomes Research) (Husereau *et al.*, 2013). Fue publicado en marzo de 2013 y está orientado a brindar herramientas amigables a los autores, editores y revisores de revistas, para facilitar el reporte de las evalua-

ciones económicas de las intervenciones en salud. El objetivo fue desarrollar recomendaciones y una lista de chequeo para facilitar un mejor reporte de las publicaciones de evaluaciones económicas. La metodología consistió en extraer una lista de ítems requerida para publicaciones económicas a partir de una guía previamente desarrollada. Se empleó un panel de expertos a través de una estrategia Delphi modificada en dos rondas para identificar un número mínimo de ítems importantes a reportar. De 42 ítems posibles candidatos, se seleccionaron 24, los que se agruparon en 6 dominios principales (título y resumen, introducción, métodos, resultados, discusión y otros) y se estructuró en una lista de chequeo acompañada de sus recomendaciones para el reporte. No se consigna asignación de puntaje.

3. El reporte ISPOR RCT-CEA, fue desarrollada por la Sociedad Internacional para la Farmacoeconomía y la Investigación de Resultados (ISPOR) (Ramsey *et al.*, 2005); debido a que un número creciente de estudios clínicos prospectivos incluyen en sus resultados análisis de costos o son terminados por razones de tipo económica. Dado que existe una gran variación en la metodología y la presentación de resultados de estos estudios; el objetivo de este documento es servir de guía de orientación para el diseño, la realización y presentación de informes de análisis de costo-efectividad dentro de los EC. El grupo de trabajo fue constituido por 2 copresidentes que fueron seleccionados por el Consejo de Administración de ISPOR. Los copresidentes invitaron a participar a los miembros del panel. A su vez los miembros del panel incluyeron a representantes académicos, de la industria farmacéutica y del seguro de salud. Se elaboró un resumen y un informe de proyecto que se presentaron en el 2004 en la reunión Internacional y Reunión europea del ISPOR, respectivamente. El manuscrito posteriormente se presentó a un grupo de referencia para la revisión y comentario. Esta pauta cuenta con 5 dominios (diseño del ensayo, elementos de información, diseño y tratamiento de la base de datos, análisis y reporte de resultados), que agrupan 26 ítems. No existe una escala de valoración numérica para esta pauta.

4. El reporte NICE-STA (National Institute for Health and Clinical Excellence, Single Technology Appraisal), fue creado por integrantes del RTI-HS (Research Triangle Institute Health Solutions), Manchester, Reino Unido (Zimovetz *et al.*, 2009). El objetivo fue entregar una lista de chequeo que evaluara la calidad de los informes económicos de salud, en especial los modelos de análisis de decisión que STA publica y además a quien NICE refiere, incorporando elementos de gran prestigio para la evaluación económica. Para su realización se realizó una RS de las actuales directrices de buenas prácticas y las listas de chequeo para la evaluación crítica de los estudios de economía de salud, usando la estrate-

gia de búsqueda de Philips *et al.* (2004). Los resultados fueron 14 listas de chequeo, que fueron analizadas críticamente. Para finalmente crear una lista de verificación con dominios similares a la lista de chequeo de Philips *et al.* (2006); pero incorporando modificaciones en los dominios de criterio para proveer mayor sensibilidad para los requisitos avanzados en el caso de las referencias de NICE. Además incorpora un dominio extra “relevancia”, en donde examina la relevancia del estudio de tasación del STA. Está compuesta por 46 ítems, divididos en 7 dominios (pertinencia con respecto a la tecnología actual de evaluación, estructura, evidencia clínica, utilidad de los datos, uso de recursos y datos de los costos, evaluación de la incertidumbre y consistencia); con 4 opciones de respuesta (sí, no, no aparece y no claro) y comentarios. Esto permite hacer un examen visual rápido de las respuestas de la lista de chequeo que indiquen la calidad de los estudios.

#### Misceláneos

Para otro tipo de escenarios (misceláneos), se recuperaron 7 documentos.

