Conducting and Reporting Systematic Reviews of Adverse Events

by

Liliane Medianeira Zorzela

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Medical Sciences - Pediatrics
University of Alberta

Abstract

Introduction

Systematic reviews (SRs) synthesize published and sometimes unpublished data and are often based on randomized controlled trials (RCTs). However, RCTs are known to be poor at identifying and reporting harms. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) Statement was published in 2009 to offer guidance on the minimum reporting standards when publishing a SR. Thus far, PRISMA has mainly focused on efficacy, but there is a need for evidence on both efficacy and harms of interventions.

Propofol is an anesthetic intervention used for pediatric sedation, but there have been several case reports of 'propofol infusion syndrome' (PRIS), a poorly understood syndrome often leading to death. In several countries, regulatory agencies have contraindicated the use of propofol infusion in pediatric intensive care units. However, propofol is still used despite the liability concerns.

The overall goals of this thesis were to improve methods of conducting and reporting systematic reviews of adverse events. More specifically, (i) to develop an extension for the PRISMA Statement, for SRs addressing adverse events (AEs): the PRISMA Harms; (ii) to identify if propofol is associated with serious AEs in children and measure if the inclusion of non-randomized studies in a SR of AEs provides further information than data from RCTs alone.

Methods

There were 2 distinct methods used in this doctoral thesis. The first was to develop the PRISMA Harms guideline. We followed the recommended steps for guideline development: 1) to document if there is need for the development of a guideline; 2) to employ a Delphi process to identify relevant items to be included in the reporting guideline; 3) to have an in-person consensus building meeting; and 4) to write the guideline.

The second was to identify if propofol infusion is associated with serious AEs in pediatric patients and to measure if the inclusion of non-randomized studies provides more relevant data than clinical trials alone, we conducted a SR of propofol infusion in pediatric patients including both clinical trials and observational studies.

Results

For the PRISMA Harms development, the first step identified 309 reviews of AEs and documented weaknesses in reporting and the need for a guideline. The second step conducted three Delphi rounds sent to 352 participants, 166 responses were received. The in-person meeting had 25 participants and the final PRISMA Harms manuscript was developed after multiple revisions containing 4 mandatory items and 14 recommended items for reviews addressing harms.

The propofol SR identified 91 serious AEs (PRIS or cardiac arrest) associated with propofol infusion, 21 identified in a single unpublished RCT and all the other serious AEs emerged from non-randomized studies. In the included studies, a total of 5633 children received propofol for

more than 60 minutes and did not have any serious events (i.e., PRIS or cardiac arrest) associated with it.

Conclusion

Through this work we developed PRISMA Harms, an international reporting guideline to improve harms reporting in SRs.

Further, we documented serious AEs associated with propofol infusion in children and the relevance of including non-randomized and unpublished studies in SRs of AEs, providing both clinical and methodological significant information.

Preface

This thesis is an original work by Liliane Medianeira Zorzela.

Some of the research conducted for this thesis was informed by an international research group, The PRISMA Harms Group: with David Moher at Ottawa Health Research Institute (OHRI); Doug Altman at University of Oxford; Jan Vandenbroucke at Leiden University; John P. Ioannidis at Stanford University; Lina Santaguida at McMaster University; Yoon Loke at University of East Anglia; Su Golder at University of York; and Sunita Vohra at University of Alberta as the group lead. As part of this thesis work, L. Zorzela developed an extension for the PRISMA statement, the PRISMA Harms. L. Zorzela collected the data, analysed the results and wrote the review basis for the development of PRISMA Harms (chapter 2, published at BMJ 2014;348:f7668 doi: 10.1136/bmj.f7668); developed, employed and analysed the results for the Delphi process and wrote the manuscript (chapter 3); organized and participated in the face to face meeting and composed the manuscript "PRISMA Harms" (chapter 4). All the PRISMA Harms related chapters were revised by the PRISMA Harms Group and by L. Zorzela's thesis committee group.

The development of the PRISMA Harms (chapters 2, 3 and 4) which is part of this thesis, received research ethics approval from the University of Alberta Research Ethics Board, Project name "PRISMA Harms Extension", Pro 00021294, May 27, 2011. The PRISMA Harms received partial funding support from "Alberta Innovates: Health Solutions".

L. Zorzela developed the protocol for the review in chapter 5 (appendix 1), published at the Cochrane Database of Systematic Reviews Volume (4), 2012, under title "Propofol infusion for paediatric sedation". L. Zorzela was the lead person for data screening, data collection, data analysis and manuscript elaboration for the final Cochrane review (chapter 5). The protocol and final review were revised by L. Zorzela's thesis supervisor, thesis steering committee group and all the co-authors.

Dedication

I would like to dedicate this work to my new family, my beloved daughter Isabel and my life partner, Robert.

Acknowledgments

First of all I would like to thank my parents, Ivo and Laureci, without them this work would not be possible.

I would like to thank my thesis steering committee members, Lisa Hartling, Ari Joffe and Yoon Loke for their expertise, guidance and treasured support during all these years.

I would like to thank the PRISMA Harms Group, Sunita Vohra, David Moher, Doug Altman, John P. Ioannidis, Lina Santaguida, Yoon Loke, Jan Vandenbroucke and Su Golder for their expertise and guidance on the development of the PRISMA Harms. I also would like to thank all the 25 PRISMA Harms meeting participants that came to Banff- Alberta to discuss and help in the development of the PRISMA Harms.

I would like to thank Katherine Pohlman, Yali Liu and Salima Punja for their persistence and help during the systematic review's development.

I would like to thank my daughter, Isabel, for her good behavior in utero and during her first months of life, allowing me to complete this work.

Mainly, I would like to thank my supervisor, Dr. Sunita Vohra, for her patience, teaching and support, but foremost for transforming me into a research believer. I learned how research can be exciting and fulfilling when done with ethics and passion.

Table of Contents

List of Tables

List of Figures

Table of Abbreviations

Abbreviations	Meaning
AE(s)	Adverse Event(s)
CONSORT	Consolidated Standards of Reporting Trials
ECMO	Extra-Corporeal Membrane Oxigenation
PRIS	Propofol Infusion Syndrome
PICU	Pediatric Intensive Care Unit
PPV	Positive Predictive Value
NPV	Negative Predictive Value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
SR(s)	Systematic Review (s)
RCT(s)	Randomized Controlled Trial(s)

Glossary of terms

ADVERSE DRUG REACTION[*]	An adverse effect specific to a drug.
ADVERSE EFFECT[*,¥]	An unfavorable outcome that occurs during or after the use of a drug or other intervention but is not necessarily caused by it.
ADVERSE EVENT [*,¥]	An unfavorable outcome that occurs during or after the use of a drug or other intervention and the causal relation between the intervention and the event is at least a reasonable possibility.
COMPLICATION [*]	An adverse event or effect following surgical and other invasive interventions.
HARM [π]	The totality of possible adverse consequences of an intervention or therapy; they are the direct opposite of benefits.
SAFETY [π]	Substantive evidence of an absence of harm. The term is often misused when there is simply absence of evidence of harm.
SIDE EFFECT [*]	Any unintended effect, adverse or beneficial, of a drug that occurs at doses normally used for treatment.
TOXICITY [π]	Describes drug-related harms. The term may be most appropriate for laboratory-determined measurements, although it is also used in relation to clinical events. The disadvantage of the term 'toxicity' is that it implies causality. If authors cannot prove causality, the terms 'abnormal laboratory measurements' or 'laboratory abnormalities' are more appropriate to use.

Adapted from: (π) Ioannidis et al. Better Reporting of Harms in Randomized Trials: An Extension of the CONSORT Statement ; (¥) Chou R, Aronson N, Atkins D, et al. AHRQ Series paper4: Assessing harms when comparing medical interventions: AHRQ and the Effective Health care Program; (*) Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions.

Chapter 1

Thesis Background

Thesis Background

Evidence-based medicine seeks high quality, unbiased evidence to make appropriate decisions to improve patient care. For optimal care, accurate knowledge about both benefits and harms is needed.

Systematic reviews

Systematic reviews of the literature synthesize the published, and sometimes unpublished, data on a given topic and are considered the top of the evidence 'pyramid' (1). Systematic reviews are used by decision-makers at every level – patients, providers, health care organizations and associations, provincial and federal health agencies, industry, insurance companies and the legal system. Recent data indicate that 11 new systematic reviews are published daily (2). Given their use in decision-making, it is critical that systematic review methods and results are reported clearly and transparently (3).

Historically, systematic reviews have focused on efficacy (4), most often based on the best evidence available, e.g., randomized controlled trials (RCTs). RCTs are usually designed to evaluate treatment efficacy/effectiveness and are often done over a short period of time, with a small number of participants and rarely assessing harms as a primary outcome. Overall, RCTs are known to be poor at identifying and reporting harms, which can lead to a misconception that a given intervention is safe, when its safety (see Glossary of terms) is actually unknown (5).

Due to poor reporting of harms data in primary studies, systematic reviews compound this omission by failing to report on harms or do so inadequately. Such reviews can be misleading as they do not represent a balanced risk-to-benefit perspective of a given treatment. The Cochrane Adverse Effects Methods Group emphasizes that when doing a systematic review of harms, the authors should include non-RCT data, as those data are more likely to capture adverse event reports (6). However, the heterogeneity of non-RCT data poses a challenge for reviewers and clear guidance on how to synthesize and report these findings is lacking.

The need to improve knowledge synthesis of harms

Several studies have identified a number of methodological challenges when conducting a systematic review of harms (7-11). These challenges include: the poor quality information on adverse events reported on original studies, difficulties in identifying relevant studies when using standard systematic searches techniques, and the lack of a specific methodological guidance to perform a systematic reviews of adverse events.

Improving reporting in systematic reviews

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) Statement was published in 2009 (12) to offer guidance to review authors and journal editors on the minimum reporting standards when publishing a systematic review or meta-analysis. The PRISMA Statement was developed following three main steps: (i) a systematic review examining the quality of reporting of systematic reviews, to assess the need for a guideline; (ii) an international survey of review authors, consumers and groups commissioning or using systematic reviews, to add or refine an existing reporting guideline - QUOROM (5); (iii) a face-to-face meeting, with 29 participants, including review authors, clinicians, methodologists, journal editors and a consumer to revise and expand QUOROM. After those steps, the meeting participants developed and approved the final document including a 27-item checklist and flow diagram - the PRISMA Statement. PRISMA has been widely endorsed by multiple international organizations, including: the Cochrane Collaboration; the Council of Science Editors; the World Association of Medical Editors; and 176 medical journals (http://www.prisma-statement.org/endorsers.htm). Thus far, PRISMA has focused mainly on efficacy.

Why a PRISMA Harms extension is important

There is a need for high quality evidence when assessing adverse effects of interventions, whether diagnostic or therapeutic, drugs, natural health products, vaccines, biologics or devices (13). Systematic reviews evaluating harms can provide valuable information to help describe adverse events (frequency, nature, seriousness), but they are complicated by lack of standardized methods to report these events at the primary study level and the fact that harms are not usually the primary outcome of included studies (14). As adverse events are often rare, systematic reviews may be the best approach to combine these rare events and present results in a single document.

Clinical context

Propofol is an anesthetic agent frequently used in adult and pediatric sedation. Since 1992, there have been several case reports of 'propofol infusion syndrome' (PRIS), a poorly understood syndrome associated with prolonged propofol infusion in high doses (15-18). In several countries, including Canada and the United States, regulatory agencies have contraindicated the use of propofol infusion in pediatric intensive care. However, propofol is still used as a continuous infusion in pediatric intensive care, pediatric operating rooms and emergency rooms despite the liability concerns and regulatory agency's contraindication (19-20). There is a need to identify the serious adverse events associated with propofol infusion in pediatric patients to support clinicians and regulatory bodies to promote evidence-based decision-making when evaluating the best drug choices for continuous sedation in pediatric patients in any setting.

Goal of this doctoral thesis

The overall goals of this thesis were to assess and if needed, improve methods of conducting and reporting systematic reviews of adverse events.

We set out to achieve these goals through the following specific objectives:

(1) To assess if there is need of a reporting guideline for systematic reviews measuring adverse events.

(2) To develop a reporting guideline for systematic reviews measuring adverse events.

(3) To identify if propofol infusion is associated with serious adverse events in pediatric patients.

(4) To assess if the inclusion of nonrandomized data in a systematic review of adverse events would provide additional information that would be otherwise overlooked.

Thesis Outline

Improving the report of adverse events in systematic reviews

The steps followed as part of this doctoral thesis for the development of a guideline for reporting adverse events in systematic reviews were: first, to document if there was room for improvement on reporting in those reviews; second, to employ a modified Delphi process to identify relevant items to be included in the reporting guideline; and third, a consensus building in-person meeting. The in-person meeting included experts in systematic reviews, journal editors, methodologists and members of the PRISMA and CONSORT group with previous experience in reporting guideline development (21-26). Together, those at the meeting would define the final set of items to be included in the guideline. The final step, also as part of this doctoral work, was the actual development of PRISMA Harms, an extension of the PRISMA Statement, which would guide how harms are reported in systematic reviews.

Improving the estimate of adverse events

To identify if propofol infusion is associated with serious adverse events in pediatric patients, as part of this doctoral thesis, a systematic review of propofol infusion in pediatric patients was conducted. The review included both clinical trials and observational studies. We also assessed if systematic reviews using multiple study designs would have a better estimate of adverse events than systematic reviews including only RCT data, evaluating if randomized and non-randomized data should be included if the outcome of a systematic review was to identify adverse events.

Manuscript based thesis chapters

This thesis is presented as four independent manuscripts. Each manuscript includes an abstract, introduction, methods, results, discussion and conclusion sections, as well as relevant figures, tables, references and appendices. The chapters are divided as follow:

Chapter 2: Quality of reporting in systematic reviews of adverse events: systematic review

This is the review used to identify the need for the development of the PRISMA Harms. It is the first step in the guideline development (27). This manuscript has been published in the British Medical Journal in 2014 (28).

Chapter 3: The development of PRISMA Harms using a modified Delphi technique

This manuscript reports the methods and findings of the modified online Delphi method used to rate the potential PRISMA Harms items. It is the second step for development of the reporting guideline.

Chapter 4: PRISMA Harms

This document is the PRISMA Harms, a reporting guideline with a main goal to improve harms reporting in systematic reviews. It is an extension of the PRISMA Statement.

Chapter 5: Propofol infusion for paediatric sedation

This systematic review includes both randomized and non-randomized data to assess if propofol infusion is associated with serious adverse events in pediatric patients and also to assess if the inclusion of non-randomized data provides additional information that would be otherwise overlooked if only clinical trials data would to be included. The protocol has been published with the Cochrane Anesthesia Review Group (CARG) (29).

References

1. Phillips B, Ball C, Badenoch D,Straus S, Haynes B ,Dawes M. Oxford Centre for Evidence-based Medicine. Levels of Evidence Model. 1998. Updated by Jeremy Howick March 2009. Oxford Centre for Evidence-based Medicine - Levels of Evidence (March 2009) http://www.cebm.net/index.aspx?o=1025
2. Bastian H, Glasziou P, Chalmers I (2010) Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up? PLoS Med 7(9): e1000326. doi:10.1371/journal.pmed.1000326 29(2): 175–181
3. The PLoS Medicine Editors (2011) Best Practice in Systematic Reviews: The Importance of Protocols and Registration. PLoS Med 8(2): e1001009. doi:10.1371/journal.pmed.1001009
4. Ernst E, Pittler M. Assessment of therapeutic safety in systematic reviews: literature review. BMJ 2001;323:546
5. Loke Y, Golder S, Vandenbroucke J. Comprehensive evaluations of the adverse effects of drugs: importance of appropriate study selection and data sources. Ther Adv Drug Saf 2011; 2(2) 59-68.
6. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
7. Derry S, Loke YK, Aronson JK. Incomplete evidence: the inadequacy of databases in tracing published adverse drug reactions in clinical trials. BMC Med Res Methodol 2001; 1: 7.
8. Golder S, Loke Y, McIntosh HM. Room for improvement? A survey of the methods used in systematic reviews of adverse effects. BMC MedRes Methodol 2006; 6: 3.
9. Golder S, Loke Y, McIntosh HM. Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. J Clin Epidemiol 2008; 61(5): 440–448.
10. Golder S, McIntosh HM, Loke Y. Identifying systematic reviews of the adverse effects of health care interventions. BMC Med Res Methodol 2006; 6: 22.

11. McIntosh HM, Woolacott NF, Bagnall AM. Assessing harmful effects in systematic reviews. BMC Med Res Methodol 2004; 4: 19.
12. Moher D, Liberati A, Tetzlaff J, Altman D. The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. July 2009; 6 (7). http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1000097
13. Loke YK, Derry S. Reporting of adverse drug reactions in randomised controlled trials—a systematic survey. BMC Clin Pharmacol. 2001;1:3. [PMID: 11591227]
14. Ioannidis JP. Reporting of safety data from randomised trials [Letter]. Lancet. 1998;352:1752-3. [PMID: 9848355]
15. Parke TJ, Stevens JE, Rice AS, Greenaway CL, Bray RJ, Smith PJ, et al. Metabolic acidosis and fatal myocardial failure after propofol infusion in children: five case reports. BMJ 1992;305(6854):613-6. [MEDLINE: 1393073]
16. Hanna JP, Ramundo ML. Rhabdomyolysis and hypoxia associated with prolonged propofol infusion in children. Neurology 1998;50(1):301-3. [MEDLINE: 9443502]
17. Bray RJ. Propofol infusion syndrome in children. Paediatric Anaesthesia 1998;8(6):491-9. [MEDLINE: 9836214]
18. Wolf A, Weir P, Segar P, Stone J, Shield J. Impaired fatty acid oxidation in propofol infusion syndrome. Lancet 2001;357(9256):606-7. [MEDLINE: 11558490]
19. FDA /Astra Zeneca Pharmaceuticals. Dear health care provider letter. www.fda.gov/medwatch/safety/2001/diprivan deardoc.pdf 2001.
20. Health Canada. Drugs and Health Products. Update on the use of Propofol. http://www.hc-sc.gc.ca/dhp-mps/medeff/advisories-avis/prof/_2002/propofol_pediatric_2_nth-ah-eng.php July10, 2002
21. Liberati A, Altman D, Tetzlaff J, Mulrow C, Gotzsche P, Ioannidis J, Clarke M, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and Meta-Analyses of Studies that evaluate health care intereventions: Explanation and Elaboration. July 2009; 6 (7) PLoSMedicine. www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1000100

22. Ioannidis J, Evans S, Gotzsche P, O'Neill R, Altman D, Schulz K, Moher D. Better Reporting of Harms in Randomized Clinical Trials: an Extension of the CONSORT Statement. Annals of Internal Medicine, 2004; 141 (10):781-788.
23. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D for the CONSORT and Pragmatic Trials in Healthcare (Practihc) group. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008; 337;a2390
24. Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. Ann Intern Med 2006; 144(5):364-367
25. Boutron I, Moher D, Altman DG, Schulz K, Ravaud P, for the CONSORT group. Methods and Processes of the CONSORT Group: Example of an Extension for Trials Assessing Nonpharmacologic Treatments. Ann Intern Med. 2008:W60-W67
26. MacPherson H, Altman DG, Hammerschlag R, Youping L, Taixiang W, White A, Moher D; STRICTA Revision Group. Revised STandards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement. PLoS Med. 2010 Jun 8;7(6):e1000261
27. Moher D, Schulz K, Simera I, Altman D. Guidance for Developers of Health Research Reporting Guidelines. PloS Medicine, 2010; 7 (2) e1000217.
28. Zorzela L, Golder S, Liu Y, Pilkington K, Hartling L, et all. Quality of reporting in systematic reviews of adverse events: systematic review. BMJ 2014; 348:f7668.
29. Zorzela LM, Punja S, Joffe A, Hartling L, Vandermeer B, Loke Y, Vohra S. Propofol infusion for paediatric sedation. Cochrane Database of Systematic Reviews 2012, Issue 4. Art. No.: CD009813. DOI: 10.1002/14651858.CD009813.

Chapter 2

Quality of reporting in systematic reviews of adverse events: systematic review

Quality of reporting in systematic reviews of adverse events: systematic review

BMJ 2014;348: f7668 doi: 10.1136/bmj.f7668 (Published 8 January 2014)

Authors: Liliane Zorzela [1], Su Golder [2], Yali Liu lecturer[3], Karen Pilkington [4], Lisa Hartling [5], Ari Joffe [6], Yoon Loke [7], Sunita Vohra [8]

[1] Department of Pediatrics, 4-548 Edmonton Clinic Health Academy, University of Alberta, Edmonton, Canada;

[2] Centre for Reviews and Dissemination, University of York, York, UK;

[3] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Chengguan District, Lanzhou, Gansu, People's Republic of China;

[4] School of Life Sciences, University of Westminster, London, UK;

[5] Department of Pediatrics, 4-472 Edmonton Clinic Health Academy, University of Alberta, Canada;

[6] Stollery Children's Hospital, Department of Pediatrics, University of Alberta, Edmonton, Canada;

[7] University of East Anglia Medical School, Norwich, UK;

[8] Department of Pediatrics, University of Alberta, Canada, T5K 0L4

Abstract

Objectives: To examine the quality of reporting of harms in systematic reviews, and to determine the need for a reporting guideline specific for reviews of harms.

Design: Systematic review.

Data sources: Cochrane Database of Systematic Reviews (CDSR) and Database of Abstracts of Reviews of Effects (DARE).

Review methods: Databases were searched for systematic reviews having an adverse event as the main outcome, published from January 2008 to April 2011. Adverse events included an adverse reaction, harms, or complications associated with any healthcare intervention. Articles with a primary aim to investigate the complete safety profile of an intervention were also included. We developed a list of 37 items to measure the quality of reporting on harms in each review; data were collected as dichotomous outcomes ("yes" or "no" for each item).

Results: Of 4644 reviews identified, 309 were systematic reviews or meta-analyses primarily assessing harms (13 from CDSR; 296 from DARE). Despite a short time interval, the comparison between the years of 2008 and 2010-11 showed no difference on the quality of reporting over time (P=0.079). Titles in fewer than half the reviews (proportion of reviews 0.46 (95% confidence interval 0.40 to 0.52)) did not mention any harm related terms. Almost one third of DARE reviews (0.26 (0.22 to 0.31)) did not clearly define the adverse events reviewed, nor did they specify the study designs selected for inclusion in their methods section. Almost half of reviews (n=170) did not consider patient risk factors or length of follow-up when reviewing harms of an intervention. Of 67 reviews of complications related to surgery or other procedures, only four (0.05 (0.01 to 0.14)) reported professional qualifications of the individuals involved. The overall, unweighted, proportion of reviews with good reporting was 0.56 (0.55 to 0.57); corresponding proportions were 0.55 (0.53 to 0.57) in 2008, 0.55 (0.54 to 0.57) in 2009, and 0.57 (0.55 to 0.58) in 2010-11.

Conclusion: Systematic reviews compound the poor reporting of harms data in primary studies by failing to report on harms or doing so inadequately. Improving reporting of adverse events in systematic reviews is an important step towards a balanced assessment of an intervention.

Introduction

A balanced assessment of interventions requires analysis of both benefits and harms. Systematic reviews or meta-analyses of randomised controlled trials are the preferred method to synthesise evidence in a comprehensive, transparent, and reproducible manner. Randomised controlled trials rarely assess harms as their primary outcome; therefore, they typically lack the power to detect differences in harms between groups (table1). Usually designed to evaluate treatment efficacy or effectiveness, randomised controlled trials are often done over a short period of time, with a relatively small number of participants. These trials are known to be poor at identifying and reporting harms, which can lead to a misconception that a given intervention is safe, when its safety is actually unknown (1-8). Systematic reviews with a primary objective to assess harms represent fewer than 10% of all systematic reviews published yearly (9, 10) Systematic reviews of harms can provide valuable information to describe adverse events (frequency, nature, seriousness), but they are hampered by a lack of standardised methods to report these events and the fact that harms are not usually the primary outcome of included studies (9-28). Several studies have identified challenges when developing a systematic review of adverse events (9-28). These include: the poor quality information on harms reported in original studies, difficulties in identifying relevant studies on adverse events when using standard systematic search techniques, and the lack of a specific guideline to perform a systematic review of adverse events. The need for better reporting on harms in general (1-8) and in systematic reviews (9-28) in particular—has been voiced. In a previous review (28) of systematic reviews from the Cochrane Database of Systematic Reviews (CDSR) and Database of Abstracts of Reviews of Effects (DARE), our team identified a significantly increased number of reviews of adverse events over the past 17 years ($P<0.001$); however, the proportion of these reviews out of the total number of reviews was unchanged at 5% (11, 14, 28). Some positive points were noted—for example, the increased number of databases searched per review and

the reduction in number of systematic reviews limiting their search strategies by date or language—but appropriate reporting of search strategies was still a problem (28).

The PRISMA (30) (preferred reporting items for systematic review and meta-analysis) statement was developed to deal with suboptimal reporting in systematic reviews. Thus far, PRISMA has mainly focused on efficacy and not on harms. A reporting guideline specific for systematic reviews of harms is crucial to provide a better assessment of adverse events of interventions. The first step for successful guideline development is to document the quality of reporting in published research articles to justify the need for the guideline (30). The goal of this review was to determine whether there is a need for a guideline specific for reviews of harms (30) through assessment of the quality of reporting in systematic reviews of harms published between January 2008 and April 2011, from two major databases.

Methods

Development of the checklist items

To assess the quality of reporting in systematic reviews of harms, we developed a set of items (Table 2) to be reported in these reviews.

The items were originally based on a draft generated from analysis of a systematic review of harms conducted previously (33). During the development of the data extraction form for this current review, several items were added. The wording and content were further refined over telephone meetings, by a group of experts in systematic reviews and guideline development (8, 31, 32, 34-39).

Not every PRISMA item has a corresponding harms item, and a few have more than one suggestion per item. PRISMA items 15, 19-25, and 27 did not have any specific harms related items.

Search strategy

We searched the CDSR (via the Cochrane Library) and DARE (via the Centre for Reviews and Dissemination and the Cochrane Library) databases for systematic reviews having an adverse event as the primary outcome measured. DARE is compiled through rigorous weekly searches

of bibliographic databases (including Medline, Embase, PsycINFO, PubMed, and the Cumulative Index to Nursing and Allied Health Literature). It also involves less frequent searches of the Allied and Complementary Medicine Database and the Education Resource Information Center, hand searching of key journals, grey literature, and regular searches of the internet. CDSR includes all the systematic reviews published by the Cochrane Collaboration. We selected the combination of these two databases because they are likely to represent the most comprehensive collection of systematic reviews published in healthcare (28).

The search was limited to a 40 month period between 1 January 2008 and 25 April 2011. The web appendix shows the search strategy used. The dates were selected to include recent reviews in order to describe the current state of reporting in systematic reviews of harms.

Eligibility criteria

Reviews were selected if the primary outcome investigated was exclusively an unintended effect or effects of an intervention. It could be an adverse event, adverse effect, adverse reaction, harms, or complications (Table 1) associated with any healthcare intervention (such as pharmaceutical interventions, diagnostic procedures, surgical interventions, or medical devices). Articles with a primary aim to investigate the complete safety profile of an intervention were included. Reviews were not excluded on the basis of their results or conclusions.

We excluded reviews assessing both beneficial and harmful effects (reviews of both efficacy and harms), reviews of desirable side effects of drugs, or reviews of prevention or reduction of unintended or adverse effects. No limitations on interventions, patient groups, or language were applied.

Screening and data extraction

Relevant studies were screened by title and abstract (when available) independently by two authors (LZ and SG). Any disagreements were resolved by consensus; if consensus could not be reached, disagreements were resolved with a third author (SV). Abstracts that were identified as potentially meeting the DARE criteria but were not assessed were called “provisional abstracts.” The full text was retrieved for these abstracts.

The data extraction was based on the items developed, piloted, and refined (table 2). Each field received a "yes" if the item was reported as defined, or "no" if not reported. The data were extracted by one author (LZ) and verified by a second author (YL). Disagreements were resolved by consensus.

Outcome

The main outcome assessed was the quality of reporting in reviews of harms for each of the 37 items. We also measured the proportion of each "yes" response for each year of the search: 2008, 2009, and 2010-11 (12 reviews published in the first four months of 2011 were combined with the reviews published in 2010). The quality of reporting was compared between the earliest and latest years reviewed (2008 v 2010-11) to assess any improvement in quality of reporting during the study period. We deliberately decided not to include an intermediate category (such as "unclear"). If the item was not clearly reported, it was considered as a "no" response and the unclear category would simply be duplication.

The present review did not aim to evaluate the reason behind the review author's decisions to examine harms. Our goal was to measure the quality of reporting in those reviews— answering the question "is the item clearly reported?" The goal was not to judge whether a methodologically appropriate decision was made (for example, statistical tests used, data extraction, data pooling), but to ensure clarity in reporting regarding the choices made. Information on the types of interventions, nature of included study designs, as well as search strategies and databases searched in systematic reviews published between 1994 and 2011 was previously reported by our team (28).

This study does not intend to measure the effect of the PRISMA statement, for two reasons. Firstly, PRISMA focuses on efficacy; thus measurement of its effect would reasonably focus on systematic reviews of efficacy and not specifically of harms. Secondly, the 37 items measured in the present review were new items and not those found in PRISMA.

Data analysis

Data were collected as dichotomous outcomes ("yes" or "no") for each item, and presented as proportions of reviews for each category (reported from 0 to 1) and as

proportions divided by the database that the reviews were identified from (CDSR and DARE). We also provided 95% confidence intervals for each proportion based on methods described by Wilson and Newcombe (40, 41) using a correction for continuity.

An overall reporting quality assessment was provided through an unweighted average of proportions of items with good reporting. P≤0.05 was considered statistically significant. We did statistical calculations using StataIC-13.

Results

The search yielded 4644 unique references. After screening and retrieving full text articles, an extra 14 papers were excluded as they did not fulfil the inclusion criteria. A total of 309 reviews were identified as systematic reviews of meta-analyses primarily assessing harms, of which 13 were identified at CDSR and 296 at DARE (Figure 1). Disagreements regarding inclusion/exclusion were discussed by LZ and SG and consensus was reached after discussion. Three of the included papers were published in Chinese, two in Spanish, and one in Portuguese. Table 3 provides detailed information on the reporting of each item.

The 309 systematic reviews and meta-analyses with harms as a primary outcome focused on the following interventions:

- Drugs (223 studies, proportion of reviews 0.72 (95% confidence interval 0.66 to 0.77))
- Surgery or other procedures (67 studies, 0.21 (0.17 to 0.26))
- Devices (13 studies, 0.04 (0.02 to 0.07))
- Blood transfusion (two studies, 0.006 (0.001 to 0.022))
- Enteral nutrition (two studies, 0.006 (0.001 to 0.022))
- Isolation rooms (one study, 0.003 (0.001 to 0.01))
- Surgical versus medical treatment (one study, 0.003 (0.001 to 0.01)).

Titles and abstracts

Titles in close to half of included systematic reviews and meta-analyses of harms did not mention any harm related terms (proportion of reviews 0.46 (95% confidence interval 0.40 to 0.52)) and had no report of a patient population or condition under review (0.42 (0.36 to 0.48)). Twenty five reviews (0.08 (0.05 to 0.11) used the word "safety" to identify a review of harms. In the abstract section, reviews often used harm related terms (0.84 (0.79 to 0.88)). One in every 6.5 reviews did not have any harms related word in the abstract, and half of reviews (0.5 (0.44 to 0.56)) did not report the study designs sought or included.

Introduction and rationale

Introductions were well written overall, explaining the rationale for the review and providing information on harms being reviewed. Fifty one reviews (proportion of reviews 0.16 (95% confidence interval 0.12 to 0.21)) were performed to investigate any adverse event associated with an intervention, rather than focusing on a specific event. As per our definition, these reviews provided "an explicit statement of questions being asked with reference to harms" as this broad goal was reported.

Methods

Protocol and registration

Consistent with Cochrane requirements for authors, all included systematic reviews conducted through the Cochrane Collaboration (Cochrane reviews) had a protocol. Cochrane reviews did not refer to a protocol in their full text reviews, but this item was considered "yes" for the reviews for which a protocol could be found. By contrast, only 22 reviews (proportion of reviews 0.07 (95% confidence interval 0.04 to 0.11)) identified through the DARE database reported the use of a protocol. Reporting clinical expertise was not deemed necessary to receive "yes" for this item (Table 2).

Eligibility criteria

Almost one third of DARE reviews (proportion of reviews 0.26 (95% confidence interval 0.22 to 0.31)) did not have a clear definition of the adverse events reviewed, nor did they specify the study designs selected for inclusion in their methods section.

All Cochrane reviews had a clear report of their search strategy, eligible study designs, and methods of data extraction.

Information sources

Authors did not usually search outside the peer reviewed literature for additional sources of adverse events; for example, only 54 reviews (proportion of reviews 0.17 (95% confidence interval 0.13 to 0.22)) searched databases of regulatory bodies or similar sources. Seven of 13 Cochrane reviews (0.53 (0.25 to 0.80)) searched for data from regulatory bodies or industry, compared with 47 of 296 (0.15 (0.11 to 0.20)) DARE reviews.

Study selection and data items

At the screening phase, most reviews only included studies if the harms being searched were reported (proportion of reviews 0.70 (95% confidence interval 0.65 to 0.75)). Of 67 reviews of complications related to surgery or procedures, only four (0.05 (0.01 to 0.14)) reported professional qualifications of the individuals involved.

Study characteristics

Reports of any possible patient related risk factors–such as age, sex, or comorbidities–were sought in less than half of reviews (proportion of reviews 0.41 (0.36 to 0.47)). Furthermore, only 10 reviews adjudicated whether the adverse event could be biologically, pharmacologically, or temporally caused by the intervention, as measured by item 12a.

Results

Fewer than half the reviews (132 of 309; proportion of reviews 0.42 (95% confidence interval 0.37 to 0.48)) only included controlled clinical trials (randomised or not). Only 59 reviews (0.19 (0.14 to 0.23)) included both clinical trials and observational studies. Only observational studies (prospective or retrospective) were included in 109 reviews (0.35 (0.29 to 0.40)), case series or

case reports were included in 24 (0.07 (0.05 to 0.11)). Three of 13 Cochrane reviews (0.23 (0.05 to 0.53)) included observational studies compared with 165 of 296 DARE reviews (0.55 (0.49 to 0.61)). After reviewing the full text, nine reviews did not report the designs included anywhere in the text. Length of follow-up or patient demographics were only reported in just over half of reviews (170 of 309 reviews; 0.55 (0.49 to 0.60)).

In a retrospective analysis, we compared the quality of reporting between the years of 2008 and 2010-11 to measure any possible improvement over time on the quality of reporting. There were no significant differences (p=0.079) in the proportion of good reporting between the earlier and later years.

The 13 CDSR reviews had overall better reporting than DARE reviews in the abstracts, methods, and results sections; the other categories had similar levels of reporting quality. Almost half of DARE reviews (proportion of reviews 0.46 (95% confidence interval 0.40 to 0.51)) had poor reporting in the results section. Because the number of reviews was considerably different between databases, we considered it inappropriate to proceed with any formal tests to compare CDSR and DARE reviews.

Figure 2 provides a graphic trend in the proportion of reviews with good reporting over time (2008, 2009, and 2010-2011), by each item reviewed. Reviews had poor quality of reporting for methods and results, with an average of half of items being poorly reported on those sections.

The overall proportion of reviews with good reporting was 0.56 (95% confidence interval 0.55 to 0.57). The corresponding proportions were 0.55 (0.53 to 0.57) in 2008, 0.55 (0.54 to 0.57) in 2009, and 0.57 (0.55 to 0.58) in 2010-11.

Discussion

Principal findings

We conducted this systematic review to assess the quality of reporting in systematic reviews of harms as a primary outcome using a set of proposed reporting items. This is the first step for the development of PRISMA Harms, a reporting guideline specifically designed for reviews of harms (30).

There is a substantial difference between the number of systematic reviews measuring an adverse event as the main outcome identified through DARE and those identified through CDSR. CDSR reviews comprised only a small fraction of the total reviews included. This distinction may compromise any direct comparison between the two databases. Reviews of harms as a secondary outcome or a co-primary outcome were not included in this review, which could be one reason for the large dissimilarity. Despite the small number of reviews of harms published in the CDSR, they were better reported overall than DARE reviews, probably owing to the clear guidelines provided by the Cochrane Collaboration (29) and the more flexible word limits allowed than those in other peer reviewed journals.

Several items were poorly reported in the included reviews, but a few are especially important when reviewing harms, because the lack of reporting on these could lead to misinterpretations of findings. In a systematic review, the screening phase is crucial and the exclusion of studies due to the absence of harms could overestimate the events and perhaps generate a biased review. Two thirds of reviews of harms only included studies if at least one adverse event was reported in the included studies. In a review of harms, "zero" is an important value, and studies with "no adverse events" are possibly as relevant to the review as those with reported adverse events. Nevertheless, zero events in studies require careful interpretation, because the lack of reported harms may have different reasons: they may not have occurred (that is, a zero event), they may not have been investigated (that is, unknown if zero or no events occurred), or they may have been detected but not reported (that is, unknown if zero or no events occurred). The lack of reporting can be thought of as a measurement bias or reporting bias and should be considered as such (1-12 24-26).

All these scenarios have different implications for readers who need to judge whether an intervention may cause harm. Almost half of the included reviews of harms as a primary outcome did not consider patient risk factors or length of follow-up when reviewing adverse events of an intervention. Readers cannot properly judge whether there is an association between intervention and harms if these critical data are not reported. Most of the poorly reported items were identified at methods and results. The clarity on methods and results are

essential to provide a clear picture of author's intentions and limitations of findings and transparency is warranted.

Strengths and limitations

This review was unique by including more than 300 reviews, from two major databases, that looked at harms as a main outcome; each review was evaluated in depth using a novel set of 37 items to measure the quality of reporting. A limitation of this review was the lack of a reporting guideline specific for systematic reviews of harms; different formats of reporting were found, and assessing whether the reporting was adequate was challenging. The reviewers were generous in their assessment, accepting a range of reports as a "yes" to the presence of an item, which could have underestimated the degree of the problem.

Our review was limited exclusively to systematic reviews where harms were the primary focus. We believed that measuring the quality of reporting of adverse events in reviews specifically designed to evaluate adverse events would generate a pure sample focusing on the reporting of such events. The documenting of poor reporting in these reviews would imply poor reporting on adverse events in general. At this stage, we also decided to be inclusive and measure all potentially relevant items. In the future PRISMA Harms extension, we will limit these to the minimum set of essential items for reporting harms in a systematic review.

Comparison with other studies

Hopewell and colleagues (14) reviewed a sample of 59 Cochrane reviews, of which only one was a harms review. The remaining 58 reviews focused primarily on benefit, with adverse events as secondary or tertiary considerations. Hopewell and colleagues reported that 32 (54%) reviews had fewer than three paragraphs of information on adverse events, and 11 (19%) had fewer than five sentences on adverse events.

Hammad and colleagues (42) reviewed a sample of 27 meta-analyses primarily assessing drug safety and identified that more than 85% of the PRISMA items were reported in the majority of the reviews. However, most reviews did not report the 20 items specifically developed by the authors to address drug safety assessment. Several of the items considered important by

Hammad and colleagues where similar to the ones developed by our team. The present review adds to the voice of many authors highlighting the poor reporting in reviews of adverse events (9-28, 42).

Policy implications

Systematic reviews may compound the poor reporting of harms data in primary studies by failing to report on harms or doing so inadequately (9-28). We recognise the need to optimise quality of reporting on harms in primary studies, and attempts to enhance it have already been made (6). At this phase, we measured the report quality in systematic reviews. As a future step, we intend to compare the quality of reporting in included studies with the reporting in the systematic review.

Despite their status as the preferred method for knowledge synthesis, systematic reviews can present an incomplete picture to readers by not representing a reliable assessment of a given treatment. Authors of the systematic review have a unique vantage point and can evaluate the entire evidence base under review, including deficits in reporting harms at the primary study level. This vantage point should be used to flag deficits in primary reporting. The goal over time should be to improve the quality and clarity of reporting in systematic reviews as well as the primary studies they evaluate. Although we are glad to see increasing numbers of systematic reviews of harms being published, we are also aware that the validity of the results can be heavily influenced by reviewers' decisions during conduct of the review–even more so than in reviews of benefit because we are dealing with sparse data and secondary outcomes (43-45). Hence, we emphasise here the crucial importance of transparent reporting of the methods used in systematic reviews of harms.

Conclusions

Systematic reviews of interventions should put equal emphasis on efficacy and harms. Improved reporting of adverse events in systematic reviews is one step towards providing a balanced assessment of an intervention. Patients, healthcare professionals, and policymakers should base their decisions not only on the efficacy of an intervention, but also on its risks.

Guidance on a minimal set of items to be reported when reviewing harms is needed to improve transparency and informed decision making, thereby greatly enhancing the relevance of systematic reviews to clinical practice.

This review used a set of proposed reporting items to assess reporting, and the findings indicate that specific aspects of harms reporting could be improved. The items will be further refined with the aim of developing a final set of criteria that would constitute the PRISMA Harms. The PRISMA statement (30) is a living document, open to criticism and suggestions; these same principles will be shared by the PRISMA Harms. The development of a standardised format for reporting harms in systematic reviews will promote clarity and help ensure that readers have the basic information necessary to make an informed assessment of the intervention under review.

Contributors: All authors participated in the study conception and design, and analysis and interpretation of data; drafted the article or revisited it critically for important intellectual content; and approved the final version to be published. SV is the study guarantor.

Funding: No specific funding was given for this study.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: No ethics approval was necessary because this was a review of published literature. No patient data or confidential information was used in this manuscript.

Data sharing: The dataset is available from the corresponding author at svohra@ualberta.ca.

This manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

What this paper adds

There is room for improved reporting in systematic of harms, including clear definition of the events measured, length of follow-up, and patient's risk factors.

Comparisons of reviews of harms have shown no improvement in the quality of reporting over time.

Lack of detail or transparency in the reporting of systematic reviews of harms could hinder proper assessment of validity of findings.

What is already known on this topic

The number of systematic reviews of adverse effects has increased significantly over the past 17 years.

Harms are poorly reported in randomised controlled trials, but it is unclear whether there are weaknesses in the reporting of systematic reviews of harms.

Although the PRISMA statement aims to provide guidance on transparent reporting in systematic reviews, its recommendations are mainly focused on studies of beneficial outcomes.

References

1 Ioannidis JP. Adverse events in randomized trials: neglected, restricted, distorted, and silenced. *Arch Intern Med* 2009;169:1737-9.

2 Sigh S, Loke YK. Drug safety assessment in clinical trials: methodological challenges and opportunities. *Trials* 2012;13:138.

3 Papanikolau PN, Ioannidis JP. Availability of large-scale evidence on specific harms from systematic reviews of randomized trials. *Am J Med* 2004;117:582-9.

4 Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. *Arch Intern Med* 2009;169:1756-61.

5 Loke YK, Derry S. Reporting of adverse drug reactions in randomized controlled trials–a systematic survey. *BMC Clin Pharmacol* 2001;1:3.

6 Ioannidis JP, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. *JAMA* 2001;285:437-43.

7 Yazici Y. Some concerns about adverse event reporting in randomized clinical trials. *Bull NYU Hosp Jt Dis* 2008;66:143-5.

8 Ioannidis JP, Evans SJ, Gøtzsche PC, O' Neill RT, Altman DG, Schultz K, et al. Better

reporting of harms in randomized clinical trials: an extension of the CONSORT statement. *Ann Int Med* 2004;141:781-8.

9 McIntosh HM, Woolacott NF, Bagnall AM. Assessing harmful effects in systematic reviews. *BMC Med Res Methodol* 2004;4:19.

10 Ernst E, Pittler MH. Assessment of therapeutic safety in systematic reviews: literature review. *BMJ* 2001;323:546.

11 Ernst E, Pittler MH. Systematic reviews neglect safety issues. *Arch Intern Med* 2001;161:125 - 6.

12 Aronson JK, Derry S, Loke YK. Adverse drug reactions: keeping up to date. *Fundam Clin Pharmacol* 2002;16:49-56.

13 Moseley AM, Elkins MR, Herbert RD, Maher CG, Sherrington C. Cochrane reviews used more rigorous methods than non-cochrane reviews: survey of systematic reviews in physiotherapy. *J Clin Epidemiol* 2009;62:1021-30.

14 Hopewell S, Wolfenden L, Clarke M. Reporting of adverse events in systematic reviews can be improved: survey results. *J Clin Epidemiol* 2007;61:597-602.

15 Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PloS Med* 2007;4:e78.

16 Golder S, Loke YK, McIntosh HM. Room for improvement? A survey of the methods used in systematic reviews of adverse effects. *BMC Med Res Methodol* 2006;6:3.

17 Vandermeer B, Bialy L, Johnston B, Hooton N, Hartling L, Klassen TP, et al. Meta-analyses of safety data: a comparison of exact versus asymptotic methods. *Stat Methods Med Res* 2009;18:421-32.

18 Loke YK, Golder S, Vandenbroucke J. Comprehensive evaluations of the adverse effects of drugs: importance of appropriate study selection and data sources. *Ther Adv Drug Saf* 2011;2:59-68.

19 Loke YK, Price D, Herxheimer A. Systematic reviews of adverse effects: framework for a structured approach. *BMC Med Res Methodol* 2007;7:32.

20 Golder S, Loke YK, McIntosh HM. Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. *J Clin Epidemiol* 2008;61:440-8.

21 Golder S, McIntosh HM, Loke YK. Identifying systematic reviews of the adverse effects of health care interventions. *BMC Med Res Methodol* 2006;6:22.

22 Derry S, Loke YK, Aronson JK. Incomplete evidence: the inadequacy of databases in tracing published adverse drug reactions in clinical trials. *BMC Med Res Methodol* 2001;1:7.

23 Golder S, Loke Y. Is there evidence for biased reporting of published adverse effects data in pharmaceutical industry-funded studies? *Br J Clin Pharmacol* 2008;66:767-3.

24 Chou R, Helfand M. Challenges in systematic reviews that assess treatment harms. *Ann Intern Med* 2005;142:1090-9.

25 Chou R, Aronson N, Atkins D, Ismaila AS, Santaguida P, Smith DH, et al. AHRQ series paper 4: assessing harms when comparing medical interventions: AHRQ and the effective health care program. *J Clin Epi* 2010;63:502-12.

26 Cornelius VR, Perrio MJ, Shakir SA, Smith LA. Systematic reviews of adverse effects of drug interventions: a survey of their conduct and reporting quality. *Pharmacoepidemiol Drug Saf* 2009;18:1223-31.

27 Warren FC, Abrams KR, Golder S, Sutton AJ. Systematic review of methods used in meta-analyses where a primary outcome is an adverse or unintended event. *BMC Med Res Methodol* 2012;12:64.

28 Golder S, Loke YK, Zorzela L. Some improvements are apparent in identifying adverse effects in systematic reviews from 1994 to 2011. *J Clin Epidemiol* 2013;66:253-60.

29 Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011]. Cochrane Collaboration, 2011. www.cochranehandbook.org.

30 Moher D, Liberati A, Tetzlaff J, Altman D. The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PloS Med* 2009;6:e1000097.

31 Moher D, Schulz K, Simera I, Altman D. Guidance for developers of health research reporting guidelines. *PloS Med* 2010;7:e1000217.

32 Hill AB. The environment and disease: association or causation? *Proc Roy Soc Med* 1965;58:295-300.

33 Pilkington K, Boshnakova A. Complementary medicine and safety: a systematic investigation of design and reporting of systematic reviews. *Complement Ther Med* 2012;20:73-82.

34 Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012;345:e5661.

35 Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of non inferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006;295:1152-60.

36 Zwarenstein M, Treweek S, Gagnier JJ, Altman DG,Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;337:a2390.

Table 1.Glossary

ADVERSE DRUG REACTION[*]	An adverse effect specific to a drug.
ADVERSE EFFECT[*,¥]	An unfavorable outcome that occurs during or after the use of a drug or other intervention but is not necessarily caused by it.
ADVERSE EVENT [*,¥]	An unfavorable outcome that occurs during or after the use of a drug or other intervention and the causal relation between the intervention and the event is at least a reasonable possibility.
COMPLICATION [*]	An adverse event or effect following surgical and other invasive interventions.
HARM [π]	The totality of possible adverse consequences of an intervention or therapy; they are the direct opposite of benefits.
SAFETY [π]	Substantive evidence of an absence of harm. The term is often misused when there is simply absence of evidence of harm.
SIDE EFFECT [*]	Any unintended effect, adverse or beneficial, of a drug that occurs at doses normally used for treatment.
TOXICITY [π]	Describes drug-related harms. The term may be most appropriate for laboratory-determined measurements, although it is also used in relation to clinical events. The disadvantage of the term 'toxicity' is that it implies causality. If authors cannot prove causality, the terms 'abnormal laboratory measurements' or 'laboratory abnormalities' are more appropriate to use.

Adapted from: (π) Ioannidis et al. Better Reporting of Harms in Randomized Trials: An Extension of the CONSORT Statement ; (¥) Chou R, Aronson N, Atkins D, et al. AHRQ Series paper4: Assessing harms when comparing medical interventions: AHRQ and the Effective Health care

Program; (*) Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions.

Table 2. Definition of Reporting Items.

Title	
Item	Definition of "yes"
1a) Specifically mention 'harms' or other related term.	It should contain a word or phrase related to harm: "adverse effect", "adverse event", "complications", "risk", etc.
1b) Clarify if both benefits and harms are examined or only harms.	Mention if benefits are reviewed or not.
1c) Mention the specific intervention being reviewed.	Mention the intervention being reviewed.
1d) Refer to specific patient group and/or conditions in which harms have been assessed.	Clearly state the specific group of patients or conditions being reviewed.
Abstract	
2a) Specifically mention 'harms' or other related terms.	It should contain a word or phrase related to harm: "adverse effect", "adverse event", "complications", "risk", etc.
2b) Clarify if both benefits and harms are examined or only harms.	Mention if benefits are reviewed or not.
2c) Refer to a specific harm assessed.	Clearly state the adverse event being reviewed. General descriptions are accepted, for example "cardiovascular events" or "maternal complications". Also acceptable is if report is searching for any adverse event and it does not refer to a specific harm.
2d) Specify what type of data was sought.	Clearly name the kind of studies searched (e.g. randomized controlled trials, cohort studies, case reports only, all types of data).
2e) Specify what type of data was included.	Clearly name the kind of studies included in the analysis or results (i.e. randomized controlled trials, cohort studies, case reports, all types of data).
2f) Specify how each type of data has been appraised.	Clearly state the method of quality appraisal for included studies (i.e. Jadad score; the Cochrane risk of bias tool, Newcastle-Ottawa scale).
Introduction	
3a) Explains rationale for addressing specific harm(s), condition(s) and patient group(s).	Report reasons for proceeding with the systematic review.
3b) Clearly define what events or effects are	Clearly state the adverse event being reviewed.

considered harms in the context of the intervention(s) examined.	General definitions are acceptable, for example "cardiovascular events" or "maternal complications". Also acceptable if report is searching for any adverse events, as in a scoping review, and does not refer to a specific harm.
3c) Describes the rationale for type of harms systematic review done: hypothesis generating vs hypothesis testing.	Clearly state if a specific adverse event is being reviewed (hypothesis testing review) (e.g. cardiovascular deaths) or the authors are searching for any adverse events related to the intervention (hypothesis generating).
3d) Explains rationale for selection of study types or data sources and relevance to focus of the review.	Clearly state which study designs are included (e.g. randomized controlled trials, cohorts, case reports). Rationale is not necessary for "yes".

Objectives	
4a) Provide an explicit statement of questions being asked with reference to harms.	Clearly defined, preferable at the end of introduction section, the intervention and the adverse events being reviewed. For this review, it was considered acceptable if intervention and outcome are clearly stated.
Methods	
Methods: Protocol and registration	
5a) Describe if protocol was developed in collaboration with someone with clinical expertise for field/intervention under study.	Mentioned if a protocol was developed previously. For this review is not necessary to state the clinical expertise of who developed it.
Methods: Eligibility criteria	
6a) Clearly define what events or effects are considered harms in the context of the intervention(s) examined.	Provide a clear definition of the adverse event being reviewed. If previously a general description was provided (i.e. "cardiovascular events") now the specific events need to be defined.
6b) Specify type of studies on harms to be included	Define which kind of study designs will be included (e.g. RCTs, cohort studies, case-control, etc.
Methods: Information sources	
7a) State if additional sources for adverse events were searched (e.g regulatory bodies, industry) if so, describe the source, terms used and dates searched.	Report if any other sources of adverse events are searched, different than regular peer reviewed journals. For this review, it is not necessary to have the terms used and dates searched.
Methods: Search	
8a) Present the full search strategy if additional searches were used to identify adverse events	Report search strategy used. Adverse events terms and databases searched.
Methods: Study selection	
9a) State what study designs were eligible and provide rationale for their selection	Report the study designs included in the review (e.g. RCTs, cohort studies, case-control). No rationale is required.