1. La declaración STROBE (STrengthening the Reporting of OBservational studies in Epidemiology), se creó en 2004 por iniciativa de académicos de las universidades de Bern, Oxford y otros centros (von Elm *et al.*, 2007). El objetivo fue desarrollar una lista de chequeo para el reporte de resultados de investigación clínica realizada mediante estudios de cohortes, de casos y controles o de corte transversal. Metodológicamente se trabajó organizando un seminario en la ciudad de Bristol, en el que participaron 23 expertos (editores de revistas [Annals of Internal Medicine, BMJ, Bulletin of the World Health Organization, International Journal of Epidemiology, JAMA, Preventive Medicine y The Lancet], epidemiólogos, especialistas en metodología, estadísticos y clínicos, tanto europeos como norteamericanos. El instrumento generado fue revisado en varias reuniones del grupo y en discusiones electrónicas con los principales participantes. De este modo, se generó una lista de chequeo compuesta por 6 dominios (título y resumen, introducción, metodología, resultados, discusión y otros apartados relevantes), que integran 22 ítems, sin asignación de puntaje. Se proporcionan versiones diferentes (ya sean totales o parciales) para estos puntos según el diseño específico (cohortes, casos y controles y corte transversal).

2. La lista de verificación COREQ (COnsolidated criteria for REporting Qualitative research), se generó por iniciativa de investigadores de la universidad de Sydney y otros centros (Tong *et al.*, 2007). Su objetivo, fue elaborar una lista de verificación para el reporte explícito y exhaustivo de resultados de estudios cualitativos (entrevistas y grupos focales). Metodológicamente, se trabajó de la siguiente for-

ma: Se realizó una búsqueda exhaustiva en las bases de datos Cochrane Central, MEDLINE, CINAHL; en los protocolos Campbell, en RS de estudios cualitativos, en guías para los autores o revisores de las principales revistas médicas para evaluar los estudios cualitativos. De este modo, se compilaron 76 artículos de 22 listas de verificación. Los elementos así identificados se agrupan en tres dominios (equipo de investigación y su reflexión, diseño del estudio y análisis de los datos), eliminando ítems duplicados y ambiguos. La lista de verificación quedó compuesta por 3 dominios y 32 ítems, sin asignación de puntaje.

3. La recomendación GRADE (Grading of Recommendations Assessment, Development and Evaluation), fue desarrollada por un grupo multidisciplinario de investigadores, metodólogos y médicos involucrados en la práctica clínica de las universidades de McMaster, Basel, del Norwegian Knowledge Centre for the Health Services, y de otros centros (Guyatt *et al.*, 2008). Su objetivo fue generar un instrumento que permitiera clasificar la calidad de la evidencia contrapesando beneficios y riesgos de una intervención sanitaria; de modo tal de poder reconocer valores y preferencias de las recomendaciones; y si la intervención en evaluación constituye un uso racional de los recursos. Metodológicamente la recomendación fue generada a través de paneles de discusión, jerarquizando los aspectos relevantes e integrando finalmente las recomendaciones principales. Conjugó la calidad de las pruebas, evaluando si los efectos deseables superan claramente los efectos indeseables, y además indaga sobre si existe un equilibrio cercano o incierto; permitiendo finalmente clasificar las recomendaciones en 4 niveles de calidad de la evidencia (alta calidad, moderada, baja y muy baja); y la fuerza de la recomendación en 4 (fuerte a favor de la intervención, débil a favor de la intervención, débil en contra de la intervención y fuerte en contra de la intervención). La ventaja del sistema GRADE es que puede ser utilizado tanto por pacientes, clínicos y autoridades sanitarias responsables de la toma de decisiones en salud.