9.b) Define if studies were screened based on the presence or absence of harms related terms in title or abstract.	Report if the study screening is based on the report of adverse events or not. (e.g. review of mortality associated with anti-glycemic drugs; only studies reporting on mortality were included).
Methods: Data Collection process	
10a) Describe method of data extraction for each type of study or report.	State if a data extraction form is used, how it is done (in duplicate, checked by a second author).
Methods: Data items	
11a) List and define variables for which data were sought for individual therapies.	Report variable(s) sought for the intervention reviewed. E.g. if the outcome is kidney failure, what variable was used to define kidney failure-creatinine level, creatinine clearance, need of dialysis, etc.).
11b) List and define variables for which data were sought for patient underlying risk factors.	State if any potential patient risk factors or confounders are sought. E.g. age, sex, co-morbidities, previous events).
11c) List and define variables for which data were sought for practitioner training/qualifications.	State any variable(s) sought for health care personal (practitioner). E.g. relevant degree, years of experience, other qualifications.
11d) List and define harms for individual therapies.	Provide the definition for the harm(s) sought. E.g. side effects of propranolol and atenolol are defined as severe bradycardia (heart rate 45 beats per minute or less) and severe hypotension (Systolic blood pressure <80mmHg).
Methods: Risk of Bias in Individual studies	
12a) For uncontrolled studies, describe if causality between intervention and adverse event was adjudicated and if so, how.	Specify if review authors adjudicate if the intervention can cause the harm. E.g. Use of Bradford Hill criteria for causality[45]
12b) Describe risk of bias in studies with incomplete or selective report of adverse events.	Describe how studies not reporting AEs are handled regarding risk of bias. E.g. RCT investigating glucose control in type 2 diabetes: AEs reported included hypoglycemia but not mortality. Did the authors assess whether the event did not occur (zero deaths), was not measured (may have occurred) or was not reported (it was considered related to the intervention).
Methods: Summary measures	
13a) If rare outcomes are being investigated, specify which summary measures will be used (e.g. event rate, events/person-time).	Define the summary measures used for rare events.
Methods: Synthesis of Results	
14a) Describe statistical methods of handling with the zero events in included studies.	Clearly define how studies with no AEs reported are reported and analysed. E.g. when zero was an outcome for a 2x2 table, reviewers add a 0.5 value

	to it.
Methods: Additional analysis	
16a) Describe additional analysis with studies with high risk of bias.	Describe if studies with high risk of bias are analyzed separately. E.g. subgroup analysis with high and low risk of bias studies analyzed separately.
Results	
Study selection	
17a) Provide process, table or flow for each type of study design.	Provide a table (or text) containing included studies and a clear reason for exclusion of studies.
Study characteristics	
18a) Report study characteristics such as patient demographics or length of follow-up that may have influenced the risk estimates for the adverse outcome of interest.	Report any potential confounder or patient risk factors that can affect the outcome (AE).
18b) Describe methods of collecting AEs in included studies (e.g.: patient report, active search).	Report how AEs are investigated in included studies (e.g. voluntary report, active search)
18c) For each primary study, list and define each AE addressed and how it is identified.	Clear definition of AE under investigation and how it is identified (method of measurement, assessment or identification).
Conclusion	
26a) Provide balanced discussion of benefits and harms with emphasis on study limitations, generalizability, and other sources of information on harms.	Clearly discuss the benefits of the interventions and the harms identified in the review.

Table 3 Summary of Findings

	(N=309) YES %	CDSR YES % (N=13)	DARE YES % (N=296)
Title			
(#1) 1a) Specifically mention 'harms' or other related term.	53.3	53.8	53.3
(#2) 1b) Clarify if both benefits and harms are examined or only harms.	75.7	92.3	75.0
(#3) 1c) Mention the intervention being reviewed.	98.0	100	98.0
(#4) 1d) Refer to specific patient group and/or conditions in which harms have been assessed.	57.6	92.3	56.1
Abstract			
(#5) 2a) Specifically mention 'harms' or other related terms.	84.4	84.6	84.5
(#6) 2b) Clarify if both benefits and harms are examined or only harms.	92.2	92.3	92.2

(#7) 2c) Refer to a specific harm assessed.	96.7	100	97.4
(#8) 2d) Specify what type of data was sought.	50.4	92.3	48.6
(#9) 2e) Specify what type of data has been included.	48.8	92.3	47.0
(#10) 2f) Specify how each type of data has been appraised.	6.1	23.0	5.4
Introduction			
(#11) 3a) Explain rationale for addressing specific harm(s), condition(s) and patient group(s).	99.3	100	99.3
(#12) 3b) Clearly define what events or effects are considered harms in the context of the intervention(s) examined.	93.5	92.3	93.6
(#13) 3c) Describes the rationale for type of harms systematic review done (hypothesis generating) vs (hypothesis proving).	99.0	100	98.9
(#14) 3d) Explains rationale for selection of study types or data sources and relevance to focus of the review.	38.8	53.8	38.2
Objective: (#15) 4a) Provide an explicit statement of questions being asked with reference to harms.	85.7	100	85.1
Methods			
Protocol and Registration (#16) 5a) Describes if protocol was developed.	11.3	100	7.4
Eligibility criteria (#17) 6a) Clearly define what events or effects are considered harms in the context of the intervention(s) examined.	74.7	92.3	74.0
(#18) 6b) Specify type of studies on harms to be included.	74.4	92.3	73.6
Information sources (#19) 7a) State if additional sources for adverse events were searched (e.g. regulatory bodies, industry).	17.4	53.8	15.9
Search (#20) 8a) Present full search strategy if additional searches were used to identify adverse events.	69.2	100	67.9
Study selection (#21) 9a) State what study designs were eligible.	74.1	100	73.0
(#22) 9b) State if studies were screened based on presence/absence of adverse events.	70.5	92.3	69.6
Data collection process (#23) 10a) Describe method of data extraction for each type of study or report.	64.4	100	62.8
Data Items (#24) 11a) List and define data variables for which data were sought for individual therapies.	49.8	23.0	51.0
(#25) 11b) List and define variables for which data were sought for patient underlying risk factors. (#26) 11c) List and define variables for which data were sought for practitioner training/qualifications. (#27) 11d) List and define harms for individual therapies.	41.7 1.2 67.9	23.0 0 69.2	42.6 1.3 67.9
Risk of Bias in Individual studies (#28) 12a) For uncontrolled studies, describe if causality	3.2	23.0	2.3

between intervention and adverse event was adjudicated.			
(#29) 12b) Describe risk of bias in studies with incomplete or selective reporting of adverse events.	1.9	0	2.0
Summary measures (#30) 13a) If rare outcomes are being investigated, specify which summary measures will be used (e.g. event rate, events/person time).	73.4	100	72.3
Synthesis of results (#31) 14a) Describe statistical methods of handling data with zero events in included studies.	13.2	46.1	11.8
Additional analysis (#32) 16a) Describe additional analysis for studies with high risk of bias.	19.7	76.9	17.2
Results			
Study selection (#33) 17a) Provide process, table or flow for each type of study design.	78.9	92.3	78.4
Study characteristics (#34) 18a) Present study characteristics such as patient demographics or length of follow-up that may have influenced the risk estimates for the adverse outcome of interest.	55.3	84.6	54.0
(#35) 18b) Describe methods of collecting adverse events in included studies (e.g.: patient report, active search).	61.5	84.6	60.5
(#36) 18c) For each primary study, list and define each harm addressed and how harm is identified.	22.0	23.0	22.0
Discussion			
(#37) 26a) Provide balanced discussion of benefits and harms with emphasis on study limitations, generalizability, and other sources of information on harms.	83.1	92.3	82.8

Figure 1.PRISMA Flow Diagram

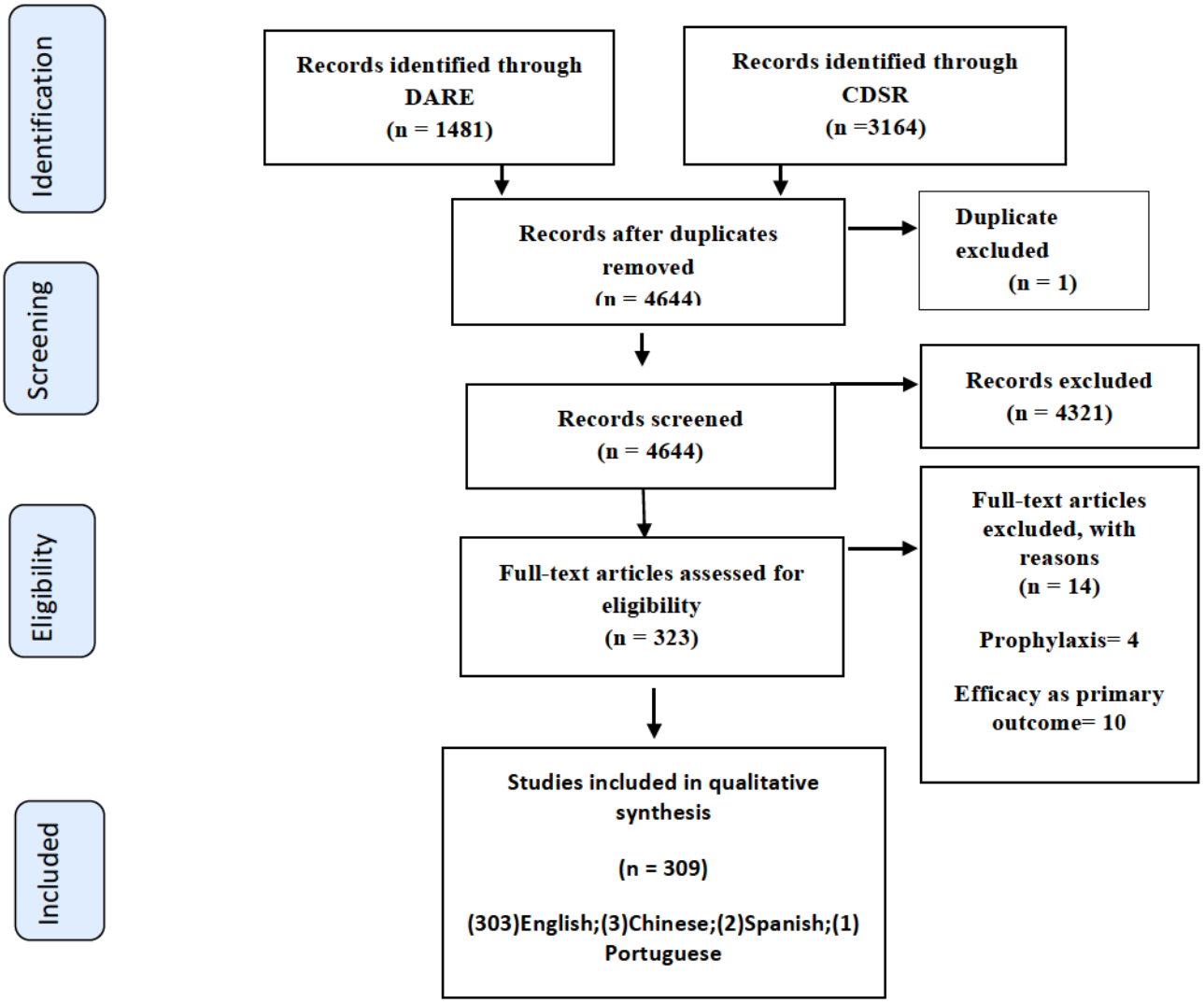
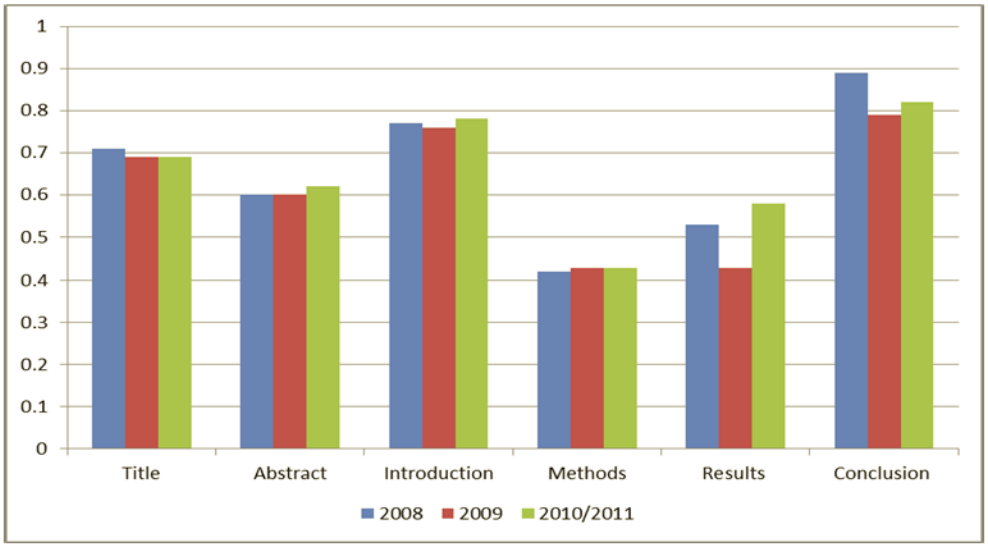


Figure 2. Good reporting by subheading



Annex 1: Search strategy

Database of Abstracts of Reviews of Effects (DARE) via the Centre for Reviews and Dissemination Internal Administration Database

Searched: 25/04/11

The following terms were searched for in the title and 'authors objective' section of the abstract:

'safe' or 'safety' or 'adverse' or 'tolerability' or 'toxicity' or 'toxic' or 'adrs' or 'adr' or 'tolerance' or 'tolerate' or 'harm' or 'harms' or 'harmful' or 'complication$' or 'risk' or 'risks' or 'side effect$' or 'undesirable effect$' or 'treatment emergent'

The full abstract was not searched as part of the DARE process of evaluation includes assessing whether adverse effects were included in the review so many DARE abstracts include sentences such as 'adverse effects were not considered' and terms such as risk would have identified lots of reviews included relative risk ratios. Any duplicates were removed from the results from the two DARE searches.

Database of Abstracts of Reviews of Effects (DARE) via Cochrane Library (Wiley interface) Issue 2 of 4 April 2011

Searched: 25/04/11

#1 Any MeSH descriptor with qualifier: AE in MeSH products

#2 Any MeSH descriptor with qualifier: DE in MeSH products

#3 Any MeSH descriptor with qualifier: CO in MeSH products

#4 Any MeSH descriptor with qualifier: PO in MeSH products

#5 Any MeSH descriptor with qualifier: TO in MeSH products

#6 Any MeSH descriptor with qualifier: CI in MeSH products

#7 MeSH descriptor Drug Hypersensitivity explode all trees in MeSH products

#8 MeSH descriptor Drug Toxicity explode all trees in MeSH products

#9 MeSH descriptor Product Surveillance, Postmarketing explode all trees in MeSH products

#10 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9

Cochrane Database of Systematic Reviews (CDSR) via Cochrane Library

(Wiley interface) Issue 4 of 12, April 2011

Searched: 25/04/11

This database uses automatic pluralisation, so terms “harms and risks” are not required as the terms “harm and risk” are already in the strategy.

#1 Any MeSH descriptor with qualifier: AE in MeSH products

#2 Any MeSH descriptor with qualifier: DE in MeSH products

#3 Any MeSH descriptor with qualifier: CO in MeSH products

#4 Any MeSH descriptor with qualifier: PO in MeSH products

#5 Any MeSH descriptor with qualifier: TO in MeSH products

#6 Any MeSH descriptor with qualifier: CI in MeSH products

#7 MeSH descriptor Drug Hypersensitivity explode all trees in MeSH products

#8 MeSH descriptor Drug Toxicity explode all trees in MeSH products

#9 MeSH descriptor Product Surveillance, Postmarketing explode all trees in MeSH products

#10 (safe or safety or adverse or tolerability or toxicity or toxic or adrs or adr or tolerance or tolerate or harm or harms or harmful or complication* or risk or risks) near/20 objective* in Abstract in all products

#11 (side next effect*) near/20 objective* in Abstract in all products

#12 (undesirable next effect*) near/20 objective* in Abstract in all products

#13 (treatment next emergent) near/20 objective* in Abstract in all products

#14 (safe or safety or adverse or tolerability or toxicity or toxic or adrs or adr or tolerance or tolerate or harm or harms or harmful or complication* or risk or risks) in Record Title in all products

#15 (side next effect*) in Record Title in all products

#16 (undesirable next effect*) in Record Title in all products

#17 (treatment next emergent) in Record Title in all products

#18 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18

Chapter 3

The development of PRISMA Harms using a modified Delphi technique

The development of PRISMA Harms using a modified Delphi technique

Zorzela L (1), Hartling L (1), Loke Y (2), Joffe A (1), Ioannidis JP (3), Golder S (4), Santaguida PL (5), Altman D (6), Vohra S(1).

(1) University of Alberta; (2) University of East Anglia; (3) Stanford University; (4) University of York; (5) McMaster University; (6) University of Oxford

Abstract

Introduction: Reporting on systematic reviews of harms carries unique challenges. The PRISMA Statement offers guidance on systematic review reporting in general, but there are important peculiarities to consider for systematic reviews measuring adverse outcomes of interventions.

The objective of this paper is to report the methods and results of a modified Delphi study used during the development of the PRISMA Harms, a guideline for reporting in systematic reviews assessing adverse events.

Methods: A modified Delphi technique was used with three rounds of questions administered online. During round one, participants rated the 37 potential PRISMA Harms items with respect to relevance/importance using a 10-point Likert scale. During round two, participants received their own previous ratings and an aggregate score for the group; each participant had a chance to re-score each item. Round three was focused on obtaining feedback from pharmaco-epidemiologists, a group that was under-represented during the first two rounds, not seeking for rating based on 10-point Likert scale.

Results: The online survey was sent to 324 participants. The total number of responses was 153 for round one and two. Delphi round one had 81 responses (25%) and round two had 72 (22.8%) responses. Fifty-six round one participants also completed round two (69.1%). Sixteen participants joined the Delphi exercise exclusively for round two. In round one, 19 items were rated 8 or higher by more than 70% of participants. Eighteen items received scattered votes and no items were voted 'not important, it should be discarded'. Post round two, of the 18 items with scattered votes in round one, nine items were rated eight or higher by more than 70% of participants and were selected to be kept; one item (11c) received a rating less than 5 by more than 50% of participants and was selected to be excluded; eight items remained with scattered votes. Thirteen (46.4%) pharmaco-epidemiologists responded to the third round. Round three generated valuable comments, mainly on methodological issues of reviews of harms.

Conclusion: Through this Delphi process, we were not able to condense the number of items to a minimum relevant set as only one item could be excluded. The remaining 36 items would be taken for discussion to the in-person meeting, which was the next step in the guideline development.

Introduction

In 2009, the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) Statement was published to offer guidance to review authors and journal editors on the minimum reporting standard when publishing a systematic review or meta-analysis (1). The majority of systematic reviews have a focus on beneficial outcomes (efficacy) of interventions (2). As a result, PRISMA is largely focused on reporting items that are relevant to systematic reviews of efficacy. Report of harms is an important issue in systematic reviews. However, adequate reporting of harms faces unique issues that differ from SRs of efficacy. Among these, some of the most limiting are the poor quality information on harms reported in original studies and the lack of a specific guideline to report on systematic reviews of harms (2-15).

As endorsed by the EQUATOR Network (16), the first step for a successful reporting guideline development is to determine if one is needed by assessing the quality of reporting on published manuscripts. As the first step for the PRISMA Harms development, a systematic review of 309 reviews with adverse events as a main outcome were evaluated by our team to measure quality of reporting on harms (17). In order to generate a conservative estimate of the quality of reporting, we chose to focus on SRs in which the primary objective was to assess harms. In this review we identified that one quarter of the included manuscripts exclusively designed to measure adverse events did not have a definition of the event under review. Fifty nine percent (183/309) did not extract any data on patient risk factors and 45% (138/309) did not collect data on length of follow up. From the 67 reviews that evaluated complications related to surgeries or procedures, only four reviews reported on professional skills and qualifications of the surgeons performing the intervention (17).

The second step for a reporting guideline development is to employ a Delphi exercise to survey content experts, journal editors, review authors and consumers, to receive feedback on the items to be included in the new guideline (16, 18).

The objective of this study is to report the modified Delphi technique used during the development of the PRISMA Harms, a guideline specific to improve harms reporting in systematic reviews.

Methods

A modified Delphi technique was employed. We used a three phase online Delphi process.

This study was approved by the University of Alberta Health Research Ethics Board.

Development of the Delphi items

The initial set of items was obtained from a systematic review done by Pilkington et al (19). During the design of the review used as step one in this guideline development (17) we identified multiple additional relevant items, such as: (i) how to deal with zero events, (ii) methods of screening studies based on the presence of adverse events, (iii) the measurement of patient risk factors, and (iv) measurement of length of follow up. These items and a few others were added to the checklist of possible items. Wording and content were further refined by the PRISMA Harms steering committee (Appendix 1) over multiple phone meetings.

The initial PRISMA Harms checklist set included a total of 37 items and it was submitted to the Delphi participants. Not every PRISMA Statement item had a suggested harms extension item and a few had more than one suggestion per item. The PRISMA statement items number 15 (*risk of bias across studies*), 19 (*risk of bias within studies*), 20 (*results of individual studies*), 21 (*synthesis of results*), 22 (*risk of bias across studies*), 23 (*additional analysis*), 24 (*summary of evidence*), 25 (*limitations*) and 27 (*funding*) did not have any dedicated PRISMA Harms item proposed. Input from Delphi participants was sought for these specific items.

Identification of Delphi Participants

Online Delphi participants were selected based on their expertise in systematic reviews methodology and in particular for their expertise in reviews focusing on adverse events, also content experts, including statisticians, epidemiologists, clinicians, journal editors, a consumer representative and a federal health regulatory agency representative complemented the list.

Potential Delphi participants were identified from a list of authors of relevant manuscripts on methodological and reporting challenges in systematic reviews in general and review of adverse events specifically. Editors from journals where the PRISMA Statement (1) was published were also invited. The health regulatory agency perspective was provided by a representative from the Marketed Health Products Directorate (MHPD) and Health Canada, who had expertise in systematic reviews of harms;

these reviews are frequently used by regulatory agencies to evaluate health interventions. The steering committee reviewed the potential participant list to ensure completeness. The majority of those invited to participate in the Delphi were corresponding authors of included studies in the systematic reviews conducted in step 1 (17). As these authors have firsthand experience with the challenges of developing a systematic review of adverse events, their input was considered valuable in defining items to promote adequate reporting of harms in systematic reviews.

The Online Delphi process

Delphi participants were invited by e-mail to participate in a web-based survey. The electronic address was retrieved from personal lists or publicly available websites (e.g. affiliated universities). The invitation email contained the purpose for the Delphi exercise, a brief explanation of what was in the survey, the estimated time commitment to complete the survey and the relevance of their participation. It was also highlighted in the invitation email that the survey goal was to select the minimum set of items to be reported in systematic reviews of harms, rather than the ideal set which would have encompassed more items. Invitations to participate were sent up to three times to non-responders.

The survey asked the participants for their input on the importance/relevance of the checklist items by ranking each item on a 10-point Likert scale, with higher numbers meaning more relevant items. The participants had a comment box in each page and opportunity to nominate any new items, if desired.

Delphi Rounds

Three Delphi rounds were planned to be sent to the participants in 4 weeks intervals. Each participant would have two weeks to complete the survey; the study coordinator (LZ) would have two weeks to combine ratings and comments and send back to the participants for the next round: (i) a graph with group scores response distribution; (ii) their previous individual rating for each item, (iii) a summary of comments received for each item. Respondents were given the opportunity to re-rate each item in light of ratings from the previous round. The online survey was deployed using the 'Checkbox Survey Server'.

Items remaining after the online Delphi process were to be used as an initial list of items for discussion in the step 3 of the guideline development process, a face-to-face meeting.

Inclusion/exclusion of items

At the end of the round two, items rated eight or higher by 70% of participants were considered 'important, it should be included'; items rated five or less by more than 50% of panellists were considered 'not important, it should be discarded' and items rated in between (i.e. '6' or '7') would be debated during the next step in the guideline development, the in-person meeting. These thresholds were arbitrarily chosen *a priori* in an attempt to reduce the number of items to the 'minimum set'. It would be easier to discard an item (a low score by more than 50% of participant) than to keep an item (a high score by more than 70% of participants).

Results

The online Delphi round one was sent to 338 potential respondents on February 13th, 2012; of those, 14 emails did not reach the intended recipient. A total of 324 potential participants received the survey link. Of these, 286 (88.3%) were corresponding authors of systematic reviews on harms previously identified by our team (17).

The Delphi round one had a total of 81 respondents (25% response rate). All respondents rated all 37 items. Of these, respondents were self-described as: research methodologist (31.6%); health care provider (26.3%); trialist (7.9%); journal editor (6.7%); researcher (5.3%); statistician (1.2%); regulatory agency representative (1.2%); health economist (1.2%); and other (18.4%); 5 participants did not check the self-described box. Forty-four of 81 respondents in round one were corresponding authors of systematic reviews identified by our team in step 1 (17).

At the end of round one, of the 37 rated items, 19 items received ratings of 8 or higher by more than 70% of participants. Eighteen items received scattered votes. No items were voted 'not important; it should be discarded' using our *a priori* criteria.

The online Delphi round two was sent on March 26th, 2012 to all 324 potential participants. The response rate per item declined, such that there were 72 responses (22.2%) for the first 4 items, but only 63 participants (19.4%) rated all 37 items.

Of the 18 items with scattered votes in round one, nine items were rated higher than eight by more than 70% of participants in round two and were therefore retained. One item was rated less than 5 by more than 50% of participants and was therefore discarded. Eight items remained with scattered votes after round two. All the other items that received high rating in round one, also received a high rating in round two.

Fifty-six round one participants also completed round two (69.1%). Eighteen participants joined the Delphi exercise exclusively for round two, 16 of those were authors of systematic reviews of harms and two were systematic reviews specialists. The response-rate for corresponding authors of systematic review of harms included in step 1 (17) was 16% (46/286) and for round two was 21.7% (62/286). If these review authors would be excluded, the response rate for round one would be 92% (35/38) and for round two 34/38 (89.4%).

The PRISMA Harms steering committee decided that proceeding with a third round with these respondents would likely yield similar results. To gain additional perspective, the PRISMA Harms steering committee recommended expanding the Delphi respondent pool. As such, a third round was conducted with new panelists to see if they agreed, or disagreed, with the items rated thus far. The new pool of participants was selected by the steering committee based on their experience in pharmaco-epidemiology, as they were found to be under-represented in the initial list of participants. The steering committee reasoned that their input would enhance the heterogeneity of the respondent pool and increase the relevance of the findings to a broader audience.

The Delphi round three was sent to 28 pharmaco-epidemiologists for comments and feedback. During this third round, the pharmaco-epidemiologists were asked to provide feedback on the items voted "important, it should be kept" after round two. Round three had an objective of pursuing feedback and comments on the items, it did not seek for rating based on 10-point Likert scale. Thirteen (46.4%) pharmaco-epidemiologists responded in this round.

The comments originated in round three provided additional input to the comments received in the previous two rounds. Some comments received in round three are presented here: (i) on study selection: "It will be interesting to see what people will list as rationale for excluding studies …there are good RCTs and bad ones and there are good cohort studies and bad ones."; (ii) on reporting rare events: "I'd suggest systematic reporting of all adverse events" and "there is a major difference between zero events that are unmeasured or unreported outcomes vs adjudicated outcomes with no events."; (iii) on causality assessment: "the reviewer should review the arguments for and against a causal interpretation."; (iv) on risk of bias: "For studies of harms, the standard checklists for bias are not very helpful, and many studies exploring harm will be at "high" or "unclear" risk of bias, and this terminology really means very little when studying rare, unexpected adverse events using many data sources and with an unknown denominator."; (v) on a summary comment: "I feel the feedback provided while very

thoughtful has originated in minds that are more concerned about conducting a review for harms outcomes than reporting of it".

Table 4 presents the scores for round one and two and summarize comments made by participants in all three rounds. Table 5 presents the opinion and suggestions on harms reporting in PRISMA items without a specific harms modification item suggested.

Many respondents felt that the original PRISMA statement would suffice for 11 items (i.e. PRISMA items 4 to 11, item 13, item 16 and item 26). Delphi participants also frequently commented under different items that while a potential new item was interesting/relevant, limited space in journals would preclude its inclusion.

Discussion

Our study reports the use of a modified Delphi technique to support the development of a reporting guideline, the PRISMA Harms. After the initial two online rounds, the majority of items were voted as being 'important' and it was felt that they should be kept and only one item was considered 'not important' and that it should be deleted. A significant volume of comments were received from all three rounds. A key characteristic of the Delphi method is to reach, anonymously, a group of individuals across diverse locations and areas of expertise, thus providing opportunity for free contribution and open feedback among participants. After each round the group participants received the summary of the results of the previous round and they had a chance to report their opinion or change their response in the light of the group opinion. The process continues until group consensus is reached (35-41).

During the second online Delphi round, the individual rates and the group score display helped in defining relevant items. The high scored items in round one received higher scores in round two. Many items with scattered votes in round one were defined as relevant in round two. Unfortunately, the Delphi exercise failed in narrowing items to the minimum set of items to be reported in a systematic review assessing harms, as only one item from 37 was excluded in this process. This high acceptance rate may reflect the importance of the topic and the relevance of items suggested, but also the difficulty for the respondents to focus and select the 'minimum' set of items. It is possible that instead, they focused on identifying an 'ideal' set of items. Some possible reasons for the high consensus rate could be that: (i) the respondent pool was very homogeneous (hence the similar scores) and (ii) it is not their expertise to create a "minimum set", but rather what is relevant and/or important. In contrast, it is the mandate of the in-person meeting (who have different, complementary expertise, as well as drawing

from the Delphi respondent pool), to take the Delphi responses and generate a minimum set that is usable in a guideline (16). The third round helped in emphasizing the need for reporting and methodological guidance in reviews of adverse events.

A weakness of our Delphi exercise is the fact that the response rate was lower than experienced by the other Delphi exercises used in systematic reviews guidelines. The PRISMA for Abstracts (18) had a response rate of 68% for round one; and the Delphi used during the development of the PROSPERO database for systematic reviews protocols had a 45% response rate of invited participants in the initial round (143/315) (42). We suspect the length of our survey, which required 45 minutes to 1 hour to rate all the 37 items, likely influenced the response rate. One participant mentioned it took him four hours to complete the survey and make comments for each item. Also, the majority of invited participants for the PRISMA Harms Delphi were correspondent authors of systematic reviews of adverse events, selected based on their experience conducting a review of harms, but perhaps these authors were not comfortable with the development of a reporting guideline and preferred not to participate in the Delphi; if this group of review authors had not been invited to participate, our response rate would be higher than 90%.

A strength of our Delphi process was the valuable comments received, demonstrating the interest among participants in this topic and the recognized need for change in reporting harms in systematic reviews. Reporting of harms in general is a challenging topic. The comments received were not limited to reporting, but also the need for methodological improvements for systematic reviews of adverse events. The comments reflected the view of experts in the field and they bring awareness to methodological weaknesses that have been neglected in reviews of harms. The comments presented here could also be used as foundation for the development of a harms extension of a methodological guideline for systematic reviews (whereas PRISMA Harms focuses on transparent and complete reporting, as opposed to methodological recommendations). The online Delphi was also a venue to explore new concepts, for example, the screening of studies based on the presence of the adverse events and the relevance of zero events which could represent an event that did not occur, but also an event that was not captured and or not reported in the primary study. Risk of bias had no specific harms items suggested and multiple comments were received in all three rounds highlighting the need of a risk of bias tool for harms. Respondents felt that the risk of bias for harms is 'overlooked' and 'fundamentally different than risk of bias for effectiveness'. While this is an important issue that needs further examination, like PRISMA, PRISMA Harms remains focused on improving reporting (rather than

trying to influence conduct of reviews). We believe that the summary of participant comments received will help support further methods research to refine how risk of bias can be assessed for harms.

Conclusion

The Delphi consensus exercise has been considered an important tool for a development of a guideline, not only to reach experts and achieve consensus, but also to assure guideline applicability and endorsement, this makes this type of endeavor a highly effective form of integrated knowledge translation and exchange. PRISMA represents a minimum set of items to be reported in systematic reviews. Through our Delphi process we were not able to condense the number of items to a minimum relevant set, as only one item was excluded. The remaining 36 items were taken forward for discussion at the in-person meeting, the next step in the guideline development.

References

1. Moher D, Liberati A, Tetzlaff J, Altman D. The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. July 2009; 6 (7). http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1000097
2. McIntosh HM, Woolacott NF, Bagnall AM. Assessing harmful effects in systematic reviews. BMC Med Res Methodol 2004; 4: 19
3. Ernst E, Pittler M. Assessment of therapeutic safety in systematic reviews: literature review. BMJ 2001;323:546
4. Cornelius V, Perrio M, Shakir A, Smith L. Systematic reviews of adverse effects of drug interventions: a survey of their conduct and reporting quality. Pharmacoepidemiology and Drug Safety 2009; 18: 1223-1231.
5. Loke Y, Golder S, Vandenbroucke J. Comprehensive evaluations of the adverse effects of drugs: importance of appropriate study selection and data sources. TherAdv Drug Saf 2011; 2(2) 59-68.
6. Loke Y, Price D, Herxheimer A. Systematic reviews of adverse effects: framework for a structured approach. BMC Medical Research Methodology. July 2007 7:32. http://biomedcentral.com/1471-2288/7/32
7. Golder S, Loke Y, McIntosh H. Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. Journal of Clinical Epidemiology 2008; 61: 440-448.
8. Derry S, Loke Y, Aronson J. Incomplete evidence: the inadequacy of database in tracing published adverse drug reaction in clinical trials. BMC Medical Research Methodology 2001,1:7. http://www.biomedcentral.com/1471-2288/1/7

9. Loke YK, Aronson JK. Incomplete evidence: the inadequacy of databases in tracing published adverse drug reactions in clinical trials. BMC Med Res Methodol 2001; 1: 7.
10. Golder S, Loke Y, McIntosh HM. Room for improvement? A survey of the methods used in systematic reviews of adverse effects. BMC MedResMethodol 2006; 6: 3.
11. Golder S, Loke Y, McIntosh HM. Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. J ClinEpidemiol 2008; 61(5): 440–448.
12. Golder S, McIntosh HM, Loke Y. Identifying systematic reviews of the adverse effects of health care interventions. BMC Med Res Methodol 2006; 6: 22.
13. Golder S, Loke YK. Is there evidence for biased reporting of published adverse effects data in pharmaceutical industry-funded studies? Br JClinPharmacol 2008; 66(6): 767–773.
14. Chou R, Fu R, Carson S, Saha S, Helfand M. Methodological shortcomings predicted lower harm estimates in one of two sets of studies of clinical interventions. J ClinEpidemiol 2007; 60(1): 18–28.22).
15. Hammad T, Neyarapally G, Pinheiro S, et al. Reporting of meta-analyses of randomized controlled trials with a focus on drug safety: An empirical assessment. *Clin Trials* 2013 10: 389 DOI: 10.1177/1740774513479467
16. Moher D, Schulz K, Simera I, Altman D. Guidance for Developers of Health Research Reporting Guidelines. PloS Medicine, 2010; 7 (2) e1000217
17. Zorzela L, Golder S, Liu Y, Pilkington K, Hartling L, et all. Quality of reporting in systematic reviews of adverse events: systematic review. *BMJ* 2014;348:f7668 doi: 10.1136/bmj.f7668.
18. Beller EM, Glasziou PP, Altman DG, Hopewell S, Bastian H, et al. PRISMA for Abstracts: Reporting Systematic Reviews in Journal and Conference Abstracts. 2013, PLoS Med 10(4): e1001419. doi:10.1371/journal.pmed.1001419
19. Pilkington K, Boshnakova A. Complementary medicine and safety: A systematic investigation of design and reporting of systematic reviews. Complement Ther Med(2011), doi:10.1016/j.ctim.2011.10.002
20. Liberati A, AltmanD, Tetzlaff J, Mulrow C, Gøtzsche P, Ioannidis J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700
21. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, Schulz KF and the CONSORT Group (2008) CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. PLoS Med 5(1): e20. doi:10.1371/journal. pmed.0050020

22. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, et al. (1994) Improving the quality of reporting of meta-analysis of randomized controlled trials: The QUOROM statement. Lancet 354: 1896–1900.
23. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D for the CONSORT and Pragmatic Trials in Healthcare (Practihc) group. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008; 337;a2390
24. Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. Ann Intern Med 2006; 144(5):364-367
25. Boutron I, Moher D, Altman DG, Schulz K, Ravaud P, for the CONSORT group. Methods and Processes of the CONSORT Group: Example of an Extension for Trials Assessing Nonpharmacologic Treatments. Ann Intern Med. 2008:W60-W67
26. MacPherson H, Altman DG, Hammerschlag R, Youping L, Taixiang W, White A, Moher D; STRICTA Revision Group. Revised STandards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement. PLoS Med. 2010 Jun 8;7(6):e1000261
27. Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA. 2001;285:1992-5. [PMID: 11308436]
28. Ioannidis J, Evans S, Gotzsche P, O'Neill R, Altman D, Schulz K, Moher D. Better Reporting of Harms in Randomized Clinical Trials: an Extension of the CONSORT Statement. Annals of Internal Medicine, 2004; 141 (10):781-788.
29. Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. BMJ 2004; 328(7441):702-708
30. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJW. Reporting of non inferiority and equivalence randomized trials: An extension of the CONSORT statement. JAMA. 2006; 295:1152-1160.
31. Vandenbroucke JP; von Elm E; Altman D; Gøtzsche P; Mulrow C ; Pocock S; et al. InitiativeStrengthening the Reporting of Observational Studies in Epidemiology (STROBE). Guidelines for reporting observation studies. PLoS Med. 2007 DOI: 10.1371/journal.pmed.0040297
32. Santaguida PL, Raina P. The development of the McHarm quality assessment scale for adverse events: Delphi consensus on important criteria for evaluating harms. Available at http://hiru.mcmaster.ca/epc/mcharm.pdf2008;Accessed on May 14, 2008

33. Chou R, Aronson N, Atkins D, Ismaila AS, Santaguida P, Smith DH, et al. AHRQ Series paper4: Assessing harms when comparing medical interventions: AHRQ and the Effective Health care Program. J Clin Epi. 2010, 63:502-12.
34. Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, et al. (2012) PRISMA-Equity 2012 Extension: Reporting Guidelines for Systematic Reviews with a Focus on Health Equity. PLoS Med. 9(10):e1001333.
35. Hasson F, Keeney S, McKenna. Research guidelines for the Delphi survey technique. journal of Advanced Nursing, 2000, 32(4), 1008-1015
36. Goodman C.M. The Delphi technique: a critique, Journal of Advanced Nursing, 1987, 12, 729-734.
37. McKenna H. The Delphi technique: a worthwhile research approach for nursing? Journal of Advanced Nursing, 1994, 19, 1221-1225
38. Lynn M.R., Laman E.L. & Englebardt S.P. Nursing administration research priorities: a national Delphi study, Journal of Nursing Administration, 1998, 28(5), 7-11.
39. Boulkedid R, Abdoul H, Loustau M, et al. Using and Reporting the Delphi Method for Selecting Healthcare Quality Indicators: A Systematic Review, 2011,PLoS ONE 6(6): e20476. doi:10.1371/journal.pone.0020476
40. Powell C. The Delphi technique: myths and realities. J Adv Nursing, 2003, 41: 376–382.
41. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CFB, et al. (1998) Consensus development methods, and their use in clinical guideline development. Health Technol Assess 2: 1–88.
42. Booth A, Clarke M, Ghersi D, Moher D, Petticrew M, et al. (2011) Establishing a Minimum Dataset for Prospective Registration of Systematic Reviews: An International Consultation. PLoS ONE 6(11): e27319. doi:10.1371/journal.pone.0027319

Appendix 1: PRISMA Harms Steering committee members

Sunita Vohra (Convenor) (University of Alberta);

David Moher (Ottawa Health Research Institute);

Yoon Loke (University of East Anglia);

John Ioannidis (Stanford University);

Doug Altman (University of Oxford);

Pasqualina Santaguida (McMaster University);

Su Golder (University of York);

Jan Vandenbroucke (Leiden University).

Table 4. Delphi Rounds Results

PRISMA Item #1: Title. "Identify the report as a systematic review, meta-analysis or both"	Online round 1 N=81		Online round 2 N=72	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
1a. *Specifically mention 'harms' or other related term (adverse effect, adverse event, risk, etc).*	9 (11.1)	62 (76.5)	5 (6.9)	62 (86.1)
1b. *Clarify if both benefits and harms are examined or only.*	14 (17.3)	47 (58.0)	11 (15.3)	52 (72.2)
1c. *Mention the specific intervention being reviewed.*	4 (4.9)	76 (93.8)	1 (1.4)	69 (95.8)
1d. *Refer to specific patient group and/or conditions in which harms have been assessed.*	10 (12.3)	56 (77.7)	9(12.5)	53 (73.6)

Comments on Title: (i) It is important for the title to have these terms (harms, population, intervention). (ii) These can be done quite concisely. (iii) Including harms in the title can help to differentiate it from other systematic reviews. (iv) It needs to be clear what is meant by "harms". (v) It may be helpful to mention in the title if the review is on ADRs of RCTs or observational studies.

PRISMA item #2: Abstract. "Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number."	First round scores N=81		Second round responses N=69	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N(%)
2a. Specifically mention 'harms' or other related term (adverse effect, adverse event, risk, complications).	2 (2.5)	41 (50.6)	0 (0.0)	66 (95.6)
2b Clarify if both benefits and harms are examined or only harms.	7 (8.6)	55 (70.4)	4 (5.8)	61 (88.4)
2c. Refer to specific harms assessed.	8 (9.8)	61 (75.3)	4 (5.8)	59 (85.5)
2d. Specify what type of data was sought (i.e. clinical trials only, case reports only, all types of data).	3 (3.7)	67 (82.7)	1 (1.4)	62 (89.9)

2e. Specify how each type of data has been included.	6 (7.4)	59 (72.8)	1 (1.4)	61 (88.4)
*2f. Specify how each type of data has been appraised.***	20 (24.7)	30 (37.0)	21(30.4)	23 (33.3)

Comments on Abstract: (i) The abstract is all that many people searching PubMed will see. (ii) 2c if "all harms" are being captured this should be specified in the abstract. (iii) 2.c It is important that all the primary endpoint(s). (iv) Point 2f is already covered well enough in "study appraisal and synthesis methods". (v) A harm is an increase in an adverse outcome or a decrease in a favorable outcome. The distinctions among favorable, adverse and neutral outcomes are usually clear and universally shared, but not always, as they depend on values and circumstances. (vi) In the case of RCTs it may be helfpul to present % of ADRs with exposure vs placebo could be presented.

PRISMA Item #3:Introduction. "Describe the rationale for the review in the context of what is already known."	First round scores N=81		Second round responses N=69	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
3a. Explain rationale for addressing specific harm(s), condition(s) and patient group(s).	4 (4.9)	68 (83.9)	1 (1.4)	62 (89.8)
3b. Clearly define what events or effects are considered harms in the context of the intervention(s) examined.	6 (8.7)	59 (72.8)	2 (2.9)	62 (89.8)
*3c. Describe the rationale for type of harms systematic review done (hypothesis generating) vs (hypothesis testing).***	17(20.9)	49 (60.5)	23 (33.3)	41 (59.4)
*3d. Explain rationale for selection of study types or data sources and relevance to focus of the review.***	10 (12.3)	56 (69.1)	13 (18.8)	46 (66.6)

Comments on Introduction: (i) Strong opposition against the term "hypothesis generation": data do not generate hypotheses, people do. (ii) A better term would be "scoping" (as proposed by Yoon Loke): the activity of trying to list all possible adverse effects that are known for a particular drug. (iii) Flexibility in the study of harms is VERY important. (iv) It should address methodologic quality and clinical importance of the harm under question, should put it in context with benefit. (v) If not based only on RCTs, should include a clear statement on why reviewing lower quality of evidence was necessary. (vi) If the review covers randomized controlled human trials, "the context of what is known" should include all relevant research apart from RCT. Assessing adverse events is regulatory obligation; it is also good practice.

PRISMA item #4: Objective. "Provide an explicit statement of questions being addressed with reference to Participants, Interventions, Comparisons, Outcomes, and Study design (PICOS)."	First round scores N=81		Second round responses N=69	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
4a. Provide an explicit statement of questions being asked with reference to harms.	5 (6.1)	65 (80.2)	3 (4.3)	63 (91.3)

Comments on Objective: (i) The existing statement would cover it. (ii) A clear research question using the PICO (T = timeline) format plus study design is essential. (iii) PICOs not a good fit, doesn't really make sense for harms, would delete. (iv) PICO is just fine. If a class of drugs is studied than that will be listed as the intervention.

PRISMA item #5: Protocol Registration. "Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number."	First round scores N=81		Second round responses N=69	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
*5a. Describe if protocol was developed in collaboration with someone with clinical expertise for field/intervention under study.***	23 (28.4)	41 (50.6)	26 (37.7)	31 (44.9)

Comments on Protocol and Registration: (i) Existing PRISMA wording is sufficient. (ii) Perhaps the issue with specifying the expertise of those developing the study protocol might be important because the definition of an adverse event is variable (i.e. what patient's perceive as AE versus clinicians vs other). (iv) Authors are usually aware of some, if not much, of what are the "right" (clinically and/or scientifically) questions to be addressed by an systematic review. (v) There should be a rationale for everything that was done, whether it is in a "protocol" or elsewhere.

PRISMA item #6: Eligibility criteria. "Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale."	First round scores N=81		Second round responses N=67	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
6a. Clearly define what events or effects are considered harms in the context of the intervention(s) examined.	6 (7.4)	66 (58.1)	2 (3.0)	59 (88.0)
6b. Specify type of studies on harms to be included.	7(8.6)	67 (82.7)	2 (3.0)	63 (94.0)

Comments Eligibility criteria: (i) Covered by the standard PRISMA. (ii) It is hard to pre-specify known harms of interventions (because a lot is unknown) as compared to benefits where easier to pre-specify events. (iii) One does not have to pre-specify all possible harms, but those that can be anticipated should be specified. (iv) It's important to include events that patients consider harms. (v) It is important that systematic review authors consider measurement or ascertainment issues very carefully. Studies may report harms obtained, categorized and also analyzed by different means (such as blinded versus unblinded data collectors, or blinded data monitoring committee). For some systematic reviews these are very critical issues as the numerical data available may not reflect the true prevalence of the harm under consideration. (vi) Outcomes should usually not be part of eligibility criteria (if one wants to evaluate the impact of reporting bias).

PRISMA item #7: Information sources. "Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched."	First round scores N=81		Second round responses N=67	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
7a. State if additional sources for adverse events were searched (e.g regulatory bodies, industry) if so, describe the source, terms used and dates searched.	3 (3.7)	63 (77.7)	5(7.4)	57 (85.1)

Comments on Information sources: (i) This is covered anyway under "Describe all information sources". (ii) This is useful as it may be a prompt for authors to consider additional sources for adverse events. (iii) Strong disagreement that it would not be important to distinguish between sources for benefits and sources for harms, especially if harms is observational. (iv) Sources of benefits and sources of harm may differ in kind and quality. All sources should be clearly specified. If some sources were not used, explain why.

PRISMA item #8: Search. "Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated."	First round scores N=81		Second round responses N=67	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
8a. Present the full search strategy if additional searches were used to identify adverse events.	9 (11.1)	54 (66.6)	5 (7.4)	49 (73.1)

Comments on Search: (i) This is covered by the standard PRISMA. (ii) Many journal articles without supplementary on-line options would not permit this due to word limits. (iii) There should be no reason to want to keep the strategy hidden, and it's useful for future readers who may wish to replicate the search. (iv) Full search strategy should be available for all data sources. (v) Extremely important point.

PRISMA item #9: Study selection. "State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)."	First round scores N=81		Second round responses N=66	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
9a. State what study designs were eligible and provide rationale for their selection.	8 (9.8)	64 (79.0)	4 (6.1)	61 (92.4)
9b. Define If studies were screened based on the presence or absence of harms related terms in title or abstract.	14 (17.3)	51 (62.9)	12 (18.2)	47 (71.2)

Comment on Study selection: (i) 9.a This is covered by the standard PRISMA. (ii) 9.b is very important because is an important source of bias. (iii) Harms are usually not mentioned in the abstract, so this should not be a criterion. (iv) 9b is obviously not good practice, but if it is done like this it should be clear to readers. (v) It is quite important that full papers for trials that DONT report adverse events in the abstract are still screened for adverse event reporting. (vi) It will be interesting to see what people will list as rationale for excluding studies. It should say good studies will be considered and bad studies not. Well, there are good RCTs and bad ones and there are good cohort studies and bad ones.

PRISMA item #10: Data collection process. "Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators."	First round scores N=81		Second round responses N=66	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
10a. Describe method of data extraction for each type of study or report	14 (17.3)	53 (65.4)	10 (15.1)	47 (71.21)

Comments on Data Selection process: (i) Covered by standard PRISMA. (ii) Level of detail suggested above may be excessive. (iii) It might be a bit tricky the extraction from grey literature sources such as those from regulatory bodies (i.e. FDA).

PRISMA item #11: Data items. List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made."	First round scores N=81		Second round responses N=65	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
*11a. List and define variables for which data was sought for individual therapies. This item received scattered votes.***	13 (16.0)	45 (39.1)	15 (23.1)	41 (63.1)
*11b. List and define variables for which data was sought for patient underlying risk factors. This item received scattered votes.***	13 (16.0)	38 (46.9)	17 (26.1)	34 (52.3)
11c. List and define variables for which data was sought for practitioner training/qualifications ◊◊	29 (44.6)	23 (28.4)	34 (52.3)	13 (20.0)
11d. List and define harms for individual therapies	14 (17.3)	47 (58.0)	11 (16.9)	47 (72.3)

Comments on Data Items: (i) The existing item 11 is enough. (ii) It is limited to expected, known adverse events. (iii) The critical variables required to analyze harms include nature of intervention; nature of harm; dose of intervention; time course and susceptibility of those harmed.

PRISMA item #12: Risk of bias in individual studies. "Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis."	First round scores N=81		Second round responses N=64	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
12a. For uncontrolled studies, describe if causality between intervention and adverse event was adjudicated and if so, how.	10 (12.3)	58 (71.6)	11 (17.2)	45 (70.3)
12b. Describe risk of bias in studies with incomplete or selective reporting of adverse events.	10 (12.3)	55 (67.9)	10 (15.6)	48 (75.0)

Comments on Risk of Bias in individual studies: (i) Causality assessment is very important, but impractical (and not validated) for the most settings. (ii) Assessing causality in studies without controls, assumes subject matter knowledge. However, knowing what the bias is in studies with incomplete reporting is indeed nigh to impossible. (iii) It's not clear how authors will know if there was incomplete or selective reporting of adverse events. (iv) Risk of bias in this PRISMA item reflects predominantly allocation concealment or overall quality of studies. The equivalent of 'concealed randomization' in

observational research on harms is the question to what extent there was allocation bias for this particular harmful outcome due to confounding by indication or contra-indication. (v) Main bias is under-reporting (information bias). (vi) Systematic reporting of all adverse events (vii) The risk of bias concept should be replaced with a concept focusing on magnitude and direction of biases.