4. La guía SQUIRE (Standards for Quality Improvement Reporting Excellence), nace en 1999 y por iniciativa de los editores de la revista *Quality in Health Care*; con el objetivo de mejorar los reportes de información científica biomédica (Ogrinc *et al.*, 2008). Estas normas proporcionan un marco explícito para compartir el conocimiento adquirido examinando aquellas intervenciones de manera estrecha, cuidadosa y detallada. La metodología utilizada parte con un panel de expertos, luego la publicaron para discusión pública y posteriormente fue analizada en una conferencia de consenso, obteniendo la guía definitiva en 2007. Está compuesta por 19 ítems que se agrupan en 6 dominios (título y resumen, introducción, método, resultados, discusión y otra información), sin asignación de puntaje.

5. La declaración STREGA (Strengthening the Reporting of Genetic Association studies), es una extensión de la declaración STROBE. Comienza en junio de 2006, en Ottawa, Ontario; a partir de una iniciativa de académicos e investigadores del Canada Research Chair in Human Genome Epidemiology, MRC Biostatistics Unit of Cambridge y otros centros (Little *et al.*, 2009). Su objetivo fue complementar y fortalecer la declaración STROBE proporcionando 12 ítems adicionales a los 22 ítems de la lista STROBE; con la idea de aportar ítems propios de estudios de asociación genética. Metodológicamente, se trabajó con un grupo multidisciplinario integrado por 33 expertos (epidemiólogos, genetistas, estadísticos, editores de revistas y estudiantes de posgrado), el que desarrolló la propuesta inicial mediante revisión de la literatura, posterior a lo cual se desarrollaron talleres y discusión electrónica. El instrumento así generado no es otra cosa que una columna adicional que se agrega a la lista STROBE, que se denomina “extensión para estudios de asociación genética”. En esta columna se adicionaron 20 ítems relacionados con la población, el genotipado, el modelo del haplotipo, fundamentos de para la selección de genes, etc., sin establecer ni sugerir cómo diseñar un estudio de asociación genética, sino que tratando de mejorar la transparencia de sus informes. No se consigna asignación de puntaje.

La declaración ORION (Guidelines for transparent reporting of Outbreak Reports and Intervention studies Of Nosocomial infection), es una iniciativa de la Infectious Disease Research Network (IDRN) (Stone *et al.*, 2007); cuyo objetivo es elevar el nivel de la investigación y la publicación en epidemiología hospitalaria relacionada con infecciones nosocomiales; para así facilitar la síntesis de la evidencia y promoción de la transparencia de la información; de tal modo que, los lectores asocien los estudios con su propia experiencia y puedan evaluar el grado en que los resultados pueden generalizarse en otros entornos. Fue desarrollada siguiendo el espíritu de las declaraciones CONSORT y TREND, mediante consensos del grupo de trabajo y ulteriores sesiones múltiples vía correo electrónico y teléfono; posterior a las cuales, la propuesta fue consultada con sociedades científicas, editores de revistas y diversos investigadores. ORION, fue diseñada especialmente para valorar estudios cuasi-experimentales comunes en epidemiología hospitalaria: series de tiempo interrumpido con y sin grupos de control; e informes de brotes epidémicos. Consiste en una lista de chequeo compuesta por 22 ítems, agrupados en 5 dominios (título/resumen, introducción, métodos, resultados y discusión), y una tabla resumen; y no se consigna asignación de puntajes.

6. La iniciativa MINCIR-EOD (Metodología e Investigación en CIRugía), para el reporte de estudios observacionales



descriptivos; fue creada a partir de un proyecto de investigación financiado por la Universidad de La Frontera (Manterola & Astudillo, 2013). El objetivo fue diseñar y construir un sistema de chequeo para el reporte de resultados con estudios observacionales descriptivos. La metodología utilizada fue cuantitativa con estrategias cualitativas para la generación de ítems y construcción del instrumento. En la primera etapa se diseñó una propuesta, mediante la recopilación de ítems y dominios a partir de una extensa revisión de la literatura. En la segunda, se construyó el instrumento, aplicando reducción de ítems y dominios a través de un panel de expertos compuesto por epidemiólogos clínicos, estadísticos, académicos clínicos, miembros de paneles de revisores de revistas biomédicas y editores de revistas biomédicas. A estos, se les consultó a través de un cuestionario semiestructurado y autoadministrado, compuesto por 18 ítems provenientes de la literatura; a los que los expertos podían agregar otros ítems según su experiencia o