PRISMA item #13: Summary Measures. "State the principal summary measures (e.g., risk ratio, difference in means)."	First round scores N=81		Second round responses N=63	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
13a. If rare outcomes are being investigated, specify which summary measures will be used (e.g. event rate, events/person-time).	4 (6.5)	67 (82.7)	3 (4.7)	58 (95.1)

Comments on Summary Measures: (i) This is important, but is covered in the original #13. (ii) Since this is observational research, better to stress that the RR or OR presented are adjusted for covariates/confounders. (iii) You may also need to make some specificatons of "composite harm outcomes". In many studies because the events are very rare, the event rates reported represent several harms a combined event rate that is reported. (iv) The addition actually has a deeper level of questioning what is now in item 13: it does specifically demand absolute risks. See also STROBE item 16. This does not only refer to rare outcomes. (v) Potential item and selected comments pertain to conduct/quality of a harms review not reporting. Consider these for AMSTAR harms extension.

PRISMA item #14: Synthesis of results. "Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2) for each meta-analysis."	First round scores N=81		Second round responses N=63	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
14a. Describe statistical methods of handling data with zero events in included studies.	8 (9.8)	62 (76.5)	8 (12.7)	52 (82.5)

Comments Synthesis of Results: (i) The handling of 'zero' events, whatever the 'zero' means (real zero, or not reported) is very important for adverse effects. (ii) Zero event studies usually have zero power. Statisticians are purists and ask for this though, clinically irrelevant. (iii) Not always zero event studies have zero power It depends from what we are looking for. Zero event tell you about the absence of harm (assuming complete reporting). (iv) In harms studies, this is extremely important, since a 'zero' may mean any of three things: not reported, not deemed important (= 'non-significant'), or really zero. (v) Zero events have a confidence interval of zero to three, so it is important to know how they were handled.

PRISMA item #16: Additional analyses. "Describe methods of additional analysis (e.g.; sensitivity or group analysis, meta-regression), if done, indicate which were pre-specified	First round scores N=81		Second round responses N=63	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
*16a. Describe additional analysis for studies with high risk of bias.***	24 (29.6)	45 (55.5)	17 (26.9)	39 (61.9)

Comments on Additional analysis: (i) Covered in original #16. (ii) Subgroup and meta-regression analysis should always be done, if possible. These are not "additional analyses." They are indispensable to a well-done meta-analysis.

PRISMA item #17: Study Selection. "Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram"	First round scores N=81		Second round responses N=63	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
*17a. Provide process or flow for each type of study design.***	18 (22.2)	47 (58)	23 (36.5)	33 (52.4)

Comments on Study Selection: (i) I believe that one flow-chart is enough, if a meta-analysis investigates different study designs, but it is important to separate different study designs in the bottom of the flow diagram. (ii) Not necessary to be subdivided by study type. (iii) Disagree strongly.

PRISMA item #18: Study characteristics. "For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations."	First round scores N=81		Second round responses N=63	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
18a. Present study characteristics such as patient demographics or length of follow-up that may have influenced the risk estimates for the adverse outcome of interest.	5 (6.2)	65 (80.2)	2 (3.2)	56 (88.8)
18b. Describe method of collecting adverse events in included studies (e.g.: patient report, active search).	3 (3.7)	68 (83.9)	2 (3.2)	57 (90.5)
18c. For each primary study, list and define each harm addressed and how harm is identified.	12 (14.8)	48 (59. 2)	11 (17.4)	46 (73.0)

Comments on Study Characteristics: (i) 18a covered by standard PRISMA. (ii) 18c is quite burdensome and requires space. (iii) Given that sources of heterogeneity might be the most important characteristics to study in a meta-analysis of harms, extensive reporting is necessary. (iv) 18b is particularly important. Many studies don't make this clear, however. (v) These are more to do with conduct and quality assessment of the review.

PRISMA item #26: Conclusions. "Provide a general interpretation of the results in the context of other evidence, and implications for future research."	First round scores N=81		Second round responses N=63	
	≤5 N (%)	≥8 N (%)	≤5 N (%)	≥8 N (%)
26a. Provide balanced discussion of harms with emphasis on study limitations, generalizability, and other sources of information on harms	9 (11.1)	64 (79.0)	2 (3.2)	58 (92.0)

Comments on Conclusion: (i) The current PRISMA item is already quite clear. (ii) I think this is important as it prompts consideration of other sources of information on harms.

** scattered votes; ◊◊ excluded item.

Table 5. Comments for PRISMA Items

PRISMA item #15: Risk of bias across studies ***"Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies)"***	(i)Maybe nothing "new" for harms, but we've got to be better at thinking about pub bias and other issues as they relate to harms and not just to benefits. (ii)Risk of bias is very important and cannot be overlooked. Further thought is needed, to build an extension from the existing RoB tool. (iii)Poorly designed but well-reported studies can score well on these instruments and be over-valued in an systematic reviews, when perhaps they should be discounted or ignored. (iv)Component-by-component assessment of magnitude and direction of bias is where we should be heading. (v)Is specific to and important for harms: the type of bias across studies (confounding) is a different type than with RCTs (where it usually is selective publication and covered partially by funnel plot).
PRISMA item #19: Risk of bias within studies ***"Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12)."***	(i)New items are required here. (ii)Risk of bias for harms is fundamentally different than risk of bias re: effectiveness and this must be elaborated in this PRISMA Extension. (iii)The risk of bias will differ for previously recognized and communicated harms and newly reported suspected or proven harms, so 'new' ones need to be labeled as such. (iv)Conflict of interest is important, but different from bias; they should not be mixed. One invalidates the study; the other raises questions about motives.
PRISMA item #20: Results of individual studies ***"For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot."***	(i)Potential confounders also need to be addressed. (ii)Not for 'all harms". (iii)New items are required here. (iv) What about mechanistic evidence, or the lack thereof.
PRISMA item #21: Synthesis of results ***"Present results of each meta-analysis done, including confidence intervals and measures of consistency."***	(i)When a random-effects summary is produced, report the estimated spread of the distribution of true values in an informative manner. (ii)Merely reporting an estimated value of tau-squared, with or without a confidence interval, is no more informative than it would be to report the estimated value of sigma-squared in a single study.
PRISMA item #22 Risk of bias across studies	(i)Might be most important item, and might lead to stopping the study at the systematic review stage.

"Present results of any assessment of risk of bias across studies (see item 15)."	(ii)New items are required here. (iii)Risk of bias wrt harms is substantially different than wrt effectiveness, and this must be elaborated on, in this Extension. (iv)Delete the words "of risk." Write the guideline to portray some kind of bias analysis as essential, not optional.
PRISMA item #23: Additional analysis *"Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression"*	(i)I would add a comparison that examines industry funding vs. government (grants, etc) funding. (ii)Here's where we've not been as thorough for harms as we have been for benefits. why? (iii)Subgroup analyses by stratifying for quality assessment may be useful to incorporate in the prisma statement (iv) Subgroup analyses, contrasting studies in which particular biases were addressed vs. those in which they were not addressed, might be the most important aspect of a meta-analysis of observational studies on harm.
PRISMA item #24: Summary of evidence *"Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers)."*	(i)Some have advocated for downgrading harm outcomes because they are not prespecified (ii)Limitations of evidence presented in primary studies must be explicitly discussed. Summarize the main findings according to the time as acute, accumulative or delayed harm. (iii)The key findings of a meta-analysis of harms may have to do with the assessment of causality in observational studies. (iv)Most of the time, more research is needed. Most of the time, therefore, researchers are a very important key group. (v)The main findings should be summarized by outcome types that are clinically/epidemiologically relevant. In this context, if harm is delayed in onset or due to a cumulative exposure, this can be specified. (vi)Inconsistencies across studies should be described with explanations for the inconsistencies if possible. Limitations of the contributing studies, especially those of sufficient impact to undermine or negate the value of a given study, should be described.
PRISMA item #25: *Limitations* *"Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias)."*	(i)Discuss potential conflicts of interest for drug and device studies. (ii)Harms consideration is very important here, since type II error is very different than type I error. (iii)The most important discussion of the limitations of the systematic review and meta-analysis might be to address the question whether important counterarguments against accepting causality have been addressed in individual studies, and whether they have been addressed in a sufficiently credible way - so that their results bear on all other studies. (iv)Make sure this gets in the abstracts and when industry funds all or most of the RCTs in a review don't put that fact under the title "Limitations". Instead put under this title: Caveats. (v)I am wondering if the PRISMA statement can be modified to suggest some typical limitations encountered when assessing harms. Only those that are unique to harms might be identified.

	Should be a different convention for significance for harms - do we really need to prove at the alpha = 0.05 level that a treatment is associated with a serious harm? From my experience dealing with meta-analyses of harms, all have presented babble on the limitations of the study, but none have adequately addressed the more serious issues of multiple indications, dose response, and potential confounders. When these issues have been noted, they have typically been dismissed as not being statistically significant, even though the statistical analyses were not adequately powered. (vi) The concept of a low type I error to guard against false positives is important in the controlled trials world where the purpose of most studies is to demonstrate a differential effect. I agree with the comment that p<0.05 is too rigid a bar for safety, especially as a policy decision-threshold. In the world of safety, I have argued (unsuccessfully) that type II error, to guard against a false negative, should be <0.05, thereby affording symmetric handling of benefits and harms. (vii)Nothing is proven, either way, at any alpha level. (viii)The systematic review authors should be obliged to make a general comment on the nature of AA reporting - vague statements about no harms?? Explicit reporting of numbers?? The more this is discussed, the more cinical trialists will start to record and report them better. (ix)The most important limitations should be discussed and their potential impact on results explored.
PRISMA item #27: FUNDING ***"Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review."***	(i)Be sure to put this in the abstract because that's all most clinicians will read. (ii)This is a bit of challenge for systems that undertake pharmacovigilance. Although studies will indicate the source of funding...information from some of these databases is a bit convoluted (as some are indeed industry funded). Some explanation about this exception may be helpful. (iii)If possible, even more important than for RCTs, given that the potential for 'fudging' about harms, either in RCTs or in observational research is even larger than for RCTs.

Chapter 4

PRISMA Harms: improving harms reporting in systematic reviews

PRISMA Harms: improving harms reporting in systematic reviews

Zorzela L (1), Loke YK (2), Ioannidis JP (3), Golder S (4), Santaguida LP(5), Altman DG (6), Moher D(7) , Vohra S(1).

(1)University of Alberta; (2)University of East Anglia; (3) Stanford University; (4) University of York; (5) McMaster University; (6) University of Oxford; (7) Ottawa Health Research Institute.

Abstract

Introduction: For any health intervention accurate knowledge of both benefits and harms is needed. Systematic reviews may compound poor reporting of harms in primary studies by failing to report harms or doing so inadequately. While the PRISMA Statement helps systematic review authors ensure complete and transparent reporting, it is focused mainly on effectiveness. Thus, PRISMA Harms has been developed to improve harms reporting in SRs, promoting a balanced assessment of benefits and harms.

Methods: We followed the development strategy endorsed by the EQUATOR Network and existing reporting guidelines including the PRISMA Statement and PRISMA for Abstracts. After the development of a draft checklist of items, a modified Delphi process was initiated. The Delphi consisted of three rounds of electronic feedback followed by an in-person meeting.

Results: The PRISMA Harms checklist contains 4 essential reporting elements to be added to the original PRISMA Statement to improve harms reporting in systematic reviews. The PRISMA Harms items are: Title: "Specifically mention 'harms' or other related terms, or the harm of interest in the review."; Synthesis of results: "Specify how zero events were handled, if relevant."; Study Characteristics: "For each study, present the harm addressed, how and when it was ascertained (e.g. passive report, active search, use of validated tools)"; and Synthesis of results "Describe any assessment of possible causality".

PRISMA Harms identifies a minimal set of items to be reported when reviewing harms. This guideline extension is intended to improve harms reporting in systematic reviews, whether harms are a primary or secondary outcome.

Conclusion: A balanced assessment of benefits and harms is needed. The PRISMA Harms is a step towards better reporting of harms in reviews. The endorsement of the PRISMA Harms by peer review journals, the Cochrane Collaboration and review authors is paramount for effective implementation.

Introduction

Evidence-based medicine seeks the best, unbiased evidence to make appropriate decisions to improve patient care. For optimal care, accurate knowledge of both benefits and harms is needed.

Well designed, conducted, and reported randomized controlled trials (RCTs) can paint an incomplete and potentially biased picture due to their emphasis on efficacy results combined with their inadequate reporting of harms (1-10). A single trial is usually not powered for the assessment of adverse events; unless this is explicitly acknowledged, a misconception may be perpetuated that a given intervention is safe, when its safety is actually unknown (4-10).

Systematic reviews compound poor reporting of harms in primary studies by failing to report harms or doing so inadequately (11-19). Such reviews can be misleading as they do not represent the true risk-to-benefit assessment of a given treatment (4, 8, 14-17). As many harms are rare and not typically the primary outcome of included studies, the search strategy, eligibility of study designs and statistical methods may differ from reviews of efficacy.

Harms are the primary outcome assessed in less than 10% of reviews (11-14). In 1994, only five reviews retrieved from Database for Abstracts of Reviews of Effects (DARE) and the Cochrane Database of Systematic Reviews (CDSR) were specifically designed to address an unintended effect of an intervention. This number has increased over time. In 2010, 104 reviews retrieved from CDSR and DARE evaluated adverse events exclusively; this increase may reflect the higher number of systematic reviews in general as the proportion of reviews of harms in comparison to efficacy reviews remained stable at 5% during this period (13).

A recent review (12) of systematic reviews from CDSR and DARE identified 296 DARE reviews and only 13 Cochrane reviews with a singular primary intent to measure adverse events of interventions. Even though systematic reviews increasingly try to consider all outcomes (both beneficial and harmful), data on adverse events may be more fragmented and incomplete, and given more cursory treatment than efficacy data. Some reporting deficiencies identified in reviews exclusively designed to measure harms included: lack of a clear definition of the adverse event reviewed; lack of specification regarding study designs selected for inclusion; and no report on length of follow up or measurement of any associated participant risk factors (12).

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) Statement (1) was published in 2009 to offer guidance to review authors, peer reviewers, and journal editors when reporting a systematic review or meta-analysis of RCTs. PRISMA helps authors ensure they report in a complete and transparent fashion. PRISMA for Abstracts have been completed (21) and PRISMA for protocols of systematic reviews and meta-analyses and network meta-analyses are under publication consideration. Other extensions are underway, including individual participant data meta-analyses.

Development of PRISMA Harms

We followed the strategy developed by the EQUATOR Network (20) and used for previous reporting guidelines including PRISMA Statement (1) and PRISMA for Abstracts (21). The primary step in developing a reporting guideline is documenting the current reporting and possible room for improvement, supporting the need for a guideline, usually through surveys of the published literature. To document the need for a reporting guideline specific for reviews measuring harms, our team completed a review to evaluate the reporting characteristics in reviews of adverse events a primary outcome and found several areas (tile, abstract, methods, results, and conclusion) that could benefit from more transparent reporting (12).

A list of potential new items was developed based on preliminary findings of previous systematic reviews (12, 25). These new items were then compared against the original PRISMA Statement to assess overlap and refinement. The wording and content were further refined by the PRISMA Harms steering committee.

After the development of a draft list of potential items, a modified Delphi process was initiated to obtain feedback from a broad spectrum of stakeholders. The Delphi method is an iterative process of obtaining information from a group of experts (26-28). We surveyed 352 people through an online survey consisting of three rounds of participant feedback. A total of 166 participants responses were received among the three rounds (24). The in-person two-day consensus meeting was held in Banff, Canada, May 2012. It included 25 experts from seven countries and with extensive experience in systematic reviews, harms, and guideline development. Meeting participants received the results of the review-base identifying the current state of reporting in reviews of harms (12), the results of the Delphi technique (26) and it was followed by the discussion of relevant items to be included in this guideline.

Scope of PRISMA Harms

A goal of the in-person consensus meeting was to define the guideline applicability. After discussion, it was agreed upon that the aim of this guideline is to improve harms (see Glossary of terms) reporting in systematic reviews. This report is an extension of the existing PRISMA Statement (1) and, as such, should be used in addition to the original PRISMA statement; its main goal is to bring attention to a minimal set of items to be reported when reviewing harms. This guideline extension is intended to improve harms reporting in systematic reviews, whether harms are a primary or secondary outcome. In this extension, the PRISMA Harms, the terms "harm" and "adverse event" are used interchangeable throughout the manuscript.

In order to provide rationale and promote understanding, we have identified examples of good reporting for each item from published systematic reviews.

How to use PRISMA Harms Checklist

The PRISMA Harms checklist contains four items to be described when reporting harms in systematic reviews. The additional items are added to the original PRISMA Statement, such that a systematic review addressing adverse events should report the PRISMA Statement items *and* the PRISMA Harms. Additional explanation and elaboration is provided for the 4 new harm items, as well as the remaining original PRISMA checklist items.

Item 1-Title

PRISMA Statement: ***"Identify the report as a systematic review, meta-analysis, or both."***

PRISMA Harms: ***"Specifically mention 'harms' or other related terms, or the harm of interest in the review."***

Examples:

"Perinatal mortality and other severe adverse pregnancy outcomes associated with treatment of cervical intraepithelial neoplasia: meta-analysis". (a.1)

"Benefits and harms of statin therapy for persons with chronic kidney disease: a systematic review and meta-analysis."(a.2)

The title should clearly reflect the objectives of the review; it is accessible to all readers and it should provide a "one line summary" of author's intention. If adverse events are part of the review's objective, the title should state this clearly. It can specify the adverse event under review or any other harms-related terms, for example: *risk*, *complication*, *adverse effects*, *adverse reaction*, etc.

It must be clear to the reader that the review includes evaluation of harms. If the harm is a co- primary outcome, for example the measurement of efficacy and harms, the title should state so.

It is important to recognize that some terms, for example 'mortality' or 'morbidity' can be used as an adverse event or a measurement of an intended effect (e.g. mortality reduction). In this situation, review authors' intention should be made clear to the reader. Intention is important, not only to promote clarity, but also because reviews of adverse events have different search methods and study selection criteria than reviews of efficacy. Reviews of efficacy have their search strategy and screening based on the efficacy question [Does the intervention provide the intended effect? (e.g., cure the disease)] and usually cannot address harms adequately (e.g., what are the unintended effects (harms) associated with the intervention?). Reviews of harms may require a distinct database filter and data sources searches (16, 28-32).

Item 2-Abstract

PRISMA Statement: "*Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.*"

Example:

Background: Bisphosphonates are used for the prevention and treatment of osteoporosis, but there have been concerns about a potential link between bisphosphonate therapy and atrial fibrillation. Data on the effects of bisphosphonate on the risk of atrial fibrillation are conflicting and the association of serious atrial fibrillation (defined as events resulting in hospitalization or disability or judged to be life-threatening) with the use of bisphosphonates is uncertain. Hypothesis: We aimed to systematically evaluate the association of bisphosphonate use with the risk of atrial fibrillation. Methods: We performed a systematic literature search for clinical trials using bisphosphonates and providing data on the outcome of atrial fibrillation. Four randomized controlled trials and 3 population based case control

studies were included in the final analysis. A meta-analysis was performed with the randomized controlled trials to determine the risk of serious atrial fibrillation. Results: For the purpose of meta-analysis, the studies were homogenous; therefore the Mantel-Haenszel fixed-effect model was used to calculate combined relative risk (RR). A two-sided alpha error of less than 0.05 was considered to be statistically significant (p<0.05). Four studies with 26126 postmenopausal women were included in the meta-analysis. Meta-analysis revealed that serious atrial fibrillation occurred more frequently in the bisphosphonate group compared to the placebo group (RR 1.525; 95% CI, 1.166 to 1.997; p=0.002). Two out of 3 observational studies indicated a statistically significant increase in the risk of atrial fibrillation with bisphosphonate therapy. Conclusions: Bisphosphonate use is associated with a significant increase in the risk of serious atrial fibrillation in postmenopausal women. (b)

Recommendations for reporting harms in reviews: Abstracts are often the only text of a review accessible to all readers. The abstract of a systematic review should be easy to read as they reach a broad audience (33). The PRISMA for Abstracts (21) recommends reporting results of the primary outcome measured (benefits or harms). To this, we recommend reporting if adverse events are primary, co-primary, or secondary outcome measure, as well as the present the findings for the adverse outcome. If a broad review has been undertaken of harms associated with the intervention, it also should be clearly presented.

Item 3-Introduction

PRISMA Statement: "***Describe the rationale for the review in the context of what is already known.***"

Example:

Epilepsy has a prevalence of 5–10 persons/1000. During pregnancy, women with epilepsy cannot generally safely discontinue their antiepileptic therapy, and the risks to the unborn child from maternal antiepileptic medication need to be balanced against the risk of uncontrolled epilepsy both to the mother and the baby. In the last decade, an increasing number of studies have addressed the long-term safety of these drugs on child development, with conflicting results. Synthesizing these data into an overall risk assessment is critical for clinical counselling of women with epilepsy and their families. The objective of this study was to perform a systematic review of the literature pertaining to long-term neurodevelopment after in utero exposure to antiepileptic drugs (AEDs) and to conduct a meta-analysis to allow overall risk estimation. (c)

Recommendations for reporting harms in reviews: The introduction should be written to inform the reader of the reviews' overall goal and to provide justification for the approach taken (1). Systematic reviews of harms can be designed with a narrow focus evaluating a specific type of adverse event or with a broad focus to evaluate all adverse events associated with a given intervention. In either case, the introduction should inform and explain the reason for the approach chosen (27).

The review should clearly describe what events are considered harms and provide a clear rationale for the specific harm(s), condition(s) and patient group(s) included in the review.

Item 4-Objective

PRISMA Statement: "Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS)."

Example:

Our objective was to systematically determine the comparative effects of rosiglitazone and pioglitazone on cardiovascular outcomes (myocardial infarction and congestive heart failure) and mortality from observational studies in patients with type 2 diabetes. (d)

Recommendations for reporting harm in reviews: The objective of a review should be clearly stated, preferably at the end of the introduction (1). The PRISMA statement (1) suggests the PICOS format (patient, intervention, comparison, outcomes and study design). Overall, the PICOS format should be filled in, although in systematic reviews of harms the selection criteria for P, C, and O may be very broad. For example, the same intervention may have been used for heterogeneous indications in a diverse range of patients so that different comparisons may be used. Similarly, if a review is attempting to evaluate any or all possible harms (including new or unexpected events) associated with a given intervention, the potential outcomes (= the harm) cannot be completely defined *a priori.*

Item 5-Protocol and Registration

PRISMA Statement*: "Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number."*

Example*:*

With more vaccines becoming available worldwide, vaccine research is on the rise in developing countries. To gain a better understanding of safety reporting from vaccine clinical research in developing countries, we conducted a systematic review in Medline and Embase (1989–2011) of published randomized clinical trials (RCTs) reporting safety outcomes with ≥50% developing country participation (PROSPERO systematic review registration number: CRD42012002025). (e)

No specific additional information is required for reviews of adverse events.

Item 6-Eligibility criteria

PRISMA Statement*: "Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale."*

Example:

Our inclusion criteria for RCTs were: (1) parallel-group RCT of at least 24 weeks duration; (2) participants with COPD of any severity; (3) fluticasone or budesnonside (which are licensed for use in COPD) as the intervention vs a control treatment, in which the comparison groups consisted of ICS vs placebo, or ICS in combination with LABA vs a LABA alone; and (4) outcome data (including zero events) on fracture adverse events. We also evaluated controlled observational studies (case control, prospective cohort or retrospective cohort) reporting on the risk of fractures with any ICS exposure compared to those without ICS exposure in COPD. Eligible studies had to present odds ratios, or relative risk/hazard ratios, or sufficient data to enable us to calculate the odds ratio. (f)

<u>*Recommendations for reporting harms in reviews:*</u> Population characteristics are important when considering harms and should be clearly reported. Special attention should be given to the report of the eligibility criteria based on the adverse event. The review author should report how they handled relevant studies when the outcomes of interest were not reported (i.e. the primary study did not mention adverse events, but it was a relevant study based on the population, intervention and comparator. Explicit review methods will provide readers important information regarding if the review might be affected by missing outcome data or missing studies. In general, it is important to know how many studies were excluded because of missing information.

Also relevant is transparent reporting regarding the review author's rationale for choosing specific study designs (e.g., limiting the review to RCTs only vs. including other study designs). Depending on the adverse event under investigation, different study designs have different strengths and weaknesses, and appropriate study designs should be reported for the particular outcome of interest (16).

Whatever methods are chosen by review authors to determine which studies are included should be explicit, allowing readers to better understand how adverse events were identified and measured.

Review authors should also define if studies were selected based on the length of follow-up. The outcome under consideration may take longer to occur than the usual time required to measure efficacy of an intervention (e.g., hospital readmission for total hip replacement). An appropriate interval for follow-up should be specified *a priori* for each type of adverse event evaluated by the review author. If the timing of the collection of outcomes in the primary study is insufficient (relative to the time interval necessary for the outcome to occur), then there is a risk of underestimating the number of events (1, 12).

Item 7-Information sources

PRISMA Statement: ***"Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched."***

Example:

Additional data were requested from the following reporting schemes: Adverse Drug Reactions Advisory Committee (ADRAC), Australia; Bundesinstitut fur Arzneimittel und Medizinprodukte (BfArM), Germany; US Food and Drug Administration (FDA); and the Medicine and Healthcare products Regulatory Agency (MHRA), UK. The WHO Collaborating Centre for International Drug Monitoring, Uppsala, Sweden (WHO-UMC) was also requested to provide the total numbers of adverse events reports received up until September 2007 involving the use of S. repens. Twenty-four manufacturers/distributors of S. repens preparations were identified from a review, standard text and from Internet searches. They were contacted and asked for adverse event reports and any other safety information held on file. Four herbalist organizations (British Herbal Medicine Association, UK; European Herbal & Traditional Medicine Practitioners Association, UK; European Scientific Cooperative on Phytotherapy, UK; National Institute of Medical Herbalists, UK) were also contacted for relevant information. (g)

Recommendations for reporting harms in reviews: Review authors should be explicit if they only searched for published data, or also sought data from unpublished sources, from authors, companies or regulatory agencies. Published data can differ significantly from unpublished data for various reasons, especially in relation to harms (34-37). If a review includes unpublished data, a clear description should be provided of the source of the data and the process of obtaining them; if any databases were searched, this should be stated along with the dates, and the search terms used, in sufficient detail that the search could be replicated.

Item 8-Search

PRISMA Statement: ***"Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated."***

Example:

We reviewed all adverse drug events reported through MedWatch or those submitted by the manufacturer from November 1997 to April 2008 through the Freedom of Information Act (FOIA) request. The FDA provided a full-text summary of 5944 reports involving oral, intramuscular and IV use of haloperidol. The FDA data were transferred to a Microsoft Access database and screened for the key terms torsade, QT, prolongation, wave. Incident report number, date of report, age, gender, origin of report, medication name, role of drug as categorized by the FDA (e.g., suspect, concomitant, primary suspect, secondary suspect), route, dose, units, duration, symptoms and FDA outcome category (e.g., death, life threatening, hospitalization initial or prolonged, disability, congenital anomaly, required intervention to prevent permanent damage, other) were recorded. Only those reports in which IV haloperidol was considered by the reporter to be the primary causative agent for the adverse event were reviewed. Available information included diagnosis, laboratory parameters, QTc measurement, cardiac symptoms, outcomes and a description of recovery. No peer review was applied to the MedWatch reports and the data reported in this publication reflect the original information from the FDA MedWatch database. (h)

Recommendations for reporting in harm reviews: If additional searches were used specifically to identify adverse events (e.g., regulatory bodies, industry) present the full search process, so it can be replicated.

Item 9-Study selection

PRISMA Statement: ***"State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)."***

Example*:*

The titles and abstracts of all retrieved articles were read (D.A.), and those of no clear relevance were eliminated. Full copies of all the remaining studies were obtained and read independently (D.A. and W.R.). In case of interrater variability, D.A. and W.R. discussed the differences and made a joint decision. Those reporting quantitative data for serious adverse events such as infection using the infection criteria used by the original authors (deep and superficial), bleeding, neurologic injury (persistent, transient, and of unknown duration), granuloma, or death associated with the intrathecal catheter system were included in an initial list of studies. We then selected those reporting at least 20 patients, with a median catheter duration of at least 7 days, quantitative data for the previously mentioned serious adverse events, and an external catheter system. Port-a-cath based systems were excluded. (i)

Recommendations for reporting harms in reviews: Review authors should specify if only studies reporting on adverse events were included in the review. There may be many studies that fulfill the criteria for population, intervention and comparisons, but harms (i.e., the outcome of interest) are not reported, as noted previously in item 6. Lack of reporting, as opposed to lack of occurrence, is important information and should be clearly reported by the reviewer.

If the review only includes studies reporting on adverse events of interest, it should be clearly defined if screening was based on adverse event reporting in title/abstract or full text. It is important to note that harms reporting are especially poor in titles and abstracts, leading to the potential exclusion of many relevant studies. Moreover, studies may not report harms in the full text, even though such data have been collected, and may be obtained upon request.

Item 10-Data collection process

PRISMA Statement: ***"Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators."***

Example:

Data were extracted using a standardised form by one author (ARW). Beside general information about the article, the study design was noted, including randomised controlled trials, prospective observational studies, retrospective studies or studies using national databanks. The assessed outcomes, the characteristics of the populations studied including disease activity and severity, rheumatoid factor and anti-CCP antibodies, comorbidities such as diabetes mellitus, chronic lung diseases or history of severe infections were collected. The proportion of patients receiving LD-CT in studies were assessed with the current daily dose used in milligrams of prednisone, and if mentioned, the cumulative dose in grams of prednisone and the duration of LD-CT therapy. Concomitant therapy including, DMARDs or biologic agents were noted. (j)

No specific additional information is required for reviews of adverse events.

Item 11-Data items

PRISMA Statement: ***"List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made."***

Example:

Anti-TNF groups were divided into drug and dose categories. Dose was defined according to the recommended maintenance dose from the product labelling. Only recommended and high doses were considered in the analysis. The number of subjects experiencing death or at least one serious adverse event or serious infection was extracted for each treatment group. Extracting malignancy data from published clinical trial manuscripts requires caution as there is considerable variation in reporting, especially in the reporting of carcinoma in situ and non melanoma skin cancers. As manuscripts may aggregate malignancies differently, malignancies were allocated to three classes allowing for comparisons of similar outcomes: lymphomas; non-melanoma skin cancers and the composite endpoint of non-cutaneous cancers and melanomas. If a subject presented with two types of cancer, the cancers were allocated as a single event in the following order of priority: lymphoma, non-cutaneous cancer/melanoma, nonmelanoma skin cancer. When the number of events instead of the number of subjects experiencing an event was reported, an assumption of one event per subject was made. All data were abstracted as reported in the publications. If an event described in a publication could not be

allocated to a particular time or treatment group other sources of information were used. All data were compiled by two authors (TRE and JPL) and disagreements were resolved by consensus. (k)

Recommendations for reporting harms in reviews:

Categorization: Harms maybe categorized in a heterogeneous fashion by primary study authors. For example, when investigating hemorrhagic stroke, some primary studies may report a combination of events under 'neurologic events', others may report under 'cardiovascular' events, a few may report as 'stroke' but not subdivide further (i.e., hemorrhagic or ischemic). These issues should preferably be considered *a priori.* All operational definitions used to classify adverse events under review should be explicitly identified by review authors. For example, while harms can have different intensities, the definitions of seriousness used can vary significantly between studies. Review authors should explain how they extracted data and, if relevant, combined events.

Events/Participants: Studies usually report on the number of events but these may not accurately reflect the number of participants experiencing the event. For example, the same patient may have experienced angina, then myocardial infarction and finally cardiovascular related death. Studies may report these three events as isolated findings (angina, MI, death), but they all occurred in one single participant during the course of the study. Participants may also experience the same event multiple times.

Factors associated with the event: When reporting on adverse events, it is important to consider if any risk factors associated with participants (e.g., participants age, gender, use medications) or provider (e.g., years of practice, level of training) were taken into consideration. Review authors should identify the specific information sought, the definition used for risk factors/professional skills, and what was done if this information was not reported within the original study.

The method of harms measurement may lead to different results. Active methods (actively seeking confirmation for harms) are associated with more reported events than passive methods (waiting for patients to report them) (23). The timing and frequency of adverse events measurement is also important (e.g., if done at only at the end of the study intervention (when participants may not accurately recall how they felt during the entire course of the study) vs. at regular intervals throughout the treatment period) (4, 9).

Poor/unclear reporting in primary studies should be anticipated and the approach used to overcome them reported in the systematic review protocol.

Item 12-Risk of bias in individual studies

PRISMA Statement: ***"Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis."***

Example:

Two reviewers independently assessed the reporting of blinding, allocation concealment, withdrawals and the loss to follow-up in RCTs. In accordance with recommendations on assessing adverse effects, we extracted information on participant selection, ascertainment of exposure and outcomes, and methods of addressing confounding in observational studies. (l)

Recommendations for reporting harms in reviews: Studies that are well designed to assess efficacy of an intervention may not preserve the same qualities when assessing harms (23, 27). The risk of bias assessment should be considered separately for outcomes of benefit and harms.

Item 13-Summary measures

PRISMA Statement: ***"State the principal summary measures (e.g., risk ratio, difference in means)."***

Example:

We initially calculated the required information size to detect or reject an intervention effect of a 10% relative risk (RR) increase, with a risk of type II error of 20% and power of 80%. We chose a 10% RR increase because even this small increase in cancer risk is likely to be clinically meaningful. If the accumulated information size exceeded the required information size, we then tested for a 90% power and then for a 7·5% or 5·0% RR increase. Our intention was to have sufficient power to test for even a very small increase in risk.(m)

No specific additional information is required for reviews of adverse events.

Item 14-Synthesis of results

PRISMA Statement: "Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2) for each meta-analysis."

PRISMA Harms: *Specify how zero events were handled, if relevant.*

Example:

In cases when only one study reported an adverse event, relative risks (RR) were calculated with the corresponding 95% confidence intervals (95% CI) using the data extracted. When in one cell there were zero cases, 0.5 was added to all four cells of a 2x2 table. When more than one study described an adverse event, the association estimators obtained in each of the studies was summarized via meta-analysis. (n)

Recommendations for reporting harms in reviews: As major harms are often rare events (especially serious events, such as suicide in trials of depressed patients), it is common to find studies having "zero" as the reported frequency of the specific harm. This situation would require systematic reviewers to consider relevant statistical issues, ideally *a priori* and documented in the systematic review protocol. The review authors should clearly report the steps taken to overcome problems associated with reporting "zero" events.

Harms are not typically the primary outcome of primary studies of efficacy. Review authors should plan and specify, preferably documenting in a registered protocol, how they will deal with: (i) studies not reporting on harms of interest, (ii) studies reporting a general statement indicating the absence of the event (e.g., 'no serious harms were identified in any group', but with definitions of seriousness not provided), or (iii) no report on adverse events was provided. The review author should specify if the situation of "no events reported" was treated as "zero events" or handled in some other manner and whether any effort was made to clarify this ambiguity with the authors of the primary studies or others having access to the original data. Consideration should also be given to whether reports of "no events" can be combined with studies that clearly reported on the harms or their absence. The chosen approach and rationale should be clearly reported in the methods of the systematic review.

Zero events- review authors should report how the absence of adverse event is dealt with in the included studies. Lack of reporting should not be interpreted or presented as equivalent to the absence of adverse events. Clarification should be sought and reported to define whether a specific adverse event did not occur, was not adequately investigated, or was identified but not reported (1-7). The review author should clearly state how studies not reporting on the events of interest were handled, any assumptions made when the event was not reported and if further attempts were made to collect data from authors of included studies.

The statistical approach on how to deal with "zero" events should be clearly reported and anticipated, preferably in a registered protocol.

Item 15-Risk of bias across studies

PRISMA Statement: "Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies)."

Example:

The studies were generally well protected against bias. Allocation concealment and sequence generation did not present undue risk of bias in the included studies, and results on serious adverse events have been obtained from all the studies. Aalbers 2004 and Busse 2008 were open studies, and Aalbers 2004 had a withdrawal rate of over 20%. Sensitivity analysis using only the blinded studies was carried out and there was still no significant difference between the treatments. Consideration of asthma related adverse events might have been subject to bias as none of the trials used independent outcome assessment for causation of adverse events. (o)

Recommendations for reporting harms in reviews: Instead of reporting a significant harm, it is possible that some investigators may choose to downplay estimates of harm and emphasize the safety of the intervention instead. Selective outcome reporting and publication bias could therefore work in different directions to that seen with efficacy trials where significant benefit is emphasized (39). Whenever statistical approaches are used to probe for the possibility of bias, they should be explicitly described and used with caution. It is a very common misconception that tests (e.g., inverted funnel plots) can confirm that publication bias does not exist (42, 43). Application of these tests for harms outcomes creates a false impression, as bias may be in the opposite direction (in favor of "negative" results).

Assessments of risk of bias across studies may need to focus more on presenting the extent of missing information (studies without harms outcomes) and any clues as to the reasons for that missingness and whether these reasons may be related to the results of these outcomes.

Item 16-Additional analyses

PRISMA Statement: "Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified."

Example:

The risk of bias in studies included in the analyses was taken into account in the interpretation of the review's results. For all outcomes for which pooling was possible we performed sensitivity analyses for all risk of bias criteria separately. We excluded studies with a high risk of bias and studies for which the presence of bias was unclear and compared the results of studies with a low risk of bias with the results of all available studies. It was our intention to perform subgroup analyses for children and adults and for leukaemias and solid tumours, but unfortunately this was not possible. (p)

No specific additional information is required for reviews of adverse events.

Item 17-Study selection

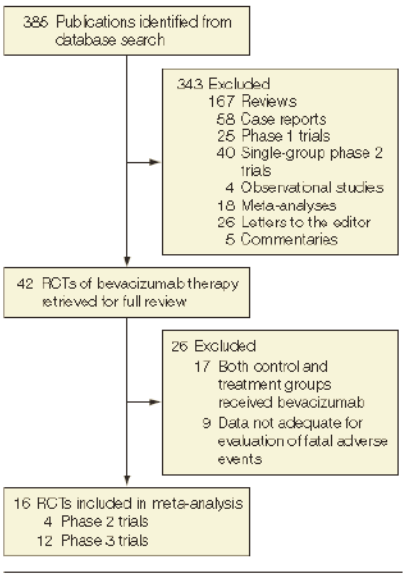
PRISMA Statement: *"Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram."*

Example:

Our search yielded a total of 275 potentially relevant studies with sorafenib or sunitinib. Initially, 175 trials were excluded for at least one of the following reasons: duplicate trials, phase I trials, review articles, observational studies, case reports, editorials, letters or commentaries. Figure 1 shows the selection process of the randomized controlled trials. After reviewing the remaining 100 publications, 10 trials were considered to be highly relevant for the meta-analysis (phase II and III trials and expanded access programs).(q1)

Example 2: (q2)

Figure 1. Selection of Randomized Controlled Trials (RCTs) Included in the Meta-analysis



No specific additional information is required for reviews of adverse events.

Item 18-Study characteristics

PRISMA Statement: ***"For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations."***

PRISMA Harms Extension: ***"Present the harms addressed, how and when it was ascertained (e.g. passive report, active search, use of validated tools)."***

Example: *(r)*

First author, year[a] [reference]	Design	Country	Type 1 diabetes			Controls		Ascertainment of Caesarean section	Available confounders[b]					
			Ascertainment method	Age at diagnosis	n	Source (matching criteria)	n		GA	MA	BW	BO	BF	MD
Dahlquist, 1992 [42]	C-C	Sweden	Swedish Childhood Diabetes Registry	0-14 years	2,757	Swedish Medical Birth Registry (year of birth, delivery unit)	8,271	Maternity record						
Patterson, 1994 [24]	C-C	Scotland	Hospital admission / childhood diabetes register	0-14 years	271	Scottish maternal discharge records (age, sex and area)	1,355	Maternity record	✓	✓	✓	✓		✓
McKinney, 1997 [25]	C-C	England	Yorkshire Childhood Diabetes Register	0-15 years	220	General practitioner's records (age and sex)	433	Questionnaire	✓	✓	✓	✓	✓	✓
Tai, 1998 [34]	C-C	China	Taipei type 1 Diabetes Registry	0-29 years (mean= 8 years)	117	Classmates or colleagues[c] (age, sex, parental and individual education)	193	Questionnaire						
Rami, 1999 [27]	C-C	Austria	Vienna type 1 diabetes register	0-14 years	114	Schools (age and sex)	495	Maternity record	✓	✓	✓	✓	✓	✓[d]
Bashe, 1999 [36]	C-C	Denmark	Hospital admission	0-14 years	839	Medical birth register (age, sex and district)	1,687	Maternity record						
	C-C	Bulgaria	West Bulgaria type 1 diabetes register	0-14 years	176	Schools and policlinics (age)	562	Maternity record	✓	✓	✓	✓	✓	✓[d]
	C-C	Latvia	Latvian type 1 diabetes register	0-14 years	143	Population register (age)	410	Maternity record	✓	✓	✓	✓	✓	✓[d]
EURODIAB, 1999 [15]	C-C	Lithuania	Lithuanian type 1 diabetes register	0-14 years	124	Policlinics (age)	369	Maternity record	✓	✓	✓	✓	✓	✓[d]
	C-C	Luxembourg	Luxembourg type 1 diabetes register	0-14 years	59	Pre-schools and schools (age)	188	Maternity record	✓	✓	✓	✓	✓	✓[d]
	C-C	Romania	Bucharest type 1 diabetes register	0-14 years	111	Pre-schools and schools (age)	342	Maternity record	✓	✓	✓	✓	✓	✓[d]
Visalli, 2003 [43]	C-C	Italy	Lazio type 1 diabetes register	0-14 years	150	Schools (age)	750	Questionnaire	✓	✓	✓		✓	✓
Stene, 2003 [17]	Cohort	Norway	Norwegian Childhood Diabetes Registry	0-14 years	1,824	Norwegian medical birth registry	1,380,778	Maternity record	✓	✓	✓	✓		✓
Stene, 2004 [30]	C-C	Norway	Norwegian Childhood Diabetes Registry	0-14 years	545	Norwegian population registry	1,668	Maternity record	✓	✓	✓	✓	✓	✓[d]
Cardwell, 2005 [32]	Cohort	Northern Ireland	Northern Ireland type 1 diabetes register	0-14 years	991	Northern Ireland Child Health Register	447,633	Maternity record	✓	✓	✓	✓	✓	✓
Šipetić, 2005 [38]	C-C	Serbia	Belgrade Hospital admission	0-16 years	105	Hospital outpatients with skin disease[c] (age, sex and area)	210	Questionnaire	✓	✓	✓	✓	✓	✓[d]

(a)Year of publication; (b) Tick denotes data recorded in study and available for analysis; (c) Not randomly selected or not population-based; (d) Maternal type 1 diabetes used in analyses

BF, Breast-feeding; BO, birth order; BW, birthweight; C–C, case–control; GA, gestational age; MA, maternal age; MD, maternal diabetes.

Reporting the characteristics of included studies is important to allow the reader to gauge the validity and generalizability of the results. As harms are not reported or measured in a standardized format, we suggest reporting the following for every included study: a) definitions for specific adverse event, b) the method of adverse events ascertainment if passive methods (patients reported a harms as emerged), or active methods (harms actively searched), c) the method of measurement (i.e., if any validated tool was used to measure them, along with appropriate reference to its validation) and d) how intensity/seriousness was measured. This information could be combined with the study characteristics or presented separately.

Recommendations for reporting harms in reviews: The PRISMA Statement suggests following the PICOS format for this item. For reviews of AEs, we would add additional characteristics as follows:

"P"- Population - add the identification of assessed patient risk factors that could lead to the outcome (harm).

"I"– Intervention - add professional expertise/skills if appropriate (for example if the intervention is a procedure)

"T" - Time - add timing of all assessments and the length of follow-up. The timing of outcome (harm) assessment will vary across studies and these differences are important to note.

Item 19-Risk of bias within studies

PRISMA Statement: ***"Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12)."***

Example:

Table 2. Quality Assessment of Included RCTs of Inhaled Corticosteroids in COPD[a]

Source	Sequence Generation	Allocation Concealment	Monitoring of AEs	Drug (No. of Subjects)	No. (%) Withdrawal Rates	No. (%) Lost to Follow-up
Aaron et al,[20] 2007	Adequate, central allocation	Adequate	Captured through monthly telephone interviews and checklist; pneumonia recorded only as SAE leading to hospitalization or death	SFC (145)	15 (10.3)	2 (1.4)
				Sal (148)	20 (13.5)	2 (1.4)
Burge et al,[21] 2000	Adequate, computer generated, stratified by center	Adequate	AEs and SAEs recorded throughout study	Flu (372)	160 (43.0)	16 (4.3)
				Placebo (370)	195 (52.7)	18 (4.9)
Calverley et al,[22] 2003	Adequate computer generated	Adequate	AE or SAE occurring during therapy	SFC (368)	89 (24.9)	8 (2.2)
				Sal (372)	119 (32.0)	8 (2.2)
				Flu (374)	108 (29.0)	8 (2.1)
				Placebo (361)	140 (38.8)	6 (1.7)
Calverley et al,[23] 2003	Unclear	Unclear	AEs recorded at 1, 2, 3, 6, 9, and 12 mo of treatment	For/Bud (254)	74 (29.1)	4 (1.6)
				For (255)	111 (43.5)	4 (1.6)
				Bud (257)	102 (40.0)	4 (1.6)
				Placebo (256)	106 (41.4)	6 (2.3)
Calverley et al,[5] 2007[b]	Adequate, schedule generated by system for central allocation	Adequate	AEs reviewed at each visit; no prospective confirmation by radiographs; pneumonia recorded as subset of exacerbations	SFC (1533)	522 (34.1)	29 (1.9)
				Sal (1521)	561 (36.9)	15 (1.0)
				Flu (1534)	587 (38.3)	24 (1.6)
				Placebo (1524)	673 (44.2)	21 (1.4)
Ferguson et al,[24] 2008	Unclear	Unclear	AEs collected at study start and end	FSC (394)	117 (29.7)	10 (2.5)
				Sal (388)	149 (38.4)	10 (2.6)
FLTA3025,[25] 2005	Unclear	Unclear	AEs and SAEs recorded at each visit	Flu (434)	147 (33.9)	NA
				Placebo (206)	79 (38.3)	NA
Hanania et al,[26] 2003	Unclear	Unclear	AE reporting at each visit	SFC (178)	53 (30.0)	NA
				Sal (177)	57 (32.2)	NA
				Flu (183)	49 (26.8)	NA
				Placebo (185)	59 (31.9)	NA
Kardos et al,[27] 2007	Adequate, centrally generated block	Adequate	AEs and SAEs recorded during trial and follow-up	SFC (507)	99 (19.5)	4 (0.8)
				Sal (487)	103 (21.1)	3 (0.6)
Mahler et al,[28] 2002	Unclear	Unclear	AEs and SAEs documented	SFC (165)	52 (31.5)	NA
				Sal (160)	45 (28.2)	NA
				Flu (168)	68 (40.5)	NA
				Placebo (181)	69 (38.1)	NA
Paggiaro et al,[29] 1998	Adequate, computer generated	Adequate	AE defined as untoward medical occurrence during treatment	Flu (142)	19 (13.3)	0
				Placebo (139)	27 (19.4)	2 (1.4)
SCO100250,[30] 2008	Unclear	Unclear	AEs and SAEs recorded after study medication administration but no later than last date after study medication administration	SFC (394)	125 (31.7)	NA
				Sal (403)	156 (38.7)	NA
SCO100470,[31] 2006	Unclear	Unclear	AEs and SAEs recorded at each study visit	SFC (518)	59 (11.4)	NA
				Sal (532)	74 (13.9)	NA
SCO40041,[32] 2008	Unclear	Unclear	AEs and SAEs monitored during therapy	SFC (92)	36 (39.1)	NA
				Sal (94)	39 (41.5)	NA
SFCT01/ SCO30002,[33] 2005	Unclear	Unclear	All AEs occurring after subject consented to participate until end of follow-up	Flu (131)	34 (26.0)	NA
				Placebo (125)	40 (32.0)	NA
van der Valk et al,[34] 2002	Adequate, permuted blocks, stratified	Adequate	3- and 6-mo follow-up	Flu (123)	1 (0.8)	0
				Placebo (121)	1 (0.8)	0
Vestbo et al,[35] 1999	Adequate, computer generated	Adequate	Participants seen every 3 mo	Bud (145)	36 (24.8)	0
				Placebo (145)	51 (35.2)	0
Wouters et al,[36] 2005	Adequate	Adequate	AE collected at start and end of treatment	SFC (189)	34 (18.0)	0
				Sal (184)	46 (25.0)	0

Abbreviations: AE, adverse event; Bud, budesonide; COPD, chronic obstructive pulmonary disease; Flu, fluticasone propionate; For, formoterol; For/Bud, combination of formoterol and budesonide; NA, not available; RCT, randomized controlled trial; SAE, serious AE; Sal, salmeterol xinafoate; SFC, combination of salmeterol and fluticasone.

[a] All RCTs were double blind.

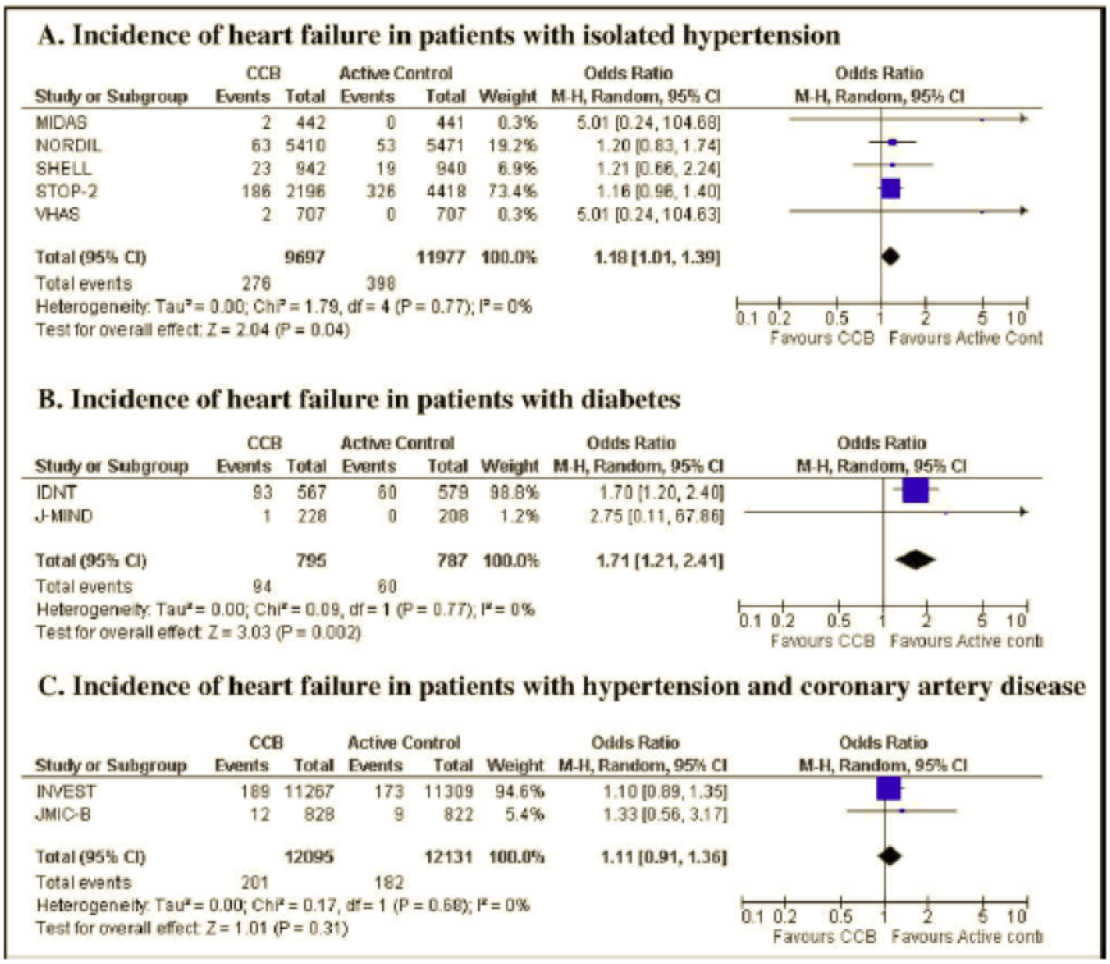
[b] Data on pneumonia were extracted from the US Food and Drug Administration presentation[37,38] because the published version provided information on the probability of pneumonia and not the actual number of events.

Recommendations for reporting harms in reviews: Review authors should consider the possible sources of systematic biases that could affect the specific harm under consideration within the review. The study designs may be 'ideal' for efficacy measurement (e.g., in terms of some standard risk of bias indicators such as concealment of allocation, sequence generation and double-blinding), but consideration of the sample selection, drop-outs and measurement of adverse events should be evaluated separately from the outcomes of benefit as described in item 12. Depending on the nature of the harms, their measurements, data collection and study designs, risk of bias may need to consider different indicators and their choice should be justified beforehand.

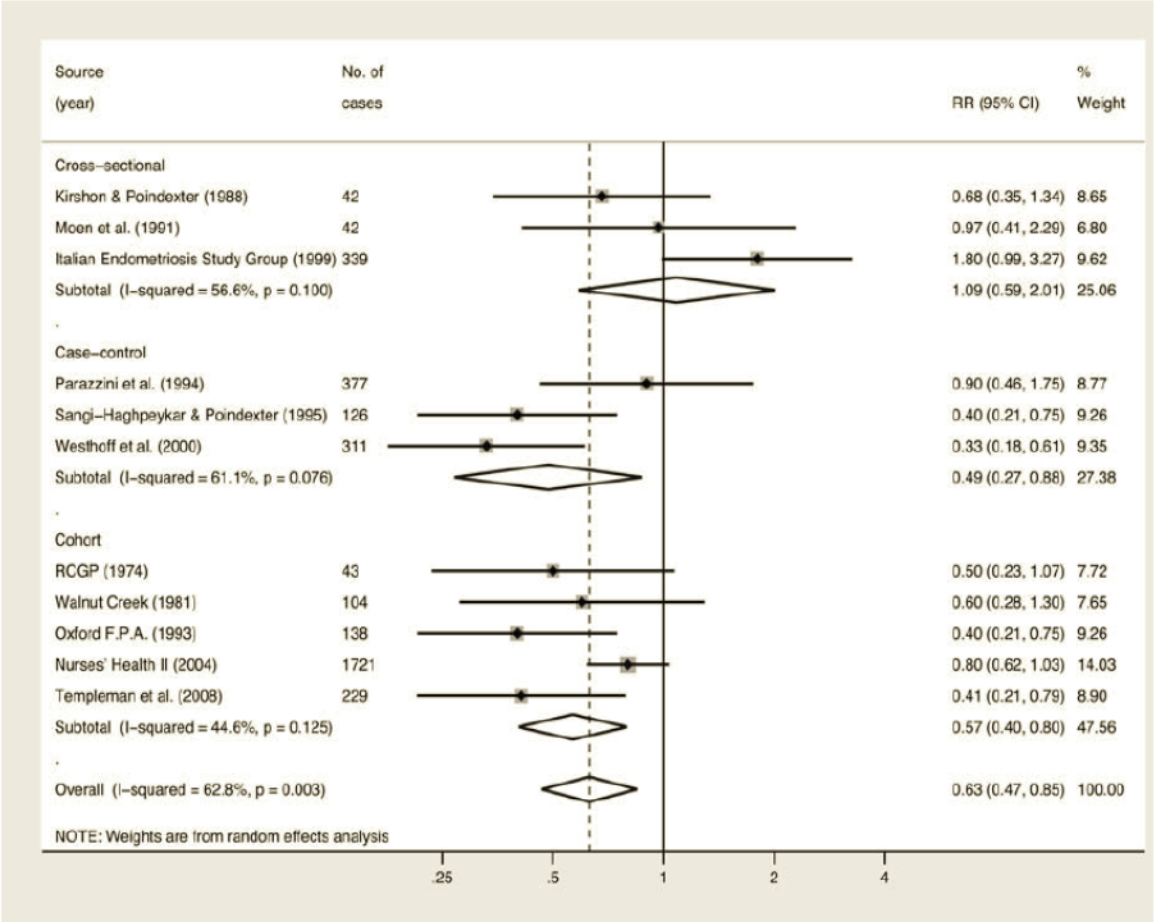
Item 20-Results of individual studies

PRISMA Statement: "For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot."

Example 1: (Results of Individual Studies) (t1)



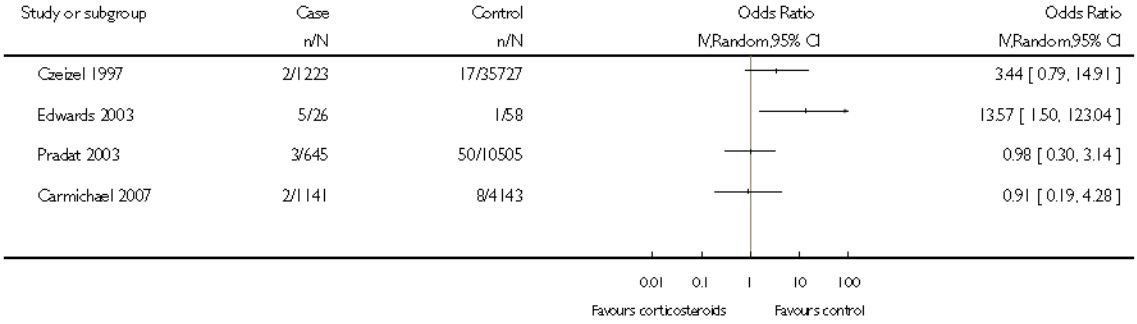
Example 2: (t2)



Example 3: (t3)

Analysis 1.5. Comparison 1 Topical corticosteroids vs no topical corticosteroids, Outcome 5 Cleft lip ± palate (case-control studies).

Review: Safety of topical corticosteroids in pregnancy
Comparison: 1 Topical corticosteroids vs no topical corticosteroids
Outcome: 5 Cleft lip palate (case-control studies)



Recommendations for reporting harms in reviews: Because adverse events are often rare, it is especially important that review authors should report the actual numbers of adverse events in each study, separately by intervention group if relevant. This information can be combined with information for each study on how adverse events were assessed (item 18). Effect estimates on their own, such as relative risk (RR), may be very misleading summaries of data on harms. Graphical display in a forest plot is often useful, although meta-analysis is not always appropriate.

Item 21-Synthesis of results

PRISMA Statement: "Present results of each meta-analysis done, including confidence intervals and measures of consistency."

PRISMA Harms: "*Describe any assessment of possible causality".*

Example 1:

When Bradford-Hill criteria for causality was applied to the reports summarized here, we found that many criteria were fulfilled, such as: consistency (the reports were very consistent in their clinical and laboratory findings); biological gradient (there was predominantly an association between higher doses and longer durations of propofol with PRIS); plausibility and coherence (mitochondrial toxicity and oxidative phosphorylation effect in energy production has been reported previously) and temporality (all patients developed the syndrome after been exposed to propofol).(u)

The assessment of causality is important when reviewing harms. The author should report the causality assessment in the light of the dataset obtained, and if it was limited to cases where relationship between intervention and adverse reaction was judged as "related", "probable" or "possible". Review authors should report if causality was assessed, how it was determined, and the definitions used to establish causality (e.g. Bradford Hill's principles of causality).

Example 2:

'There was no increased risk for all atrial fibrillation (AF) adverse events in each of the four randomized trials. Meta-analysis of the four randomized control trials (RCT) revealed that serious atrial fibrillation occurred more frequently in the bisphosphonate group compared to the placebo group (relative risk, 1.525; 95% CI, 1.166 to 1.997; p=0.002). The results are shown in the Fig. 2.

Among the observational studies, Heckbert et al. in their case-control study, found that more AF case patients than controls had ever used alendronate [6.5% (n=47) versus 4.1% (n=40); p=0.03]. Ever use of alendronate was associated with a higher risk of incident atrial fibrillation after adjustment for the matching variables, a diagnosis of osteoporosis and any cardiovascular disease [odds ratio (OR): 1.86; confidence interval: 1.09–3.15]. The OR was slightly higher for past users than for current users, but this difference did not reach significance (p=0.08). The risk of sustained AF associated with alendronate ever use was higher than the risk of transitory or intermittent AF (sustained AF: OR, 5.75 [95% CI, 2.50–13.25]; transitory AF: OR, 1.93 [95% CI, 0.95–3.92]; and intermittent AF: OR, 1.25 [95% CI, 0.64–2.44]; test of difference in OR across strata, p=0.005).

Sorenson et al. in their case-control study, with almost equal use of etidronate and alendronate among cases and controls, found a similar adjusted relative risk of AF among ever and never users of bisphosphonate. Although both Heckbert et al. and Sorenson et al. gave conflicting results, in both instances, the association was not statistically significant when restricted to current users.

Abrahamsen et al. found increased risk of AF in fracture patients prescribed with oral bisphosphonates compared with matched fracture patients who did not receive bisphosphonates. An unadjusted hazard ratio (HR) of 1.29 (1.17–1.41) and 1.22 (1.09– 1.36) was found for probable and hospital-treated AF, respectively. The effect size was reduced to fully adjusted (adjustment for age, sex, use of cardiovascular medications, number of co-medications and co-morbidity) HR of 1.18 (1.08–1.29) and 1.13 (1.01–1.26) for probable and hospital-treated AF respectively. A multivariable Cox regression analysis identified the following factors as associated with increased risk of incident AF after exposure to bisphosphonates: male sex, increasing age, number of co-medications and use of anti-hypertensive drugs. In a subgroup analyses, according to bisphosphonate type, alendronate use was indicative of significantly increased risk of AF after adjustment for co-morbidity and co-medications, whereas this was not the case for etidronate. (u2)

<u>*Recommendations for reporting harms in reviews:*</u> The synthesis of results should report the findings of all included studies. A summary of findings sub-grouped by study designs (if more than one design was included) with consideration of the limitations of each type should be considered. Meta-analysis may not be advised and a qualitative report of findings may be the best approach for rare events. If meta-analysis is undertaken, the estimates should also be sub-grouped by the different study designs (see example 2, item 20). When a rare outcome is being reviewed, the synthesis of results may only include a

few cases and the choice of statistical methods may make a major difference to the results and inferences drawn (23, 44, 45).

Item 22-Risk of bias across studies

PRISMA Statement: ***"Present results of any assessment of risk of bias across studies (see Item 15)."***

Example*:*

The included studies were broadly similar in terms of ascertainment of drug use and cardiovascular outcomes; they relied mainly on computerised diagnostic codes, pharmacy claims databases, and retrospective chart reviews. Few researchers made attempts to verify drug history directly with the patients or to check the validity of the prescriptions data source; only one study was deemed to have specific validation of drug use. Most studies reported the accuracy of outcome ascertainment on the basis of historical validation studies; only two studies specifically cross checked or validated outcomes for this analysis. Both of these studies showed a limited degree of misclassification. None of the studies provided details about the severity and consequences of the cardiac adverse events. Almost all the studies used a wide variety of variables to adjust for potential confounders. Two cohort studies checked specifically for similarities between the rosiglitazone and pioglitazone populations and did not find any major differences in the characteristics evaluated. We were able to use adjusted risk estimates for most studies, except for two studies for which we calculated odds ratios from the raw data. (w)

No specific additional information is required for reviews of adverse events.

Item 23-Additional analysis

PRISMA Statement: ***"Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16])."***

Examples:

Comparison of study designs: "*We performed subgroup analysis on the 39 studies reporting risk of incident Clostridium difficile infection (CDI) by restricting the studies according to design (case-control or cohort), presence of laboratory verification of CDI, or to hospital setting only. The significant association*

between PPI use and CDI was present in all the subgroup analyses, *even when we restricted it to studies that provided adjusted data and had laboratory diagnostic confirmation" (v1)*

High-quality vs. low-quality cohorts: *There were two cohorts with a quality score > 5 (presumed high quality) and three with a respective score of ≤ 5 (presumed low quality). The OR (random effect model) for the frequency of gastro-esophageal reflux disease (GERD) the Hp eradicated group vs. the Hp persistent group was 1.85 (95 % CI: 0.25 – 13.99; P = 0.55) in high-quality studies. Test of heterogeneity was significant for the meta-analysis (P = 0.05, χ 2 = 3.73, d.f.: 1, I 2 = 73.2 %). The OR (fixed effect model) for the frequency of GERD in the Hp eradicated group vs. the Hp persistent group was 1.59 (95 % CI: 0.80– 3.15; P = 0.14) in high-quality studies. Test of heterogeneity was not significant for the meta-analysis (P = 0.48, χ 2 = 1.47, d.f.:2, I2 = 0.0%).* (v2)

No specific additional information is required for reviews of adverse events.

Item 24-Summary of evidence

PRISMA Statement: "Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy *makers).*"

Example:

This meta-analysis, based on 1,530 PPI-exposed and 133,410 non-exposed controls, showed that the first-trimester use of PPIs does not seem to be associated with an increased risk for major congenital malformations. The CIs are tight, suggesting that it is unlikely that a β -error may contribute to the lack of significant effect. Furthermore, secondary analysis revealed no apparent increased risk for spontaneous abortions or preterm delivery. On the basis of 1,341 omeprazole-exposed and 120,137 non-exposed controls, the use of omeprazole does not seem to be associated with an increased risk for congenital malformations.

The funnel plot did not reveal a potential publication bias, and the included studies were not heterogeneous. On the basis of the overall acceptable quality of the studies included in this meta-analysis, the results obtained offer reassurance with respect to the safety of the use of PPIs in pregnancy; similarly, the same can be said about the use of omeprazole in pregnancy. Our results should be

interpreted with some caution, as they may not be generalizable to all populations; however, it is encouraging that our results are consistent with earlier studies. (x)

No specific additional information is required for reviews of adverse events.

Item 25-Limitations

PRISMA Statement: ***"Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias)."***

Example:

This systematic review and meta-analyses demonstrated that testosterone therapy in men was associated with significant increases in hemoglobin and hematocrit. There also was a statistically significant, but small reduction in HDL cholesterol. However, testosterone therapy had no significant effects on all-cause mortality, prostatic or urological outcomes, cardiovascular events, or cardiovascular risk factors. Several subgroup interactions were found and can partially explain the heterogeneity associated with the analysis. However, caution should be exercised in interpreting these analyses because they are considered observational in nature (despite the fact that the original studies were randomized) and the associations found can be attributable to chance due to the multiple simultaneous comparisons. Subgroup interactions generated by study-level meta-analyses are considered hypothesis-generating and should be confirmed at a patient-level (in a large trial or individual patient meta-analysis) before clinical implications are inferred.

The strengths of this study include the comprehensive literature search, the application of bias protection measures in the selection of the studies, and the evaluation of their methodological quality. Nevertheless, the quality of the evidence varied from low to medium considering the imprecision (small number of events), heterogeneity (for the outcomes of cardio metabolic risk factors, hemoglobin and hematocrit), and methodological limitations of the included trials. In particular, the brief duration of most testosterone trials limited inferences about the long term safety of this treatment. In addition, publication and reporting biases likely affected the inferences in this review because not all studies reported the outcomes of interest.(y)

Recommendations for reporting harms in reviews: The review author should recognize the findings of a possible association may not necessarily indicate a definitive link between intervention and adverse

event. It is important to recognize limitations of meta-analysis for rare adverse events (issues noted previously related to collection and reporting).

Item 26-Conclusions

PRISMA Statement: ***"Provide a general interpretation of the results in the context of other evidence, and implications for future research."***

Example:

Further complicating the interpretation of our findings is that few high-quality studies have compared the safety profiles of other induction agents. None of the induction agents used in the ED are devoid of potential adverse effects. Our study did not examine immediate adverse events such as hemodynamic and cardio depressant adverse effects and factors such as ease of tracheal intubation and simplicity of use that also need to be taken into consideration when choosing the most appropriate induction agent for a given patient.

The available evidence suggests that etomidate suppresses adrenal function transiently, without demonstrating a significant effect on mortality. However, to our knowledge no studies to date have been powered to detect a difference in mortality or time in the hospital, the ICU, or receiving ventilator support. According to robust evidence that etomidate transiently decreases adrenal function and that a significant effect on mortality cannot be excluded, alternate induction agents may be considered for use in rapid sequence intubation, particularly for septic patients. More data are needed to determine etomidate's effect on mortality. (z)

Recommendations for reporting harms in reviews: It is important to state conclusions in coherence with the review findings. Not infrequently, the review author states the limitations and weaknesses of the included studies in the discussion (e.g., high risk of bias, poor reporting of adverse events). Despite this, the systematic review conclusion may state that the adverse events were not identified and the intervention is therefore considered 'safe', when, in reality, its safety is actually unknown.

Review authors should be careful not to conclude safety from an absence of evidence of harm.

Item 27-Funding

PRISMA Statement: ***"Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review."***

Example:

Mr. Edwards is an employee of Astra Zeneca UK Ltd. (manufacturer of Seroquel [quetiapinej). Dr. Smith has attended educational meetings sponsored by AstraZeneca UK Ltd. and Sanofi-Synthelabo Ltd., and was previously an employee of janssen-Cilag (manufacturer of Risperdal [risperidone]) and Astra Zeneca UK Ltd. Alex Mitchell, PhD, of Complete Medical Communications, Macclesfield, United Kingdom, provided medical writing support funded by AstraZeneca. (ab)

No specific additional information is required for reviews of adverse events.

Discussion

Poor reporting in medical research has serious implications for health care. It wastes funds invested as it requires duplication of experiments and exposes patients to treatments with unknown effects (48). Reporting guidelines offer guidance to authors and journal editors on a minimal set of items to be reported to assure transparency. Transparent reporting improves patient care and supports proper allocation of heath care investment. Health care professionals, police makers and the general public rely on systematic reviews to provide a summary of outcomes of a given intervention.

Systematic reviews are the ideal method to combine information on rare outcomes. Harms are often infrequent and individual studies may not have the power to detect differences between interventions (22). The review author has a unique opportunity to access and evaluate the entire evidence including strengths and deficits in harms reporting in included studies (12). Clear reporting of harms is critical. Systematic reviews can present a misleading picture to readers if the lack of evidence of harms is presented as evidence of safety. Every intervention carries the potential for both benefits and harms and both should be clearly reported.

The PRISMA Harms is an extension of the PRISMA Statement and it should be used in every review assessing adverse events. The checklist is a *minimum* set of items to be reported to provide transparency. The sections "recommendations for reporting harms in reviews" highlights important concepts to consider when reviewing adverse events as they can be particularly different than reviews of efficacy. PRISMA is a reporting guideline and does not provide methodological advice, but strongly seeks the transparency in reporting the reasons for any methodological choices made.

Conclusion

A balanced assessment of benefits and harms is needed. The PRISMA Harms is a step towards better reporting of harms in reviews. The endorsement of the PRISMA Harms by peer review journals, the Cochrane Collaboration and review authors is paramount for effective implementation.

Examples references:

Title	(a.1) Arbyn M, Kyrgiou M, Simoens C, Raifu O, G Koliopoulos, et al. Perinatal mortality and other severe adverse pregnancy outcomes associated with treatment of cervical intraepithelial neoplasia: meta-analysis BMJ 2008;337:a1284. (a.2) Palmer SC, Craig JC, Navaneethan SD, Tonelli M, Pellegrini F, Strippoli GF. Benefits and harms of statin therapy for persons with chronic kidney disease: a systematic review and meta-analysis. Ann Intern Med 2012 Aug 21;157(4):263-75.
Abstract	(b) Bhuriya R, Singh M, Molnar J, Arora R, Khosla S. Bisphosphonate use in women and the risk of atrial fibrillation: A systematic review and meta-analysis. International Journal of Cardiology 2010;142:213–217.
Introduction	(c) Banach R, Boskovic R, Einarson, Koren G. Long-Term Developmental Outcome of Children of Women with Epilepsy, Unexposed or Exposed Prenatally to Antiepileptic Drugs. A Meta-Analysis of Cohort Studies. Drug Saf 2010;33(1):73-79.
Objective	(d) Loke YK, Kwok CS, Singh S. Comparative cardiovascular effects of thiazolidinediones: systematic review and meta-analysis of observational studies. BMJ 2011;342:d1309.
Protocol and Registration	(e) Muehlhans S, Richard G, Ali M, et al. Safety reporting in developing country vaccine clinical trials—A systematic review. Vaccine, 2012;22(30):3255-65.
Eligibility criteria	(f) Loke YK, Cavallazzi R, Singh S. Risk of fractures with inhaled corticosteroids in COPD: systematic review and meta-analysis of randomised controlled trials and observational studies. Thorax. 2011 Aug;66(8):699-708.
Information sources	(g) Agbabiaka T, Pittler M, Wider B, Edzard Er. Serenoa repens (Saw Palmetto):A Systematic Review of Adverse Events. Drug Safety 2009;32(8):637-647.
Searches	(h) Meyer-Massetti C, Cheng C, Sharpe B, Meier C, Guglielmo J. The FDA Extended Warning for Intravenous Haloperidol and Torsades de Pointes: How Should Institutions Respond? Journal of Hospital Medicine 2010;5:E8–E16.

Study selection	(i) Aprili D, Bandschapp O, Rochlitz C, Urwyler A, Ruppen W. Serious Complications Associated with External Intrathecal Catheters Used in Cancer Pain Patients A Systematic Review and Meta-analysis. Anesthesiology 2009;111:1346–55.
Data collection process	(j) Ruyssen-Witranda A, Fautrelb B, Sarauxc A, Le-Loëtd X, Phame T. Infections induced by low-dose corticosteroids in rheumatoid arthritis: A systematic literature review. Joint Bone Spine 2010;77:246–251.
Data Items	(k) Leombruno J, Einarson T,1 Keystone E. The safety of anti-tumour necrosis factor treatments in rheumatoid arthritis: meta and exposure-adjusted pooled analyses of serious adverse events. Ann Rheum Dis 2009;68:1136–1145.
Risk of bias in individual studies	(l) Loke YK, Kwok CS, Singh S. Comparative cardiovascular effects of thiazolidinediones: systematic review and meta-analysis of observational studies. BMJ 2011;342:d1309.
Summary measures	(m) Bangalore S, Kumar S, Wetterslev J, Messerli F. Angiotensin receptor blockers and risk of myocardial infarction: meta-analyses and trial sequential analyses of 147 020 patients from randomised trials. BMJ 2011;342:d2234.
Synthesis of results	(n) Albavera-Hernández C, Rodríguez J, Idrovo A. Safety of botulinum toxin type A among children with spasticity secondary to cerebral palsy: a systematic review of randomized clinical trials. Clin Rehabil 2009;23:394-407.
Risk of bias across studies	(o) CatesCJ, Lasserson TJ, Jaeschke R. Regular treatment with salmeterol and inhaled steroids for chronic asthma: serious adverse events. Cochrane Database of Systematic Reviews 2009, Issue 3. Art. No.: CD006922. DOI: 10.1002/14651858.CD006922.pub2.
Additional analyses	(p) vanDalen EC,Michiels EMC,CaronHN,Kremer LCM. Different anthracycline derivates for reducing cardiotoxicity in cancer patients. Cochrane Database of Systematic Reviews 2010, Issue 3. Art.No.: CD005006. DOI: 10.1002/14651858.CD005006.pub3.
Study selection	(q1) Choueiri T, Schutz F, Je Y, Rosenberg J, Bellmunt J. Risk of Arterial Thromboembolic Events With Sunitinib and Sorafenib: A Systematic Review and Meta-Analysis of Clinical Trials. J Clin Oncol 2010;28:2280-2285. (q2) Ranpura V, Hapani S, Wu S. Treatment-Related Mortality With Bevacizumab in

	Cancer Patients. A Meta-analysis. JAMA. 2011;305(5):487-494.
Study characteristics	(r) Cardwell C, Stene L, Joner G, Cinek O, Svensson J, et al. Caesarean section is associated with an increased risk of childhood-onset type 1 diabetes mellitus: a meta-analysis of observational studies. Diabetologia 2008;51:726–735.
Risk of bias within studies	(s) Singh S, Amin A , Loke YK. Long-term Use of Inhaled Corticosteroids and the Risk of Pneumonia in Chronic Obstructive Pulmonary Disease A Meta-analysis. Arch Intern Med. 2009;169(3):219-229.
Results of individual studies	(t 1) Shibata M, León H, Chatterley T, Dorgan M, Vandermeer B. Do Calcium Channel Blockers Increase the Diagnosis of Heart Failure in Patients With Hypertension? Am J Cardiol 2010;106:228–235. (t2) Vercellini P, Eskenazi B, Consonni D, Somigliana E,Parazzini F, et al. Oral contraceptives and risk of endometriosis: a systematic review and meta-analysis. Human Reproduction Update 2011;17(2):159–170. (t3) Chi CC, Lee CW, Wojnarowska F, KirtschigG. Safety of topical corticosteroids in pregnancy. Cochrane Database of Systematic Reviews 2009, Issue 3. Art. No.: CD007346. DOI:10.1002/14651858.CD007346.pub2.
Synthesis of results	(u1) Zorzela L, Pohlman K, Punja S, Joffe A, Hartling L, Vandermeer B, Loke L, Vohra S. Propofol infusion for paediatric sedation. A systematic review of serious adverse events. Chapter 5. Unpublished. Cochrane Library. Protocol. April 2012 DOI: 10.1002/14651858.CD009813. (u2) Bhuriya R, Singh M, Molnar J, Arora R, Khosla S. Bisphosphonate use in women and the risk of atrial fibrillation: A systematic review and meta-analysis. International Journal of Cardiology 2010;142:213–217.
Risk of bias across studies	(w) Loke YK, Kwok C, Singh S. Comparative cardiovascular effects of thiazolidinediones: systematic review and meta-analysis of observational studies. BMJ 2011;342:d1309.
Additional analysis	(v1) Kwok CS, Arthur AK, Anibueze CI, Singh S, Cavallazzi R, Loke YK. Am J Gastroenterol. Risk of Clostridium difficile infection with acid suppressing drugs and antibiotics: meta-analysis 2012 Jul;107(7):1011-9. doi: 10.1038/ajg.2012.108. (v2) Yaghoobi M, Farrokhyar F, Yuan Y, Hunt R. Is There an Increased Risk of GERD After Helicobacter pylori Eradication?: A Meta-Analysis. Am J Gastroenterol

	2010;105:1007–1013; doi: 10.1038/ajg.2009.734;
Summary of evidence	(x) Gill S, O' Brien L, Einarson T, Koren G. The Safety of Proton Pump Inhibitors (PPIs) in Pregnancy: A Meta-Analysis. Am J Gastroenterol 2009;104:1541–1545; doi: 10.1038/ajg.2009.122
Limitations	(y) Fernandez-Balsells M, Murad M, Lane M, Lampropulos J, Albuquerque F, et al. Adverse Effects of Testosterone Therapy in Adult Men: A Systematic Review and Meta-Analysis. J Clin Endocrinol Metab 2010;95:2560–2575.
Conclusions	(z) Hohl C, Kelly-Smith C, Yeung T, Sweet D, Doyle-Waters M, Schulzer M. The Effect of a Bolus Dose of Etomidate on Cortisol Levels, Mortality, and Health Services Utilization: A Systematic Review. Annals of Emergency Medicine 2010; 56(2):105-113.
Funding	(ab) Edwards S, Smith C. Tolerability of Atypical Antipsychotics in the Treatment of Adults With Schizophrenia or Bipolar Disorder: A Mixed Treatment Comparison of Randomized Controlled Trials. Clinical Therapeutics 2009;31: 1350-1359.

References

1. Moher D, Liberati A, Tetzlaff J, Altman D. The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PloS Med2009;6:e1000097

2. Muir Gray JA. Evidence-based in Healthcare. How to make health policy and management decisions. Churchill Livingstone,2001.

3. Sackett DL, Wennberg JE. Choosing the best research design for each question. BMJ1997; 315(7123):1636.

4. Ioannidis JP. Adverse events in randomized trials: neglected, restricted, distorted, and silenced. Arch Intern Med2009;169:1737-9

5. Sigh S, Loke YK. Drug safety assessment in clinical trials: methodological challenges and opportunities. Trials2012;13:138

6. Papanikolau PN, Ioannidis JP. Availability of large-scale evidence on specific harms from systematic reviews of randomized trials. Am J Med2004;117:582-9

7. Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. Arch Intern Med2009;169:1756-61.

8. Loke YK, Derry S. Reporting of adverse drug reactions in randomized controlled trials—a systematic survey. BMC Clin Pharmacol2001;1:3

9. Ioannidis JP, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. JAMA2001;285:437-43.

10. Yazici Y. Some concerns about adverse event reporting in randomized clinical trials. Bull NYU HospJt Dis2008;66:143-5

11. Ernst E, Pittler MH. Assessment of therapeutic safety in systematic reviews: literature review. BMJ2001;323:546.

12. Zorzela L, Golder S, Liu Y, Pilkington K, Hartling L, et all. Quality of reporting in systematic reviews of adverse events: systematic review. BMJ 2014;348:f7668

13. Golder S, Loke YK, Zorzela L. Some improvements are apparent in identifying adverse effects in systematic reviews from 1994 to 2011. J Clin Epidemiol2013;66:253-60.

14. McIntosh HM, Woolacott NF, Bagnall AM. Assessing harmful effects in systematic reviews. BMC Med Res Methodol2004;4:19

15. Cornelius VR, Perrio MJ, Shakir SA, Smith LA. Systematic reviews of adverse effects of drug interventions: a survey of their conduct and reporting quality. Pharmacoepidemiol Drug Saf2009;18:1223-31

16. Loke YK, Golder S, Vandenbroucke J. Comprehensive evaluations of the adverse effects of drugs: importance of appropriate study selection and data sources. TherAdv Drug Saf2011;2:59-68.

17. Chou R, Fu R, Carson S, Saha S, Helfand M. Methodological shortcomings predicted lower harm estimates in one of two sets of studies of clinical interventions. J ClinEpidemiol 2007;60(1): 18–28

18. Chou R, Helfand M. Challenges in systematic reviews that assess treatment harms. Ann Intern Med2005;142:1090-9

19. Hopewell S, Wolfenden L, Clarke M. Reporting of adverse events in systematic reviews can be improved: survey results. J Clin Epidemiol2007;61:597-602.

20. Moher D, Schulz K, Simera I, Altman D. Guidance for developers of health research reporting guidelines. PloS Med2010;7:e1000217.

21. Beller EM, Glasziou PP, Altman DG, Hopewell S, Bastian H, et al. PRISMA for Abstracts: Reporting Systematic Reviews in Journal and Conference Abstracts. PLoS Med 2013, 10(4): e1001419. doi:10.1371/journal.pmed.1001419

22. Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011]. Cochrane Collaboration, 2011. www.cochrane-handbook.org.

23. Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schultz K, et al. Better reporting of harms in randomized clinical trials: an extension of the CONSORT statement. Ann Int Med2004;141:781-8.

24. Hernandez AV, Walker E, Ioannides-Demos LL, Kattan MW. Challenges in meta-analysis of randomized clinical trials for rare harmful cardiovascular events: the case of rosiglitazone. Am Heart J2008;156:23-30

25. Pilkington K, Boshnakova A. Complementary medicine and safety: a systematic investigation of design and reporting of systematic reviews. Complement Ther Med2012;20:73-82.

26. Zorzela L, Hartling L, Loke Y, Joffe A, Ioannidis JP, Golder S, Santaguida PL, Altman D, Vohra S. The Development of PRISMA Harms using a Modified Delphi Technique. Thesis Chapter 5. Unpublished.

27. Murphy M, Black N, Lamping D, McKee C, Sanderson C, et al. Consensus Development Methods and their use in clinical guideline development. Health Technol Assess 1998; 2: 1-88.

28. Loke YK, Price D, Herxheimer A. Systematic reviews of adverse effects: framework for a structured approach. BMC Med Res Methodol2007;7:32.

29. Golder S, Loke YK, McIntosh HM. Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. J Clin Epidemiol2008;61:440-8.

30. Derry S, Loke YK, Aronson JK. Incomplete evidence: the inadequacy of databases in tracing published adverse drug reactions in clinical trials. BMC Med Res Methodol2001;1:7

31. Golder S, Loke Y. The contribution of different information sources for adverse effects data. Int J Technol Assess Health Care 2012;28(2):133-7.

32. Golder S, Loke Y. Failure or success of electronic search strategies to identify adverse effects data. J Med Libr Assoc. 2012;100(2):130-4.

33. Hartley J. Clarifying the abstracts of systematic literature reviews. Bull Med LibrAssoc 2000; 88(4): 332–337.

34. Hammad T, Neyarapally G, Pinheiro S, Iyasu S, Rochester G, Dal Pan G. Reporting of meta-analyses of randomized controlled trials with a focus on drug safety: An empirical assessment. Clin Trials 2013;10:389-97.

35. Michele TM, Pinheiro S, Iyasu S. The safety of tiotropium–the FDA's conclusions. N Engl J Med 2010; 363: 1097–99.

36. Anello C, O'Neill RT. Does research synthesis have a place in drug regulatory policy? Clin Res RegulAff 1996;13:13–21.

37. Golder S, Loke YK, McIntosh HM. Room for improvement? A survey of the methods used in systematic reviews of adverse effects. BMC Med Res Methodol2006;6:3.

38. Golder S, Loke Y, Bland M. Unpublished data can be of value in systematic reviews of adverse effects: methodological overview. Journal of Clinical Epidemiology 2010; (63) 1071e1081

39. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291:2457e65.

40. Scharf O, Colevas AD. Adverse event reporting in publications compared with sponsor database for cancer clinical trials. J ClinOncol 2006;24:3933e8.

41. Mahoney MR, Sargent DJ. Adverse-event rates: journals versus databases. Lancet 2007;369:171e2.

42. Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. CMAJ 2004; 171:735e40.

43. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, et al.Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011;343:d4002

44. Lau J, Ioannidis JP, Terrin N, et al. The case of the misleading funnel plot. BMJ. 2006;333(7568):597-600.

45. Shuster JJ, Jones LS, Salmon DA. Fixed vs random effects meta-analysis in rare event studies: the rosiglitazone link with myocardial infarction and cardiac death. Stat Med. 2007 Oct 30;26(24):4375-85.

46. Shuster JJ.Empirical vs natural weighting in random effects meta-analysis. Stat Med. 2010 May 30;29(12):1259-65. doi: 10.1002/sim.3607.

47. Ioannidis JP, Hozo I, Djulbegovic B. Optimal type I and type II error pairs when the available sample size is fixed. J ClinEpidemiol. 2013 Aug;66(8):903-910.e2. doi: 10.1016/j.jclinepi.2013.03.002. Epub 2013 May 9.

48. Shamseer L, Galipeau J, Turner L, Moher D. Improving the reporting and usability of research studies. Can J Anesth 2013; 60:337–344.

Table 6. PRISMA Harms checklist

Section/topic	#	Checklist item	Harms Extension	Check if done
TITLE				
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Specifically mention "harms" or other related terms, or the harm of interest in the review.	
ABSTRACT				
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.		
INTRODUCTION				
Rationale	3	Describe the rationale for the review in the context of what is already known.		
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).		
METHODS				
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.		
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.		
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.		
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.		
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).		
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.		
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.		
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.		
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).		
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	Specify how zero events were handled, if relevant.	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).		
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.		

RESULTS				
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.		
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Define each harm addressed, how it was ascertained (e.g. patient report, active search), and over what time period.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).		
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.		
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Describe any assessment of possible causality.	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).		
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).		
DISCUSSION				
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).		
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).		
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.		
FUNDING				
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.		

Appendix 1:

PRISMA Harms Group

Sunita Vohra (Convenor) (University of Alberta)

David Moher (Ottawa Health Research Institute; University of Ottawa)

Doug Altman (University of Oxford)

Yoon Loke (University of East Anglia)

John Ioannidis (Stanford University)

Pasqualina Santaguida (McMaster University)

Su Golder (University of York)

Jan Vandenbroucke (Leiden University)

Chapter 5

Propofol infusion for paediatric sedation

Propofol infusion for paediatric sedation. A systematic review of serious adverse events.

Cochrane Library. Protocol. April 2012 DOI: 10.1002/14651858.CD009813.

Authors: Liliane M Zorzela[1], Katherine A Pohlman[1], Salima Punja[2], Ari Joffe[3], Lisa Hartling[1], Ben Vandermeer[1], Yoon Loke[4], Sunita Vohra[1]

[1]Department of Pediatrics, University of Alberta, Edmonton, Canada

[2]Department of Medicine, University of Alberta, Edmonton, Canada

[3]Department of Pediatrics, Division of Pediatric Intensive Care, University of Alberta and Stollery Children's Hospital, Edmonton, Canada

[4]School of Medicine, University of East Anglia, Norwich, UK

Abstract

Background

Propofol is an anaesthetic agent frequently used in paediatric sedation. Since 1992, there have been several case reports of metabolic acidosis, rhabdomyolysis, liver dysfunction, cardiac arrhythmias and cardiac arrest associated with a prolonged propofol infusion in high doses, this phenomenon became referred to as 'propofol infusion syndrome' (PRIS). There is need to review the current knowledge of serious adverse events of propofol infusion in paediatric patients.

Objectives

To identify if propofol infusion is associated with serious adverse events (PRIS or cardiac arrest) in paediatric patients; to identify if there is a propofol dose, duration of infusion, patient characteristics and patient care setting associated with PRIS or cardiac arrest. We aimed to identify, if available, the mode of surveillance of adverse events in different study designs (active surveillance, passive or specific monitoring procedures for collecting serious adverse events). Also, we intended to identify if the inclusion of non-randomized studies and grey

literature in a review of adverse events provide relevant data that would be otherwise overlooked if only clinical trials would be included.

Search methods

We used the PRISMA Harms as a reporting guideline for this review of adverse events.

We searched the Cochrane Central Register of Controlled Trials (CENTRAL) (*The Cochrane Library*, latest version); Ovid MEDLINE; and Ovid EMBASE (1988 to date). There were no language restrictions. We searched the web sites of regulatory pharmaceutical authorities. Date of last search was January 22nd, 2014.

Selection criteria

We did not limit our inclusion criteria by study design.

The target population consisted of paediatric patients (age ranging from 28 days to 19 years) receiving sedation using propofol for 60 or more minutes in a hospital (intensive care unit, operating room, emergency room or any other location within the hospital) or other medical setting (for example dental offices).

The intervention studied was the use of propofol as an anaesthetic drug given by continuous infusion.

Data collection and analysis

Two authors (LZ and KP) independently extracted data using standardized data extraction forms developed *a priori*. Data was entered into RevMan Version 5.3. We used two different data extraction forms, one for randomised or quasi- randomised trials and a second for the non-randomised studies.

Main results

The search identified 8801 papers from electronic database and 13 publications from other sources. The screening of title and abstract for the population of interest (age between 28 days

and 19 years) and intervention of interest (propofol infusion for 60 minutes or more) generated 883 papers. A total of 104 studies were included: 34 clinical trials (total 2194 participants and 1089 receiving propofol); 7 cohorts (1466 patients and 1027 receiving propofol); 36 case series (total of 3540 patients) and 27 case reports reporting 33 paediatric cases. We did not identify any case-control studies.

A total of 91 serious adverse events (PRIS and /or cardiac arrest) associated with propofol infusion were identified in this review; 21 were identified in an unpublished RCT comparing propofol 1%, propofol 2% and standard sedative agents (SSA) in patients in the paediatric intensive care unit (PICU). Participants received propofol infusion for 81 hours (propofol 1% group); 74 hours (propofol 2% group) and the comparison group received SSA for 65 hours. The dose of propofol used in this trial was 4 to 6 mg/kg/hr. The mortality of participants receiving propofol 1% was 8.2% (9/109) (p=0.17); propofol 2% was 10.7% (12/112) (p=0.05) compared to SSA which was 3.8% (4/105). One cohort study reported an increased number of cardiac arrest, 9.9% (7/71) (p=0.02) and increased mortality, 5.6% (4/71) (p=0.08), in patients receiving propofol. The comparison group had 2% (3/151) of patients experiencing cardiac arrest, of which 2 died (1.3%). Two other cohorts reported on mortality, a total of 15 patients died using propofol, the controls had 22 deaths. A meta-analysis for the cohorts was considered but, excluded due high heterogeneity (I^2= 83%). The case series reported a total of 11 deaths out of 3540 (0.3%) in patients receiving propofol. All patients experiencing cardiac arrest were admitted to intensive care and an individual dose or duration of infusion of propofol could not be obtained despite attempts to contact authors.

The case reports had a total of 53 reports as possible cases of PRIS. Of these, 20 cases did not fulfill our criteria. Of 33 case reports of patients fulfilling PRIS criteria, 20 died; 6 of the 13 survivors used renal replacement therapy (two used ECMO with hemofilter) as part of the treatment. All the patients/participants experiencing PRIS were admitted to intensive care units. The reasons for admission were: neurologic in 18 patients (54.5%), respiratory 11 patients (33.3%), post operatory 4 patients (12.1%). Patients received propofol for a median

duration of 50.5 hours (IQR 36-91), range (5-144) median dose of 7.95 mg/kg/h (IQR 6.45-11.77), range (1.8-27).

The patients with PRIS who developed a fatal outcome received propofol for a median of 67.5 hours (IQR 37.5- 96.5) range (19-144) in a median dose of 8 mg/kg/h, (IQR 7.05-10.5) range (4.5-27). Thirty of 33 reports of PRIS used propofol for 12 or more hours. One patient developed PRIS with as propofol dose of 1.8mg/kg/h and two others with an infusion as short as 5 hours, these three patients were post operatory and all survived.

Propofol infusion used for 12 or more hours had sensitivity= 90.91 % (95%CI: 75.64%- 97.98%), specificity= 73.68 % (95%CI:48.80%-90.75%) of being indicative of PRIS. Propofol dose >4 mg/kg/h had sensitivity= 93.75 % (95%CI:79.16%-99.05 % and specificity= 18.75 %(4.27 % to 45.66 % of being indicative of PRIS.

Of the studies identified in this review, a total of 5633 children received propofol for 60 or more minutes and did not have any serious events (PRIS and cardiac arrest) associated with it.

The mode of surveillance of adverse events was not clearly reported for any study design.

Authors' conclusions

In this review of serious adverse events associated with propofol infusion in children, we identified 91 cases of serious adverse events (PRIS or cardiac arrest). We also could identify the median dose of propofol and duration of infusion associated with PRIS as 7.95mg/kg/h and 50.5 hours. We identified a propofol dose of 4mg/kg/h and duration of propofol infusion longer than 12 hours as associated with the majority of the cases of PRIS and cardiac arrest. We could extrapolate these findings and suggest that a propofol use of less than 4mg/kg/h for a length of no more than 12 hours could be used by clinicians with caution, being mindful of a few non-fatal cases of PRIS associated with lower doses and shorter duration. All patients with serious adverse events were admitted to intensive care units, the majority of these due to neurological or respiratory disorder.

The data on the primary outcomes for this review originated exclusively from non-randomised studies in combination with unpublished data. The inclusion of non-randomised studies and grey literature may be the cornerstone for data collection in reviews of rare events, or perhaps for any review of adverse events, given the poor reporting of adverse events in general, especially in clinical trials.

This review is important to combine the information on PRIS, highlighting patient characteristics, propofol dose and length of propofol infusion published to date. Since the majority of data on PRIS was found in non-randomized studies, we cannot set a definitive causative link between propofol and the development of PRIS. It is also difficult to interpret data from case series and case reports due to a lack of a comparison group. However, the case reports are very consistent in presenting similar clinical findings, as patients used propofol for a prolonged period of time and/or in high doses, which could be seen as a strong association between propofol and PRIS.

The overall GRADE score was "moderate" for propofol and cardiac arrest, with one RCT and 3 cohort studies reporting on it. The overall GRADE score for PRIS was "low" with only case reports identifying it.

We would like to encourage the publication of institutional experiences in use of propofol infusion in pediatric intensive care units, if any, to better understand this poorly understood syndrome.

Plain language summary

Serious adverse events of propofol infusion in children

Propofol is a drug used for paediatric sedation. There were reports of paediatric patients receiving propofol in high doses and for a more than 2 days that developed signs of organ damage and sometimes death, this was called PRIS (Propofol Infusion syndrome). This review is to investigate if there are serious adverse events associated of propofol infusion in children.

We included different study designs to obtain information from various sources. We identified 91 serious adverse events (PRIS or cardiac arrest) associated with propofol infusion in the

included studies. One trial, done by the industry, found an increased number of deaths, 9 of 109 (8.2%) patients receiving propofol in a concentration of 1% and 12 of 112 (10.7%) patients receiving propofol 2%, died. The comparison group used standard sedative agents and 4 of 105 (3.8%) died. A different study, a cohort, found an increased number of cardiac arrest, 7 of 71 (9.9%) and death, 4 of 71 (5.6%) in patients receiving propofol in comparison with standard sedative drugs where 3 of 151 (2%) of patients had a cardiac arrest, and 2 died (1.3%). In series of patients receiving propofol, out of 3540 children, 11 died. There was also, total of 33 reported cases of children that used propofol in a high dose and for multiple hours and developed PRIS, 20 of them died. A total of 5633 children received propofol infusion and did not have any serious events (PRIS and cardiac arrest) associated with it.

In this review we could identify a few cases of serious events associated with propofol, we could also see that more than five thousand children received propofol infusion and did not have any serious events related to it. It seems that a high dose of propofol (4mg/kg/h or higher) and a long duration (12 hours or more) are associated with PRIS. Also, it was difficult to interpret data from case series and case reports, as they do not have a comparison group.

We encourage the publication of institutional experiences in use of propofol infusion in paediatric patients, as the majority of information obtained in this review originated from spontaneous case reports and a single unpublished trial; more information is needed to better understand this syndrome.

Background

Propofol is an anaesthetic agent frequently used in paediatric sedation. Some of the valuable features of this drug include its quick onset of action and its short half-life, such that the patient is fully awake a few minutes after the infusion is stopped. Propofol can play an important role in allowing neurologic assessment of a sedated patient and also during the weaning of mechanical ventilation.

Since 1992, there has been several case reports of metabolic acidosis, rhabdomyolysis, liver dysfunction, cardiac arrhythmias and cardiac arrest associated with a prolonged propofol

infusion in high doses. This phenomenon became referred to as 'propofol infusion syndrome' (PRIS). It is a poorly understood syndrome, usually causing death. PRIS was originally only reported in paediatric patients (1-3), but several more recent case reports of adult patients (4-6) with the same signs and symptoms suggest it is not exclusive to the paediatric population.

As PRIS is a rare disorder, it is unlikely to be identified in randomised clinical trials. These studies are usually powered to detect the efficacy of interventions, and therefore do not have sufficient data to confirm their safety (7). The Cochrane Adverse Effects Methods Group suggests the inclusion of non-randomised studies when reviewing adverse events (8). Although, a multicentre adult cohort study found a 1.1% incidence of PRIS in critically ill adults receiving propofol (9), the only comparable paediatric study was stopped early and has never been published. Partial data reported a comparison of the brand name of propofol Diprivan 2%, Diprivan 1% versus standard sedative agents (lorazepam, chloral hydrate, fentanyl, ketamine, morphine or phenobarbital) in paediatric critical care patients and identified an increase in mortality in patients receiving propofol, the first group had 11% mortality; second group had 9% mortality; and the standard sedation group had only 4% mortality (10). Unfortunately, this study was never published in a peer reviewed journal and many uncertainties regarding the incidence of PRIS in paediatric patients remain. While available reports discuss an association between PRIS and propofol infusion, causation has not been clearly determined; nor has the effect of dose, duration or setting, should these relationships exist.

In several countries, including Canada and the United States, regulatory agencies (11-12) have “re-emphasized that propofol is currently not approved for sedation in paediatric ICU patients”; however, propofol is used freely as a continuous infusion in paediatric operating rooms and emergency rooms. There is an urgent need to review the current knowledge of serious adverse events of propofol infusion in paediatric patients to support clinicians and regulatory bodies to promote evidence-based decision-making when evaluating the best drug choices for continuous anaesthesia in paediatric patients.

Description of the condition

Propofol infusion syndrome (PRIS) has been described mostly by case reports or case series (1-6) of patients receiving prolonged infusions of propofol (more than 48 hours) and in high doses (more than 4 mg/kg/hr). Patients with PRIS develop metabolic acidosis, rhabdomyolisis, liver dysfunction and arrhythmias, often followed by cardiac arrest and death.

Description of the intervention

The intervention evaluated in this systematic review was the use of propofol as a continuous infusion for 60 minutes or longer in paediatric patients (aged more than 28 days and less than 19 years old).

The length of propofol infusion necessary to cause PRIS is unclear. We assumed in this review that induction doses of propofol do not cause PRIS. There is an association of PRIS with longer duration and higher doses of propofol in paediatric studies (1-3).

Why it is important to do this review

Propofol is an anaesthetic agent with short onset of action and short half-life. These two characteristics make the drug clinically useful. However, there have been case reports of an association between propofol used in prolonged infusions and in high dosages with metabolic acidosis, liver dysfunction, arrhythmias and death. Regulatory bodies, such as FDA and Health Canada (11,12) contraindicate the use of propofol as a continuous infusion in critically ill paediatric patients, mostly based on case reports (1-3), but propofol is freely used in paediatric emergency rooms and operating rooms. We understand that PRIS in not exclusive to the paediatric population (4-6), but due to limitations of propofol use in critically ill children and no restriction of its use in other paediatric populations, we believe it is important to systematically review the use of propofol infusion in children in an attempt to identify risk factors and address its safety concerns.

The goal of this review was to identity if there is an association of PRIS or cardiac arrest with propofol use as a continuous infusion in paediatric patients, and therefore identify, if possible, any risk factors associated with these events, such as patient characteristics, dosage and

duration of infusion. By doing a systematic review of all study designs we intended to obtain all the evidence available in an attempt answer those questions.

Objectives

1. To identify if propofol infusion is associated with serious adverse events (PRIS or cardiac arrest) in paediatric population.
2. To identify the propofol dose and duration of infusion associated with PRIS or cardiac arrest.
3. To describe the patient characteristics associated with PRIS or with cardiac arrest.
4. To identify a possible association between PRIS or cardiac arrest and patient care setting (paediatric intensive care, operating room, paediatric emergency or other).
5. To identify the mode of surveillance for adverse events in different study designs (active surveillance, passive or any specific monitoring procedures for collecting serious adverse events).
6. To identify if the inclusion of non-randomized studies and grey literature in a review of adverse events provide relevant data that would be otherwise overlooked if only clinical trials would be included.

Methods

We used the PRISMA Harms as a reporting guideline for this review of adverse events (13).

Criteria for considering studies for this review

Types of studies

We did not limit our inclusion criteria by study design. The Cochrane Adverse Effects Methods Group recommends the use of non-randomised studies for reviews of adverse events (8).

We deliberately excluded surveys from our review as they would represent professional opinion, rather than providing actual patient outcome data.

Types of participants

The target population consisted of paediatric patients (age ranging from 28 days to 19 years) receiving sedation using propofol for 60 or more minutes in a hospital (intensive care unit, operating room, emergency room or any other location within the hospital) or other medical setting (for example dental offices). The indication for sedation was not a restriction; instead, it is the subject of a subgroup analysis. As a review of serious adverse events of propofol infusion in paediatric patients, it has broad inclusion criteria, i.e. any children receiving propofol infusion for sedation.

Types of interventions

The intervention being studied was the use of propofol as an anaesthetic drug given by continuous infusion. It is not clear whether the risk of PRIS is related to the duration of infusion. The case reports are usually associated with prolonged use (more than 48 hours) and in high doses (more than 4 mg/kg/hr). Based on information from these anecdotal reports, we chose to focus our review on patients receiving infusions lasting 60 minutes or longer. Further, we excluded studies which had a comparison group using propofol bolus or infusion, in order to have a propofol-free comparison group and to limit the exposure of interest to one single arm. As propofol does not provide analgesia, the group receiving propofol can also receive other analgesic or sedative drugs concomitantly.

The comparison group, when applicable, was any sedative or analgesic agent different from propofol, alone or in combination.

Types of outcome measures

Primary outcomes

The primary outcome was the report of PRIS or any other serious adverse events resulting in cardiac arrest (with return to spontaneous circulation or use of extra-corporeal life support) or death associated with propofol use as a continuous infusion in paediatric patients. The follow-up time to measure the primary outcome for this study was the development of the outcome of interest (PRIS or cardiac arrest) or time of hospital discharge for patients who did not develop the primary outcome.

The primary outcomes were defined as the following:

1. Cardiac arrest, a reduction of cardiac output requiring any of the following: chest compressions; defibrillation; epinephrine boluses; or cardiac mechanical support (extra-corporeal life support (ECLS), or ventricular assist device (VAD)). The patient may experience return to spontaneous circulation or not (death).