particular visión de la situación. De este modo, se construyó un instrumento que quedó compuesto por 19 ítems (problema en estudio, objetivos, diseño del estudio, lugar donde se desarrolló la investigación, características de los participantes, muestreo, variables, seguimiento, estadísticas, ética, análisis de grupos y subgrupos, novedad de la propuesta, comentarios de los resultados, limitaciones del estudio y conclusiones), agrupados en 4 dominios: Introducción, metodología, resultados y discusión. No hay asignación de puntaje.

## DISCUSIÓN

La publicación de los documentos antes mencionados y descritos (Tabla I), demuestra la creciente tendencia al reconocimiento, de que la valoración exitosa de las inter-

Tabla I. Resumen de la información recopilada.

Iniciativas	Tipos de escenarios				
	Terapia o procedimientos	Diagnóstico	Pronóstico	Evaluaciones económicas	Misceláneos
CONSORT	X				
QUOROM	X				
MOOSE	X				
STRICTA	X				
TREND	X				
MINCIR-T	X				
RedHot	X				
REHBaR	X				
PRISMA	X				
REFLECT	X				
Ottawa	X				
SPIRIT	X				
STARD		X			
QUADAS		X			
QAREL		X			
GRRAS		X			
MINCIR-D		X			
REMARK			X		
MINCIR-P			X		
GRIPS			X		
NHS-HTA				X	
CHEERS				X	
ISPOR				X	
NICE-STA				X	
STROBE					X
COREQ					X
GRADE					X
SQUIRE					X
STREGA					X
ORION					X
MINCIR					X

venciones en salud, implica necesariamente la planificación concienzuda de los estudios basados en los distintos tipos de diseños de investigación existentes (Kirkwood, 2004).

Por otra parte, son evidencia concreta de la corriente orientada al reporte adecuado de resultados, utilizando para ello estas denominadas listas de chequeo, de verificación o de comprobación; que son instrumentos que contienen criterios o indicadores a partir de los cuales se evalúan ciertas características de un manuscrito. Constituyen una forma objetiva de valorar el cumplimiento de atributos mínimos establecidos; lo que se logra en base al carácter cerrado de las respuestas y al limitado número de ítems que se evalúan. Estas listas de comprobación son además útiles para servir de guía y recordar los puntos que deben ser inspeccionados.

No obstante ello, es importante señalar que las listas de verificación no fueron diseñadas para valorar CM de los artículos, la que se evalúa con escalas ad-hoc como algunas de las que fueron mencionadas y descritas.

La CM de los artículos, se ha de entender como un constructo multidimensional, en el que es posible evaluar múltiples aspectos de un artículo (una de las cuales podría ser la calidad del reporte, la que mejora con la utilización de listas de chequeo), de las que por el momento no existe un consenso (Manterola *et al.*, 2006); razón por la que parece relevante que al considerar un artículo, se apliquen diferentes herramientas que permitan valorarlo de forma global, de manera tal de obtener la mayor cantidad de datos referentes a su validez interna y externa.

---

MANTEROLA, C.; OTZEN, T.; LORENZINI, N.; DÍAZ, A.; TORRES-QUEVEDO, R. & CLAROS, N. Initiatives for reporting biomedical research results with different types of designs. *Int. J. Morphol.*, 31(3):945-956, 2013.