2. Propofol infusion syndrome (PRIS) was defined as metabolic acidosis, arterial pH ≤ 7.3 along with a serum bicarbonate ≤ 18 mg/dL; plus the presence of any sign in any of the below categories (adapted from Roberts (9)):

a) rhabdomyolisis (the breakdown of muscle fibres resulting in the release of muscle fibre contents (myoglobin) into the bloodstream), defined as creatine phosphokinase (CPK) ≥ 10,000 IU/L or positive serum or urine myoglobin test or positive urinary casts for haemoglobin;

b) hypotension (initiation of a vasopressor agent or increase of ≥ 20% from baseline);

c) hepatic transaminitis (increase in aspartate aminotransferase or alanine aminotransferase, or both, ≥ 3 times above baseline);

d) hypertriglyceridaemia (serum triglyceride concentration ≥ 400 mg/dL);

e) hypoxia (partial pressure of arterial oxygen ≤ 60 mm Hg);

f) hyperthermia (temperature ≥ 38.3 °C);

g) cardiac dysfunction that includes asystole, pulseless electrical activity, ventricular fibrillation, sustained ventricular tachycardia of 30 seconds or longer, myocardial failure (ejection fraction ≤ 40%) or bradycardia (heart rate ≤ 60 bpm);

h) renal failure that includes oliguria (urine output ≤ 0.5 mL/kg/hr for ≥ 6 hours), anuria (urine output ≤ 10 mL/hr for ≥ 6 hours), elevation in serum creatinine (increase of ≥ 1 mg/dL from baseline) or hyperkalemia (serum K^+ ≥ 6 mg/dL (excluding other known causes or hemolyzed specimens)).

We also captured related adverse events that do not clearly fulfil the above inclusion criteria. These events were reported as "not fulfilling PRIS criteria" and were the subject of a sensitivity analysis.

Search methods for identification of studies

Electronic searches

We searched the Cochrane Central Register of Controlled Trials (CENTRAL) (*The Cochrane Library*, latest version); Ovid MEDLINE; and Ovid EMBASE (1988 to date).

The search strategy is given in Appendix 1 (MEDLINE), Appendix 2 (EMBASE) and Appendix 3 (CENTRAL). The date of last search was January 22, 2014 for all databases. We screened references of included articles to identify additional publications. The search was not limited by study type. There were no language restrictions.

Searching other resources

We searched the web sites of regulatory pharmaceutical authorities such as Current Problems in Pharmacology (www.mhra.gov.uk); the Australian Adverse Drug Reaction Bulletin (www.tga.gov.au/adr/aadrb.htm); the European Public Assessment Reports from the European Medicine Evaluation Agency (www.emca.eu) and the Food and Drug Administration FDA Medwatch (www.fda.gov/medwatch).

Selection of studies

All studies meeting the following criteria were included.

1. Paediatric population: age ranging from 28 days to 19 years, exclusively.
2. Use of propofol as a continuous infusion for 60 or more minutes.
3. The comparison group (when applicable) was any other sedative or analgesic agents (alone or in combination) different from propofol.

There were no study design restrictions. Case reports and conference or meeting abstracts were included if they met the above criteria and also reported on the primary outcome (PRIS or cardiac arrest).

One author (LZ) carried out the searches. Two review authors (LZ and KP) independently screened all titles and abstracts and classified them as either:

- relevant (meeting all the inclusion criteria);
- possible (meeting some of the inclusion criteria or unclear whether meets inclusion criteria);
- rejected (failing to meet the inclusion criteria).

The studies classified as relevant or possible had their full text reviewed by both review authors and any disagreement were resolved by a third person (SV). For studies not published in English, a translator provided the translation of titles and abstracts for screening and the full text if the study met the inclusion criteria.

Data extraction and management

Two authors (LZ and KP) independently extracted data using standardized data extraction forms (Appendix 4) developed a priori by the authors. Data was entered into RevMan Version 5.3. We used two different data extraction forms, one for randomised or quasi- randomised clinical trials and a second for the non-randomised studies. The data extraction forms were piloted using two different studies of different designs.

We extracted the following data:

• study information, including type of study, inclusion and exclusion criteria;

• study population characteristics, baseline data such as age and gender, reason for hospital admission, setting, history of organ dysfunction prior to propofol initiation;

• medication details, including name of drug, dose, route, frequency and duration of treatment;

• definitions given for the reported adverse events (AE) and methods used to monitor and report these.

Assessment of risk of bias in included studies

Clinical trials

Two review authors (LZ and KP) independently assessed the methodological quality of each clinical trial using the Cochrane 'Risk of bias' tool (8). We assessed the following sources of bias:

1. Selection bias, including randomisation: described the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups. Described the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.

2. Performance bias: described all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provided any information relating to whether the intended blinding was effective.

3. Detection bias: described all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provided any information related to whether the intended blinding was effective.

4. Attrition bias: described the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. Stated whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomised participants), reasons for attrition or exclusions where reported, and any re-inclusions in analyses performed by the review authors.

5. Selective reporting bias: stated how the possibility of selective outcome reporting was examined by the review authors, and what was found. We developed definitions of reporting bias specific to this review (see below).

6. Other source of bias: stated any important concerns about bias not addressed in the other domains in the tool.

Non-randomized studies

We assessed the risk of bias of cohort studies using the Newcastle-Ottawa Quality Assessment scale (14). We rated studies based on selection of participants, comparability between groups and exposure to the intervention and outcome. The authors assessed risk of bias in studies, using a 'star' if the option (a) or for some factors also (b) are selected (Appendix 4). The more stars a study received, the lower the risk of bias. Studies with seven or more stars were considered to have a low risk of bias.

Each domain could only receive one star, with the exception of the field comparability, which could receive two stars.

Case series and case reports were clinically assessed by the report, if in accordance with the PRIS definition adopted for this review.

We reported the 'Risk of bias' table as part of the table 'Characteristics of included studies' and present a 'Risk of bias summary' figure, which detailed all of the judgements made for all included studies in the review.

Measures of treatment effect

As the primary outcomes are PRIS or cardiac arrest, these events were reported as number of events for all study designs. In case series or case reports, the events were reported as a "narrative report".

Unit of analysis issues

Patients may experience a serious adverse event more than once (e.g. cardiac arrest that responds to resuscitation and has a second event. For studies reporting in more than one serious event per participant/patient, we used one event per patient (the first event).

Dealing with missing data

We attempted to obtain any relevant missing data from the study authors. We contacted authors, through the corresponding author's e-mail provided and sought individual patient data or any clarification needed. We received several responses, described in the results section.

Assessment of heterogeneity

Heterogeneity in this review could originate from differences in comparison groups across trials or cohorts (standard of care may involve different drugs). In the analysis phase they were treated as a single 'control' group. A subgroup analysis was done if more than two studies were found using the same comparison drug (e.g. ketamine) and reporting on the main outcome (PRIS or cardiac arrest).

As we included non-randomised studies in this review, heterogeneity was expected to be greater than if RCTs alone were included. We analysed studies grouped by design (RCTs, cohort, case control, case series and case reports). Statistical heterogeneity would be assessed using the I^2 statistic and the following as a guide for interpretation: 0% to 30% might not be important, 30% to 70% may represent moderate heterogeneity, 70% to 100% may represent substantial heterogeneity. Any analysis resulting in substantial heterogeneity (I^2 =70% to 100%) was considered too heterogeneous and it was not presented as it could be misleading (8).

Assessment of reporting biases

We developed an assessment criterion on reporting bias specific for this review. For the randomised studies fulfilling our inclusion criteria, but not reporting on the main outcomes, we developed a table to identify the possibility of selective outcome reporting. The studies were classified as “likely", “possible”, “unlikely” or “uncertain” to have a selective report of adverse events.

A study was classified as “likely” to have a selective report of adverse events if referred in the methods section an intention to measure it, but the presence or absence of these events was not reported in the results.

A study was classified as "possibly" to have a selective report of adverse events if serious events were reported but no detailed information on resolution or treatment needed for those events were provided (e.g.: patients experienced bradycardia, but no clarification on treatment needed or resolution of the event). Studies were also classified as "possibly" to have a selective report of adverse events if reported adverse events as a general statement, for example "the incidence of adverse events was similar in both groups", but no number of events or seriousness is provided.

A study was classified as "unlikely" to have selective report of adverse events if the authors reported their intention to measure serious adverse events, which were reported as a number of events or number of patients per group, as well as the need for intervention and the resolution of the event (if treatment was needed).

A study was classified as "uncertain" if the authors had no intention to measure adverse events (as determined by their reported methods) and no adverse events were reported in the manuscript.

Data synthesis

Propofol dose was recorded as mg/kg/hr. For studies presenting the dose in mcg/kg/min a conversion of 1mcg/kg/min = 0.06mg/kg/hr was used. For studies only reporting a dose range (i.e., dose ranged between x and y), we attempted to contact author for clarification of the dose, if this was not successful, we used the mean between the two values provided. We addressed any issues related to the unit of analysis as recommended in the Cochrane Handbook (8).

The main comparison was propofol infusion and control. In this review, the control group is called standard of care and it includes the most common sedative agents used for paediatric sedation. All analyses were performed using the Review Manager software RevMan 5.3. Data from case series and case reports were descriptively reported (8).

We intended to pool data from studies that were sufficiently homogenous. Data across different study designs were therefore not pooled.

We calculated sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) using cases of PRIS and cases not fulfilling PRIS criteria to find if a dose of 4mg/kg/h and length of infusion of 12hours could be indicative of PRIS.

Subgroup analysis and investigation of heterogeneity

We intended to perform subgroups for the primary outcome (PRIS and cardiac arrest) based on the following:

1. Comparison group: if two or more trials were found using the same comparison group (e.g., ketamine), these studies were subgrouped in an attempt to identify any risk increase or reduction for developing the primary outcome.

2. Dosage of propofol: we intended to subgroup studies using propofol infusion at equal to or less than 4 mg/kg/hr or more than 4 mg/kg/hr as this seemed to be the dosage cut-off for reporting PRIS in children, but it is not clear what dose of propofol is necessary to cause PRIS. This subgroup analysis would be done as an attempt to measure a dose-effect relationship between propofol and the development of PRIS.

3. Duration of propofol infusion: we intended to subgroup studies with propofol infusion lasting less than or equal to 12 hours, between 12 and 24 hours and more than 24 hours. This subgroup analysis was chosen as the duration of exposure of propofol required to cause PRIS is unclear and as an attempt to determine if there is any relationship between duration of infusion of propofol and the development of PRIS.

4. Indication for sedation: studies done under similar settings, for example sedation for mechanical ventilation in intensive care unit or sedation procedures in the operating room, these studies were subgrouped in an attempt to identify patient setting (location within the healthcare facility) and indication in an association with the primary outcome.

Sensitivity analysis

We proceeded with a sensitivity analysis comparing the cases that fulfilled the PRIS definition and progressed to a fatal outcome with cases that did not progress to a fatal outcome in order

to identify possible differences in propofol dose and length of infusion and treatment used after PRIS was identified.

We also conducted a sensitivity analysis comparing the cases of PRIS and the cases not fulfilling the PRIS criteria.

Summary of findings table

We summarized the evidence in a 'Summary of findings' table, as recommended by the Cochrane Anaesthesia Review Group (CARG) and described by Guyatt using the GRADE approach (15).

GRADE appraises the quality of a body of evidence based on the extent to which one can be confident that an estimate of effect or association reflects the item being assessed. The quality of a body of evidence considered within study risk of bias (methodological quality), the directness of the evidence, heterogeneity of the data, precision of effect estimates and risk of publication bias.

We used GRADEpro system to assess the quality of the body of evidence associated with the specific outcomes below:

1. Cardiac arrest.
2. Propofol infusion syndrome.

We examined five main domains: “study limitations” (or risk of bias), “directness” “consistency”, “precision” and “reporting bias”.

“Study limitation” is the likelihood of adequate protection against bias. It was assessed according to study designs (RCTs or observational studies) and study conduct (based on ratings of risk of bias of the individual studies). It was scored as “no”, “serious” or “very serious”.

“Directness” relates to (i) whether evidence links interventions directly to a health outcome of specific importance for the review and (ii) for comparative studies, whether the comparisons

are based on head-to-head studies. It was scored as “no” (no indirectness), “serious” or “very serious”.

“Consistency” is the degree to which included studies find either the same direction or similar magnitude of effect. For this review we used “direction of effect” (when the effects have the same sign, or direction). It was scored as “no” (no inconsistency), “serious” or “very serious”.

“Precision” is the degree of certainty of an effect estimate, based on the sufficiency of sample size and events. A body of evidence was considered “precise” if the minimum number of patients or events needed for an evidence base be considered powered is met, or, if a meta-analysis was judged inappropriate, we considered the significance level of p-values in the individual studies. It was scored as “no” (no imprecision), “serious” or “very serious”.

“Reporting bias” is the result from selectively publishing or reporting research findings based on the favorability of direction or magnitude of effect. We considered “study publication bias” if the full study was not reported; “selective outcome reporting bias”, if a study did not report (or incomplete report) planned outcomes or reported unplanned outcomes, and “selective analysis reporting bias”, if a study reported favorable analyses for a given outcome while not reporting others. It was scored as “undetected” or “strongly suspected”.

We also examined the additional domains “dose-response association”,“plausible confounding” and “strength of association and magnitudes of effect” (16).

“Dose-response association” was used either across studies associated with a pattern of a larger effect with greater exposure (dose, duration, and adherence). It was scored as “detected” or “undetected”.

“Plausible confounding that would decrease observed effect” was used when a possible confounder worked in opposite direction to direction to the effect, if the confounders were not present, the observed effect would have been even larger than the one observed. It was scored as “present” or “absent”.

"Strength of association or magnitude of effect" refers to the likelihood that the observed effect is large enough that it cannot have occurred solely as a result of bias from potential confounding factors. It was scored as "no", "large" or "very large".

We developed domain scores and strength of evidence grades separately for RCT evidence and observational study evidence. If studies reporting on the same outcome have significant differences in risk of bias, so that it could compromise the strength of evidence for the outcome, this particular study, or studies, were not included in the GRADE table.

The overall strength of evidence was graded as "high" if further research is very unlikely to change our confidence in the estimate of effect; "moderate" if further research may change our confidence in the estimate of effect and may change the estimate, "low" if further research is likely to change our confidence in the estimate of effect and is likely to change the estimate, or "insufficient" if evidence either is unavailable or does not permit estimation of an effect.

Two review authors (LZ,KP) graded the body of evidence using GRADE guidance (GRADEpro 2008) and discrepancies were resolved through consensus.

Results

Description of studies

Results of the search

The database search result had 8801 papers from electronic database and 13 reports from other sources (i.e. regulatory pharmaceutical web pages); once duplicates were removed, there were 5011 unique reports. The screening of title and abstract for the population of interest (age between 28 days and 19 years) and intervention of interest (propofol infusion for more than 60 minutes) yielded to 883 papers. Case reports and conference abstracts were included if reporting on the primary outcome (PRIS or cardiac arrest). Reports of adverse events were not used as selection criterion for other manuscripts (figure 3). If study population or propofol length of infusion were not clearly reported in the abstract, the full text was reviewed to assess

inclusion criteria. Search of regulatory agencies databases yield 13 publications. They included the drug warnings letters and medication labels. No new patient data were identified from this data source.

A total of 34 clinical trials; 7 cohort studies; 36 case series and 27 case reports (reporting 33 paediatric cases) were included in the analysis (table 7).

A total of 780 studies were excluded from this analysis. The reasons for exclusions were: infusion was used for less than 60 minutes in 238 studies or propofol was not used as an infusion (used as a bolus) in 35 studies; and 28 studies had included participants aged 19 years old or higher. Other publications like conference abstracts, letter or comments not adding clinical data (e.g., reviews, comments or letter to editors) were the reason to exclude 295 retrieved papers. Twenty studies did not fulfil PRIS criteria (table 12) and 88 were case reports of non-serious events. General characteristics of excluded studies can be found in table 8.

Serious adverse events

A total of 91 serious adverse events (PRIS and cardiac arrest) associated with propofol infusion were identified in this review; 21 were identified in an unpublished RCT (10) comparing propofol 1%, propofol 2% and standard sedative agents (SSA). The mortality of participants receiving propofol 1% was 8.2% (9/109) (p=0.17); propofol 2% was 10.7% (12/112) (p=0.05) compared to SSA which was 3.8% (4/105). One cohort study (17) reported an increased number of cardiac arrest 9.9% (7/71) (p=0.02) and increased mortality 5.6% (4/71) (p=0.08) in patients receiving propofol, the comparison group had 2% (3/151) of patients experiencing cardiac arrest, of which 2 died (1.3%). Two other cohorts (18,19) reported on a 13 and 2 deaths in patients using propofol, the comparison groups had 14 and 8 deaths, respectively. A meta-analysis was considered but excluded due high heterogeneity (I^2>80%).The case series reported a total of 11 deaths out of 3540 (0.3%) in patients receiving propofol. The case reports had a total of 53 cases selected as possible cases of PRIS. Of these, 33 cases fulfilled our definition of PRIS and 20 cases did not. Of 33 case reports of patients fulfilling PRIS criteria, 20 died; 6 of the

13 survivors used renal replacement therapy (two used ECMO with hemo filter) as part of the treatment. We did not identify any case-control studies fulfilling our inclusion criteria.

Details of the included studies can be found below, divided by study designs.

Included studies

Studies characteristics:

The majority of the included clinical trials were written in English (28 trials); in addition, there were one each in Chinese, French, German, Spanish, Russian and Turkish.

For non-randomised studies, the publication language was English (n=61), German (n=2), Spanish (n=2), and one each of Dutch, French, Italian, Polish and Russian. The characteristics of included studies can be found in table 7.

One author was contacted by e-mail to clarify if two studies (20,21) conducted in the same centre and in similar population (use of propofol compared to sevoflurane in the similar age group, but reporting different outcomes) had any duplication of data (i.e. if any patients were enrolled in both studies). No reply was received. As the studies report on different outcomes, they were treated as two independent populations.

Clinical trials:

The 34 trials presented in this review had a total of 2194 participants, of which 1089 received propofol. One trial (22) did not report on the number of patients per group; only the total number of enrolled patients and percentages of events were reported, for purposes of this review the total number of patients were divided equally between groups. Sample size ranged from 18 (in a cross-over study with 18 enrolled participants) to 326 participants (an unpublished RCT). The median age of the participants was 60.8 months (IQR 39.2-86.0).

The reasons for sedation were: elective general surgery (9 studies, 748 patients); ophthalmologic procedure (strabismus repair or eye exam) (5 studies, 200 patients); ENT procedure (prominent ears, cleft lip, cochlear implantation) (3 studies, 109 patients); correction

for congenital cardiac disease under bypass (6 studies, 186 patients); magnetic resonance imaging (3 studies, 239 patients); urologic surgery (2 studies, 167 patients); mechanically ventilated PICU patients (1 study, 326 patients); cardiac catheterization (1 study, 20 patients); endoscopy (1 study, 50 patients); cardiac ablation (1 study, 60 patients); cranial-facial surgery (1 study, 44 patients) and orthopedic procedure (1 study, 45 patients).

Settings:

Sedations were done in: operating room 26 studies (1443 participants); diagnostic imaging in 3 studies (239 participants); 2 in the cardiac catheterization suite (62 participants) and 1 in electrophysiology lab (60 participants). Two studies were conducted in paediatric intensive care units (370 participants).

Intervention:

The median dose of propofol used was 9.0mg/kg/hr (IQR 7.0-12.7). Ten of the 34 included studies did not present a precise propofol dose use (10, 20, 23-30), but a range. These authors were contacted for clarification on the propofol dose used. Four studies did not have a contact e-mail (23, 26-28) therefore we could not reach the authors. One had a "fail to deliver message" (29). One answer was received clarifying that the dose reported of 4-6 mg/kg/min in the manuscript (25) was a typo and it should be read as "mg/kg/h" and also confirmed that no individual doses used were recorded during the study, also a reply from the author (24) was received clarifying the mean propofol dose used. No answers were received for the remaining studies (10, 20, 30). The mean between the two values in the range was used for data analysis.

Five studies (23, 29, 31-33) used propofol bolus in a dose 1-5 mg/kg prior to propofol initiation. Ten studies (30,32,34-41) used an opioid with propofol infusion (remifentanil, alfentanil or fentanyl). Nine studies (20-22;24,27,28,36,43,44) used propofol in association with an inhaled agent.

The median duration of propofol use in included trials was 78.5 minutes (IQR 68-107.2). Two trials (10, 45), involving a total 370 patients used propofol for more than 12 hours. There were no reports of any adverse events by Peeters (45). Blumer (10) reported and increased mortality

in patients receiving propofol compared to standard of care (control group mainly used midazolam).

Blumer (10) is an unpublished industry sponsored trial designed to measure the "safety and efficacy of propofol compared with standard sedative regimens in mechanically ventilated PICU patients". The pharmaceutical company was contacted to obtain further data on this particular study and an abstract was obtained. Several unsuccessful attempts, by phone and e-mail, were done to obtain more data on this study. Based on the abstract provided and the e-mail received from the Global Safety Physician for Diprivan, we obtained the following data: "a draft CSR* was produced but an approved report was never produced". Through the abstract we could retrieve the following information: the study enrolled 348 patients, 22 were excluded prior to the trial initiation, 326 were randomised to receive propofol 1% (109 participants); propofol 2% (112 participants) and standard sedative agents (SSA) (105 participants). The median duration of sedation was 81, 74 and 65 hours, respectively. This study identified a higher number of deaths in children receiving propofol: 9 (8.2%) deaths in propofol 1% (p=0.17); 12 (10.7%) deaths in propofol 2% (p=0.05) compared to 4 (3.8%) deaths in the SSA group. They report that "no patient experienced events consistent with PRIS". *CSR definition was not provided; it is likely an abbreviation for "corporate social responsibility". The p values were not provided in the abstract, but calculated for this review. This trial would not fulfil our inclusion criteria, but due to relevant data reporting on our primary outcomes we added the data as a *post hoc.*

Six studies (30,34,41,45-47) did not report clearly the length of propofol infusion, but reported duration of surgery or procedure and this value was used as a surrogate measure of duration of anaesthesia.

Comparison:

The comparison is called "standard of care" in this review and it is any sedative agent (or combination of agents) used in comparison with propofol. If a trial did not report a drug dosage used in the results section, but did report in the methods section, this value was used.

The included studies were divided in subgroups according to the drug being studied in comparison to propofol. Four studies used midazolam infusion for sedation: Deng (35) and Xia

(30) used 1 mcg/kg/min, Salvitano (39) reports as dose consumed for midazolam and reports as 9.47 mcg/kg/hour. Peeters (45) compared propofol to midazolam and it was adjusted to maintain COMFORT score higher than 17, no midazolam dose was provided. One study used ketamine (46) as a continuous infusion for children undergoing cardiac catheterization at dose 7.7 mg/kg/hr. Pentobarbital bolus (4.1 mg/kg) associated with midazolam (0.9 mg/kg) bolus and fentanyl bolus was used by Pershad (48) for MRI procedures. LeBlanc (36) used similar protocol for both groups: isoflurane and fentanyl for the surgical procedure and one group received propofol until cerebral electrical silence; no propofol was added to the comparison group. Anesthetic gas was used in all other 26 studies. The most common gas used was sevoflurane in 14 studies with a concentration ranging between 1.5% to 8%. The second most frequent gas used was halothane in 6 studies, followed by isoflurane in four studies and enflurane in one study, one study used nitrous oxide and another a combination of three different gases.

Outcome:

Only one trial reported on the primary outcome. The trial reported on increased mortality in children receiving propofol compared to standard sedatives agents (10). Children receiving propofol 1% had 8.2% mortality (9/109) (p=0.17), propofol 2% had 10.7% (12/112) (p=0.05) and were compared to standard sedatives (3.8% mortality (4/105)). This study was done in mechanically ventilated PICU patients.

Risk of bias in included RCTs

The risk of bias in the included trials was high overall. Figure 4 and figure 5 represent the risk of bias table and graph for the RCTs included in this review.

Allocation (selection bias)

An adequate sequence generation was reported in 40% of trials.

Blinding (performance bias and detection bias)

Blinding of outcome assessors was reported in 45% of trials.

Incomplete outcome data (attrition bias)

Attrition bias was difficult to assess as it was often not reported in the included trials.

Selective reporting (reporting bias)

The analysis of reporting characteristics of the 34 included trials showed that three trials included in this review were "likely" to have reporting bias, as they intended to measure serious events, but did not report in their results.

Fourteen trials were "possibly" to have reporting bias for serious adverse events, as other events were measured and reported, but the presence/absence of serious events were not mentioned or mentioned through general statements.

Ten trials reported enough information on the adverse events and were unlikely to have a reporting bias.

Other potential sources of bias

The only trial reporting in increased mortality in patients receiving propofol was a pharmaceutical sponsored trial and never fully published in a peer review journal. The reason to stop the trial prematurely is not clear and it was not provided by the pharmaceutical company contact person, for all these reasons we consider this trial as "high risk of bias".

Strength of Evidence for RCTs (GRADE)

The GRADE score for RCTs was based on a single trial (10) that reported on the primary outcome. It scored "serious" for "study limitation" on risk of bias, as this trial was stopped early and not published in a peer reviewed journal, also no detailed information on selection of participants, randomization, concealment, blindness, measurement methods and attrition were provided in the abstract. The trial had a strong outcome reported (mortality) and head to head comparisons, it was considered "no" on "directness". Consistency score was "no". It was rated

"no" on "precision" (no imprecision") with a statistical significant difference in mortality between one interventional arm and control. As for reporting bias, it was considered "strongly suspected" as the trial was stopped early and the findings are only partially reported through an abstract. The RCT was also upgraded for dose-response gradient and very large effect. The overall RCT GRADE score was "high".

As only one trial reported on primary outcomes, the measurement of the: (i) association of propofol dosage and duration of infusion with PRIS or cardiac arrest; (ii) patient characteristics associated with serious adverse events related to propofol use; and (iii) association between incidence of serious adverse events related to propofol use with patient location were not ascertained.

The mode of surveillance for adverse events in clinical trials was not clearly reported. No trial reported on passive report of adverse events (spontaneous reports).

Non-Randomized studies

Cohorts:

Seven cohorts included a total of 2073 patients, 1027 of them received propofol infusion. The reason for sedation were diverse, 3 studies had general paediatric intensive care patients, 1 was exclusively post-operative cardiac surgery in PICU, 1 study included only patients on status epilepticus in PICU; 2 studies were done in diagnostic imaging. The most frequent comparison was with other sedative drugs with midazolam as the drug of choice in four of these cohorts. One study did not report clearly what the comparison group had used, only stated it was compared to "randomised controls", but we identified this study as being a cohort study and not an RCT despite the term "randomised controls".

The seven included cohorts are reported in more detail below:

Agudelo (17) is a retrospective descriptive observational study of all patients admitted to the intensive care unit during the course of one year who required sedation. The authors compared participants that used propofol infusion (n=71) to participants that did not use propofol

(n=151). Participants that used propofol had a median age of 24 months (IQR 6-65); they were mostly post-operative (62%), mainly post heart surgery (44%); the mean propofol dose used was 2.1 mg/kg/hr (varying from 1 to 4 mg/kg/hr) and for a mean length of 6.7 days (range 0.5 to 40 days). The number of participants in the propofol group needing vasoactive drugs and mechanical ventilation was statistically significant higher than in the control group ($p < 0.05$). A total of 7 patients of 71 (9.8%) receiving propofol had a cardiac arrest (p=0.02) and 4 (5.6%) died (p=0.08). The comparison group had 3 patients that experienced cardiac arrest of 151 included (2.0%) patients and 2 experienced fatal outcomes (1.3%). The comparison group had more metabolic acidosis (20 patients), in comparison to propofol (6 patients). Patients receiving propofol had higher triglycerides and similar creatine kinase in relation to the comparison group. Authors report the program of pharmacovigilance did not identify any of these events as being related to propofol and it was not characterized as PRIS. We contacted the authors seeking individual patient data for the seven patients experiencing cardiac arrest. The author replied stating that he did not have current access to the data, but he recalls that "none of the seven patients who experienced cardiac arrest were associated to administration of propofol".

Amundsen (49) is a prospective cohort study of 108 children with autism spectrum disorder or with idiopathic developmental delay submitted to neuroimaging; the comparison group was 26 children with typical development. Mean age of the children was 47.6 +/- 4.5 months. The propofol initial dose was 12 mg/kg/hr (dose used ranged from 10.8 to 13.2 mg/kg/hr) and titrated based on usual clinical indicators such as blood pressure, pulse rate, respiratory rate and movement. The duration of exam ranged from 90 to 120 minutes. There were a total of 4 adverse events: airway obstruction in two participants, treated by insertion of an oral airway; bradycardia in one participant not requiring treatment; and premature termination of the study in one participant due to excessive salivation. There was no occurrence of any other adverse events such as hypoxia, hypotension, apnea, cyanosis or nausea. No adverse events were reported in the comparison group.

Cray (50) is a prospective cohort study of 91 paediatric patients with complex congenital heart defects, undergoing cardiopulmonary bypass that were suitable for a fast track extubation protocol. The comparison group was a historical cohort of 135 similar patients, matched for age

and cardiac lesion, which underwent surgery in the previous year. The median duration of propofol infusion was 3 hours 35 minutes (range 50 minutes to 15 hours). The mean dose of propofol infusion was 4.2 mg/kg/hr (range 1.5-13.8 mg/kg/hr). Children also received morphine infusion 10-40 mcg/kg/hr. The patients receiving propofol had a length of stay in PICU of 1.7 days compared to 2.6 days in the control group ($p<0.05$). One patient had to be re-intubated after extubation due to excessive sedation in the propofol group. No reports of other adverse events in either group.

Kiringoda (51) is a retrospective cohort study of children requiring sedation for research related imaging. There were a total of 607 patients who received 1480 propofol anaesthesia procedures. The median age of the children was 5.1 years (range 0.2-17.5 years). The only information about the controls was that they were "randomly assigned", but after careful consideration we identified this study as a retrospective cohort. The duration of first anaesthesia procedure was 122 minutes (SD 59). The dose of propofol used was not provided. Author was contacted for clarification on propofol dose. Clarification was received stating that 'propofol was titrated to maintain the patient motionless and a patient would usually receive a loading dose of 1-3 mg/kg followed by a continuous infusion that would be titrated to the patient's response. As that was a retrospective study, one of its limitations is that I do not believe we had the total propofol dose administered in all patients'. Adverse events were defined as 'all occurrences that warranted intervention and were noted on the anaesthesia records'. The report states: "There were no cardiorespiratory arrests or episodes of aspiration pneumonia. No adverse event prolonged the hospitalisation of any subject, and only one adverse event changed the hospital course (overnight intensive care unit observation)." They report 67 respiratory events and 17 cardiovascular events in patients using propofol. No mention of adverse events in the comparison group.

Pepperman (18) is a retrospective cohort study of children requiring mechanical ventilation; 106 received propofol and 92 received other sedative agents (61 receiving midazolam). Participants were also divided according to their reason for ICU admission: respiratory, cardiac or neurologic illness. The infusion of propofol had a length of 30 minutes to 156 days. The mean

infusion rate of propofol was 3.4 mg/kg/hr (range from 0.4 to 30 mg/kg/hr). A pathological metabolic acidosis was defined as pH≤7.2 with normal PaCO2 or base excess≥10. Two patients in the propofol group had pathological metabolic acidosis and 3 in the non-propofol group. Overall mortality was 27 deaths, 13 in the propofol group (7 in the cardiac group and 6 in the respiratory group). None of participants who died had lipaemic plasma.

Prins (52) followed children aged from 1 month to 2 years admitted to the paediatric surgical intensive care unit (PSICU) during the first 24 hours after elective craniofacial surgery. The median age was 10 months (IQR 3 -17 months). Parents of eligible participants were asked to give written informed consent for either propofol or midazolam. Propofol was used in 22 patients for 12 (6-17) hours for a dose of 2.4 (1-4) mg/kg/hr. The control group had 23 participants receiving midazolam only. Triglyceride level was 2.00 mmol/L in one patient during propofol infusion, but without metabolic acidosis, physiological parameters disturbances or creatine phosphokinase (CPK) increase. Four patients, three in the propofol group, had raised CPK levels ranging from 261-313 UI/L without metabolic acidosis.

Van Gestel (19) reported a total of 34 paediatric patients admitted to PICU due to status epilepticus; 22 patients used propofol and 12 thiopental, but eight propofol patients did not respond and also received thiopental. The length of propofol infusion was 57 hours (range 10 to 264 hours) and propofol dose 1-5 mg/kg/hr. Propofol was stopped in one participant due to rhabdomyolysis and in three due to hypertriglyceridemia. No report of acidosis in these patients. All recovered after propofol was stopped. Thiopental group had eight deaths, manuscript described as two of them were likely due to thiopental, one due to hepatic failure and one due to pulmonary damage. There were two deaths in patients treated for seizures with propofol.

Based on the cohorts identified in this review, we can say that propofol was used in 1027 patients and a total of 19 deaths were identified (2.53%). The control groups had a total of 24 deaths out of a 1046 controls (2.67%). All the patients experiencing serious adverse events were admitted to the paediatric intensive care unit.

A meta-analysis including the three cohorts reporting the primary outcome was attempted but excluded due high heterogeneity (I^2= 83%).

Risk of bias assessment for cohort studies

The Newcastle-Ottawa scale for risk of bias assessment in cohort studies can be found in table 9. Cohorts had overall good quality with only one study receiving four stars, all remaining studies received seven or more stars.

Strength of Evidence for Cohorts (GRADE)

The GRADE score for cohort studies (n=3) reporting on cardiac arrest received "no" for study limitations as the overall risk of based on the Newcastle-Ottawa scale being low. It was considered "no" for "directness" due the strong outcome reported and head to head comparisons. The consistency was scored "serious" as one cohort found a statically significant difference between groups on cardiac arrest, in favor of the control; the other two cohorts had more adverse events in the control groups. The cohorts were considered "no" on "precision", as one study (17) had a statistically significant difference between groups. As for reporting bias, it was considered "unsuspected" for study publication bias, selective outcome reporting bias and selective analysis reporting bias.

For the additional domains, the dose-response association was considered "detected" as propofol was associated with cardiac arrest when used for prolonged periods of time. We considered the possibly of a confounding "absent" that would decrease observed effect and we considered the strength of association "large" especially due the statiscally significant findings in one study (17). The overall GRADE score for cohort studies is "moderate".

Case series (more than 5 patients)

There were a total of 36 case-series (53-88) (table 10) of 3540 patients aged 1 to 15 years who received propofol infusion for more than 60 minutes. The reason for sedation was divided among: elective surgery in 19 studies; diagnostic imaging in 4 studies; general PICU admission in

6 studies; cardiac procedures in 5 studies; 1 study with patients having an anticipated difficult weaning off mechanical ventilation in PICU (patients with chromosomal anomalies, laryngo/tracheamalacia or pulmonary hypertension) and 1 study with burn patients. Duration of propofol infusion varied from 60 minutes to 2880 minutes (median 85.5; IQR 68.5 - 424.7). Details of each case series can be found in table 10.

Five case series (58, 74, 79, 83, 84) used propofol for more than 12 hours, including a total of 1127 patients. Eleven patients died. One study (58) reported that "ten patients died within one week of receiving propofol infusions, the cause of death was attributed to the underlying disease state in each of the ten patients" and one study (83) had one patient who used propofol for 10 hours and died due to multi-organ failure 14 hours after propofol was stopped. We attempted to contact one author (58) to obtain more information on these 10 deaths, but the email provided had a "failure to deliver" message. None of these case series report on patients developing PRIS or the death as related to propofol infusion.

Based on the case series, out of 3540 patients receiving propofol infusion for more than 60 minutes, 11 had a fatal outcome, 10 reported in one study (58). All the 11 patients identified in the case series as developing the primary outcome (cardiac arrest/death) were admitted to the paediatric intensive care unit. No individual patient data were available to report on the length on propofol infusion and dose in these 11 patients.

Strength of Evidence for Case series (GRADE)

We decided to exclude case-series from the GRADE profile for the outcome "cardiac arrest" as these studies would all have reduced strength of evidence. Also, only 2 of 36 studies had fatal outcomes. One of these two studies had a single patient experiencing a fatal outcome and the other study had 10 patients, but all were described as likely due to the baseline disease.

Case reports (5 or less individual reports per manuscript):

Based on case reports, a total of 33 cases fulfilled our criteria for PRIS (1-3, 89-113) reported in detail in table 11. Twenty cases did not fulfil the case definition for PRIS, mainly due to the lack of metabolic acidosis, and they are presented separately (table 12).

One manuscript described five cases (3) of death in patients with respiratory illness, four with upper airway illness (details of each case can be found in table 11) and one with bronchiolitis. In this particular report of five patients, the propofol was used from 2 to 6 days in a dose 4.8 to 10 mg/kg/hr.

Of the 33 patients who developed PRIS, 18 had a neurological disorder: traumatic brain injury (n=7); seizures (n=4); meningitis/encephalitis (n=3); astrocytoma resection (n=2); cerebral vascular procedures (n=2). Twelve of these 18 patients had a fatal outcome. Respiratory illness was associated with 11 cases of PRIS: laryngitis (n=5); epiglottitis (n=2); foreign body aspiration (n=2 cases); subglottic stenosis (n=1), bronchiolitis (n=1), 9 patients had a fatal outcome. Post-surgical cases of PRIS were: post cardiac surgery (n= 2); cleft palate repair (n=1); post orthopedic surgery (n=1). All four post-surgical, non-neurological patients who developed PRIS survived.

For all cases of PRIS the median duration of propofol infusion was 50.5 hours (IQR 36-91) range (5-144 hours) and median dose of propofol was 7.95mg/kg/h (IQR 7.0-10.5), range (1.8-27.0). The shortest duration of propofol infusion associated with PRIS was 5 hours in two patients that received propofol at a dose of 6 and 12 mg/kg/h. These two patients survived. Thirty of 33 reports of PRIS used propofol for more than 12 hours.

One study (99) reported on two patients admitted for cardiac surgery and received propofol for 8 and 15 hours in a dose of 2.5 and 1.8 mg/kg/h respectively. These two patients developed metabolic acidosis and high lactate which improved after propofol was stopped. The patients also required epinephrine at a dose of 0.2mcg/kg/min. These patients were classified as cases of PRIS due the use of inotropic support associated with metabolic acidosis. This is the only report of a low dose of propofol (less than 4 mg/kg/h) associated with PRIS. Both patients survived.

Of the thirteen patients with PRIS whom survived, 6 had renal replacement therapy, 2 also used extra corporeal membrane oxygenation (ECMO), and 1 had exchange transfusion. One other study (98) reports use of dialysis and ECMO, but ECMO was initiated after patient was receiving chest compressions for a long period of time, this patient did not survive. Four fatal cases of PRIS also used dialysis.

In all case reports of PRIS, the patients were admitted to an intensive care unit.

Strength of Evidence for Case reports (GRADE)

The case reports of PRIS had an overall "low" GRADE score. The reports were "strongly suspected" for reporting bias, as only selected cases were reported and "serious" for study limitations. Graded "present" for plausible confounding and "no" for "strength of association" also, studies were graded "detected" for "dose-response gradient" as reports of propofol use in higher doses and prolonged time were associated with the outcome. All the other fields (directness, consistency, precision) had "no" as score.

Additional analysis

As part of our sensitivity analysis, we isolated patients with PRIS who developed a fatal outcome and found a median duration of propofol use of 67.5 hours (IQR 37.5- 96.5) range (19-144) and a median dose 8 mg/kg/h, (IQR 7.05-10.5) range (4.5-27). Two patients died using a dose of less than 5 mg/kg/hr (4.5 and 4.6mg/kg/h) and they used propofol infusion for 72 and 96 hours respectively (1, 3), also 2 fatal cases had a total duration of propofol of less than 24 hours (96, 97); these two patients had propofol used in a high dose: Propofol dose of 21 mg/kg/h for 11 hours and then 4 mg/kg/h for 8 hours (total of 19 hours) (96); Propofol dose of 20 mg/kg/h for 15 hours and then 4.2 mg/kg/h for 8 hours (total of 23 hours) (97). The non-fatal cases of PRIS had a median duration of propofol use of 48 hours (IQR 15-56) range (5-91) and a median dose 7.8 mg/kg/h, (IQR 6-12) range (1.8-15.6).

Also, as part of our sensitivity analysis, we analysed the 20 cases that did not fulfil the PRIS criteria and found only one with a fatal outcome, this report states that "patient likely died due

to the baseline disorder, a severe viral pneumonia that progressed to septic shock, not the propofol use" (130). Details of patients not fulfilling PRIS criteria can be found in table 12 the main reason for exclusion was the lack of metabolic acidosis (114-131).

The mode of surveillance of adverse events in non-randomized studies was not clearly stated.

We found that propofol infusion used for 12 hours or longer had sensitivity= 90.91 % (95%CI: 75.64%- 97.98%), specificity= 73.68 % (95%CI: 48.80% -90.75%), PPV= 85.71% (95% CI: 69.73 % - 95.14%) and NPV= 82.35 % (95% CI: 56.55 %-95.99%) of being indicative of PRIS.

For dose >4mg/kg/h we found a sensitivity= 93.75% (95%CI: 79.16%-99.05%, specificity= 18.75 %(95% CI 4.27%- 45.66 %), PPV= 69.77% (95% CI: 53.87 %-82.80%) and NPV= 60.00 % (95% CI: 15.40 %-93.51%) of being indicative of PRIS.

When we applied the Bradford-Hill (132) criteria for causality to the reports summarized here, many criteria were fulfilled, such as: consistency (the reports were very consistent in their clinical and laboratory findings); biological gradient (there was predominantly an association between higher doses and longer durations of propofol with PRIS); plausibility and coherence (mitochondrial toxicity and oxidative phosphorylation effect in energy production has been reported previously) (133-135) and temporality (all patients developed the syndrome after been exposed to propofol).

Important to note, the inclusion of non-randomized studies and grey literature were the main source of data for this review of adverse events. The only RCT reporting on the primary outcome was never published in a peer review journal and all the other included reports were non-randomized studies.

Excluded studies

The use of propofol in more than one study arm was the reason for 76 exclusions. As the only included trial (10) reporting on the increased mortality on patients using propofol would be excluded due to propofol use in more than one study arm, we decided to look into the 76 excluded studies due to propofol use in more than one study arm and found no other report of PRIS or cardiac arrest in these 76 studies.

In a case report (117) of a patient with known glycogen storage disorder submitted to an adeno-tonsillectomy received propofol for 83 minutes during surgery in a dose of 6-12 mg/kg/hr. The procedure was uneventful and on the following day (post-operative day 1) patient had abdominal pain, increasing lipase and liver enzymes and metabolic acidosis. This particular case was classified as "not fulfilling PRIS criteria" as the patient developed the symptoms many hours after propofol was stopped and the laboratory findings could be likely due to the baseline disease (glycogen storage disorder and pancreatitis).

One case report (121) of possible PRIS in a patient receiving propofol at 3mg/kg/h during 65 minutes of surgery for spasticity had a diagnosis of adrenoleukodystrophy, the reporting authors did not think the metabolic acidosis and liver dysfunction, electrolytes abnormalities and high triglycerides were caused by PRIS and this case was classified as "not fulfilling PRIS criteria".

Features of the excluded surveys can be found in the discussion section.

Quality of the evidence

The summary of findings table can be found as table 13 and summary of key outcome as table 14. The quality of evidence varies from "low" to "high", with the RCTs receiving a "high", cohorts "moderate" and case reports "low" grading. The RCT and cohort studies were upgraded due large effect size (statiscally significant findings) and case reports were upgraded due dose-response gradient. The RCT was also downgraded due the suspected publication bias.

The overall strength of evidence was considered "moderate" (further research may change our confidence in the estimate of effect and may change the estimate) for the outcome "cardiac arrest" as there are not high evidence studies published, but two studies showed statiscally significant findings for cardiac arrest when propofol was used for for multiple days.

The overall strength of evidence was considered "low" (further research will likely change our confidence in the estimate of effect and will likely change the estimate) for the outcome PRIS as the evidence originated exclusively from case reports.

For this review, all the data on the main outcome originated from non-randomized studies and grey literature. Currently, the evidence supporting the withdrawal of licensed therapeutic drugs has mainly come from passive reports (136). Usually designed to evaluate treatment efficacy/effectiveness, RCTs are often done over a short period of time, with a relatively small number of participants (7). RCTs are known to be poor at identifying and reporting harms, which can lead to a misconception that a given intervention is safe, when its safety is actually unknown (7). It has been suggested that the "evidence pyramid" be inverted when considering harms, as randomised clinical trials are insufficiently powered to detect rare events, while non-randomized studies may be extremely informative in identifying emerging harms (141).

Potential biases in the review process

As for possible biases in the review process, the trial done by Blumer (10) was initially excluded from this review as it had propofol in two study arms and it would not fulfill inclusion criteria. As a *post -hoc* approach, it was decided to include it due to the relevance of its findings. To assure no further relevant outcomes were missed due to this exclusion criteria we looked at the other excluded trials for the same reason (propofol use in more than one arm) we found zero reports of PRIS or cardiac arrest/death in those.

Effects of interventions

The beneficial effects of the intervention were not assessed in this review.

Discussion

We conducted a systematic review of serious adverse events associated with propofol infusion in paediatric patients. The main objectives were to identify cases of PRIS, a rare and poorly understood syndrome and cardiac arrest that progressed to a fatal outcome or not.

We approached the review of rare events through broad inclusion criteria, including any pediatric patient receiving propofol infusion for 60 or more minutes, for any reason, reported as an interventional or observational study, to capture all relevant events.

PRIS, as defined in this review, was only identified in case reports. There was considerable consistency of findings on these reports, such as patient's characteristics and clinical findings,

but also PRIS occurred predominantly in patients receiving more than 4 mg/kg/h of propofol for longer than 12 hours. It is also important to note that patients with concomitant respiratory illness, most frequently upper respiratory infections, and neurologic injuries are the group most reported to develop PRIS.

Among the clinical trials, we identified one unpublished industry-sponsored trial (10) that was stopped early and reported an increased mortality in critically ill paediatric patients in intensive care units receiving propofol infusion. Patients with upper airway disorders were excluded, as they were recognized as being a high risk population. The non-randomized studies provided important additional information on our main outcomes. Three cohorts reported on mortality of patients receiving propofol and one found a significant increase in cardiac arrest in patients using propofol (17). Cohorts, case series and case reports also found muscle breakdown and metabolic changes associated with propofol, which could be early signs of PRIS. The case reports of PRIS were very consistent in their findings, with metabolic and muscular disarrangements leading to death in the majority of patients. The cases that did not have metabolic acidosis and therefore did not fulfil our definition of PRIS had a better outcome with only one fatal case.

The main contribution to the outcome of interest came from uncontrolled observational studies, which means that a causal link between propofol infusion and PRIS and death could not be clearly defined. The patients who developed the fatal outcome were critically ill children and received other drugs concomitantly with propofol, factors can be seen as possible confounders. The case reports highlight the similarity among patients who developed PRIS, the constellations of findings (rhabdomyolysis, metabolic acidosis, lipaemic plasma and arrhythmias) were very similar among cases and not anticipated to be found on some of these particular patient populations.

Case reports are the cornerstone of pharmacovigilance. Regulatory agencies rely mainly on the health care providers to report unusual events, as the majority of drug warnings and drugs removed from the market are based on individual case reports of adverse events (136). Propofol is a good example of a “drug warning” issued by multiple health care agencies around

the world (11-12) based on case reports and on the unpublished randomised clinical trial (10). PRIS is still a poorly understood syndrome; it seems to be linked to unpredictable individual risk factors, but also associated with high doses and prolonged infusions of propofol. The warning issued by health regulatory agencies is exclusively focused on patients admitted to paediatric intensive care units, due to the case reports mainly describing critically ill patients admitted to PICU. Although, it is important to note that the same patient population may receive high doses of propofol for hours in a different setting, for example the operating room. This was one of the objectives of this review, but despite multiple included studies of patients in the operating room, none reported on cases of PRIS, but also none of these reports had a prolonged use of propofol infusion as seen in patients in the intensive care units.

Although, we had no intention to include surveys in this review, we would like to highlight three surveys done with paediatric intensivists in five different countries specifically asking about the use of propofol infusion in paediatric intensive care units (PICU) after the drug warnings issued by regulatory agencies. In one survey done in Australia and New Zealand PICUs (137), 33 of the 44 intensive care physicians replied representing 12 of 13 PICUs surveyed. In this report, 80% the physicians report a restriction on the propofol dose used to less than 5 mg/kg/hr and duration of less than 72 hours. One other survey done United States (n=33 PICUs) and United Kingdom PICUs (n=15 PICUs) found that propofol was used for continuous sedation in 61% of US PICUs and 47% of UK PICUs in specific patient population (as a bridge to extubation; as a fast track post-operatory recovery and for neurological assessment after head injury), also it would be limited to no more than 48 hours (US) and with a dose limit of 4 mg/kg/hr (138). A more recent survey done with intensive care units across Germany, identified that of 184 surveys (134 mixed PICU/neonatal ICUs; 28 PICUs and 22 neonatal ICUs), all used propofol infusion for ≥ 3 hours with a time limit of less than 24 hours. A dose limit was used by 51% of respondents, median dose limit of 4 mg/kg/hr, varying from 1 to 10 mg/kg/hr. The main indications for propofol application were: difficult sedation (44%), postoperative ventilation (43%), and difficult extubation (30%). Seven cases of PRIS were reported by seven centres in this German study (139).

Some of the strengths of this review, other than the inclusion of more than 100 studies of propofol use in children, are the fact that we could identify a length of propofol infusion longer than 12 hours and a dose higher than 4mg/kg/h were linked to the majority of cases of PRIS. Further, we could document that the inclusion on non-randomized and unpublished studies are key to obtain data in a review of rare events.

One of weaknesses of this review is the fact that despite 34 trials fulfilling the inclusion criteria, only one trial reported on mortality associated with propofol infusion in paediatric patients, but it was never published in a peer review journal and the data could only be retrieved from an abstract. Also, as the majority of data originated from case-reports and case series, study designs without a denominator, making it difficult to assess frequency of events.

We hope that this review provides guidance to clinicians and policy makers around the world when considering use propofol infusion in children.

Summary of main results

A total of 91 serious adverse events (PRIS and cardiac arrest) associated with propofol infusion were identified in this review; 21 were identified in an unpublished RCT (10) comparing propofol 1%, propofol 2% and standard sedative agents (SSA) (patients mainly received lorazepam in this group) in patients sedated in the paediatric intensive care unit, participants received propofol infusion for 81 hours (propofol 1% group); 74 hours (propofol 2% group) and the comparison group received SSA for 65 hours. The dose of propofol used in this trial was 4 to 6 mg/kg/hr. The mortality of participants receiving propofol 1% was 8.2% (9/109) (p=0.17); propofol 2% was 10.7% (12/112) (p=0.05) compared to SSA which was 3.8% (4/105). One cohort study (Agudelo 2011) reported an increased number of cardiac arrest 9.9% (7/71) (p=0.02) and increased mortality 5.6% (4/71) (p=0.08) in patients receiving propofol, the comparison group had 2% (3/151) of patients experiencing cardiac arrest, of which 2 died (1.3%). Two other cohorts reported on a total of 15 deaths in patients receiving propofol, the controls had 22 deaths. The case series reported a total of 11 deaths out of 3540 (3.1%) in patients receiving propofol. The case reports had a total of 53 cases selected as possible cases of PRIS. Of these, 33 cases fulfilled our definition of PRIS. Of 33 case reports of patients fulfilling

PRIS criteria, 20 died; 6 of the 13 survivors used renal replacement therapy (two used ECMO with hemo filter). We did not identify any case-control studies fulfilling our inclusion criteria.

All the patients/participants experiencing PRIS were admitted to Intensive Care Units and received propofol for a median of 50.5 hours (IQR 36-91), range (5-144)) and median dose of 7.95 mg/kg/h (IQR 6.45-11.77), range (1.8-27). The patients with PRIS who developed a fatal outcome received propofol for a median of 67.5 hours (IQR 37.5- 96.5) range (19-144) in a median dose of 8 mg/kg/h, (IQR 7.05-10.5) range (4.5-27). Thirty of 33 reports of PRIS used propofol for more than 12 hours and 31 reports used a dose higher than 4 mg/kg/h. The reasons for intensive care admission were neurologic in 18 patients (54.5%), respiratory 11 patients (33.3%), and post operatory 4 patients (12.1%). All the patients experiencing cardiac arrest were admitted to intensive care unit, and an individual dose or duration of infusion of propofol could be obtained despite attempt to contact authors.