**SUMMARY:** Quality of results reporting is not perfect, many initiatives tending to improve this aspect of clinical research have been developed in the last decade. The aim of this manuscript is to mention and describe the existent initiatives for reporting biomedical research results in different scenarios and special situations. To obtain check-lists, a search in THE COCHRANE LIBRARY, MEDLINE, SciELO y Redalyc; Clinical Evidence, TRIP database, Fistera, Rafabravo, EQUATOR Network, BIREME and HINARI Program was developed. Identified documents were grouped in relation with clinical research scenarios (therapy, diagnosis, prognosis and economic evaluations) and miscellaneous. The search allows finding 31 documents. Twelve for therapy (CONSORT, QUOROM, MOOSE, STRICTA, TREND, MINCIR-Therapy, RedHot, REHBaR, PRISMA, REFLECT, Ottawa and SPIRIT), 5 for diagnosis (STARD, QUADAS, QAREL, GRRAS and MINCIR-Diagnosis), 3 for prognosis (REMARK, MINCIR-Prognosis and GRIPS), 4 for economic evaluations (NHS-HTA, CHEERS, ISPOR RCT-CEA and NICE-STA.) and 7 miscellaneous (STROBE, COREQ, GRADE, SQUIRE, STREGA, ORION and MINCIR-EOD). Different initiatives and statements were found. These must be noted and used by writers, reviewers and editors of biomedical journals, in order to improve the quality of reporting results.

**KEY WORDS:** Research Report; Research Design; Therapeutics; Diagnosis; Prognosis; Economic evaluations.

---

## REFERENCIAS BIBLIOGRÁFICAS

- Begg, C.; Cho, M.; Eastwood, S.; Horton, R.; Moher, D.; Olkin, I.; Pitkin, R.; Rennie, D.; Schulz, K. F.; Simel, D. & Stroup, D. F. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA*, 276(8):637-9, 1996.
- Bossuyt, P. M.; Reitsma, J. B.; Bruns, D. E.; Gatsonis, C. A.; Glasziou, P. P.; Irwig, L. M.; Lijmer, J. G.; Moher, D.; Rennie, D.; de Vet, H. C. & Standards for Reporting of Diagnostic Accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann. Intern. Med.*, 138(1):40-4, 2003.
- Burgos, M. E.; Manterola, C. & Sanhueza, A. Diseño de una escala para evaluar calidad metodológica de estudios de pruebas diagnósticas. Estudio piloto. *Rev. Chil. Cir.*, 63(5):493-7, 2011.
- Campbell, M. K.; Elbourne, D. R.; Altman, D. G. & CONSORT group. CONSORT statement: extension to cluster randomized trials. *BMJ*, 328(7441):702-8, 2004.
- Chan, A. W.; Tetzlaff, J. M.; Altman, D. G.; Laupacis, A.; Gotzsche, P. C.; Krleza-Jeric, K.; Hróbjartsson, A.; Mann, H.; Dickersin, K.; Berlin, J. A.; *et al.* SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann. Intern. Med.*, 158(3):200-7, 2013.
- Dean, M. E.; Coulter, M. K.; Fisher, P.; Jobst, K. A. & Walach, H. Reporting data on homeopathic treatments (RedHot): a supplement to CONSORT. *J. Altern. Complement. Med.*, 13(1):19-23, 2007.
- Des Jarlais, D. C.; Lyles, C.; Crepaz, N. & TREND Group. Improving the Reporting Quality of Nonrandomized Evaluations of Behavioral and Public Health Interventions: The TREND Statement. *Am. J. Public Health.*, 94(3):361-6, 2004.
- Drummond, M.; Manca, A. & Sculpher, M. Increasing the generalizability of economic evaluations: recommendations for