Of the studies identified in this review, a total of 5633 children received propofol for more than 60 minutes and did not have any serious events (PRIS and cardiac arrest) associated with it.

We found that propofol infusion used for 12 hours or more hours had a sensitivity= 90.91 % (95%CI: 75.64%- 97.98%), specificity= 73.68 % (95%CI: 48.80% -90.75%), PPV= 85.71 %95% CI: 69.73 %- 95.14 % and NPV= 82.35 % (95% CI: 56.55 %-95.99%) of being associated with PRIS.

For dose >4mg/kg/h we found a sensitivity= 93.75% (95%CI: 79.16%-99.05 %, specificity= 18.75 %(95%CI 4.27%- 45.66 %), PPV= 69.77% (95% CI: 53.87 %-82.80%) and NPV= 60.00 % (95% CI: 15.40 %-93.51%) of being indicative of PRIS.

In this review of adverse events, the inclusion of non-randomized studies and grey literature provided the totality of data on the main outcomes. If only published RCTs had been included, as is common in systematic reviews, zero reports of PRIS and/or cardiac arrest associated with propofol infusion in children would have been identified.

The mode of surveillance of adverse events was not reported for any study design.

Overall completeness and applicability of evidence

To our knowledge this is the most comprehensive review of propofol infusion in paediatric patients to date, including multiple study designs and with broad inclusion criteria, we combined data of more than 100 manuscripts in one single document.

Agreements and disagreements with other studies or reviews

Previous reviews of propofol infusion have been published, but none of them had followed a broad approach as we have, searching multiple databases, using strict inclusion and exclusion criteria and as consequence they had reported on a fraction of studies presented here (133-135,140). These reviews seemed to have a different focus than this present review, they were mainly clinical reviews, with an intention to explain the syndrome (clinical features, biochemical interactions), but not to summarize the evidence in a systematic fashion as we have.

Authors' conclusions

Implications for practice

In this review of serious adverse events associated with propofol infusion in children, we identified 91 cases of serious events (PRIS or cardiac arrest). We also could identify the median dose of propofol and duration of infusion associated with PRIS as 7.95mg/kg/h and 50.5 hours, also that the dose of 4mg/kg/h and a duration of propofol infusion longer than 12 hours were indicative of PRIS, so we could extrapolate the findings and say that anything lower than these values could be used with caution. Although, it is important to note that there were also a patient who developed PRIS with as propofol dose of 1.8mg/kg/h and two others with an infusion as short as 5 hours, these three patients were post operatory and all survived. All patients with PRIS were admitted to intensive care units due to neurological or respiratory disorder or for post-operatory care.

The data obtained in this review originated mainly from non-randomised studies, case reports and data not published in a peer review journal. The inclusion of non-randomised studies and grey literature may be the cornerstone for data collection in other reviews of rare events, or perhaps for any review of adverse events, given the poor reporting of adverse events in general

and especially in clinical trials (7). The only clinical trial reporting on the main outcome for this review was an unpublished pharmaceutical sponsored trial. It is likely

Implications for research

There are suggestions, based on identified surveys, recent case series and cohorts that propofol has been used for continuous infusion in paediatric intensive care units despite the health regulatory agencies' warnings and contra-indications. There is a need to achieve more information on number of patients using propofol infusion, patient's characteristics and complication seen to better understand PRIS, improving evidence based practice and patient care. We encourage authors to report their experience, if any, in use of propofol infusion in paediatric intensive care.

Declarations of interest

None known.

Funding support

None.

Differences between protocol and review

The main deviation from protocol was the addition of an objective: "To identify if the inclusion of non-randomized studies and grey literature in a review of adverse events provide relevant data that would be otherwise overlooked if only clinical trials would be included". We added this objective due the importance of using non randomized studies and grey literature in reviews of rare events and the clear documentation of it through this review. Other deviations from protocol seen in this review are related to the lack of clinical trials presenting the main outcome of interest. The analysis plan reported in the protocol for this review could not be undertaken, namely meta-analysis and subgroup analysis.

Also, we looked into trials that were excluded due to the use of propofol in more than one study arm to check if there were reports on PRIS or cardiac arrest, none was found other than Blumer (10), which was included in this review.

PRISMA Harms checklist

The PRISMA Harms checklist for this review can be found as table 15.

Acknowledgements

We would like to thank the translators: Derek Wang (Chinese); Dogan Saglam (Turkish); Elena Taianovskaia and Tatiana Kryuchkova (Russian); Giovana Bervian (Italian); Harm Van Seijen (Dutch); Marie Jose Dumont (French); Suzanne King Jones and Jens Gerbitz (German); Liliane Zorzela (Portuguese and Spanish).

Contributions of authors

Conceiving the review: Liliane Zorzela (LZ), Katherine Pohlman (KP), Salima Punja (SP), Lisa Hartling (LH), Ari Joffe (AJ), Yoon Loke, (YL) Sunita Vohra (SV)

Co-ordinating the review: LZ, KP, SP, LH, AJ, YL, SV

Undertaking manual searches: LZ

Screening search results: LZ, KP

Organizing retrieval of papers: LZ

Screening retrieved papers against inclusion criteria: LZ, KP

Appraising quality of papers: LZ, KP

Abstracting data from papers: LZ, KP

Writing to authors of papers for additional information: LZ

Providing additional data about papers: LZ and Ben Vandermeer (BV)

Obtaining and screening data on unpublished studies: LZ, KP

Data management for the review: LZ

Entering data into Review Manager (RevMan 5.3): LZ

RevMan statistical data: LZ, BV

Other statistical analysis not using RevMan: BV, LZ

Interpretation of data: LZ, KP, SP, LH, AJ, YL, SV

Statistical inferences: LZ, KP, SP, LH, AJ, YL, SV

Writing the review: LZ, KP, SP, LH, v, YL, SV

Securing funding for the review: none

Performing previous work that was the foundation of the present study: LH, YL, SV

Guarantor for the review (one author): SV

Person responsible for reading and checking review before submission: LZ, KP, SP, LH, AJ, YL, SV

References

1. Bray RJ. Fatal myocardial failure associated with a propofol infusion in a child. Anesthesia 1995;50(1):94.
2. Hanna JP, Ramundo ML. Rhabdomyolisis and hypoxia associated with prolonged propofol infusion in children. Neurology 1998;50(1):301-303.
3. Parke TJ, Stevens JE, Rice AS, Greenaway CL, Bray RJ, Smith PJ, Waldmann CS, Verghese C. Metabolic acidosis and fatal myocardial failure after propofol infusion in children: five case reports. BMJ 1992;305(6854):613-616.
4. Cremer OL, Moons KG, Bouman EA, Kruijswijk JE, de Smet AM, Kalkman CJ. Long-term propofol infusion and cardiac failure in adult head-injured patients. Lancet 2001;357(9250):117-118.
5. Eriksen J, Povey HM. A case of suspected non-neurosurgical adult fatal propofol infusion syndrome. Acta Anaesthesioll Scand 2006;50(1):117-119. [MEDLINE: 16451160]
6. Kang TM. Propofol infusion syndrome in critically ill patients. Ann Pharmacother 2002;36(9):1453-1456.
7. Loke YK, Golder SP, Vandenbroucke JP. Comprehensive evaluations of the adverse effects of drugs: importance of appropriate study selection and data sources. Ther Adv Drug Saf 2011;2(2):59-68.

8. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0 .. The Cochrane Collaboration 2011;[updated March 2011]:Available from www.cochrane-handbook.org.
9. Roberts RJ, Barletta JF, Fong JJ, Schumaker G, Kuper PJ, Papadopoulos S, Yogaratnam D, Kendall E, Xamplas R, Gerlach AT, Szumita PM, Anger KE, Arpino PA, Voils SA, Grgurich P, Ruthazer R, Devlin JW. Incidence of propofol-related infusion syndrome in critically ill adults: a prospective, multicenter study. Crit Care 2009;13(5):R169. [MEDLINE: 19874582]
10. Blumer JL, Goodale D. Randomized controlled study of propofol sedation for mechanically ventilated infants and children. Unpublished abstract 2002.
11. Health Canada. Drugs and Health Products. Update on the use of propofol. http://www.hc-sc.gc.ca/dhp-mps/medeff/advisories-avis/prof/_2002/propofol_pediatric_2_nth-ah-eng.php July10, 2002.
12. FDA / Astra Zeneca Pharmaceuticals. Dear health care provider letter. www.fda.gov/ medwatch/safety/2001/diprivan deardoc.pdf 2001
13. Zorzela L, Loke YK, Ioannidis JP, Golder S, Santaguida LP, Altman DG, Moher D, Vohra S.PRISMA Harms: improving harms reporting in systematic reviews. Chapter 4.Unpublished.
14. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm.
15. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ. What is "quality of evidence" and why is it important to clinicians? BMJ 2008;336(7651):995-998.
16. Berkman N, Lohr K, Ansari M,McDonagh M, Balk E, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet], November 2013. http://www.ncbi.nlm.nih.gov/books/NBK174881/

17. Agudelo SC, Mencía S, Faro A, Escudero V, Sanavia E, López-Herce J. Continuous propofol perfusion in critically ill children [Propofol en perfusión continua en ninos en estado crítico]. Med Intensiva 2012;36(6):410-415.
18. Pepperman M, Macrae D. A comparison of propofol and other sedative use in paediatric intensive care in the United Kingdom. Paediatr Anaesth 1997;7(2):143-153.
19. van Gestel JP, Blusse van Oud-Alblas JH, Malingre M, Ververs FF, Braun KP, van Nieuwenhuizen O. Propofol and thiopental for refractory status epilepticus in children. Neurology 2005;65(4):591-592.
20. Cohen IT, Finkel JC, Hannallah RS, Hummer KA, Patel KM. Rapid emergence does not explain agitation following sevoflurane anaesthesia in infants and children: a comparison with propofol. Paediatr Anaesth 2003;13(1):63-67.
21. Cohen IT, Finkel JC, Hannallah RS, Goodale DB. Clinical and biochemical effects of propofol EDTA vs sevoflurane in healthy infants and young children. Pediatr Anaesth 2004;14(2):135-142.
22. Erb TO, Hall JM, Ing RJ, Kanter RJ, Kern FH, Schulman SR, Gan TJ. Postoperative nausea and vomiting in children and adolescents undergoing radiofrequency catheter ablation: a randomized comparison of propofol- and isoflurane-based anesthetics. Anesth Analg 2002;95(6):1577-1581.
23. Deramoudt V, Gaudon M, Malledant Y, Chatellier A, Saint-Marc C, Lecallonnec A. Effects of propofol on intraocular pressure in surgery of strabismus in children [Effets du propofol sur la ,pression intraoculairedans la chirurgie du strabisme de l'enfant]. Ann Fr Anesth Reanim 1990;9(1):1-5.
24. Hasani A, Gecaj-Gashi A, Llullaku S, Jashari H. Postoperative analgesia in children after propofol versus sevoflurane anesthesia. Pain Med 2013;14(3):442-446.
25. Fudickar A, Smigaj K, Ensenauer R, Fischer G, Dutschke P, Steinfath M, Bein B. Effect of propofol and sevoflurane on acid-base balance during pediatric heart catheterization. Minerva Anestesiol 2013;79(6):626-633.

26. Kaddu R, Bhattacharya D, MD, Metriyakool K, Thomas R, Tolia V. Propofol compared with general anesthesia for pediatricGI endoscopy: is propofol better? Gastrointest Endosc 2002;55(1):27-32.
27. Martin TM, Nicolson SC, Bargas MS. Propofol anesthesia reduces emesis and airway obstruction in pediatric outpatients. Anesth Analg 1993;76(1):144-148.
28. Woodward WM, Barker I, John RE, Peacock JE. Propofol infusion vs thiopentone/isoflurane anaesthesia for prominent ear correction in children. Paediatr Anaesth 1997;7(5):379-383.
29. Uezono S, Goto T, Terui K, Ichinose F, Ishguro Y, Nakata Y, Morita S. Emergence agitation after sevoflurane versus propofol in pediatric patients. Anesth Analg 2000;91(3):563-566
30. Xia WF, Liu Y, Zhou QS, Tang QZ, Zou HD. Protective effect of propofol and its relation to postoperation recovery in children undergoing cardiac surgery with cardiopulmonary bypass. Pediatr Cardiol 2011;32(7):940-946.
31. Larsson S, Asgeirsson B, Magnusson J. Propofol-fentanyl anesthesia compared to thiopental-halothane with special reference to recovery and vomiting after pediatric strabismus surgery. Acta Anaesthesiol Scand 1992;36(2):182-186.
32. Schmidt J, Fechner J, Fritsch B, Schmitz B, Carbon R, Rosch W, Albrecht S. Propofol-remifentanil versus sevoflurane-remifentanil for anesthesia for pediatric procedures in infants, children and adolescents. Anaesthesist 2001;50(10):757-766.
33. Weir PM, Munro HM, Reynolds PI, Lewis IH, Wilton NC. Propofol infusion and the incidence of emesis in pediatric outpatient strabismus surgery. Anesth Analg 1993;76(4):760-764.
34. Kobacas S, Balcioglu T, Cevik A. The comparison of the effect of inhalation and intravenous anesthesia on the neuromuscular block of rocuronium in children. Turk Anest Rean Cem Meemast 2002;30:127-132.
35. Deng XQ, Wang M, Ji Y. Clinical comparison of propofol and remifentanil anaesthesia with sevoflurane and remifentanil anaesthesia for children with cleft lip and palate repair surgery. Hua Xi Kou Qiang Yi Xue Za Zhi (West China Journal of Stomatology) 2009;27(5):531-534

36. LeBlanc JG, Blackstock D, Macnab A, Gagnon F, Gagnon R, Russell J, Ring T. Effects of propofol on cerebral oxygenation during cardiopulmonary bypass in children. Can J Anaesth 2000;47(11):1082-1089.
37. Pedersen N, Jensen A, Kilmose L, Olsen K.. Propofol-remifentanil or sevoflurane for children undergoing magnetic resonance imaging? A randomised study. Acta Anaesthesiol Scand 2013;57(8):988-995.
38. Mironov PI, Sergeev AV. Anestesic maintenance during cochlear implnatation in children. Anesteziologiya i reannimatologia 2012;1:18-20.
39. Salvitano BM, Diaz MVP, Flores MdCL, Aldana LR. Intravenous general anesthesia: Propofol-Fentanyl vs Midazolan-Fentanyl in pediatric patient [Anestesia general endovenosa: propofol-fentanyl vs. midazolam-fentanyl en el paciente pediatrico]. Rev Anest Mex 1997;9(6):204-211.
40. Steinmetz J, Holm-Knudsen R, Sorensen MK, Eriksen K, Rasmussen LS. Hemodynamic differences between propofol–remifentanil and sevoflurane anesthesia for repair of cleft lip and palate in infants. Pediatr Anaesth 2007;17(1):32-37.
41. Xia WF, Liu Y, Zhou QS, Tang QZ, Zou HD. Comparison of the effects of propofol and midazolam on inflammation and oxidase stress in children with congenital heart disease undergoing cardiac surgery. Yonsei Med J 2011;52(2):326-332.
42. Erb TO, Kanter RJ, Hall JM, Gan TJ, Kern FH, Schulman SR. Comparison of electrophysiologic effects of propofol and isoflurane-based anesthetics in children undergoing radiofrequency catheter ablation for supraventricular tachycardia. Anesthesiology 2002;96(6):1386-1394.
43. Abbasi H, Razavi S, Hajiesmaeili M, Behdad S, Ghiamat M, Eghbali A. Unwanted intra-operative penile erection during pediatric hypospadiasis repair comparison of propofol and halothane. Urol J 2013;10(3):988-992.
44. Hasani A, Ozgen S, Baftiu N. Emergence agitation in clhildren after propofol versus halothane anesthesia. Med Sci Monit 2009;15(6):CR302-CR306.
45. Peeters MY, Prins SA, Knibbe CA, DeJongh J, van Schaik RH, van Dijk M, van der Heiden IP, Tibboel D, Danhof M. Propofol pharmacokinetics and pharmacodynamics for depth of

sedation in nonventilated infants after major craniofacial surgery. Anesthesiology 2006;104(3):466-47

46. Lebovic S, Reich DL, Steinberg LG, Vela FP, Silvay G. Comparison of propofol versus ketarnine for anesthesia in pediatric patients undergoing cardiac catheterization. Anesth Analg 1992;74(4):490-494.
47. Martin PH, Murthy BV, Petros AJ. Metabolic, biochemical and haemodynamic effects of infusion of propofol for long-term sedation of children undergoing intensive care. Br J Anaesth 1997;79(3):276-279.
48. Pershad J, Wan J, Anghelescu DL. Comparison of propofol with pentobarbital/midazolam/fentanyl sedation for magnetic resonance imaging of the brain in children. Pediatrics 2007;120(3):e629-e636.
49. Amundsen LB, Artru AA, Dager SR, Shaw DW, Friedman S, Sparks B, Dawson G. Propofol sedation for longitudinal pediatric neuroimaging research. J Neurosurg Anesthesiol 2005;17(4):180-192.
50. Cray SH, Holtby HM, Kartha VM, Cox PN, Roy WL. Early tracheal extubation after paediatric cardiac surgery: the use of propofol to supplement low-dose opioid anaesthesia. Paediatr Anaesth 2001;11(4):465-471.
51. Kiringoda R, Thurm AE, Hirschtritt ME, Koziol D, Wesley R, Swedo SE, O'Grady NP, Quezado ZM. Risks of propofol sedation/anesthesia for imaging studies in pediatric research: eight years of experience in a clinical research center. Arch Pediatr Adolesc Med 2010;164(6):554-560.
52. Prins SA, Peeters MY, Houmes RJ, van Dijk M, Knibbe CA, Danhof M, Tibboel D. Propofol 6% as sedative in children under 2 years of age following major craniofacial surgery. Br J Anaesth 2005;94(5):630-635.
53. Absalom A, Amutike D, Lal A, White M, Kenny GN. Accuracy of the `Paedfusor' in children undergoing cardiacs urgery or catheterization. Br J Anaesth 2003;91(4):507-513.
54. Agzamkhodzhaev TS, Tokhirov ShM, Kholmukhamedov BM. Changes in hemodynamics with a combination of epidural anesthesia and propofol sedation in children. Anesteziol Reanimatol 2002;1:6-9.

55. Arosio EM, Gramegna M, Savoia G, Borroni M, Colombo N, Ego E, Levati A. Propofol anesthesia in spontaneous respiration in pediatric patients undergoing nuclear magnetic ressonance [Anestesia con propofol in respiro spontaneo in pazienti pediatrici sottoposti a risonanza magnetica nucleare.]. Minerva Anestesiol 1993;59(12):759-762.
56. Ayala-Sandoval S, Gonzalez-Guzman MP, Pineda-Diaz MV. Propofol infusion as induction and maintenance of anesthesia in pediatric patients [Propofol como agente de induccion y mantenimiento anestesico en infusion en el paciente pedaitrico]. Rev Mex Anest 1991;14(2):55-62.
57. Chidambaran V, Sadhasivam S, Diepstraten J, Esslinger H, Cox S, Schnell B, Samuels BM, Inge T, Vinks AA, Knibbe CA. Evaluation of propofol anesthesia in morbidly obese children and adolescents. BMC Anesthesiol 2013;13(8):1-9.
58. Cornfield D, Tegtmeyer K, Nelson M, Milla C, Sweeney M. Continuous propofol infusion in 142 critically ill children. Pedaitrics 2002;110(6):1177-1181.
59. Blussee van Oud-Alblas HJ, Peters JW, de Leeuw TG, Vermeylen KT, de Klerk LW, Tibboel D, Klein J, Weber F. A comparison in adolescents of composite auditory evoked potential index and bispectral index during propofol-remifentanil anesthesia for scoliosis surgery with intraoperative wake-up test. Anesth Analg 2008;107(5):1683-1688.
60. Cox RG, Levy R, Hamilton MG, Ewen A, Farran P, Neil SG. Anesthesia can be safely provided for children in a high-field intraoperative magnetic resonance imaging environment. Pediatr Anaesth 2011;21(4):454-458.
61. Diepstraten J, Chidambaran V, Sadhasivam S, Esslinger HR, Cox SL, Inge TH, Knibbe CA, Vinks AA. Propofol clearance in morbidly obese children and adolescents: influence of age and body size. Clin Pharmacokinet 2012;51(8):543-551. Engelhardt, 2007
62. Engelhardt T, Chan M, McCheyne A, Karsli C, Luginbuehl I, Bissonnette B.. The effect of varying continuous propofol infusions on plasma cyclic guanosine 3',5'-Monophosphate concentrations in anesthetized children. Anesth Analg 2007;105(3):616-619.
63. Engelhardt T, McCheyne AJ, Morton N, Karsli C, Luginbueh I, Adeli K, Walsh W, Bissonnette B. Clinical adaptation of a pharmacokinetic model of Propofol plasma concentrations in children. Pediatr Anaesth 2008;18(3):235-239.

64. Ivani G, Della Rocca M, De Negri P, Corca D, Cuneo GP, Caffarena PE. Continuous-flow propofol in pediatric oncologic surgery [Le propofol en debit continu en chirurgie oncologique pediatrique]. Cah Anesthesiol 1991;39(6):409-410.
65. Jeleazcov C, Ihmsen H, Schmidt J, Ammon C, Schwilden H, Schuttler J, Fechner J. Pharmacodynamic modelling of the bispectral index response to propofol-based anaesthesia during general surgery in children. Br J Anaesth 2008;100(4):509-516.
66. Jeleazcov C, Schmidt J, Schmitz B, Becke K, Albrecht S. EEG variables as measures of arousal during propofol anaesthesia for general surgery in children: rational selection and age dependence. Br J Anaesth 2007;99(6):845-854.
67. Klockars JG, Hiller A, Munte S, van Gils M, Taivainen T. Spectral entropy as a measure of hypnosis and hypnotic drug effect of total intravenous anesthesia in children during slow induction and maintenance. Anesthesiology 2012;116(2):340-351.
68. Knibbe CA, Melenhorst-de Jong G, Mestrom M, Rademaker CM, Reijnvaan AF, Zuideveld KP, Kuks PF, van Vught H, Danhof M. Pharmacokinetics and effects of propofol 6% for short-term sedation in paediatric patients following cardiac surgery. Br J Clin Pharmacol 2002;54(4):415-422.
69. Kogan A, Efrat R, Katz J, Vidne BA. Propofol-ketamine mixture for anesthesia in pediatric patients undergoing cardiac catheterization. J Cardiothorac Vasc Anesth 2003;17(6):691-693.
70. Koriyama H, Duff JP, Guerra GG, Chan AW, Sedation Withdrawal and Analgesia Team. Is propofol a friend or a foe of the pediatric intensivist? Description of propofol use in a PICU. Pediatr Crit Care Med 2014;15(2):e66-e71.
71. Krause TK, Jansen L, Scholz J, Böttcher H, Wappler F, Burmeister MA, am Esch JS. Propofol anesthesia in children does not induce sister chromatid exchanges in lymphocytes. Mutat Res 2003;542(1-2):59-64.
72. Levati A, Colombo N, Arosio EM, Savoia G, Tommasino C, Scialfa G, Boselli L. Propofol anaesthesia in spontaneously breathing paediatric patients during magnetic resonance imaging. Acta Anaesthesiol Scand 1996;40(5):561-565.

73. Maldini B, Miskulin M. Outpatient arthroscopic knee surgery under combined local and intravenous propofol anesthesia in children and adolescents. Pediatr Anaesth 2006;16(11):1125-1132.
74. Martin P, Murthy V, Petros A. Methabolic, biochemical and hemodynamics effects of infusion of propofol for long term sedation of children undergoing intensive care. British J Anesth 1997;79:276-279.
75. Murray DM, Thorne GC, Rigby-Jones AE, Tonucci D, Grimes S, Tooley MA, Wolf AR. Electroencephalograph variables, drug concentrations and sedation scores in children emerging from propofol infusion anaesthesia. Paediatr Anaesth 2004;14(2):143-151.
76. Park HJ, Kim YL, Kim CS, Kim SD, Kim HS. Changes of bispectral index during recovery from general anesthesia with 2% propofol and remifentanil in children. Pediatr Anesth 2007;17(4):353-357.
77. Patel KN, Simon HK, Stockwell CA, Stockwell JA, DeGuzman MA, Roerig PL, Rigby MR. Pedaitric procedural sedation by a dedicated nonanesthesiology pedaitric sedation service using propofol. Pediatr Emerg Care 2009;25(3):133-138.
78. Reed MD, Yamashita TS, Marx CM, Myers CM, Blumer JL. A pharmacokinetically based propofol dosing strategy for sedation of the critically ill, mechanically ventilated pediatric patient. Critl Care Med 1996;24(9):1473-1481.
79. Rigby-Jones AE, Nolan JA, Priston MJ, Wright PM, Sneyd JR, Wolf AR. Pharmacokinetics of propofol infusions in critically ill neonates, infants, and children in an intensive care unit. Anesthesiology 2002;97(6):1393-1400.
80. Sepulveda P, Cortınez LI, Saez C, Penna A, Solari S, Guerra I, Absalom AR. Performance evaluation of paediatric propofol pharmacokinetic models in healthy young children. Br J Anaesth 2011;107(4):593-600.
81. Sheridan RL, Keaney T, Stoddard F, Enfanto R, Kadillack P, Breault L. Short- term propofol infusion as an adjunct to extubation in burned children. J Burn Care Rehabil 2003;24(6):356-360.
82. Srinivasan M, Turmelle M, Depalma LM, Mao J, Carlson DW. Procedural sedation for diagnostic imaging in children by pedaitric hospitalists using propofol: analysis of nature,

frequency, and predictors of adverse events and interventions. J Pediatric 2012;160(5):801-806.

83. Svensson ML, Lindberg L. The use of propofol sedation in a paediatric intensive care unit. Nurs Crit Care 2012;17(4):198-203.
84. Teng SN, Kaufman J, Czaja AS, Friesen RH, da Cruz EM. Propofol as a bridge to extubation for high-risk children with congenital cardiac disease. Cardiol Young 2011;21(1):46-51.
85. Ture H, Mercan A, Koner O, Aykac B, Ture U. The effects of propofol infusion on hepatic and pancreatic function and acid-base status in children undergoing craniotomy and receiving phenytoin. Anesth Analg 2009;109(2):366-371.
86. Welzing L, Vierzig A, Junghaenel S, Eifinger F, Oberthuer A, Trieschmann U, Roth B. Remifentanil and propofol for weaning of mechanically ventilated pediatric intensive care patients. Eur J Pedaitr 2011;170(4):477-481.
87. Williams GD, Jones TK, Hanson KA, Morray JP. The hemodynamic effects of propofol in children with congenital heart disease. Anesth Analg 1999;89(6):1411-1416.
88. Yang W, Gu H, Chen B, Li J, Fan Q, Yuan Y, Wang X. Evaluation of SLOG/TCI-III pediatric system on target control infusion of propofol. J Transl Med 2011;9(187):1-6.
89. Barclay K, Williams A, Major E.. Letter to editor: Propofol infusion in children. BMJ 1992;305(6859):953
90. Cannon M, Glazier S, Bauman L. Metabolic acidosis, rhabdomyolysis, and cardiovascularcollapse after prolonged propofol infusion. J Neurosurg 2001;95:1053-1056.
91. Cray S, Robinson B, Cox P. Lactic acidemia and bradyarrhythmia in a child sedated with propofol. Crit Care Med 1998;26(12):2087-2092.
92. Culp KE, Augoustides JG, Ochroch AE, Milas BL. Clinical management of cardiogenic shock associated with prolonged propofol infusion. Anesth Analg 2004;99(1):221-226.
93. Da-Silva SS, Wong R, Coquillon P, Gavrilita C, Asuncion A. Partial-exchange blood transfusion: an effective method for preventing mortality in a child with propofol infusion syndrome. Pediatrics 2010;125(6):e1493-e1499.

94. Guitton C, Gabillet L, Latour P, Rigal JC, Boutoille D, Al Habash O. Propofol infusion syndrome during refractory status epilepticus in a young adult: successful ECMO resuscitation. Neurocrit Care 2011;15(1):139-145.
95. Hein S, Benken I, Quintel M, Kettler D. Propofol Infusion Syndrome [Das propofol-infusions-syndrom]. Anasth Intensivecmed 2004;45:580-584.
96. Holzki J. Fatal bradyarrhythmia following long term propofol infusion in children [Todliche bradyarrythmien nach propofoldauersedierung im kindesalter]. Journal fur anesthesie und intensivbehandlung 2002;9(1):201-203.
97. Holzki J, Aring C, Gillor A. Death after re-exposure to propofol in a 3-year-old child: a case report. Paediatr Anaesth 2004;14(3):265-270.
98. Jorens PG, Van den Eynden GG. Propofol infusion syndrome with arrhythmia, myocardial fat accumulation and cardiac failure. Am J Cardiol 2009;104(8):1160-1162.
99. Laquay N, Pouard P, Silicani MA, Vaccaroni L, Orliaguet G. Early stages of propofol infusion syndrome in paediatric cardiac surgery: two cases in adolescent girls. Br J Anaesth 2008;101(6):880-881.
100. Mehta N, DeMunter C, Habibi P, Nadel S, Britto J. Short-term propofol infusions in children. Lancet 1999;354(9181):866-867.
101. Mierzewska-Schmidt M, Arlukiewicz-Piwowar M, Rawicz M, Woloszczuk-Gebicka B. Propofol infusion syndrome in a 4 year-old girl. Case report. Anestezjologia Intensywna Terapia 2006;38(3):139-142.
102. Perrier ND, Baerga-Varela Y, Murray MJ. Death related to propofol use in an adult patient. Crit Care Med 2000;28(8):3071-3074
103. Robinson JD, Melman Y, Walsh EP. Cardiac conduction disturbances and ventricular tachycardia after prolonged propofol infusion in an infant. Pacing Clin Electrophysiol 2008;31(8):1070-1073.
104. Rosen DJ, Nicoara A, Koshy N, Wedderburn RV. Too much of a good thing? Tracing the history of the propofol infusion syndrome. J Trauma 2007;63(2):443-447.
105. Sabsovich I, Rehman Z, Yunen J, Coritsidis G. Propofol infusion syndrome: a case of increasing morbidity with traumatic brain injury. Am J Crit Care 2007;16(1):82-85.

106. Shankar V. Severe propofol infusion syndrome in a child treated with continous veno-venous hemofiltration (CVVH). Crit Care Med 2009;37(12 sup):A540.

107. Strickland RA, Murray MJ. Fatal metabolic acidosis in a pediatric patient receiving an infusion of propofol in the intensive care unit: is there a relationship? Crit Care Med 1995;23(2):405-409.

108. Veldhoen ES, Hartman BJ, van Gestel JP. Monitoring biochemical parameters as an early sign of propofol infusion syndrome: false feeling of security. Pediatr Crit Care Med 2009;10(2):e19-e21.

109. Waalkens H, Plotz F, Verkade H, Strengers J, Knoester H, Mandema J. Continuous propofol infusion contraindicated in children [Langdurige sedatie met propofol gecontraindiceerd bij kindern]. Nederlands Tijdschirift Geneeskunde 1997;141(9):440-442.

110. Westhout FD, Muhonen MG, Nwagwu CI. Early propofol infusion syndrome following cerebral angiographic embolization for giant aneurysm repair. Case report. J Neurosurg 2007;106(2 Suppl):139-142.

111. Withington DE, Decell MK, Al Ayed T. A case of propofol toxicity: further evidence fora causal mechanism. Pediatric Anaesth 2004;14(6):505-508.

112. Wolf A, Weir P, Segar P, Stone J, Shield J. Impaired fatty acid oxidation in propofol infusion syndrome. Lancet 2001;357(9256):606-607.

113. Zou H, Zhou Q, Xia W, Wu L, Cheng B.. Cardio protective effects of propofol and midazolam in children with congenital heart disease undergoing open heart surgery. Chinese Medical J 2007;87(33):2309-2312.

114. Anand K, Ramsay M, Crippin J. Hepatocellular injury following the administration of propofol. Anesthesiology 2001;95(6):1523-1524.

115. Bagshaw O. TIVA with propofol and remifentanil. Anaesthesia 1999;54(5):501-5

116. Bendiksen A, Larsen M.. Convulsions, ataxia and hallucinations following propofol. Acta Anarsthesiol Scand 1998;42(6):739-741.

117. Bustamante S, Appachi E.. Acute pancreatitis after anesthesia with propofol in a child with glycogen storage disease type IA. Paediatr Anaesth 2006;16(6):680-683.

118. Casserly B, O'Mahony E, Timm E, Haqqie S, Eisele G, Urizar R.. Propofol Infusion syndrome: an unusual cause of renal failure. Am J Kidney Dis 2004;44(6):e98-e101.

119. Chevret L, Husson B, Nguefack S, Nehlig A, Bouilleret V. Prolonged refractory status epilepticus with early and persistent restricted hippocampal signal MRI abnormality. J Neurol 2008;255(1):112-116.

120. Hussein G, Olejniczak P, Carey M, Fisch B, Casey G, Garcha T. Propofoi-induced dystonia: a case report and a review of the literature. Clin Res Regul Aff 1999;16(4):175-181.

121. Karaman Y, Goktay A, Agin H, Karaarslan U.. Propofol infusion syndrome or adrenoleukodystrophy? Paediatr Anaesth 2013;23(4):368-370.

122. Kassebaum N, Hairr J, Goldsmith W, Barwise J, Pandharipande P. Diabetes insipidus associated with propofol anesthesia. J Clin Anesth 2008;20(6):466-468.

123. Kill C, Leonhardt A, Wulf H. Lacticacidosis after short-term infusion of propofol for anaesthesia in a child with osteogenesis imperfecta. Paediatr Anaesth 2003;13(9):823-826.

124. Kirkpartrick M, Cole G. Propofol infusion in children: letter to editor. BMJ 1992;305(6863):1223.

125. Koch M, De Backer D, Vincent JL. Lactic acidosis: an early marker of propofol infusion syndrome? Intensive Care Med 2004;30(5):522.

126. Lanigan C, Sury M, Bingham R, Howard R, Mackersie A. Neurological sequelae in children after prolonged propofol infusion. Anaesthesia 1992;47(9):810-811.

127. Meyer P, Langlois C, Soete S, Leydet J, Echenne B, Rivier F, Bonafe A, Roubertie A. Unexpected neurological sequelae following propofol anesthesia in infants: three case reports. Brain Dev 2010;32(10):872-878.

128. Patermann B, Buzello S, Duck M, Paul M, Kampe S. Accidental tenfold overdose of propofol in a 6-month oldi nfant undergoing elective craniosynostosis repair. Anaesthesia 2004;59(9):912-914.

129. van Straaten E, Hendriks JJ, Ramsey G, Vos GD. Rhabdomyolysis and pulmonary hypertension in a child, possibly due to long-term high-dose propofol infusion. Intensive Care Med 1996;22(9):997-1001.

130. Ting TW, Lee JH. Acute pancreatitis after propofol infusion in a teenage patient. Anaesth Intensive Care 2012;40(3):561-562.

131. van Meurs M, van der Starre PJA. Green Urine [Groene Urine]. Nederlands Tijdschrift voor Anesthesiologie 2000;13(3):87-89.

132. Bradford Hill, A. The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine 1965;58:295-300.

133. Fudickar A, Bein B. Propofol infusion syndrome: update of clinical manifestations and pathophysiology. Minerva Anesthesiol 2009;75:339-344.

134. Papaioannou V, Dragoumanis C, Theodorou V, Pneumatikos I.. The propofol infusion 'syndrome' in intensive care unit:from pathophysiology to prophylaxis and treatment.. Acta Anesth Belg 2008;59:79-86.

135. Wong J. Propofol Infusion Syndrome. Am J Therapeut 2010;17:487-491.

136. Clarke A, Deeks JJ, Shakir SA. An assessment of the publicly disseminated evidence of safety used in decisions to withdraw medicinal products from the UK and US markets. Drug Saf 2006;29(2):175-181.

137. Festa M, Bowra J, Schell D. Use of propofol infusion in Australian and New Zealand paediatric Intensive care units. Anaesth Intensive Care 2002;30(6):786-793.

138. Playfor SD, Venkatesh K. Current patterns of propofol use in PICU in the United Kingdom and North America. Paediatr Anaesth 2004;14(6):501-504.

139. Kruessell MA, Udink ten Cate FE, Kraus AJ, Roth B, Trieschmann U. Use of propofol in pediatric intensive care units: a national survey in Germany. Pediatr Crit Care Med 2012;13(3):e150-e154.

140. Ahlen K, Buckley CJ, Goodale DB, Pulsford AH. The 'propofol infusion syndrome': the facts, their interpretation and implications for patient care. Europ J Anesthesiol 2006;23:990-998.

141. Howick J, Glasziou P, Aronson JK. The evolution of evidence hierarchies: what can Bradford Hill's ‘guidelines for causation’ contribute? J R Soc Med. May 1, 2009; 102(5): 186–194.

Table 7. Characteristics of Included RCTs

Abbasi, 2013

Methods	RCT
Participants	Inclusion criteria: Surgical hypospadiasis repair; exclusion criteria: airway abnormalities and known adverse reactions to the study drugs. Total 117 participants; propofol group 58 participants; comparison 59 participants; age average 37 months.
Interventions	Propofol 100 mcg/kg/min, duration 174.15 minutes; Comparison- Halothane (concentration not provided)
Outcomes	Main outcome measured: unwanted intraoperative penile erection
Notes	Also reported investigations of other AEs: hematoma wound infection (zero identified). No report of measurement of serious AEs.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Low (randomized using a sealed envelope)
Allocation concealment (selection bias)	Low risk	Used a sealed envelope

Blinding of participants and personnel (performance bias)	Low risk	Investigator blinded.
Blinding of outcome assessment (detection bias)	Low risk	Investigator blinded.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported if any patients lost follow-up.
Selective reporting (reporting bias)	High risk	No report of serious AEs (if measured, if identified or not).
Other bias	Unclear risk	not stated

Blumer 2002

Methods	RCT
Participants	326 randomised to receive propofol 1% (109 participants); propofol 2% (112 participants) and standard sedative agents (SSA) (105 participants). .
Interventions	The median duration of sedation was 81, 74 and 65 hours, respectively. The dose of propofol used was 4-6mg/kg/h
Outcomes	This study identified a higher number of deaths in children receiving propofol: 9 (8.2%) deaths in propofol 1%; 12 (10.7%) deaths in propofol 2% and 4 (3.8%) deaths in the SSA group
Notes	Pharmaceuticl sponsored trial, stopped early, never

published in a peer review journal.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Patients were randomized, not clearly how.
Allocation concealment (selection bias)	Unclear risk	Not reported in the abstract
Blinding of participants and personnel (performance bias)	Unclear risk	Not reported in the abstract
Blinding of outcome assessment (detection bias)	Unclear risk	Not reported in the abstract
Incomplete outcome data (attrition bias)	High risk	Data never fully reported.
Selective reporting (reporting bias)	High risk	Study never published.
Other bias	High risk	Industry sponsored reporting in higher number of deaths in patients using propofol, trial stopped early and reasons not provided.

Budic, 2011

Methods	RCT

Participants	45 patients submitted to orthopedic surgery needing bloodless limb operations.
Interventions	15 patients in each group (propofol, causal block or sevoflurane). Propofol average duration was 65 minutes.
Outcomes	Effect of different anesthesia techniques on oxidative stress and endothelial dysfunction connected with ischemia-reperfusion.
Notes	

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	Unclear risk	Not reported.
Blinding of participants and personnel (performance bias)	Unclear risk	Not reported.
Blinding of outcome assessment (detection bias)	Unclear risk	Not reported.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.

Selective reporting (reporting bias)	Unclear risk	Not reported.
Other bias	Unclear risk	Not reported.

Cohen, 2003

Methods	RCT
Participants	Inclusion criteria: ASA I and II undergoing elective ambulatory surgery. Exclusion criteria: 1) less than 40 weeks postconceptional age; 2) demonstrated signs of developmental delay; 3) surgery expected to last less than 30 min. Total: 53 participants, age average 15-16 months. Propofol group 27 participants, control 26 participants.
Interventions	Propofol 12 mg/kg/hr, duration 84 minutes; Control-Sevoflurane 1.5 to 2.5%, duration 76.5 minutes.
Outcomes	Primary outcomes: degree of agitation, time to emergence, time to recovery, time to discharge.
Notes	AEs: Severe agitation- Propofol 1 / Control 6; Pain- Propofol 2 / Control 6; Vomiting- Propofol 3 / Control 2

Risk of bias table

Bias	**Authors'**	**Support for judgement**

	judgement	
Random sequence generation (selection bias)	Unclear risk	Not reported. Stated 'randomized', not how it was done.
Allocation concealment (selection bias)	High risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	High risk	Not reported. Not likely done.
Blinding of outcome assessment (detection bias)	Low risk	Blinded observer.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Unclear risk	Some adverse events reported.
Other bias	Unclear risk	Not stated.

Cohen, 2004

Methods	RCT
Participants	Inlcusion criteria: 1) younger than 3 years of age; 2) ASA physical status I or II; 3) scheduled for ambulatory surgical procedures; 4) parental consent. Exclusion criteria: None Specified. Total 56 participants. 28 in propofol group; Average age 13-14 months.
Interventions	Propofol 225 mcg/kg/min for 61.8 minutes; Comparison

	- Sevoflurane 1.5 to 2.0%
Outcomes	1. Haemodynamic assessments 2. Ionized calcium and magnesium concentrations 3. Recovery assessments 4. Side-effects / adverse events
Notes	4 patients in propofol group experienced AEs & 3 patients in the comparison group experiences AEs. Total adverse events: Postoperative Vomiting - Propofol 3 / Comparison 2; Upper airway obstruction and / or congestion - Propofol 2 / Comparison 1; Hypotension, bradycardia, cough, breath holding - Propofol 0 / Comparison 1; Cough- Propofol 1 / Comparison 1

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	High risk	Open label.
Blinding of participants and personnel (performance bias)	High risk	Open label.

Blinding of outcome assessment (detection bias)	Low risk	Assessment was performed by investigator blinded to the anaesthetic technique.
Incomplete outcome data (attrition bias)	Unclear risk	Not clearly reported.
Selective reporting (reporting bias)	High risk	Number of adverse events and reports do not match.
Other bias	Unclear risk	Funded by Astra-Zeneca

Deng, 2009

Methods	RCT
Participants	40 pediatric patients submitted to cleft palate repair, 20 per group.
Interventions	Propofol and remifentanil for an average of 102.33 minutes; Comparison - Sevoflurane and remifentanil for an average of 101.21 minutes.
Outcomes	Clinical effect and safety of these two drugs. Bradycardia - Propofol 2 / Comparison 0 Agitation - Propofol 2 / Comparison 8
Notes	Written in Chinese.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random numbers table.
Allocation concealment (selection bias)	Unclear risk	Not reported.
Blinding of participants and personnel (performance bias)	Low risk	Patients were blinded.
Blinding of outcome assessment (detection bias)	High risk	Anesthesiologist was not blinded.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Unclear risk	Not reported.
Other bias	Unclear risk	Not stated.

Deramount, 1990

Methods	RCT
Participants	Inlcusion criteria: ASA I or II submitted to strabism repair; Total 20 participants (10 in each group), mean age of 97 months (propofol) and 74 months (comparison).
Interventions	Propofol 10-15 mg/kg/hr for an average of 75 minutes; Comparison - Halothane 1.5% and Thiopental (induction

	10 mg/kg) for an average of 83 minutes.
Outcomes	Side effects of anaesthesia, hemodynamic changes.
Notes	Written in French. Bradycardia responding to atropine - Propofol 2 / Comparison 2; Apnea - Propofol 7 / Comparison 6; Cough- Propofol 1 / Comparison 1; Flush- Propofol 2 / Comparison 0; Myoclonus - Propofol 1 / Comparison 0; Injection pain - Propofol 5 / Comparison 0; Nausea/vomit - Propofol 1 / Comparison 5; Agitation- Propofol 1 / Comparison 0.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	High risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	High risk	Not reported. Not likely done.

Blinding of outcome assessment (detection bias)	High risk	Not reported. Not likely done.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Low risk	Reported AEs.
Other bias	Unclear risk	Not stated.

Erb, 2002

Methods	RCT
Participants	Patients scheduled for a radiofrequency catheter ablation (RFCA) for supraventricular tachycardia (SVT); between 4 - 8 years old. Exclusion criteria: Participant with contraindications to the use of propofol or isoflurane-based anaesthetic. Total of 56 patients, not specified how many in each group.
Interventions	Propofol 106 mcg/kg/min for 319 minutes; Comparison - Isoflurane (no dose / concentration provided).
Outcomes	1) Incidence of postoperative nausea and retching or vomiting 2) Postprocedural vital signs 3) Postprocedural assessment of sedation scores 4) Costs
Notes	Given in percentage only:

Nausea - Propofol 7-21% / Comparison 41-63%

Vomiting - Propofol 0-6% / Comparison 26-55%

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Blocked randomization was generated by using computer random numbers.
Allocation concealment (selection bias)	High risk	Not reported.
Blinding of participants and personnel (performance bias)	Low risk	Not reported. Not likely done.
Blinding of outcome assessment (detection bias)	Low risk	Nurse or research assistant trained in the assessment of PONV symptoms and who were unaware of the patient's treatment group.
Incomplete outcome data (attrition bias)	High risk	Not defined how many patients were enrolled in each group, so can't measure incomplete outcome data.
Selective reporting (reporting bias)	High risk	Not reported number of patients per group or number of events.
Other bias	Unclear risk	Not stated.

Fudickar, 2013

Methods	RCT
Participants	Children with congenital heart disease scheduled for elective pediatric heart catheterization. Exclusion criteria: contraindications to propofol, total 42 participants (22 in propofol group, 20 in control group).
Interventions	Propofol 4000 mcg/kg/min for 75 minutes; Comparison - Sevoflurane 1.5%
Outcomes	Serum base excess, pH, and lactate.
Notes	Extremely high dose of propofol used, reported in the manuscript as 4 mg/kg/min - converted here to be comparable to other studies. Author contacted and clarified use of 4mg/kg/hr.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	Not reported. Not likely done.
Allocation concealment (selection bias)	High risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	High risk	Not reported.Not likely done.
Blinding of outcome assessment	High risk	Not reported, likely low as it was a

(detection bias)		blood gas.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	High risk	No other clinical relevant outcomes (adverse events) reported.
Other bias	Unclear risk	Not stated.

Guard, 1998

Methods	RCT
Participants	Children age 2-8 years, scheduled for elective urological surgery as outpatient. Total of 50 participants, 25 in each group.
Interventions	Propofol infusion 5-10 mg/kg/hr for 88.6 minutes; Comparison - Sevoflurane (2 to 8%) for 83.6 minutes.
Outcomes	Heart rate; systolic blood pressure; emergence from anaesthesia; recovery from anaesthesia.
Notes	Propofol was stated to have significantly ($p< 0.01$) more episodes of bradycardia (reduction of 20% from baseline). Maintenance: Signs of light of anaesthesia - Propofol 4 / Control 0 Emergence: Breath holding - Propofol 2 / Control 3

Coughing - Propofol 12 / Control 11

Laryngospasm - Propofol 0 / Control 2

Bronchospasm - Propofol 0 / Control 0

Secretions - Propofol 9 / Control 8

Excitation - Propofol 1 / Control 2

Rigidity / shivering - Propofol 1 / Control 1

SpO2 < 90% - Propofol 0 / Control 0

Vomiting - Propofol 0 / Control 2

Recovery:

Pain requiring morphine - Propofol 3 / Control 7

Vomiting before discharge - Propofol 1 / Control 3

Vomiting in first 24 hours - Propofol 6 / Control 7

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomization was achieved using random number tables prepared in advance of commencing the study.
Allocation concealment (selection bias)	Unclear risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	High risk	Not reported. Not likely done.
Blinding of outcome assessment (detection bias)	Unclear risk	The recovery assessment was done by a blind nurse, but the OR personal

		was not blind.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Low risk	Outcomes were well reported and measured.
Other bias	Unclear risk	not stated

Gurkan, 1999

Methods	RCT
Participants	Children 3-15 years, ASA physical status I admitted for strabismus surgery. Exclusion criteria: allergy to any of the drugs to be used or pt who had a symptomatic medical illness. Total of 40 participants.
Interventions	Propofol 132-214 mcg/kg/min for 71 minutes; Comparison - Incremental sevoflurane (1.5% - 2% increment every 3-4 breaths)
Outcomes	Incidence of vomiting / nausea and oculocardiac reflex.
Notes	Nausea and Vomiting - Propofol 7 / Control 14 Vomiting alone - Propofol 6 / Control 20 Episodes of oculocardiac reflex requiring atropine - Propofol 9 / Control 4

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Patients were randomly assigned using a table of random numbers.
Allocation concealment (selection bias)	Unclear risk	Not reported.Not likely done.
Blinding of participants and personnel (performance bias)	High risk	No OR personal was blinded.
Blinding of outcome assessment (detection bias)	Low risk	Blind nurse did the follow-up call.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Low risk	Overall outcomes were well reported on tables.
Other bias	Unclear risk	Not stated.

Hasani, 2009

Methods	RCT
Participants	83 children submitted to surgical interventions, such as adeno-tonsillectomy, hernia repair, orchiopexy, circumcision. Exclusion criteria: known allergies to the drugs used, any evidence of post operatory pain or

	ASA>II.
Interventions	Propofol (n=41) 6-12 mg/kg/hr for 68 min; Comparison - Halothane (n=42) 1.5-2% for an average of 79 minutes.
Outcomes	emergence agitation; parental satisfaction; operation times
Notes	Hypotension - Propofol 4 / Comparison 3 (p=0.04); Bradycardia - Propofol 5 / Comparsion 5; Cough - Propofol 2 / Comparsion 2; Laryngospasm - Propofol 2 / Comparsion 1; Hypersalivation - Propofol 3 / Comparsion 4; Nausea / Vomit - Propofol 2 / Comparsion 4.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated random number table.
Allocation concealment (selection bias)	Unclear risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	Unclear risk	Not reported. Not likely done.
Blinding of outcome assessment (detection bias)	Low risk	Two separate anesthesiologists were involved in each case. The first blinded anesthesiologist collected the

		following data: age, weight, premedication, type of anesthesia, and duration of surgery and anesthesia. In the postanesthesia care unit, the second blinded anesthesiologist observed and collected the following data: recovery time, pain, and adverse events.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Low risk	Overall outcomes were well reported on tables.
Other bias	Unclear risk	Not stated.

Hasani, 2013

Methods	RCT
Participants	Children admitted for Inguinal hernia repair. Exclusion criteria: Known allergy to any of the drugs involved in the study, showed any evidence of perioperative anxiety or postoperative agitation, or ASA>II. Randomized 92 patients; Analyzed 88 patients.
Interventions	Propofol infusion 9-15 mg/kg/hr (mean of 9.1 mg/kg/hr - obtained from author's correspondence) for an average of 68 minutes; Comparison - Sevoflurane (induction 4-

	6%) followed by maintenance of 1.5%-2% for an average duration of 59 minutes.
Outcomes	Pain score 2 hours post-operative; recovery time; AE
Notes	Hypotension - Propofol 6 / Comparison 5 (p=0.9) Bradycardia - Propofol 7 / Comparison 5 (p=0.9) Cough - Propofol 6 / Comparison 5 (p=0.9) Laryngospasm - Propofol 0 / Comparison 0 Hypersalivation - Propofol 3 / Comparison 4 (p=0.9) Postoperative nausea - Propofol 2 / Comparison 9 (p<0.01) Postoperative vomiting - Propofol 2 / Comparsion13 (p<0.01)

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	The children were randomly assigned using a computer-generated random number table.
Allocation concealment (selection bias)	Unclear risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	Unclear risk	Not reported. Not likely done.
Blinding of outcome assessment (detection bias)	Low risk	Two separate anesthesiologists were involved in each case. The first

		blinded anesthesiologist collected the following data: age, weight, premedication, type of anesthesia, and duration of surgery and anesthesia. In the postanesthesia care unit, the second blinded anesthesiologist observed and collected the following data: recovery time, pain, and adverse events.
Incomplete outcome data (attrition bias)	Low risk	Reported 4 patients excluded from sevoflurane group due emergence agitation.
Selective reporting (reporting bias)	Unclear risk	Not clear why the 4 patients were excluded due emergence agitation.
Other bias	Unclear risk	Not stated.

Kaddu, 2002

Methods	RCT
Participants	50 children between 2 and 21 months admitted for elective GI endoscopic procedure. Exclusion: History of adverse reaction to propofol or a known anatomic airway abnormality.

Interventions	Propofol (n= 25) 300-500 mcg/kg/min for an average of 107 minutes. Comparison - Isoflurane (n=25) 1.5% - 2% for an average of 139 minutes.
Outcomes	Wake-up time, total anaesthesia time and recovery time.
Notes	Patients in the propofol group were breathing spontaneously and in the isoflurane group were intubated.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	The patients were randomly assigned by a computer generated sequence.
Allocation concealment (selection bias)	High risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	High risk	Not reported. Not likely done.
Blinding of outcome assessment (detection bias)	Unclear risk	Not reported. Not likely done.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Low risk	Adverse events reported.

Other bias	Unclear risk	Not stated.

Kobacas, 2002

Methods	RCT
Participants	40 ASA I or II children admitted for elective surgery. Exclusion criteria: cardiac, renal, hepatic, metabolic or muscular disorders.
Interventions	Propofol 6-12 mg/kg/hr for 66 minutes; Comparison - Nitrous oxide 67%
Outcomes	Intubation score.
Notes	Written in Turkish. No adverse events reported.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	High risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	High risk	Not reported. Not likely done.

Blinding of outcome assessment (detection bias)	High risk	Not reported. Not likely done.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	High risk	No adverse events reported, not sure if they were measured.
Other bias	Unclear risk	Not stated.

Larson, 1992

Methods	RCT
Participants	44 children aged 1.5 - 14 years, ASA I & II and admitted for strabismus surgery.
Interventions	Propofol 5-12 mg/kg/hr for an average of 82 minutes; Comparison - Thiopental for induction (4-6 mg/kg) and Halothane for maintainence (0.5%-1.5%) for an average of 96 minutes.
Outcomes	Vomiting Degree of sedation Apprehension Postoperative pain
Notes	Bradycardia requiring intervention (intervention not defined): Propofol 4 / Comparison 1

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	High risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	High risk	Not reported. Not likely done.
Blinding of outcome assessment (detection bias)	High risk	Not reported. Not likely done.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	High risk	Intervention required for treatment of bradycardia was not reported.
Other bias	Unclear risk	Not stated.

LeBlanc, 2000

Methods	RCT
Participants	24 paediatric patients (< 16 yoa) admitted for cardiac bypass surgery for congenital heart defect. Excluded: History of seizures and down syndrom chilren.

Interventions	Propofol (n=11) for 50-350 mcg/kg/min, duration average 78 minutes prior to cooling to eliminate electrical activity versus control (n=13) without any additional intervention.
Outcomes	Near infrared spectroscopy data and changes in cytochrome aa3.
Notes	Minimal clinical data presented.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	Unclear risk	Not reported.
Blinding of participants and personnel (performance bias)	Low risk	Single blinded. Anesthesiologist not blinded.
Blinding of outcome assessment (detection bias)	Unclear risk	Not reported.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Unclear risk	Not reported

Other bias	Unclear risk Not stated

Lebovic, 1992

Methods	RCT
Participants	20 children aged 7 mo - 6 yrs, ASA II & III, scheduled for elective cardiac catheterization for evaluation of congenital heart disease. Exclusion criteria: Patients requiring mechanical ventilation or intravenous inotropic support.
Interventions	Propofol (n=10) 10 mg/kg/hr for 108 minutes; Comparison - Ketamine (n=10) 2 mg/kg/hr.
Outcomes	Hemodynamic effects Recovery times Arterial blood gases
Notes	Arterial desaturation at induction: Propofol 4/0 control HR decrease >20% compared to baseline: Propofol 4 / Comparison 0 MAP decrease >20% compared w/ baseline: Propofol 7 / Comparison 1 SPO2 decrease > 5 compared w/ baseline: Propofol 4 / Comparison 0

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**

Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	High risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	High risk	The investigators were not blinded during the procedure.
Blinding of outcome assessment (detection bias)	High risk	Independent blinded observer determined postanaesthesia recovery scores.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Unclear risk	Overall AEs were reported.
Other bias	Unclear risk	Not stated.

Martin, 1993

Methods	RCT
Participants	156 children admitted for short surgical procedures (strabismus surgery, tonsillectomy, herniorrhaphy, orchiopexy) age 1 - 7. Exclusion criteria: Known allergy to study drugs.
Interventions	Propofol (n=75) 50-400 mcg/kg/min for an average of 67 minutes; Comparison NO 70%, halothane 0.5% and

	isoflurane (n=68) for an average of 67 minutes.
Outcomes	Vomiting and airway obstruction Time intervals throughout surgical procedures Vital signs
Notes	Both the systolic and diastolic components of the blood pressure were lower in the inhalation group than in the propofol group (p<0.01). Airway obstruction with jaw trust or CPAP needed: Propofol 5 / Comparison 14; Airway obstruction with oropharyngeal airway needed: Propofol 2 / Comparison 7; Laryngospasm: Propofol 0 / Comparison 2; Vomiting episodes: Propofol 23 / Comparison 45

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random numbers table was used.
Allocation concealment (selection bias)	Unclear risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	Unclear risk	Not reported. Not likely done.

Blinding of outcome assessment (detection bias)	Low risk	A trained research assistant blinded to the anesthetic used counted these events in the PACU and in the Day Surgery Unit (DSU).
Incomplete outcome data (attrition bias)	High risk	Thirteen patients were excluded from analysis because the protocol was not followed. Data of 10 patients were not collected by the research assistant due to the late hour of the completion of the operative procedures.
Selective reporting (reporting bias)	Low risk	Serious AEs were assessed.
Other bias	Unclear risk	Not stated.

Mironov, 2012

Methods	RCT
Participants	72 children undergoing cochlear implantation
Interventions	Propofol 4.9 mg/kg/hr over 239 minutes; Comparison-Sevoflurane
Outcomes	Hemodynamic changes
Notes	Written in Russian. HR was lower in propofol group ($p<0.05$); Emergence agiattion (treated with midazolan): Propofol

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Used sealed envelopes.
Allocation concealment (selection bias)	Low risk	Used sealed envelopes.
Blinding of participants and personnel (performance bias)	High risk	Not reported. Not likely done.
Blinding of outcome assessment (detection bias)	High risk	Not reported. Not likely done.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Unclear risk	Some AEs reported.
Other bias	Unclear risk	Not stated.

Pedersen, 2013

Methods	RCT
Participants	120 children aged 1-10 yoa, scheduled for an MRI scan with an ASA I or II. Exclusion criteria: allergy to any of the

	study drugs, if any procedure other than MRI were performed, unexplained fever, potential for airway problems.
Interventions	Propofol (n=60) 56 mcg/kg/min with nasal cannulas for an average of 60 minutes; Comparison- Sevoflurane (n=60) and laryngeal mask airway
Outcomes	Emergence agitation, parental satisfaction, and operation times
Notes	Respiratory rate < 10 beats/min: Propofol 5 / Comparison 0; Desaturation to 84%: Propofol 1 / Comparison 0; Higher agitation score in comparision group (p=0.026).

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	The randomisation assignments were generated by a computer.
Allocation concealment (selection bias)	Low risk	The specially marked opaque envelopes were closed and consecutively numbered by staff not involved in the study.
Blinding of participants and personnel (performance bias)	Low risk	The envelopes were not opened until the anaesthesia was induced.

Blinding of outcome assessment (detection bias)	Low risk	The envelopes were not opened until the anaesthesia was induced. After the anaesthesia, the anaesthetic record was placed in the envelope, which again was sealed. Thus, the staff in the recovery room and in the children's ward, as well as the parents were blinded to the choice of anaesthesia.
Incomplete outcome data (attrition bias)	Low risk	'No children were excluded.' Flow chart documents no lost to F/U.
Selective reporting (reporting bias)	Low risk	Serious adverse events were reported.
Other bias	Unclear risk	Not stated.

Peeters, 2006

Methods	RCT
Participants	44 pediatric patients admitted for cranial facial surgery.
Interventions	Propofol sedation 2-4 mg/kg/hr or midazolam titrated to a COMFORT score >17
Outcomes	Developement of a pharmacokinetic and pharmacodynamic model.
Notes	Minimal clinical information provided as it is a pharmacodynamic study.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	Unclear risk	Not reported.
Blinding of participants and personnel (performance bias)	Unclear risk	Not reported.
Blinding of outcome assessment (detection bias)	Unclear risk	Not reported.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	Unclear risk	Not reported.
Other bias	Unclear risk	Not stated.

Pershad, 2007

Methods	RCT
Participants	60 children with indication for a brain MRI. Exclusion criterion: ASA>3, trauma or contra-indication to study drugs.

Interventions	Propofol 7.6 mg/kg/hr for 68 min; Comparison - Pentobarbital and Midazolam + Fentanyl.
Outcomes	Sedation efficacy, recovery time, adverse events, and patient satisfaction.
Notes	Reduced BP not required intervention: Propofol 4 / Comparison 1; Respiratory depression: Propofol 4 (1 pt need bag/mask and 1 suctioning and repositioning) / Comparison 0; Emergence agitation: Propofol 0 / Comparison 2

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Permuted blocks of size 10.
Allocation concealment (selection bias)	Low risk	Sealed numbered envelopes.
Blinding of participants and personnel (performance bias)	Low risk	The sequence was concealed from the sedating physician, until intervention was assigned.
Blinding of outcome assessment (detection bias)	Low risk	A blinded registered nurse evaluated discharge readiness.
Incomplete outcome data (attrition	Low risk	8 patients were enrolled and refused

bias)		participation prior to randomization.
Selective reporting (reporting bias)	Low risk	Measured adverse events.
Other bias	Unclear risk	Not stated.

Salvitano, 1997

Methods	RCT
Participants	60 children admitted for an elective surgery (not specified what kind of surgery); Exclusion criterion: patients with cardiac, respiratory, hepatic or renal problems; patients that used CNS depressors drugs; patients with allergy to the study drugs; patient whom received alternative anesthesic regimen; patients with bleeding superior to 30% of the total circulant volume; or patient with prior anesthestic complications.
Interventions	Propofol 90 mcg/kg/min for 67 min; Comparison- Midazolam infusion 408 mcg/kg/min.
Outcomes	Compare the wake-up time between midazolam and propofol and compare hemodynamic AEs of the study drugs.
Notes	Manuscript written in Spanish. Not clear how many patients in each group.

No statistically difference between in HR or BP;

Propofol group had: apnea for >20 sec- 44.8% (numbers of patients not provided);

1 patient had rash;

6.6% of patients had facial itchiness.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Not described as randomized. Not likely done.
Allocation concealment (selection bias)	High risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	High risk	Not reported. Not likely done.
Blinding of outcome assessment (detection bias)	High risk	Not reported. Not likely done.
Incomplete outcome data (attrition bias)	High risk	Number of participants per group, not reported.
Selective reporting (reporting bias)	High risk	Not reported the number of participants per group.
Other bias	Unclear risk	Not stated.

Schimidt, 2001

Methods	RCT
Participants	120 children submitted to minor lower abdominal surgery; Exclusion criteria: heart disease, allergies to study drugs, preterm delivery, utilizing opioids.
Interventions	Propofol 7.5 mg/kg/hr for an average of 91 minutes; Comparison- Sevoflurane (2.4%-3.3%) and Remifentanil 0.5 mcg/kg/min.
Outcomes	Peri-op hemodynamics, recovery and discharge times, and PONV.
Notes	Manuscript written in German. Described as number of events not patients: Hypotension (change of 20% in BP form baseline, treated with fluids and reduction of propofol infusion): Propofol 48 / Comparison 106; Bradycardia (<100 bpm if <4 yo; <80 bpm if 4-8 yo; <60 bpm if > 8 yo, treated with atropine): Propofol 110 / Comparison 32; Statiscaly significant more hypotension with sevoflurane during surgery and prior to extubation.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	Unclear risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	Unclear risk	Not reported.Not likely done.
Blinding of outcome assessment (detection bias)	Unclear risk	Not reported. Not likely done.
Incomplete outcome data (attrition bias)	Low risk	All participants completed the study.
Selective reporting (reporting bias)	Unclear risk	Reported number of events, but not number of patients. Good definition of adverse event.
Other bias	Unclear risk	Not stated.

Steinmetz, 2007

Methods	RCT
Participants	39 children aged 1-5 years old, undergoing cleft lip/palate surgery; Exclusion criteria: predisposition for malignant hypertermia, cerebral disease, previous

	anesthesia and any syndrome.
Interventions	Propofol (n=17) 7 mg/kg/hr for 184 min (also received remifentanil); Comparison (n=22)- sevoflurane and Fentanyl.
Outcomes	hemodynamic changes
Notes	Groups were not equal in size- reason were provided.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	The randomization was computer-generated (program developed by Eindhoven University of Technology).
Allocation concealment (selection bias)	Low risk	The allocation sequence was not known in advance
Blinding of participants and personnel (performance bias)	High risk	No report of anesthesia personal as blinded. Not likely done.
Blinding of outcome assessment (detection bias)	Low risk	Postoperative caretakers were blinded to the type of anesthesia and the anesthesia record was kept in a closed envelope, deposited in the patient's record and this remained closed until the infant was discharged from hospital.

Incomplete outcome data (attrition bias)	Low risk	It states no patients were lost to followup and only one did not receive the IV intervention due to lack of IV access.
Selective reporting (reporting bias)	Unclear risk	AEs were reported.
Other bias	Unclear risk	Not stated.

Uezono, 2000

Methods	RCT - crossover study
Participants	18 children seen for eye examination due to retinoblastoma. Exclusion criterion: neurological disorder and radiation therapy requiring sedation.
Interventions	Propofol 100-400 mcg/kg/min for over 60 minutes; Comparison: Sevoflurane 2-4%
Outcomes	Emergence agitation; parental satisfaction; operation times
Notes	2 drop outs had second examination longer than 6 months after the first. Emergence agitation: Propofol 0 / Comparison 6; Vomit: Propofol 0 / Comparison 2.

Risk of bias table

Bias	**Authors'**	**Support for judgement**

	judgement	
Random sequence generation (selection bias)	Unclear risk	Not reported
Allocation concealment (selection bias)	Unclear risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	High risk	Anesthesiologist giving intervention was not blinded.
Blinding of outcome assessment (detection bias)	Low risk	Blinded, independent anesthesiologist evaluated the recovery from anesthetics.
Incomplete outcome data (attrition bias)	Unclear risk	2 drop outs, reasons provided. Not analysed.
Selective reporting (reporting bias)	High risk	No serious events reported (as present or not).
Other bias	Unclear risk	Not stated.

Weir, 1993

Methods	RCT
Participants	78 children admitted for strabismus surgery.
Interventions	Propofol 150-200 mcg/kg/min for an average of 90 minutes; Comparison- Halothane 0.5%-1%.

Outcomes	post operatory emesis
Notes	Vomiting: Propofol 16 / Comparisonl 25; Nausea + vomiting: Propofol 17 / Comparison 27; Opioid use: Propofol 18 / Comparison 12

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	Unclear risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	Unclear risk	Not reported.Not likely done.
Blinding of outcome assessment (detection bias)	Low risk	Incidence of nausea, retching, or vomiting were recorded by a blinded observer.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	High risk	Minimal report of adverse events, only limited to nausea / vomiting.
Other bias	Unclear risk	Not stated.

Woodward, 1997

Methods	RCT.
Participants	30 children aged 4 to 14 years, admitted for correction of prominent ears. Exclusion criteria: seizure disorder, allergy to propofol or eggs, the need for inhalation induction.
Interventions	Propofol 6-10 mg/kg/hr for an average of 73 minutes. Comparison: Thiopentone induction at 5-7 mg/kg and Isoflurane maitenance at 1%-2%.
Outcomes	PONV and post op pain
Notes	Funded by Astra-Zeneca. Nausea: Propofol 1 / Comparison 8; Vomit: Propofol 3 / Comparison 10; High pain score: Propofol 2 / Comparison 3.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	Unclear risk	Not reported.

Blinding of participants and personnel (performance bias)	Low risk	Anesthesiologist was not blinded. Patients were blinded.
Blinding of outcome assessment (detection bias)	Low risk	Observers were blinded to the anaesthetic technique.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	High risk	No other AEs were measured / reported other than PONV and pain.
Other bias	Unclear risk	Not stated.

Xia, 2011

Methods	RCT
Participants	20 children, aged 1 to 9 years old with congenital heart defects admitted for bypass surgery. Exclusion criteria: liver, kidney dysfunction, inflammatory or hemostatic disease, use of anti-inflammatory/anti-oxidant drugs.
Interventions	Propofol 15-20 mg/kg/hr for an average of 71 minutes; Comparison - Enflurane 0.25%-2%
Outcomes	IL6, lactate, CK and myocardial immuno-hystoligical changes.
Notes	

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	Unclear risk	Not reported. Likely not done.
Blinding of participants and personnel (performance bias)	Unclear risk	Not reported.Likely not done.
Blinding of outcome assessment (detection bias)	Unclear risk	Not reported. Likely not done.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported. Likely not done.
Selective reporting (reporting bias)	High risk	No serious or other adverse events reported.
Other bias	Unclear risk	Not stated.

Xia, 2011b

Methods	RCT
Participants	32 children h witcongenital heart defects admitted for bypass surgery for TET, TGA,SV,PS, TA, TAPVD, PA; Exclusion criteria: liver, kidney dysfunction, hemostatic disorders inflammatory disease, use of anti-inflammatory-oxidant drugs.

Interventions	Propofol 150 mch/kg/min for an average of 64 minutes; Comparison- Midrazolam 1 mcg/kg/min and fentanyl 0.05 mcg/kg/min.
Outcomes	IL6, IL 8,SOD, myocardial immuno-hystological changes.
Notes	No other AEs reported.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported.
Allocation concealment (selection bias)	Unclear risk	Not reported. Not likely done.
Blinding of participants and personnel (performance bias)	Unclear risk	Not reported. Not likely done.
Blinding of outcome assessment (detection bias)	Unclear risk	Not reported. Not likely done.
Incomplete outcome data (attrition bias)	Unclear risk	Not reported.
Selective reporting (reporting bias)	High risk	No other AEs reported.
Other bias	Unclear risk	Not stated.

Yang, 2011

Methods	Case series
Participants	30 patients admitted for elective surgery.
Interventions	Propofol with target blood level of 1-4 mcg/ml, dose not provided; duration was an average of 197 minutes.
Outcomes	Not stated if AEs were seen.
Notes	

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	case series
Allocation concealment (selection bias)	Unclear risk	case series
Blinding of participants and personnel (performance bias)	Unclear risk	case series
Blinding of outcome assessment (detection bias)	Unclear risk	case series
Incomplete outcome data (attrition bias)	Unclear risk	case series
Selective reporting (reporting bias)	Unclear risk	case series

Other bias	Unclear risk case series

Zou, 2007

Methods	RCT
Participants	32 patients admitted for congenital heart defect repair, 3-14 years old.
Interventions	Propofol 150 mcg/kg/min and comparison propofol 1mcg/kg/min
Outcomes	to measure the cardio protective effects of both drugs.
Notes	Written in Chinese.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	Not reported. Likely not done.
Allocation concealment (selection bias)	High risk	Not reported. Likely not done.
Blinding of participants and personnel (performance bias)	High risk	Not reported.Likely not done.

Table 8. Excluded Studies

Anand 2001	
Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.
Anghelescu 2012	
Reason for exclusion	Use of propofol infusion in palliative cancer patients, excluded due all patients experiencing an anticipated fatal outcome.
Bagshaw 1999	
Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.
Bendiksen 1998	
Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS
Bustamante 2006	
Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.
Casserly 2004	
Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible

	reporting early signs of PRIS.
Chevret 2008	
Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.
Hooke 2007	
Reason for exclusion	Use of propofol infusion in palliative cancer patients, excluded due all patients experiencing an anticipated fatal outcome.
Hussein 1999	
Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.
Karaman 2012	
Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.
Kassebaum 2008	
Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.
Kill 2003	

Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.

KIRKPATRICK 1992

Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.

Koch 2004

Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.

LANIGAN 1992

Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.

Meyer 2010

Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.

Patermann 2004

Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.

Ting 2012

Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.

Van Meurs 2000

Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.

Van Straaten

Reason for exclusion	It does not fulfil PRIS criteria used in this review. Possible reporting early signs of PRIS.

Table 9.Newcastle-Ottawa

ROB Newcastle-Ottawa:	**Selection 1**	**Selection 2**	**Selection 3**	**Selection 4**	**Comparability**	**Outcome 1**	**Outcome 2**	**Outcome 3**
Agudelo 2011	b - *	a - *	a - *	a - *	same population, just different anesthesia- *	a*	a - *	a -*
Amundsen,	b - *	b	a - *	a - *	matched by	b - *	a - *	b - *

2005					age - *			
Cray, 2001	a - *	a - *	a - *	a - *	matched for age and cardiac lesion - **	b - *	a - *	a -*
Van Gestel 2005	b - *	a - *	a - *	a - *	same population, just different anesthesia- *	a*	a - *	a -*
Kiringoda, 2010	a - *	d	c	b	None	b - *	a - *	a -*
Pepperman , 1997	a - *	a - *	a - *	a - *	attempted to match on age, sex, and weight - **	b - *	a - *	a -*
Prins, 2005	b - *	a - *	a - *	a - *	same population, just different anesthesia- *	b - *	a - *	a -*

Table 10. Characteristics of included Case Series

Study ID	# patients	Duration of propofol (min)	propofol dose (mg/kg/h)	AE reported (please report the event)

Absalom, 2003	32	60 to 360 min during procedure, in PICU patients received 2-18 hours of propofol infusion- no exact value	used target control infusion of 5mcg/ml,	7 children had cardiac catheterization: 5 children with pulmonary hypertension had LMA inserted and two required intubation, all patients undergoing cardiac surgery were intubated. No AEs reported.
Agzamkhodzhaev, 2002	40	118	15	No serious AE. Reported decreased in BP but responsive to reduction to propofol infusion.
Arosio, 1993	76	67	8.2	1 patient had hypoventilation, no interventions reported.
Ayala-Sandoval, 1991	120	67	3.7- 7	No serious events reported.
Blusse Van, 2008	20	389	not provided (continous infusion at discretion of anesthesiologist)	Not stated.
Chimdambarn, 2013	20	135	11.5	Vital signs varied 20% from baseline, patients had increased sonolence and delayed

				emergence.
Cornfield, 2002	142	median 16.5 hours, longest 19 days and 11.5 hours.	3	10 had evidence of metabolic acidosis; only 3 had acidosis that persisted longer than 1 hour. The acidosis resolved without discontinuation of propofol or other pharmacological intervention.
Cox, 2011	105 infusions (13 used propofol) with 98 patients	median 439 (185-710)	not provided	AEs were noted, but not what anesthetic technique was used for the patient with the AE.
Dipestraten, 2012	20	60	15-21	Pharmacokinetic study, Minimal clincal information. No AEs reported.
Englehart, 2007	18	70	18	Pharmacokinetic study. Reported as no complications during the study period.
Englehart, 2008	17	70	18	Pharmacokinetic study, Minimal clinical information. No adverse events reported. It reports blood pressure and heart rate but no AEs.

Ivani, 1991	16	172	9	None.
Jelaskov, 2007	56	65	6 to 10	None reported. Measured heart rate and blood pressure, but no statscaly changed and no report of any treatment needed.
Jelaskov, 2008	49	80	6 to 10	
Klockars, 2012	60	108 (17-288) in children3 to 6 yearsold; 74 (24-142) in 7 to 11 years old; 72 (21-198) in 12 to 16 years old.	15.9 (aged 3-6 yearld old group and for the 6 to 11 years old group) and 15.1 in the older group	Reports use of: (1)phenylephrine: 1 patient (3 times); 1 patient (twice) and 2 patients once; Atropine: 1 patient twice; vasoactive drugs were used more 'commonly' during burst supressions.
Knibbe, 2002	6	360	3	Not stated.
Kogan, 2003	45	85.5 (38-161)	4	4 patients had Heart rate and 5 patients had mean arterial pressure change>20% from baseline in 4 and transient reduction in oxygen saturations in 3 patients responsive to oxygen and manual assisted ventilation.
Kogan, 2003	45	85.5	4	1- Heartrate decrease >20% 3- mean arterial pressure decrease >20%

				3- Oxygen saturation in arterial blood decrease >5%
Koriyama, 2014	223 infusions (w/ 210 patients)	618	2.7	4- Lipemic plasma 4- Mild metabolic acidosis 12- low blood pressure 4- Lactic acidosis
Krause, 2003	40	70	12	No serious AEs reported.
Levati, 1996	82	68	8.3	8- Hypotension (none greater than 20% from baseline) 3- EtCOs increased (2- hypoventilation; 1- mild airway obstruction)
Maldine, 2006	147	98	10	No respiratory or cardiovascular complications reported.
Martin, 1997	9	2118	2.1	None.
Murray, 2004	10	120	1-4	Not stated.
Park, 2007	30	70	Target control infusion 3-5 mcg/ml	Not stated.
Patel, 2009	210	75	8.5(1.5-23.3)	No serious events reported.
Reed, 1996	30	240	2.5	1 patient had an acute decrease in blood pressure with propofol and was removed out of the study and two other Two

				patients experienced an acute decrease in blood pressure which resolved without treatment. 1 patient undergoing a Fontan operation developed hypotension and metabolic acidosis after 5 hours of propofol use at 4 mg/kg/hr; the metabolic acidosis was present prior to propofol initiation and it was considered due poor cardiac output. Triglycerdies were normal and the hypotension and acidosis improved with fluids and inotropic support; 2 patients required reduction in propofol dose due to low blood pressure; Urea, electrolytes, liver function and triglycerides were unchanged from baseline in all patients.
Rigby-Jones, 2002	21	720 (270-1680)	4	
Sepulveda, 2011	41	99	8	Not stated.

Sheridan, 2003	11	480	3.6 (0.4-8.1)	No serious AEs reported.
Srinivasan, 2011	1649	65	12	49 airway obstructions; 26 desaturations (<90%); 5 apneas, 2 aspirations, 1 intubation; 4 hypotension (change more than 30mmHg) needing fluid.
Svensson, 2012	955	780	2.9	1 death (patient had propofol infusion for 10 hrs, death occured 14 hours after propofol was stopped and due to MOF). 2 patients with pulmonary hypertension with desaturations needing intervention
Teng, 2011	11	732 (180-2160)	0.4-5.6	1 patient required albumin (not clear if it was due to hypotension). No evidence of PRIS.
Ture, 2009	30	294 (280-950)	7.5 (7.5 to 15)	Statiscally significant increase in liver and pancreatic enzymes. Not associated with metabolic acidosis and clinical changes.
Welzing, 2010	23	420 (4-12 hours)	2.8 (2-4)	No serious AEs reported.
Williams, 1999	31	60 (55-65)	6 (3-12)	Apnea after propofol use, needed intubation due to

				worsening on pulmoanry hypertension.
Yang, 2011	30	197	not provided. Target blood level of 1-4 mcg/ml.	Not stated.

Table 11. Characteristics of included Case Reports of PRIS

Study ID	Age (years)	AE reported (please report the event)	Outcome
Barclay, 1992	1.7	Patient was sedated w/ chloral hydrate (400 mg 4 hourly) and propofol (5-10 mg/kg/hr), along with ampicillin and chloramphenicol. On day 3, patient's serum was lipaemic as she had developed metabolic acidosis, associated with bradycardic, hypotensive, and oliguric. The chloral hydrate was stopped on day 3 and the propofol infusion was stopped after 56 hours (4980 mg). She was given adrenaline and dobutamine; however, the bradycardia persisted. Veno-venous haemodiafiltration was started which reduced her temperature to normal within 2 hours, but the bradycardia persisted. Adrenaline infusion was stopped on day 5 and she continued to improve through day 6. Haemodiafiltration was stopped, but it resulted in a fever, acidosis, and oliguria, which caused the haemodiafiltration to be restarted. Her condition steadily improved from day 17. On day 28, propofol was used as an induction agent for	survived

		the insertion of a tracheostomy tube, which was uneventful. She went home on day 54, with a six month evaluation revealing no neurological deficit.	
Bray, 1995	9	Patient was first sedated with midazolam infusion plus morphine, which was replaced with propofol and fentanyl infusions. These were stopped on the 4th day. That evening he was tachypneic, which was treated with 2 boluses of fentanyl (total 250 mcg) and artificial ventilation. The medical providers then noticed T-wave inversion, followed by ventricular complexes widened, 50 mg of propofol was given and the infusion restarted. The patient heart rate rapidly decreased from 100 to 55 beat/min with complete heart block. No change occurred in his ventricular rate with atropine, but isoprenaline increased it for a time. He continued deteriorating, until death. The post-mortem report stated mild micro vesicular steatosis in the liver and some lymphocyte infiltration of the heart. Propofol average dose was 4.5 mg/kg/hr used for 72 hours.	death
Cannon, 2001	13	Patient was found at the bottom of a hill unconscious with gastric content around her face with a palpable pulse and spontaneous breathing. Admitted to ICU for mechanical ventilation, mannitol therapy, and sedation with propofol 100 mcg/kg/min. Propofol infusion was used for 4 days. The patient failed to awaken after propofol infusion was discontinued. A transfer to a PICU occurred and profound non lactate metabolic acidosis was found, along with a right frontal and temporal contusion, a small right subdural hematoma, and a right-to-left midline shift of 3mm. Patient	death

		was treated with escalating dosage of dopamine and epinephrine, intravascular volume support, and broad-spectrum antibiotic agents. Despite aggressive efforts metabolic acidosis continued, along with hypotension, leading to the patient's death.	
Cray,1998	0.8	Infant with upper respiratory obstruction secondary to an esophageal foreign body who required tracheal intubation and mechanical ventilation. The patient received a propofol infusion at a mean rate of 10 mg/kg/hr for 50.5 hrs. He developed lipemia and green urine and subsequently, a progressive severe lactic acidemia and brady arrhythmias unresponsive to conventional treatment. These abnormalities resolved with CVVH. He was encephalopathic and developed liver and muscle necrosis histologically compatible with a toxic insult. Examination of homogenized muscle tissue demonstrated a reduction in cytochrome C oxidase activity. There was no evidence of systemic infection or underlying metabolic disease. He eventually recovered completely.	survived
Culp, 2004	13	Patient had an acute intra ventricular hemorrhage, which was managed with a ventriculo peritoneal shunt with a full neurologic recovery, 6 weeks prior to patient presentation. During surgery, considerable bleeding and significant brain edema ensued, thus nitrous oxide was discontinued, the head of the bed was elevated, PCO2 of 30 mmHg maintained, and 50 mcg/kg/min propofol was infused. Due to severe bleeding the resection was abandoned with a total operative time of 17 hours. Patient was admitted to	survived

		neurosurgical ICU with elevation of the head of the patient's bed; sedation with a propofol infusion titrated between 100-190 mcg/kg/min; mechanical ventilation; maintenance of a normal mean arterial blood pressure of 55-65 mm Hg. Hemodynamically unstable polymorphic ventricular tachycardia develop 74 hours after surgery, which caused the stop of the propofol infusion. Despite cardioversion, lidocaine and amiodarone, the patient continued to deteriorate and cardiogenic shock ensued. Infusions of epinephrine and norepinephrine were started. Blood chemistry results showed severe metabolic acidosis. Mechanical circulatory support was started, despite the increased risk of intracranial bleeding from heparinization. Extracorporeal circulation with membrane oxygenation was instituted with a dialysis filter. Forty-eight hours later there was improvement with native cardiac ejection and normal cardiac rhythm. He was discharged to a rehabilitation facility 28 days post-operation.	
Da-Silva, 2010	4.75	Patient was seen in the emergency department with persistent vomiting and undocumented fever. While in the emergency department, she had an initial tonic-clonic seizure that did not require treatment. Meningoencephalitis was found with CT and lumbar puncture and treatment with cefotaxime, vancomycin, and acyclovir was started with a transfer to the PICU. Because of recurrent seizure activity, phenytoin and levetiracetam was started, but the seizure activity continued and phenobarbital was added. A 0.05 mg/kg/hr midazolam drip	survived

	was started and increased to 0.7 mg/kg/hr, but burst suppression was not achieved. Midazolam was switched to 0.6 mg/kg/h propofol and advanced until burst suppression was achieved. This was achieved and maintained with phenobarbital, levetiracetam, phenytoin, and 15.6 mg/kg/h propofol infusion. After 24 hours, propofol weaning was started. The following day, intermittent episodes of non-sinus bradycardia started. After an echocardiography, which demonstrated idioventricular versus junctional rhythm, increased QTc, and ventricular escape, PRIS to be suspected. Propofol was stopped and 20 ug/kg/min of dopamine and 0.75 mcg/kg/min of milrinone were started. Throughout the next 24 hours, multiple intermittent bradycardia episodes occurred, which responded to calcium gluconate boluses. Burst suppression was maintained after the propofol was stopped. Because of the metabolic acidosis and elevated CK levels, sodium bicarbonate drip was started. A double-lumen dialysis catheter was inserted along with a manual exchange transfusion with 600 mL of packed red blood cells. Rhabdomyolisis persisted and creatine kinase (CK) levels increased. Another round of exchange transfusion performed with 300 mL of packed red blood cells lead the CK levels to increase and more haemodynamically stable over the following 24 hours. He was weaned off milrinone and dopamine and discharged from the hospital after significant improvement in all affected organ systems.	
Guitton, 2010 17	Patient suffered 3 generalized tonic-clonic epileptic seizures	survived

after being febrile and confused, This was treated with 10 mg diazepam and 500 mg phenobarbital. This stopped the seizure activity, but the patient remained unconscious. Upon admission into the emergency room, patient was treated with ceftriaxone because bacterial meningitis was suspected. The patient was intubated with etomidate and suxamethonium, ventilated and admitted to the ICU, where she was sedated w/ 10 mg/hr of midazolam and fentanyl. Propofol, as an anticonvulsant and sedation agent, was added to prevent seizure activity EEG evidence. During the first 4 hours, 4 boluses of 100 mg of propofol was administered, followed by 7.8 mg/kg/hr (400 mg/hr) propofol to suppress burst activity on the EEG. This increase led to arterial hypotension, requiring fluid challenge and an infusion of 0.3 mcg/kg/min of norepinephrine. After 12 hours of continuous propofol administration, paroxysm activity in the left posterior regions were revealed on EEG. This required an increase in propofol dose to 450 mg/hr, along with the start of 200 mg/day carbamazepine. Propofol was reduced progressively as the EEG improved. At hour 58, patient developed severe metabolic acidosis. PRIS was suspected and propofol was stopped immediately, replaced with thiopental and the carbamazepine was replaced with levatiracetam. At hour 78 the left ventricular function was gradually decreasing and it was decided to urgently implement rescue circulatory bypass with arteriovenous extracorporeal membrane oxygenation (ECMO). This was complicated by cardio circulatory arrest with refractory ventricular fibrillation which need external

		cardiac massage for 30 minutes. Once ECMO was started the patient improved immediately and impressively. Epinephrine was infused at a rate of 2.7 mcg/kg/min and stopped 24 hours later. ECMO was maintained for 5 days. Eighteen days after cardiac arrest, the patient awakened and renal dialysis was stopped on day 26. Patient was released from the ICU on day 36 and returned home at 2 months. She returned to normal life with a left foot drop at 3 months. Neurological examination was found to be completely normal 2 years later.	
Hanna, 1998	Case 1: 17 years;	Patient had severe neurological delay and refractory epilepsy. Admitted due to persistent status epilepticus despite use of phenytoin, phenobarbital and benzodiazepines. Patient was intubated with 80 mg of propofol and initiated an infusion of propofol at 10 mg/kg/hr and titrated from 7 to 14 mg/kg/hr to maintain burst suppression. Propofol was weaned off after 18 hours, but seizures re-started and propofol was re-initiated between 9 and 18 mg/kg/hr. After 20 hours, patient developed hypotension, needing neosynephrine, dopamine and colloid. Propofol was discontinued after 44 hours due to worsening hypoxia despite normal chest x-ray and 100% oxygen. Urine became brown and metabolic acidosis, hyperkalaemia and high lactate (19mmol/l). Hypotension and hypoxaemia, refractory to treatment, continued and wide complex bradycardia with runs of asystole resistant to atropine, calcium and bicarbonate lead to a full cardiac arrest not responsive to reanimation. Autopsy revealed	death

		rhabdomyolysis of the diaphragm and all appendicular muscles. Mild acute bronchopneumonia, myoglobin casts in renal tubules.	
Hanna, 1998	Case 2: 7 year old	Patient had previous diagnosis of rolandic epilepsy of childhood, treated with carbamazepine, phenobarbital, hydrocortisone and clonazepam, along with a ketogenic diet. Transferred to PICU for treatment of focal motor epilepsy. He was intubated and initiated on propofol infusion at 2 mg/kg/hr and increased to 11 mg/kg/hr until clinical seizures ceased. The propofol was increased over 48 hours to 27 mg/kg/hr, but seizures continued. His urine became tea-coloured and tachycardia and oliguria were present after 50 hours of propofol use. Hypotension started 62 hours after propofol initiation and propofol was stopped at 63 hours of infusion. Patient became progressively acidotic and hypoxic. Normal chest x-ray. Hypotension progressed needing dobutamine, epinephrine and phenylephrine. Normal cardiac ECHO. Initiated dialysis due to anuria. Patient had temp 39.5 degrees Celsius, lactate 24.9 mmol/l, and creatine kinase (CK) 49,992 U/L. Despite medical therapy he developed a wide complex tachycardia, followed by bradycardia and asystole. Patient died 15 hours after propofol was stopped. Autopsy revealed rhabdomyolysis of limb muscles and myoglobin casts in renal tubules.	death
Hein, 2004	2	Patient was sedated with propofol and alfentanil for 8 hours during surgery. Transferred to PICU post op and was stable haemodynamically and neurologically. Placed on	death

		propofol 2% dose of 16mg/kg/h for 22 hours. Sedation was stopped and patient did not wake-up, had pinpoint pupils and absent reflexes. Eleven hours after sedation was stopped developed SVT, had a dark urine, acidosis and lactate increased to 7.4mg/dl. 46 hours post surgery had a bradycardic arrest and did not respond to reanimation.	
Holzki, 2002	3	While at day care patient became unconscious, required intubation, possible aspiration identified on chest x-ray. Initiated propofol for sedation while in PICU, dose of 21.6 mg/kg/hr. After 11 hours of propofol, patient had brady arrhythmia, hypotension and lactate increased to 5.1 mg/dl. Propofol was stopped for 11 hours and re-started at a dose of 4 mg/kg/hr. In 8 hours bradycardia re-occurred patient initiated on catecholamine, used external pacemaker. Cardiac ECHO showed normal myocardial function. Lipemic plasma, liver enlargement and metabolic acidosis were also present. Cardiac arrest occurred, which lead to patient's death. Not clear when propofol was stopped after re-initiation.	death
Holzki, 2004	3	Patient sedation and analgesia was provided through propofol, midazolam, and fentanyl. An infusion of 20 mg/hr was ordered, an infusion of 20 mg/kg/hr was accidentally given. After 15 hours of sedation, the patient developed a combined respiratory and metabolic acidosis and bronchospasm. This resolved with mechanical ventilation. The accidental propofol infusion dosage was realized to be too high and the infusion was stopped, but the fentanyl and midazolam continued. Thirteen hours after the propofol	death

		was stopped, the child was awaken and fighting the ventilator. The severe aspiration pneumonia required mechanical ventilation, thus 1.25 mg/kg/hr propofol infusion was started for sedation. The average dose was 4.2 mg/kg/hr over the next 8 hours. At the eighth hours, bradycardic dysrhythmia, ventricular ectopic, and an incomplete right bundle-branch block were revealed on ECG, which caused the propofol infusion to be stopped. Catecholamines were administrated. Despite attempts to stabilize, the child was transferred to a paediatric cardiology department, which they suspected propofol intoxication. An external cardiac pacemaker was inserted and the catecholamine infusion was increased. The metabolic acidosis was resistant to therapy and remained unchanged. Fifteen hours after the second propofol infusion was stopped the patient had a cardiac arrest. Resuscitation was carried out for 60 minutes without success.	
Jorens, 2009	12	Patient received a continuous infusion of propofol, and continuous administration of catecholamines was also necessary to obtain acceptable cerebral perfusion pressure. On day 5, a sudden nodal bradyarrhythmia with coved-type ST-segment elevation in the right precordial leads was observed, along with: Rhabdomyolisis (creatine kinase 863,000 IU/L), elevated cardiac troponin levels, lactic acidosis, and lipemic serum with a triglyceride level of 591 ng/ml. The administration of bicarbonate and dialysis did not improved the cardiovascular instability nor the rapidly	death

		evolving ventricular arrhythmias. The patient was receiving continuous heart massage while transferred to the tertiary centre. Despite cardiac extracorporeal support, the patient died in the operating room. The post-mortem report showed an accumulation of fat droplets and acute vacuolar degeneration, along with myocytolysis of skeletal cardiac muscle.	
Laquay, 2008	Case 2: 16 years	Anesthesia was induced and maintained with propofol, alfentanil or sufentanil, and vecuronium and which was successful. Patient was transferred to PICU receiving low doses of milrinone (0.3 mcg/kg/min) and epinephrine (0.2 mcg/kg/min). Propofol infusion was kept at a dose < 3mg/kg/hr (no exact value provided, based on graph it was ~ 2.5 mg/kg/hr for 8 hrs). Patient developed metabolic acidosis (no values provided) and propofol was stopped. There were no signs of myocardial or renal failure. After stopping infusion the metabolic acidosis improved and patient fully recovered.	survived
Laquay, 2008	Case 1: 12 years	Anesthesia was induced and maintained with propofol, alfentanil or sufentanil, and vecuronium and which was successful. Patient was transferred to PICU receiving low doses of milrinone (0.3 mcg/kg/min) and epinephrine (0.2 mcg/kg/min). Propofol infusion was kept at a dose < 3mg/kg/hr (no exact value provided, based on graph it was ~ 2.5 mg/kg/hr for 15 hrs). Patient developed metabolic acidosis (no values provided) and propofol was stopped. There were no signs of myocardial or renal failure. After stopping infusion the metabolic acidosis improved and	survived

		patient fully recovered.	
Metha,1999	1.5	Patient received propofol with an oral airway during surgery for 5 hours, at a rate of 6 mg/kg/hr; after surgery oxygen saturations dropped and emergent intubation and mechanical ventilations needed. Chest x-ray showed signs of aspiration and acute lung injury. Blood gas had metabolic acidosis (no values). BE- 17 mmol/L; lactate- 3.4 mmol/L. Required bicarbonate infusion for 36 hrs. Progressed with bradyarrhythmias, requiring volume and inotropic support. Triglycerides= 3.4 mmol/L. Required dialysis due oliguria. Muscle biopsy done suggestive of possible mitochondria respiratory-chain enzyme deficiency.	Not clear, likely survived.
Mierzewska-Schmidt, 2006	4	Patient required intubation due epiglottitis. Metabolic acidosis was noted after 27 hours of propofol infusion in a dose form 10-12 mg/kg/hr, but it was misdiagnosed as early signs of sepsis. After 48 hrs patients developed tachycardia, hypotension and hepatomegaly. Serum was lipemic. Propofol was stopped but severe acidosis progressed. Liver enzymes were GOT 2694 U/L and GPT 746 U/L. Potassium was 7.15 mmol/L. Patient developed ventricular fibrillation, which was not responsive to defibrillation.	death
Parke, 1992	1.5	Patient was intubated and sedated with midazolam for 2 days, she was transitioned to propofol for extubation in the following day. She developed fever, worsening ventilation and oxygenation the night prior to the planned extubation. She received propofol infusion at 10 mg/kg/hr. On day 2 of propofol infusion, she developed metabolic acidosis with	death

		BE- 9.2. On day 6, she developed bradycardia and progressed to cardiac arrest. She had returned to spontaneous circulation, but required high inotropic support. Her blood was lipemic, but chemistry results were normal. Head CT and cardiac ECHO were unremarkable. She progressed to bradycardic arrest, had pacing wires placed directly on her myocardium, she lost pacemaker capture and died 7 hours later. Postmortem reported stated fatty infiltration of the liver, bilateral pneumonia and heart was unremarkable.	
Parke, 1992	2.75	Patient was intubated and received propofol sedation at 4.8 mg/kg/hr, she improved and was extubated in 48 hours. She had recurrent stridor and had to be re-intubated, propofol was re-started. On day 4, she had nodal bradycardia and right bundle brunch block. On day 5,, she developed a fever and hepatomegaly. Her blood was lipemic and propofol was stopped. Investigation for sepsis were all negative. She required inotropic support, 100% oxygen and had progressive metabolic acidosis (pH- 7.14 and BE- 11.8). On day 6, peritoneal dialysis was initiated to treat hyperkalaemia and acidosis and haemofiltration was started. She died on day 6 due to asystole, which was not responsive to resuscitations.	death
Parke, 1992	1.5	Patient required intubation and mechanical ventilation. Propofol infusion initiated at 7.9 mg/kg/hr. Antibiotics were also started. Extubated at day 3, but re-intubated 30 minutes later due to stridor. On day 5, she became oliguric, bradycardic with a right bundle brunch block and	death

		hepatomegaly. Her blood was lipaemic and propofol was stopped. On day 6, she required dialysis (peritoneal and continuous venous venous) and multiple inotropic support infusions (epinephrine, dobutamine and isoprenaline). She progressed to a bradycardic arrest, which was responsive to resuscitation.	
Parke, 1992	0.1	Patient was intubated, initiated on propofol 7.1-10.7 mg/kg/hr; by the second day she developed a metabolic acidosis (pH- 7.05, BE- 17.3) and her serum had become lipemic. She had a bradycardic event and cardiac ECHO were unremarkable. Her acidosis worsened (pH- 6.94, BE- 22.8); propofol was then stopped. Epinephrine infusion increased heart rate increased from 60 to 90, but she became oliguric needing peritoneal dialysis and her liver was enlarged. Next morning she had fixed and dilated pupils and had life sustained therapy withdrawn.	death
Parke, 1992	6	Patient was intubated due to stridor, however only a 3.5 mm ETT was able to fit in his airway. He was transferred to another hospital, re-intubated with 4.5 mm ETT and initiated on propofol 4- 9.2 mg/kg/hr; parainfluenza virus was isolated. Next day his pH 7.24 and BE- 9.5, on day 3 he was extubated, but had to be re-intubated 30 minutes later due to stridor. The blood specimens were reported lipaemic. He had progressive metabolic acidosis with BR- 16 mmol/L. Antibiotics were changed as he was thought to be septic. Patient progressed to bradycardia and hypotension, which were not responsive to pacing or inotrope. Propofol was stopped and patient died of a bradycardic arrest 3	death

		hours after propofol was stopped on day 4.	
Perrier, 2000	18	Involved in MVC, ejected 50 feet from car. Injuries: severe closed head injury, fracture of C4 vertebra and basal skull fracture, right ocular injury, mandibular fracture, clavicular fracture, bilateral pneumothorax, lung contusion, right tibia fracture, multiple right hand fractures and first degree facial burns. Sedated with propofol 55 mg/hr. Stable until day 3 when decerebrate posturing developed, but with normal intra-cranial pressures. Creatine kinase (CK) was 20.520 U/L due to muscle injuries. Patient has had a fever. Positive endotracheal secretions for gran-negative bacilli was treated. On day 5, the patient developed atrium fibrillation with rapid ventricular escape. In 4 hours rhythm changed to right bundle bunch block with bradycardia. Patient developed progressive metabolic acidosis, serum was lipaemic and potassium was 6.4 Eq/dl. Cardiac EHCO revealed global hypokinesis and no pericardial effusion. Patient became hypotensive and bradycardic, which was not responsive to treatment. Patient developed a PEA arrest and death followed. He received a total of 98 hours of propofol at 530-700vmg/hr.	death
Robinson, 2008	0.8	Previously healthy child, admitted for focal seizures of left arm and leg with eye deviation, received phenobarbital, lorazepam and phenytoin. Electrolytes and CSF fluid were normal. CT head showed changes in right thalamus consistent with prolonged seizures and Chiari I malformation. Propofol used at day 2 to treat seizures - bolus and infusion of 14 mg/kg/hr. Clinical seizures ceased	survived

		with propofol. At day 4, EKG showed slow sinus rhythm, low amplitude p waves and first degree AV block and wide QRS with ST changes and right bundle brunch block. Propofol stopped, but EKG changes remained and seizures resumed. He was transferred to a different hospital. CK=1.794 U/L and HCO3 15 mmol/L, lactate 3.5 mmol/L. After 6 hours in the new hospital, he developed wide complex tachycardia, received lidocaine and adenosine, and synchronized cardioversion. Runs of ventricular tachycardia were noted and lidocaine infusion initiated. Over the next 48 hours, EKG gradually normalized.	
Rosen, 2007	18	Patient was involved in motor vehicle collision. GCS=3 and was intubated at the scene, with a severe traumatic brain injury, pelvic and long-bone fractures and bladder rupture. After the operating room, he was taken to the PICU and placed on propofol and morphine infusions. Propofol was titrated to keep ICP <20 mmHg and CPP >60 mg/dl. Propofol dose was increased to 7.5 mg/kg/hr, maintained for 72 hrs. He then developed oliguria and increased CK to 95,440 U/L from 1552 U/L 3 days prior. He became anuric and had a progressive metabolic acidosis and cardiac arrhythmias. He developed an asystolic cardiac arrest and resuscitation was unsuccessful.	death
Sabsovich, 2007	16	Previously healthy, admitted due closed TBI, used propofol to control ICP. Three days after propofol was started developed metabolic acidosis and rhabdomyolysis, renal failure, hyperkalaemia and wide complex tachycardia. Patient died of refractory arrhythmia 36 hours after first	death

		signs of PRIS. The infusion was between 6.7 and 8.3 mg/kg/hr and it was continued for 35 hours to treat ICP.	
Shankar, 2009	4	Patient was placed on propofol for control of status epilepticus, not responsive to lorazepam, phenytoin, pentobarbital and levetiracetan. Propofol was titrated from 6 to 18 mg/kg/hr until seizure control was achieved. After 44 hours, patient developed right bundle brunch block and ST changes and sinus bradycardia. Also metabolic acidosis, hyperlipidaemia, elevated CK and dark urine. Propofol was stopped and a double volume blood exchange done. Acidosis, heart block and renal failure progressed and he was placed on CRRT. After 24 hours of dialysis his acidosis and bradycardia improved and in 10 days he was extubated.	survived
Strikland, 1995	11	Admitted due to neurological decompensation due to astrocytoma, received propofol sedation in OR dose of 7 mg/kg/hr. Due to increase ICP, which was not responsive to clinical management. She was taken to the OR 6 hours after first procedure and received propofol 8-12 mg/kg/hr for approximately 6 hours during OR, followed for another approximately 8 hours of propofol infusion in PICU at dose 9.9 mg/kg/hr. Patient became anuric, with profound metabolic acidosis and hypotensive. In a few hours, she had progressive changes on EKG: junctional rhythm, elevated T waves and ventricular tachycardia and ventricular fibrillation, which lead to a cardiac arrest. Resuscitation was not successful. At time of death, she was in hospital for 38	death

		hours and received 9.4 mg/kg/hr of propofol.	
Veldhoen, 2009	17	Admitted due to a motor vehicle collision (MVC). Intubated due to low GCS. In PICU, patient received propofol at 8 mg/kg/hr for 14 hrs with intention to stop in the next day, but patient deteriorated with increased ICP and severe brain edema, propofol continued at 5 mg/kg/hr, electrolytes, CK, pH, lactate, triglycerides were monitored. On the fourth day, propofol was reduced to 3 mg/kg/hr due to progressive lactic acidosis (pH 7.27, lactate 5.3 mmo/l) and oliguria and CK 7476 U/L. Propofol was stopped, but progressive haemodynamic instability not responsive to resuscitation lead to death 15 hours after propofol was stopped.	death
Waalkens, 1997	6	Admitted to PICU for mechanical ventilation due to severe subglottic laryngitis, received propofol infusion for 60 hours, dose of 5-10 mg/kg/hr from a different institution. Transferred to PICU, at admission: ABG pH- 7.3; HCO3- 15; triglycerides- 24.7 mmol/l (normal <2,2). Urine myoglobin noted. Two hours after admission, patient developed bradycardia and hypotension. Propofol was stopped. Received bolus of fluids and initiated dopamine, dobutamine, isoproterenol and epinephrine. Rhythm was identified as nodal bradycardia, AV block and ventricular tachycardia. Trans-esophageal pacemaker placed and a calcium infusion started. Patient had a temp. of 41 degrees Celsius and CK=30.000 U/L. The patient then received dantrolene and cooling with normal temp. in 2 hours. Patient clinically improved, but due to serious cerebral	death

		damage had withdrawn any life sustaining therapy.	
Westhout, 2007	3	Patient admitted for an angiographic embolization of a cerebral aneurism and sedated with propofol 12 mg/kg/hr. At 5 hours, she developed metabolic acidosis (pH 7.28 HCO3=18). She was transferred to PICU and developed hypotension needing fluids and inotropic support. Lactate values were 263 U/L (normal range 22–269 U/L), CK 179 U/L (normal range 91–180 U/L). She improved in the next few days and was extubated and discharged home in one week.	survived
Withington, 2004	0.5	Admitted for cleft lip/palate repair, required progressive sedation over 48 hours. Propofol used 11-15 mg/kg/hr, mean infusion 11.7 mg/kg/hr. On the second day, patient developed metabolic acidosis (pH- 7.21 / HCO3- 18), hepatic and renal failure and multiple cardiac arrhythmias (wide complex tachycardia, right bundle brunch block, ventricular tachycardia and SVT). Patient also had sinus bradycardia, which was not responsive to inotropic support or pacing. Propofol was stopped. Patient improved after initiation of charcoal hemoperfusion. Further investigation identified abnormality on acylcarnitine metabolism.	survived
Wolf, 2001	2	Patient was admitted to PICU due to a gun shot in the head, intubated and sedated with propofol for 72 hours at a dose of 5.2 (range 4 to 5.4) mg/kg/hr. On day 4, he became oliguric and bradycardic (HR- 28 beats/min) with no obvious cause. Propofol was stopped and isoprenaline started. Transvenous pacing restored the heart rate, but metabolic	survived

		acidosis progressed, with increased lactate. Hemofiltration started based on a previous case report (Cray, 1998), maximum lactate was 4.9 mmol/L and BD- 10mmol/l; CK-879 U/L. Acidosis cleared and cardiac function normalized with haemofiltration.	