- the design, analysis, and reporting of studies. *Int. J. Technol. Assess. Health Care*, 21(2):165-71, 2005.
- Guyatt, G. H.; Oxman, A. D.; Vist, G. E.; Kunz, R.; Falck-Ytter, Y.; Alonso-Coello, P.; Schünemann, H. J. & GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336(7650):924-6, 2008.
- Husereau, D.; Drummond, M.; Petrou, S.; Carswell, C.; Moher, D.; Greenberg, D.; Augustovski, F.; Briggs, A. H.; Mauskopf, J.; Loder, E. & CHEERS Task Force. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMC Med.*, 11:80, 2013.
- Janssens, A. C.; Ioannidis, J. P.; van Duijn, C. M.; Little, J.; Khoury, M. J. & GRIPS Group. Strengthening the reporting of Genetic Risk Prediction Studies: the GRIPS Statement. *J. Clin. Epidemiol.*, 64(8):843-7, 2011.
- Kirkwood, B. Making public health interventions more evidence based. *BMJ*, 328(7446):966-7, 2004.
- Kottner, J.; Audige, L.; Brorson, S.; Donner, A.; Gajewski, B. J.; Hróbjartsson, A.; Roberts, C.; Shoukri, M. & Streiner, D. L. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J. Clin. Epidemiol.*, 64(1):96-106, 2011.
- Little, J.; Higgins, J. P.; Ioannidis, J. P.; Moher, D.; Gagnon, F.; von Elm, E.; Khoury, M. J.; Cohen, B.; Davey-Smith, G.; Grimshaw, J. *et al.* STrengthening the REporting of Genetic Association Studies (STREGA): an extension of the STROBE statement. *PLoS Med.*, 6(2):e22, 2009.
- Losada, H.; Manterola, C.; Pineda, V.; Vial, M.; & Sanhueza, A. Grupo MINCIR. Diseño de una escala para la evaluación de calidad metodológica de estudios de pronóstico. *Rev. Chil. Cir.*, 61(1):59-72, 2009.
- Lucas, N. P.; Macaskill, P.; Irwig, L. & Bogduk N. The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). *J. Clin. Epidemiol.*, 63(8):854-61, 2010.
- MacPherson, H.; White, A.; Cummings, M.; Jobst, K.; Rose, K. & Niemtzw, R. Standards for reporting intervention in controlled trials of acupuncture: the STRICTA recommendations. *Complement. Ther. Med.*, 9(4):246-9, 2001.
- Manterola, C.; Vial, M.; Pineda, V. & Losada H. Revisión sistemática de la literatura. Propuesta metodológica para su realización. *Rev. Chil. Cir.*, 55(2):204-8, 2003.
- Manterola, C.; Pineda, V.; Vial, M.; Losada, H. & MINCIR Group. Which is the methodological quality of human therapy studies in ISI surgical publications? *Ann. Surg.*, 244(5):827-32, 2006.
- Manterola, C.; Vial, M.; Pineda, V. & Sanhueza, A. Systematic Review of Literature with Different Types of Designs. *Int. J. Morphol.*, 27(4):1179-86, 2009.
- Manterola, C. & Astudillo, P. Checklist for reporting of descriptive observational studies. MINCIR initiative. *Int. J. Morphol.*, 31(1):115-20, 2013.
- McShane, L. M.; Altman, D. G.; Sauerbrei, W.; Taube, S. E.; Gion, M. & Clark, G. M; Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics. REporting recommendations for tumour MARKer prognostic studies (REMARK). *Br. J. Cancer*, 93(4):387-91, 2005.
- Moher, D.; Cook, D. J.; Eastwood, S.; Olkin, I.; Rennie, D. & Stroup, D. F. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet*, 354(9193):1896-900, 1999.
- Moher, D.; Schulz, K. F.; Altman, D. G. & CONSORT GROUP (Consolidated Standards of Reporting Trials). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann. Intern. Med.*, 134(8):657-62, 2001.
- Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D. G & PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Ann. Intern. Med.*, 151(4):264-9, W64, 2009.
- Ogrinc, G.; Mooney, S. E.; Estrada, C.; Foster, T.; Goldmann, D.; Hall, L. W.; Huizinga, M. M.; Liu, S. K.; Mills, P.; Neily, J.; *et al.* The SQUIRE (Standards for Quality Improvement Reporting Excellence) guidelines for quality improvement reporting: explanation and elaboration. *Qual. Saf. Health Care*, 17(Suppl. 1):i13-i32, 2008.
- Philips, Z.; Ginnelly, L.; Sculpher, M.; Claxton, K.; Golder, S.; Riemsma, R.; Woolacoot, N. & Glanville, J. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol. Assess.*, 8(36):iii-iv, ix-xi, 1-158, 2004.
- Philips, Z.; Bojke, L.; Sculpher, M.; Claxton, K. & Golder, S. Good practice guidelines for decision analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics*, 24(4):355-71, 2006.
- Ramsey, S.; Willke, R.; Briggs, A.; Brown, R.; Buxton, M.; Chawla, A.; Cook, J.; Glick, H.; Liljas, B.; Petitti, D. & Reed, S. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. *Value Health*, 8(5):521-33, 2005.
- Sargeant, J. M.; O'Connor, A. M.; Gardner, I. A.; Dickson, J. S.; Torrence, M. E. & Consensus Meeting Participants. The

REFLECT Statement: Reporting Guidelines for Randomized Controlled Trials in Livestock and Food Safety: Explanation and Elaboration. *Zoonoses Public Health*, 57(2):105-36, 2010.

- Stock-Schröer, B.; Albrecht, H.; Betti, L.; Endler, P. C.; Linde, K.; Lüdtke, R.; Musial, F.; van Wijk, R.; Witt, C. & Baumgartner, S. Reporting experiments in homeopathic basic research (REHBaR) - a detailed guideline for authors. *Homeopathy*, 98(4):287-98, 2009.
- Stone, S. P.; Cooper, B. S.; Kibbler, C. C.; Cookson, B. D.; Roberts, J. A.; Medley, G. F.; Duckworth, G.; Lai, R.; Ebrahim, S.; Brown, E. M.; Wiffen, P. J. & Davey, P. G. The ORION STATEMENT Guidelines for Transparent Reporting of Outbreak Reports and Intervention studies of Nosocomial infection. *J. Antimicrob. Chemother.*, 59(5):833-40, 2007.
- Stroup, D. F.; Berlin, J. A.; Morton, S. C.; Olkin, I.; Williamson, G. D.; Rennie, D.; Moher, D.; Becker, B. J.; Sipe, T. A. & Thacker, S. B. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. *JAMA*, 283(15):2008-12, 2000.
- Tong, A.; Sainsbury, P. & Craig, J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int. J. Qual. Health Care*, 19(6):349-57, 2007.
- von Elm, E.; Altman, D. G.; Egger, M.; Pocock, S. J.; Gøtzsche, P. C. & Vandenbroucke J. P. & STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann. Intern. Med.*, 147(8):573-7, 2007.
- Weijer, C.; Grimshaw, J. M.; Eccles, M. P.; McRae, A. D.; White, A.; Brehaut, J. C.; Taljaard, M. & Ottawa Ethics of Cluster Randomized Trials Consensus Group. The Ottawa Statement on the Ethical Design and Conduct of Cluster Randomized Trials. *PLoS Med.*, 9(11):e1001346, 2012.
- Whiting, P.; Rutjes, A. W. S.; Reitsma, J. B.; Bossuyt, P. M. & Kleijnen, J. The development of QUADAS: a tool for the quality assessment of Studies of diagnostic accuracy included in systematic reviews. *BMC Med. Res. Methodol.*, 3:25, 2003.
- Whiting, P. F.; Rutjes, A. W.; Westwood, M. E.; Mallett, S.; Deeks, J. J.; Reitsma J. B.; Leeflang, M. M.; Sterne, J. A.; Bossuyt, P. M. & QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy Studies. *Ann. Intern. Med.*, 155(8):529-36, 2011.
- Zimovetz, E. & Wolowacz, S. Reviewer's Checklist for Assessing the Quality of Decision Models. RTI Health Solutions, Manchester, United Kingdom. 2009. Disponible en: <http://www.rtihs.org/request/index.cfm?fuseaction=display&PID=13979>

Dirección para Correspondencia:  
Dr. Carlos Manterola  
Departamento de Cirugía  
Universidad de La Frontera.  
Casilla 54-D, Temuco  
CHILE

Teléfono: 56-45-2325760  
Fax: 56-45-2325761

Email: [carlos.manterola@ufrontera.cl](mailto:carlos.manterola@ufrontera.cl)

Recibido: 14-05-2013  
Aceptado: 14-06-2013