Table 12. Characteristics of excluded case reports. Non PRIS

Study ID	Age (years)	AE reported (please report the event)	Outcome
Anand, 2001	17	Patient had a history of bronchial asthma treated with several medication, resolved history of hepatitis A, family history of multiple drug allergies. Hernia repair anaesthesia was a propofol infused at a mean dose of 150 mcg/kg/min for a total of 682 mg. No anaesthetic complications occurred during surgery and patient had a satisfactory condition and was discharged. After discharge, patient developed severe nausea and vomiting and was readmitted to the hospital. She was diagnosed with acute hepatitis (serum aspartate aminotransferase concentration of 241 U/l, serum alanine aminotransferase concentration of 174 U/l). Intravenous rehydration was	survived

		provided and levels were normal 10 days post-surgery.	
Bagshaw, 1999	4	Patient received Propofol during surgery at 10 mg/kg/h and remifentanil at 1 mcg/kg/min. Patient's heart rate decreased over the first 30 minutes of surgery, reaching 40 beats/min, the propofol infusion were reduced to 6 mg/kg/h ECG revealed patient had developed a nodal bradycardia. Blood pressure was no affected. Three doses of atropine 0.1 mg was administered, until the heart rate increased to 130 beats/min. The heartbeat continued to decrease, but stayed above 100 beats/min for the remaining 45 minutes of surgery. Patient recovered rapidly and without incident.	survived
Bendiksen, 1998	6	Patient was on penicillin prior to surgery, which was induced with alfentanil 0.5 mg, propofol 100 mg. Propofol 200 mg/hr and 50% nitrous oxide was used to maintain anaesthesia with a total of 180 mg propofol. Surgery was uncomplicated. 44 hours after surgery, the patient experienced a grand mal convulsion lasting 6-7 minutes, which responded to 5mg diazepam. She	survived

		remained unconscious with respiratory insufficiency, she was intubated with 0.05mg fentanyl, 100 mg thiopentone and 30 mg suxamethonium. 30 minutes later she awoke and was breathing spontaneously. An infusion of 200-240 mg/hr of propofol (total 1600 mg over 6.5 hr) was started to allow for a CT scan and lumbar puncture, along with 8 mg of midazolam and 0.75mg of alfentanil divided into 4 doses. after the sedation was stopped, but prior to patient awaking the nurse observed a generalized seizure lasting 60 seconds. When patient was fully awaken, she had ataxia for the following 5 days.	
Bustamante, 2006	4	Patient maintenance of surgery was done with propofol 200 mcg/kg/min for 30 min and then decreased to 100 mcg/kg/min for the next 30 minutes. Other medications N2O, dexamethasone, dolasetron, and reversal of neuromuscular blockade with neostigmine and glycopyrrolate. Total anaesthesia lasted 83 minutes. Patient developed nausea, vomiting, intermittent episodes of diffuse abdominal pain and inability to tolerate	survived

		enteral feeding 1 day post operation, in which blood analysis confirmed metabolic acidosis. Treatments include fluid resuscitation. On day 2, patient was transferred to PICU for management of acute pancreatitis. Pain was managed with intermittent doses of morphine. On day 7, patient was discharged after enteral feeding returned to normal.	
Casserly, 2004	17	Patient had intra-abdominal haemorrhage and closed head contusion, managed with intermittent mannitol infusions and loop diuretics. On day 2, the patient underwent a right ventriculostomy with increasing dose of propofol as the sedation agent. On day 4, the propofol dosage was at 118 mcg/kg/min and was maintained for a total of 5 days, as well as phenylephrine and atracurium. On day 8, the propofol infusion was replaced with midazolam because rhabdomyolysis was suspected. Hemodialysis was started, but stopped due to worsening of the patient's intra-cranial pressure. For 3 days, continuous renal replacement therapy was substituted and tolerated. Then haemodialysis was re instituted and	survived

		discontinued 4 weeks later.	
Chevret, 2008	9	Patient had a fever with cephalalgia for 5 days when generalized tonic-clonic seizures started and patient was admitted into the hospital. Patient received intra-venous acyclovir, followed by phenobarbital and phenytoin. Propofol was started on day 2, with carbamazepine added during the first week. Upon waking, seizures were not controlled, so topiramate and levetiracetam was also added. On average immunomodulation therapy with plasmaphereses was performed between day 11 and day 57. Mechanical ventilation lasted for 20-57 days and patient was discharged between day 26 and day 115. Residual severe epilepsy and moderate disability in functional outcome. Dose of propofol and duration not provided.	survived
Hussein, 1999	16	Antiepileptic medications were withheld for EEG monitoring after the electrode implantation. Three spontaneous seizures were noted the following three days post-surgery. Phenytoin therapy was restarted for the following three days. The subdural electrode gird was	survived

		then removed with local anaesthesia supplemented with 1800 mg of propofol boluses and 860 mg of propofol as a continuous drip over 6 hours. In addition, he received 2.5 ug of fentanyl. Over the next 3 hours, the movement and confusion abated.	
Karaman, 2013	5	Patient was induced with 2.5 mg/kg propofol bolus, 0.5 mg/kg atracurium, 10mcg fentanyl. Anesthesia was maintained with 50 mcg/kg/min propofol and 1 mcg/kg/h fentanyl for the 65 minute surgery, which was uncomplicated. Five hours after surgery, the patient experienced a convulsion. Labs revealed hypertriglyceridaemia, metabolic acidosis, hyperpotassaemia, and high levels of creatinine kinase, urea, SGOT, and SGPT. The patient was stable by day 16 and discharged.	survived
Kassebaum, 2008	13	Patient was preoperative medications including bromocriptine and thyroxine. Anesthesia induction was with propofol and vecuronium, followed by infusions of 100-200 mcg/kg/min propofol and 0.05-0.15 mcg/kg/min remifentanil. One hour after induction, urinary output was noted to be about 1,000 mL in spite of	survived

		having only 400 mL of normal saline. The next hour an additional 1,500 mL urinary output was noted. Because of the polyuria a diagnosis of diabetes insipidus. Euvolemia with hypotonic crystalloids was maintained versus other treatment. The urinary output continued to be large, reaching 7,000 mL during the 5th hours; however, no other haemodynamically changes occurred. Propofol was discontinued at this time. Postoperatively, brisk diuresis continued, but maintained normal urine output soon. Patient was discharged at day 2.	
Kill, 2003	7	Surgical anaesthesia was induced with 0.1 mg fentanyl, 120 mg propofol, and 30 mg suxamethonium and maintained with 13.5 mg/kg/h propofol and repeated fentanyl. Anaesthesia duration was 150 minutes. Sixty minutes after the end of propofol infusion, the patient developed respiratory distress with tachypnoea and expiratory wheezing. Aminophylline (150 mg), methylprednisolone (250 mg), and reproterol (30 mcg) was initiated for treatment. Patient was observed for the	survived

		next 24 hours in the PICU, which was uneventful.	
Kirkpatrick, 1992	0.1	Patient was deteriorating upon admission to the hospital with increasing frequent paroxysms of coughing and become increasingly tired. She was sedated with intravenous propofol at a rate of 10 mg/h. After 4 days, she had short convulsions treated with phenobarbitone and her serum was severely lipaemic. No evidence of acidosis or haemodynamic compromise. Propofol was immediately stopped. She continued ventilation for 16 days.	survived
Koch, 2004	5	Admitted to PICU for recovery from endovascular coil embolization of a complex high output arterio-venous malformation of the right middle cerebral artery. Surgery and anaesthetic management were uneventful and complete embolization was achieved. Patient was to remain highly sedated for 24 hour to assure haemodynamic stability and coil thrombosis. Patient remained mechanically ventilated and propofol-based sedation was initiated at 15 mg/kg/hr. To maintain systolic blood pressure within a normal range,	survived

		repeated boluses of labetalol was given. After 6 hours, the patient developed lactic acidosis, serum HCO3 <19mmol/L and pH <7.31. Propofol sedation was tapered down to 6 mg/kg/hr and finally stopped.	
Lanigan, 1992	2	Patient was intubated with halothane, then sedated with morphine and midazolam infusions. Extubation was trialed on day 6, which was unsuccessful due to restlessness and stridor. Propofol was commenced and infusion was increased stepwise to 10.9 mg/kg/h over the next 5 days. Day 12-16 demonstrated fluid overload, hepatomegaly, a rise in liver enzymes, and lipemic blood, which changed sedation to morphine, midazolam, and vecuronium.No report of metabolic acidosis. All sedation was stopped on day 17, with trachea successfully extubated at day 21. From day 17 to day 33, the patient experienced functional blindness and jerky, twitch movements of the limbs, along with the restlessness that was present throughout the illness. On day 33, motor function improved to normal.	survived

Meyer, 2010	case 1- 34 days	Received propofol infusion for 150 min. 29 hours after anaesthesia had partial seizures. Used clonazepam and phosphenytoine. Did not require intubation and had no haemodynamic compromise. Follow-up showed normal physico motor achievements .Microcephaly (percentile 3.4)	survived
Meyer, 2010	case 2- 54 days	Received propofol for 180 min. Had partial secondary generalized clinical seizures. Received diazepam, phenobarbital, phenytoin. No required intubation. Hemodynamic stable. Mental deficiency with hyperkinesia. Microcephaly (percentile <0.1)	survived
Meyer, 2010	case 3- 55 days	Received propofol for 60 min. Had partial secondary generalized clinical seizures. Received diazepam, phenobarbital, phenytoin. No required intubation. Hemodynamic stable. Global developmental delay with hyperkinesia. Microcephaly (percentile <0.1)	survived
Patermann, 2004	0.5	Sedated for craniosynostosis repair, had a HR70-80 not responsive to atropine. Hypotension (SBP 40mmHg) and HR 100. Surgery stopped, received a dose of epinephrine haemodynamics improved	survived

		and dopamine and dobutamine initiated. Patient was receiving 44.7mg/kg/h of propofol. The overdose was recognized and after 4 hours and propofol was stopped. Patient tolerated the procedure well and was discharged home in a few days.	
Straaten, 1995	4	Patient was intubated and sedated with propofol infusion 8.3mg/kg/h for 3 days, progressive increased oxygen demand and high ventilator pressure (PIP 50/PEEP 12cm H20/FiO2 1.0 and no metabolic acidosis. Serum was lipaemic. Propofol infusion stopped. Cardiac ECHO showed signs of pulmonary hypertension. CK increased to 127.000U/L and LFTs were mildly elevated. Muscle biopsy showed 5% muscle fibber necrosis, no signs of mitochondrial or storage disorders. Hyper hydration and dialysis initiated and patient was successfully extubated in 11 days and had full recovery.	survived
Ting, 2012	14	Patient received propofol as it was difficult to sedate with other drugs, received 5.4mg/kg/h over two days. 48 hours after propofol was stopped, she developed pancreatitis(serum amylase	death (possible due baseline disease, not propofol)

		1055 U/L) with cholestasis (Bili 366mcmmol/L (normal 3-24), pancreatitis confirmed by CT abdomen. Pancreatitis improved, but patient progressively deteriorated due her pneumonia and died 12 days later.	
Van Meurs, 2000	14	Propofol used during surgery fro 2 hours, extubated prior to transfer to PICU and had to be re-intubated due increased work of breathing, received propofol 3.6 mg/kg/h for the first hour then 5.4 mg/kg/h for 24 hours. She had fever 38 degrees, edema in her face and urine was green. She had some abnormal movements, possible seizures and propofol was stopped. In Blood gas kept Ph 7.37; HCO3 21, lactate 6.5mmol/L and CPK= 896 U/l. Patient was extubated after 4 days and discharged home.	survived

Table 13.Summary of Findings

propofol for children

Patient or population: children
Settings:
Intervention: propofol

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk **Control**	Corresponding risk **Propofol**				
cardiac arrest (RCT)	**Study population**		Not estimable	326 (1 study)	⊕⊕⊕⊕ **high**[1,2,3,4]	
	4 per 100	**9.5 per 100**				
	Moderate					
cardiac arrest	See comment	See comment	Not estimable	454 (3 studies)	⊕⊕⊕⊖ **moderate**[5,6,7,8]	
PRIS	**Moderate**		Not estimable	33 (27 studies)	⊕⊕⊖⊖ **low**[9,10]	

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio; **HR:** Hazard ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Unpublished RCT, high risk of publication bias.

[2] single study

[3] Statiscally significant finings between propofol 2% and control.

[4] Longe duration of infusion associated with outcome.

[5] Newcastle-Ottawa scale used with low risk of bias.

[6] the 3 cohorts had inconsisten finding and high heterogenity (I2=83%)

[7] Statiscally significant difference in one cohort.

[8] Prolonged propofol infusion seemed to be assocaited with the outcome.

[9] case reports. Only positive findings are reported.

[10] Propofol dose and duration of infusion were strongly reported in these cases.

Table 14. Summary of key outcomes, findings and strength of evidence

Outcome	Study design {Number of studies (n)}	Findings and direction of effect	Strength of Evidence
Cardiac arrest	RCT {1 (326)} Cohorts {3 (454)	The RCT and one cohort had statiscally significant increase in number of death (RCT) and cardiac arrest (cohort) in patients receiving propofol in comparison to non-propofol group. Large magnitude of effect.	Moderate to high.
PRIS	Case reports (27)	Single and multiple case reports (total of 33 cases), with similar clinical findings and important dose-effect gradient.	Low.

Table 15. PRISMA Harms Checklist

Section/topic	#	Checklist item	Harms Extension	Check if done
TITLE				
Title	1	Identify the report as a systematic review, meta-analysis, or both.	*Specifically mention "harms" or other related terms, or the harm of interest in the review.*	X x
ABSTRACT				
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.		X
INTRODUCTION				
Rationale	3	Describe the rationale for the review in the context of what is already known.		X
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).		X
METHODS				
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.		X
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.		X
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.		X
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.		X
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).		X
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.		X
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.		X
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.		X
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).		X
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	*Specify how zero events were handled, if relevant.*	X x
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication		X

		bias, selective reporting within studies).		
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.		X
RESULTS				
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.		X
Study characteristics	18	18a) For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	*Define each harm addressed, how it was ascertained (e.g. patient report, active search), and over what time period.*	X X
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).		X
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.		X
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	*Describe any assessment of possible causality.*	X X
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).		X
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).		X
DISCUSSION				
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).		X
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).		X
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.		X
FUNDING				
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.		X

Figures

Figure 3. Study Flow Diagram

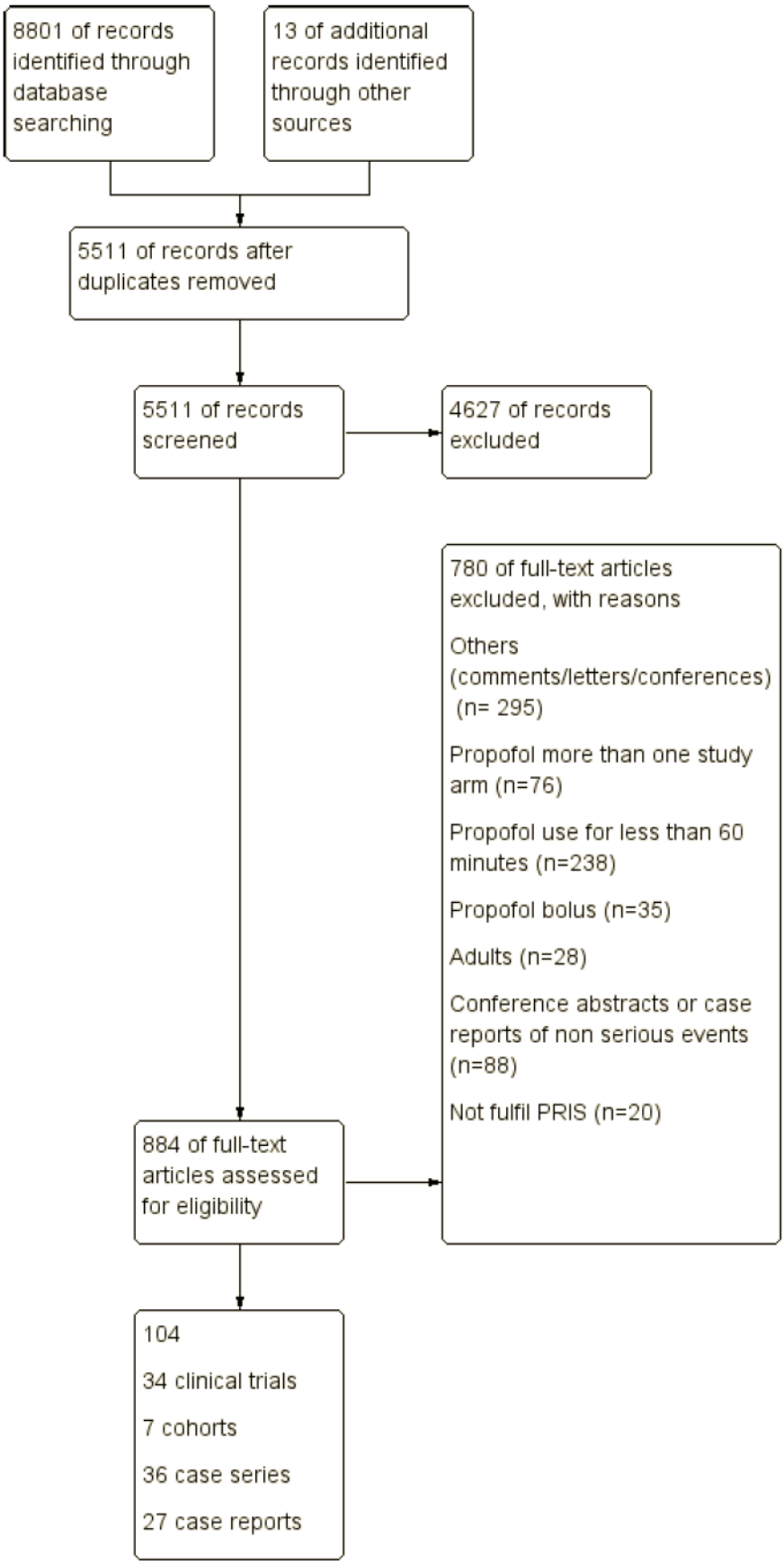

8801 of records identified through database searching
13 of additional records identified through other sources
5511 of records after duplicates removed
5511 of records screened
4627 of records excluded
780 of full-text articles excluded, with reasons
Others (comments/letters/conferences) (n= 295)
Propofol more than one study arm (n=76)
Propofol use for less than 60 minutes (n=238)
Propofol bolus (n=35)
Adults (n=28)
Conference abstracts or case reports of non serious events (n=88)
Not fulfil PRIS (n=20)
884 of full-text articles assessed for eligibility
104
34 clinical trials
7 cohorts
36 case series
27 case reports

Figure 4.Risk of Bias Summary for included RCTs

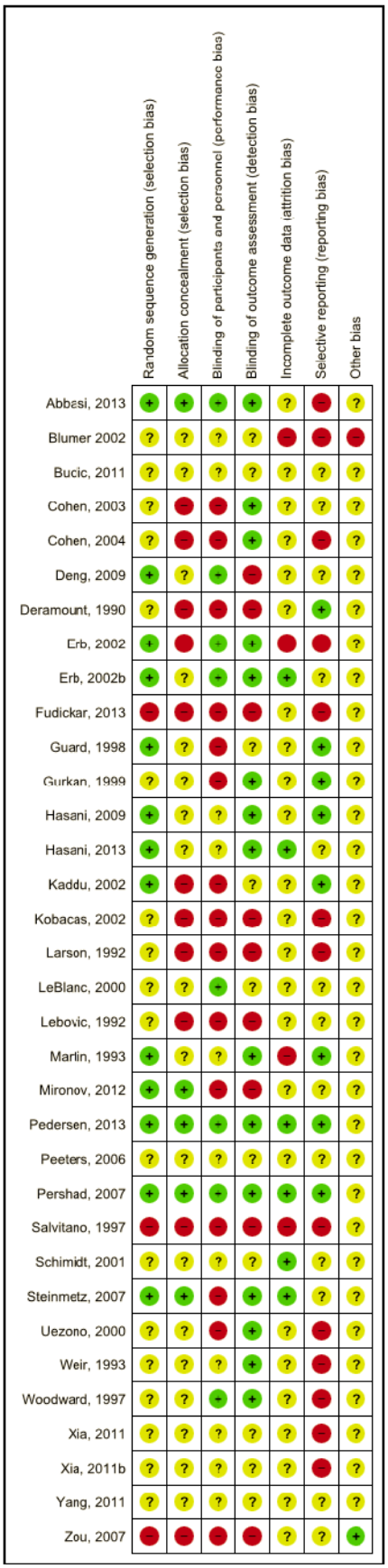
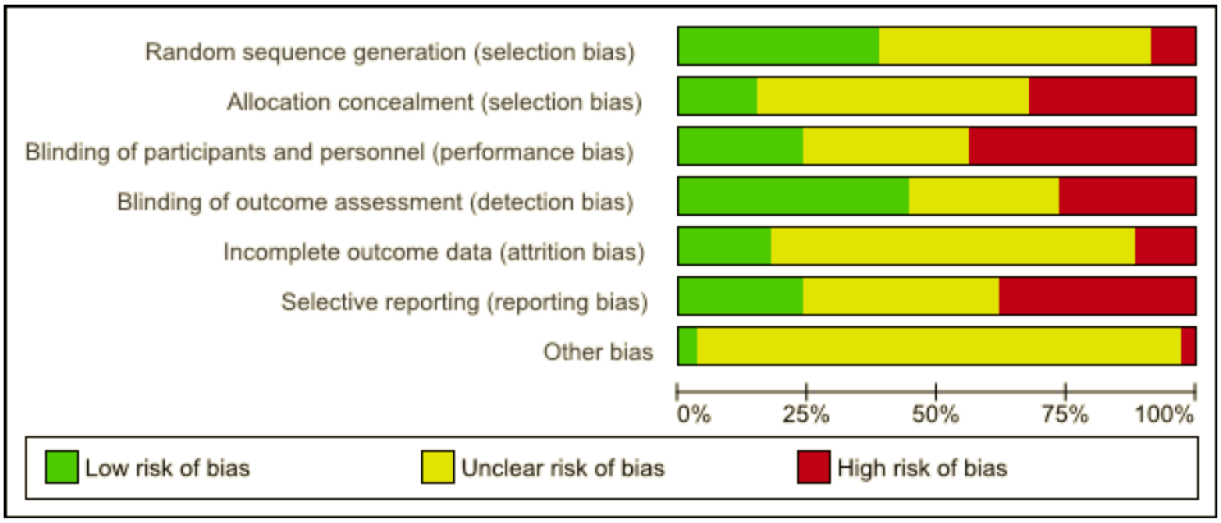

Random sequence generation (selection bias)
Allocation concealment (selection bias)
Blinding of participants and personnel (performance bias)
Blinding of outcome assessment (detection bias)
Incomplete outcome data (attrition bias)
Selective reporting (reporting bias)
Other bias
Abbasi, 2013
Blumer 2002
Bucic, 2011
Cohen, 2003
Cohen, 2004
Deng, 2009
Deramount, 1990
Erb, 2002
Erb, 2002b
Fudickar, 2013
Guard, 1998
Gurkan, 1999
Hasani, 2009
Hasani, 2013
Kaddu, 2002
Kobacas, 2002
Larson, 1992
LeBlanc, 2000
Lebovic, 1992
Marlin, 1993
Mironov, 2012
Pedersen, 2013
Peeters, 2006
Pershad, 2007
Salvitano, 1997
Schimidt, 2001
Steinmetz, 2007
Uezono, 2000
Weir, 1993
Woodward, 1997
Xia, 2011
Xia, 2011b
Yang, 2011
Zou, 2007

Figure 5. Risk of Bias Graph for included RCTs



Appendices

1. MEDLINE search strategy

1. (propofol* or 2,6-diisopropylphenol* or disoprofol or disoprivan or Diprivan or fresofol or recofol or ivofol or aquafol or rapinovet).mp. or exp Propofol/

2. (child* or p?ediatric* or infant* or teen* or youth*).mp. or exp child/ or infant/ or exp Pediatrics/ or exp Minors/ or Adolescent/

3. 1 and 2

2. Ovid EMBASE search strategy

1. (propofol* or 2,6-diisopropylphenol* or disoprofol or disoprivan or diprivan or fresofol or recofol or ivofol or aquafol or rapinovet).mp. or exp PROPOFOL/

2. (child* or p?ediatric* or infant* or teen* or youth*).mp. or exp child/ or infant/ or exp PEDIATRICS/ or exp juvenile/ or adolescent/

3. 1 and 2

3 CENTRAL search strategy

#1 MeSH descriptor Propofol explode all trees

#2 propofol* or diisopropylphenol* or disoprofol or disoprivan or diprivan or fresofol or recofol or ivofol or aquafol or rapinovet

#3 (#1 OR #2)

#4 MeSH descriptor Child explode all trees

#5 MeSH descriptor Infant explode all trees

#6 MeSH descriptor Minors explode all trees

#7 MeSH descriptor Pediatrics explode all trees

#8 MeSH descriptor Adolescent explode all trees

#9 child* or p?ediatric* or infant* or teen* or youth*

#10 (#4 OR #5 OR #6 OR #7 OR #8 OR #9)

#11 (#3 AND #10)

4 Data extraction form

Propofol Infusion for paediatric sedation

CARG 239

Clinical trials

Reviewer Name:

Review Date:

Study ID (last name/year):

Title:

Language:

Country:

Inclusion criteria (must fulfil all brackets)

( ) only paediatric patients (28 days to 19 years old)

( ) used propofol as a continuous infusions (mg/kg/hr or mcg/kg/min) for more than 60 min

( ) propofol used in only one study arm

( ) it is a controlled clinical trial (there is a comparison group)

Definitions:

Cardiac arrest will be defined as significant reduction on cardiac output requiring any of the following: chest compressions; defibrillation, epinephrine boluses or cardiac mechanical support (Extra corporeal life support –ECLS; or ventricular assist device-VAD).

Propofol infusion syndrome (PRIS) will be defined as metabolic acidosis, defined as arterial pH ≤ 7.30 along with a serum bicarbonate ≤ 18 mg/dL; plus the presence of any signs in any the below categories. A patient was deemed to have experienced a particular manifestation category if they experienced any manifestation within the category (adapted from Roberts 2009)

1) rhabdomyolysis (CPK ≥ 10,000 IU/L) or positive serum or urine myoglobin test or positives urinary casts for haemoglobin;

2) hypotension (initiation of a vasopressor agent or increase of ≥20% from baseline vasopressor infusion prior to propofol initiation);

3) hepatic transaminitis (increase in the aspartate aminotransferase and/or alanine aminotransferase ≥ 3 times above baseline);

4) hypertriglyceridaemia (serum triglyceride concentration ≥ 400 mg/dL);

5) hypoxia (partial pressure of arterial oxygen ≤ 60 mmHg);

6) hyperthermia (temperature ≥ 38.3°C);

7) cardiac dysfunction that included asystole, pulseless electrical activity, ventricular fibrillation, sustained ventricular tachycardia of 30 seconds or longer, myocardial failure (ejection fraction ≤40%), or bradycardia (heart rate ≤ 60 bpm)

8) renal failure that included oliguria (urine output ≤ 0.5 mL/kg/hr for ≥ 6 hours), anuria (urine output ≤ 10 mL/hr for ≥ 6 hours), elevation in serum creatinine (increase of ≥ 1 mg/dL from baseline), or hyperkalaemia (serum K+ ≥ 6 mg/dL (excluding other known causes or haemolyzed specimens)

Study participants inclusion criteria:

Study participant exclusion criteria:

1) Patient demographics:

Number of patients enrolled (___ Females)

Age (average):_____ Range:_____ (months/years)

Patient location:

( ) intensive care

( ) diagnostic imaging (MRI/ CT/other:_____)

( ) operating room due to:_____

( ) other:_____

Reason for hospital admission:

Primary outcome being measured:

2) Study drug: PROPOFOL

Number of patients enrolled (___ Females)

Age (average):_____ Range:_____ (months/years)

Are any signs of organ dysfunction prior to propofol infusion? ( ) Yes ( ) No

( ) cardiac : on inotropic support; previous arrhythmia; previous reduced ejection fraction by ECHO;

( ) hepatic: elevated liver enzymes

( ) hypoxia (pao2 <60)

( ) renal: oliguria, anuria or elevation of serum creatinine

2.1) Propofol used as an infusion at average dose of _____ mg/kg/h (or) mcg/kg/min

minimum of ____ and maximum of ____ mg/kg/h (or) mcg/kg/min

2.2) Duration of propofol infusion ___ hours (min:_____/ max:_____)

2.3) Other drugs used in association with propofol (name and dose)

a) prior to propofol infusion:

b) during propofol infusion:

c) after propofol infusion was stopped:

2.4) Any report of PRIS or cardiac arrest (events of interest) in this group? ( ) yes ( ) no

How were these reports collected? ( ) spontaneous report ( ) active surveillance ( ) not clear

2.5) How many events of interest in total?

In how many unique patients?

2.6) Report the AE found (kind of event, methods of monitoring and diagnosing the AE, treatment and outcome):

2.7) Possible confounding factors or biases as per:

Authors:

SR extractor:

2.8) If PRIS or cardiac arrest were identified, what is it relationship with propofol infusion? (time from initiation/ending of infusion to the event)

2.9) How long was the follow-up for?

2.10) Number of patients that lost follow-up and reasons:

2.11) Any other events of interest reported, not fulfilling the PRIS inclusion criteria?

2.12) Were any patients excluded from the adverse effect analysis? Why?

3) Comparison group:

Number of patients enrolled (___ Females)

Age (average):_____ Range:_____ (months/years)

Are any signs of organ dysfunction prior to initiation of sedation infusion? ( ) Yes ( ) No

( ) cardiac : on inotropic support; previous arrhythmia; previous reduced ejection fraction by ECHO;

( ) hepatic: elevated liver enzymes

( ) hypoxia (pao2 <60)

( ) renal: oliguria, anuria or elevation of serum creatinine

3.1) drug (s) dosage

drug (1): ____ dose: ____ duration: ____

drug (2): ____ dose: ____ duration: ____

drug (3): ____ dose: ____ duration: ____

drug (4): ____ dose: ____ duration: ____

3.2) Any serious adverse event (PRIS or cardiac arrest) reported on this group? ( ) yes ( ) no

How were these reports collected? ( ) spontaneous report ( ) active surveillance ( ) not clear

3.4) How many AEs in total?

In how many unique patients .....

3.5) Report the AE found (kind of event, methods of monitoring and diagnosing the AE, treatment and outcome):

3.6) Possible confounding factors or biases as per:

Authors:

SR extractor:

3.7) If PRIS or cardiac arrest was identified, what is it relationship the drugs been used? (time from initiation/ending of drugs been used to the event):

3.8) How long was the follow-up for?

3.9) Number of patients that lost follow up and reasons

3.10) Any other events of interest reported, not fulfilling the PRIS inclusion criteria?

3.11) Were any patients excluded from the adverse effect analysis? Why?

4) Study bias:

4.1) Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence. Low ( ) High ( ) Unclear ( )

4.2) Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment. Low ( ) High ( ) Unclear ( )

4.3) Performance bias due to knowledge of the allocated interventions by participants and personnel during the study. Low ( ) High ( ) Unclear ( )

4.4) Detection bias due to knowledge of the allocated interventions by outcome assessors. Low ( ) High ( ) Unclear ( )

4.5) Attrition bias due to amount, nature or handling of incomplete outcome data. Low ( ) High ( ) Unclear ( )

4.6) Reporting bias due to selective outcome reporting. Low ( ) High ( ) Unclear ( )

4.7) Other biases: Bias due to problems not covered elsewhere in the table. Low ( ) High ( )

Unclear ( ) Why?

Overall risk of bias: Low ( ) High ( ) Unclear ( )

Propofol Infusion for paediatric sedation

CARG 239

Non- randomized studies

Reviewer Name:

Review Date:

Study ID (last name/year):

Title:

Language:

Country:

Kind of study:

( ) case report (number of patients:____)

( ) case-control

( ) cohort study

( ) other:

Propofol:

Patient's characteristics:

Number of patients enrolled (___ Females)

Age (average):_____ Range:_____ (months/years)

Reason for hospital admission:

Patient location:

( ) intensive care

( ) diagnostic imaging (MRI/ CT/other: _____)

( ) operating room due

( ) other:

Are any signs of organ dysfunction prior to propofol infusion? ( ) Yes ( ) No

( ) cardiac : on inotropic support; previous arrhythmia; previous reduced ejection fraction by

ECHO

( ) hepatic: elevated liver enzymes

( ) hypoxia (pao2 <60)

( ) renal: oliguria, anuria or elevation of serum creatinine

Propofol dose and duration of infusion:

Other drugs used (name and dosage):

- Prior to propofol infusion:

- During propofol infusion:

- After propofol infusion was stopped:

Adverse events reported (kind of event, duration, treatment used and outcome):

If PRIS or cardiac arrest was identified, what is it relationship the drugs been used? (time from initiation/ending of drugs been used to the event):

Comparison group (if applicable):

Drug(s) used for comparison (name and dosage):

Number of patients enrolled (___ Females)

Age (average):_____ Range:_____ (months/years)

Reason for hospital admission:

Patient location:

( ) intensive care

( ) diagnostic imaging (MRI/ CT/other: _____)

( ) operating room due

( ) other:

Personal delivering the intervention: Pediatric trained: Yes ( ) No ( ) ( ) Not clear

( ) anaesthesiologist

( ) intensivists

( ) emergency physician

( ) other:

How was the control selected?

Were the patients similar between intervention and control?

Adverse events reported (kind of event, duration, treatment used and outcome):

If PRIS or cardiac arrest was identified, what is it relationship the drugs been used? (time from initiation/ending of drugs been used to the event)

Possible confounding factors:

By the study authors:

By the reviewer:

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE

CASE CONTROL STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and

Exposure categories. A maximum of two stars can be given for Comparability.

Selection

1) Is the case definition adequate?

a) yes, with independent validation

b) yes, e.g. record linkage or based on self reports

c) no description

2) Representativeness of the cases

a) consecutive or obviously representative series of cases

b) potential for selection biases or not stated

3) Selection of Controls

a) community controls

b) hospital controls

c) no description

4) Definition of Controls

a) no history of disease (endpoint)

b) no description of source

Comparability (This criteria could be modified to indicate specific control for a second important factor.)

1) Comparability of cases and controls on the basis of the design or analysis

a) study controls for _____ (Select the most important factor.)

b) study controls for any additional factor

Exposure

1) Ascertainment of exposure

a) secure record (e.g. surgical records)

b) structured interview where blind to case/control status

c) interview not blinded to case/control status

d) written self report or medical record only

e) no description

2) Same method of ascertainment for cases and controls

a) yes

b) no

3) Non-Response rate

a) same rate for both groups

b) non respondents described

c) rate different and no designation

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE

COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and

Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort

a) truly representative of the average _____ (describe) in the community

b) somewhat representative of the average _____ in the community

c) selected group of users e.g. nurses, volunteers

d) no description of the derivation of the cohort

2) Selection of the non exposed cohort

a) drawn from the same community as the exposed cohort

b) drawn from a different source

c) no description of the derivation of the non exposed cohort

3) Ascertainment of exposure

a) secure record (e.g. surgical records)

b) structured interview

c) written self report

d) no description

4) Demonstration that outcome of interest was not present at start of study

a) yes

b) no

Comparability (This criteria could be modified to indicate specific control for a second important factor.)

1) Comparability of cohorts on the basis of the design or analysis

a) study controls for _____ (select the most important factor)

b) study controls for any additional factor

Outcome

1) Assessment of outcome

a) independent blind assessment

b) record linkage

c) self report

d) no description

2) Was follow-up long enough for outcomes to occur

a) yes (select an adequate follow up period for outcome of interest)

b) no

3) Adequacy of follow up of cohorts

a) complete follow up - all subjects accounted for

b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost)

c) follow up rate < ____% (select an adequate %) and no description of those lost

d) no statement

Chapter 6

Conclusion

Conclusion

This doctoral work emerged from the need for a balanced assessment of benefits and harms when making decisions around the use of therapeutic interventions for patient care. Every intervention carries potential for benefit as well as harm, and informed decision-making requires both to be measured and reported with equal clarity.

Improving report in reviews of adverse events and the PRISMA Harms

Systematic reviews are the gold standard for knowledge synthesis; there is broad consensus that systematic reviews should be conducted and reported in a transparent fashion (1). The assessment of harms through a systematic review with meta-analysis is potentially valuable as individual studies are usually underpowered to detect differences in rare outcomes, whereas a meta-analysis of many studies may have more power to investigate the incidence of the rare events associated with therapeutic interventions(2). The difficulties in developing a review of adverse events and the lack of guidance in reporting led to the development of the PRISMA Harms, a reporting guideline to improve harms reporting in systematic reviews.

We followed the recommended steps for a guideline development (3). First, we documented the weaknesses in reporting adverse on reviews exclusively designed to measure these events, as a conservative estimate of the magnitude of the problem (4). The second step in the guideline development was to seek expert opinion in an anonymous fashion through a modified Delphi exercise. We found the respondents very engaged and it was difficult for them to select items that could be considered “not important”. To our surprise, only one proposed item was excluded after two Delphi rounds. We added pharmaco-epidemiologists to the Delphi study, and submitted the potential items for their consideration. After three Delphi rounds, we collected unique comments that could form the basis of future methodological research (e.g. the difference when assessing “risk of bias” in harms vs. when assessing “risk of bias” for studies assessing effectiveness). During the in-person consensus meeting, mandatory items were differentiated from those that are recommended to improve clarity and completeness in harms reporting.

After multiple revisions by the highly experienced PRISMA Harms steering committee, the PRISMA Harms guideline comprised 4 mandatory items as well as 14 "recommendations for harms reporting in reviews".

The ultimate goal of PRISMA Harms, like the PRISMA Statement itself, is to improve the quality of reporting in systematic reviews, so that both benefits and harms are well reported, improving patient safety and promoting knowledge to all decision-makers, whether they are patients, clinicians, policy-makers, or government.

Improving conduct of reviews of adverse events

In the systematic review of propofol infusion in pediatric patients, we documented that zero serious adverse events would have been identified if only published clinical trial data were examined. The use of propofol infusion in pediatric patients could be mistakenly labelled a "safe" intervention. In comparison, by including non-randomized studies and grey literature, a number of patients experiencing fatal events associated with use of propofol infusion were identified, with some studies identifying a statiscally significant difference in the number of serious events. Further, we were able to estimate a dose of propofol and a length of infusion associated with these serious adverse events, providing both clinically and methodologically relevant information.

Methodologically relevant information

The inclusion of non-randomized studies in reviews of adverse events has been endorsed by the Adverse Effects Methods Group at the Cochrane Collaboration (2), but to our knowledge the difference between including only randomized clinical trials vs. including both randomized and nonrandomized designs has not been documented in a Cochrane systematic review. This impact became clear in the example of the propofol infusion in pediatric patients, when a review of only published randomized studies would generated zero data and the review with observational studies and grey literature included more than 30 different publications reporting data on propofol infusion syndrome and cardiac arrest.

Clinically relevant information

Based on the systematic reviews of propofol infusion in pediatric patients including randomized and non-randomized studies, we identified a series of patients with a constellation of symptoms (metabolic acidosis, rhabdomyolisis, liver dysfunction, cardiac arrhythmias). The majority of the patients who developed these symptoms had a fatal outcome. This constellation of signs and symptoms has been previously identified as a "syndrome" associated with propofol infusion and called "Propofol Infusion Syndrome" or "PRIS" and has led to a prohibition, by health regulatory agencies around the world, of the use of propofol infusion for sedation in pediatric critical care. Despite not being part of our review inclusion criteria, we also identified surveys done with physicians working in pediatric intensive care units in Germany, United Kingdom, and United States and some recent case series and cohort studies conducted in Spain and Canada reporting on the continuous use of propofol infusion in pediatric critical care, despite the prohibition issued by health regulatory agencies to be used in this setting.

Every intervention carries benefits and harms. Propofol is known to be a clinically useful drug due the fast effect and very short half-life which allows the patient to be fully awake just a few minutes after the infusion is stopped. We identified that pediatric patients admitted to intensive care units with respiratory or neurologic disorders were most frequently reported to develop PRIS; also a propofol dose of 4mg/kg/h and duration of propofol infusion longer than 12 hours were associated with the majority of the cases of PRIS and cardiac arrest. We could extrapolate these findings and suggest that a propofol use of less than 4mg/kg/h for a length of no more than 12 hours could be used by clinicians with caution, being mindful of a few non-fatal cases of PRIS associated with lower propofol doses and shorter duration. It is also important to note that the totality of evidence for PRIS emerged from case reports, making the lack of a comparison group difficult to draw definitive conclusions and impossible to measure its incidence. There is need for more evidence to better understand PRIS and we encourage further publications of current experience on propofol infusion used in intensive care, if any.

The above information combined in one single document, as a systematic review of rare events, has never been done and it will guide health care professionals in making an informed decision on the use of propofol infusion in pediatric patients.

This doctoral work succeeded in accomplishing its main goals, developed the PRISMA Harms to improve reporting in systematic reviews measuring adverse events and, documented the relevance of including non randomized studies and grey literature when reviewing adverse events, improving methods of conducting systematic review of adverse events.

References

1. Moher D, Liberati A, Tetzlaff J, Altman D. The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PloS Med2009;6:e1000097
2. Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011]. Cochrane Collaboration, 2011. www.cochrane-handbook.org.
3. Moher D, Schulz K, Simera I, Altman D. Guidance for developers of health research reporting guidelines. PloS Med2010;7:e1000217.
4. Zorzela L, Golder S, Liu Y, Pilkington K, Hartling L, et all. Quality of reporting in systematic reviews of adverse events: systematic review. BMJ 2014;348:f7668

Thesis References (Alphabetical order)

1. Abbasi H, Razavi S, Hajiesmaeili M, Behdad S, Ghiamat M, Eghbali A. Unwanted intra-operative penile erection during pediatric hypospadiasis repair comparison of propofol and halothane. Urol J 2013;10(3):988-992.
2. Absalom A, Amutike D, Lal A, White M, Kenny GN. Accuracy of the `Paedfusor' in children undergoing cardiacs urgery or catheterization. Br J Anaesth 2003;91(4):507-513.
3. Agudelo SC, Mencía S, Faro A, Escudero V, Sanavia E, López-Herce J. Continuous propofol perfusion in critically ill children [Propofol en perfusión continua en ninos en estado crítico]. Med Intensiva 2012;36(6):410-415.
4. Agzamkhodzhaev TS, Tokhirov ShM, Kholmukhamedov BM. Changes in hemodynamics with a combination of epidural anesthesia and propofol sedation in children. Anesteziol Reanimatol 2002;1:6-9.
5. Ahlen K, Buckley CJ, Goodale DB, Pulsford AH. The 'propofol infusion syndrome': the facts, their interpretation and implications for patient care. Europ J Anesthesiol 2006;23:990-998.
6. Amundsen LB, Artru AA, Dager SR, Shaw DW, Friedman S, Sparks B, Dawson G. Propofol sedation for longitudinal pediatric neuroimaging research. J Neurosurg Anesthesiol 2005;17(4):180-192.
7. Anand K, Ramsay M, Crippin J. Hepatocellular injury following the administration of propofol. Anesthesiology 2001;95(6):1523-1524.
8. Anello C, O'Neill RT. Does research synthesis have a place in drug regulatory policy? Clin Res RegulAff 1996;13:13–21.
9. Aronson JK, Derry S, Loke YK. Adverse drug reactions: keeping up to date. Fundam Clin Pharmacol 2002;16:49-56.
10. Arosio EM, Gramegna M, Savoia G, Borroni M, Colombo N, Ego E, Levati A. Propofol anesthesia in spontaneous respiration in pediatric patients undergoing nuclear magnetic

ressonance [Anestesia con propofol in respiro spontaneo in pazienti pediatrici sottoposti a risonanza magnetica nucleare.]. Minerva Anestesiol 1993;59(12):759-762.

11. Ayala-Sandoval S, Gonzalez-Guzman MP, Pineda-Diaz MV. Propofol infusion as induction and maintenance of anesthesia in pediatric patients [Propofol como agente de induccion y mantenimiento anestesico en infusion en el paciente pedaitrico]. Rev Mex Anest 1991;14(2):55-62.
12. Bagshaw O. TIVA with propofol and remifentanil. Anaesthesia 1999;54(5):501-5
13. Barclay K, Williams A, Major E.. Letter to editor: Propofol infusion in children. BMJ 1992;305(6859):953
14. Bastian H, Glasziou P, Chalmers I (2010) Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up? PLoS Med 7(9): e1000326. doi:10.1371/journal.pmed.1000326 29(2): 175–181
15. Beller EM, Glasziou PP, Altman DG, Hopewell S, Bastian H, et al. PRISMA for Abstracts: Reporting Systematic Reviews in Journal and Conference Abstracts. PLoS Med 2013, 10(4): e1001419. doi:10.1371/journal.pmed.1001419
16. Bendiksen A, Larsen M.. Convulsions, ataxia and hallucinations following propofol. Acta Anarsthesiol Scand 1998;42(6):739-741.
17. Blumer JL, Goodale D. Randomized controlled study of propofol sedation for mechanically ventilated infants and children. Unpublished abstract 2002.
18. Blussee van Oud-Alblas HJ, Peters JW, de Leeuw TG, Vermeylen KT, de Klerk LW, Tibboel D, Klein J, Weber F. A comparison in adolescents of composite auditory evoked potential index and bispectral index during propofol-remifentanil anesthesia for scoliosis surgery with intraoperative wake-up test. Anesth Analg 2008;107(5):1683-1688.
19. Booth A, Clarke M, Ghersi D, Moher D, Petticrew M, et al. (2011) Establishing a Minimum Dataset for Prospective Registration of Systematic Reviews: An International Consultation. PLoS ONE 6(11): e27319. doi:10.1371/journal.pone.0027319
20. Boulkedid R, Abdoul H, Loustau M, et al. Using and Reporting the Delphi Method for Selecting Healthcare Quality Indicators: A Systematic Review, 2011,PLoS ONE 6(6): e20476. doi:10.1371/journal.pone.0020476

21. Boutron I, Moher D, Altman DG, Schulz K, Ravaud P, for the CONSORT group. Methods and Processes of the CONSORT Group: Example of an Extension for Trials Assessing Nonpharmacologic Treatments. Ann Intern Med. 2008:W60-W67
22. Bradford Hill, A. The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine 1965;58:295-300.
23. Bray RJ. Fatal myocardial failure associated with a propofol infusion in a child. Anesthesia 1995;50(1):94.
24. Bray RJ. Propofol infusion syndrome in children. Paediatric Anaesthesia 1998;8(6):491-9. [MEDLINE: 9836214]
25. Bustamante S, Appachi E.. Acute pancreatitis after anesthesia with propofol in a child with glycogen storage disease type IA. Paediatr Anaesth 2006;16(6):680-683.
26. Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. BMJ 2004; 328(7441):702-708
27. Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. BMJ 2012;345:e5661.
28. Cannon M, Glazier S, Bauman L. Metabolic acidosis, rhabdomyolysis, and cardiovascularcollapse after prolonged propofol infusion. J Neurosurg 2001;95:1053-1056.
29. Casserly B, O'Mahony E, Timm E, Haqqie S, Eisele G, Urizar R. Propofol Infusion syndrome: an unusual cause of renal failure. Am J Kidney Dis 2004;44(6):e98-e101.
30. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291:2457e65.
31. Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. CMAJ 2004; 171:735e40.
32. Chevret L, Husson B, Nguefack S, Nehlig A, Bouilleret V. Prolonged refractory status epilepticus with early and persistent restricted hippocampal signal MRI abnormality. J Neurol 2008;255(1):112-116.

33. Chidambaran V, Sadhasivam S, Diepstraten J, Esslinger H, Cox S, Schnell B, Samuels BM, Inge T, Vinks AA, Knibbe CA. Evaluation of propofol anesthesia in morbidly obese children and adolescents. BMC Anesthesiol 2013;13(8):1-9.
34. Chou R, Aronson N, Atkins D, Ismaila AS, Santaguida P, Smith DH, et al. AHRQ Series paper4: Assessing harms when comparing medical interventions: AHRQ and the Effective Health care Program. J Clin Epi. 2010, 63:502-12.
35. Chou R, Fu R, Carson S, Saha S, Helfand M. Methodological shortcomings predicted lower harm estimates in one of two sets of studies of clinical interventions. J ClinEpidemiol 2007; 60(1): 18–28.22).
36. Chou R, Helfand M. Challenges in systematic reviews that assess treatment harms. Ann Intern Med 2005;142:1090-9.
37. Clarke A, Deeks JJ, Shakir SA. An assessment of the publicly disseminated evidence of safety used in decisions to withdraw medicinal products from the UK and US markets. Drug Saf 2006;29(2):175-181.
38. Cohen IT, Finkel JC, Hannallah RS, Goodale DB. Clinical and biochemical effects of propofol EDTA vs sevoflurane in healthy infants and young children. Pediatr Anaesth 2004;14(2):135-142.
39. Cohen IT, Finkel JC, Hannallah RS, Hummer KA, Patel KM. Rapid emergence does not explain agitation following sevoflurane anaesthesia in infants and children: a comparison with propofol. Paediatr Anaesth 2003;13(1):63-67.
40. Cornelius VR, Perrio MJ, Shakir SA, Smith LA. Systematic reviews of adverse effects of drug interventions: a survey of their conduct and reporting quality. Pharmacoepidemiol Drug Saf2009;18:1223-31
41. Cornfield D, Tegtmeyer K, Nelson M, Milla C, Sweeney M. Continuous propofol infusion in 142 critically ill children. Pedaitrics 2002;110(6):1177-1181.
42. Cox RG, Levy R, Hamilton MG, Ewen A, Farran P, Neil SG. Anesthesia can be safely provided for children in a high-field intraoperative magnetic resonance imaging environment. Pediatr Anaesth 2011;21(4):454-458.

43. Cray S, Robinson B, Cox P. Lactic acidemia and bradyarrhythmia in a child sedated with propofol. Crit Care Med 1998;26(12):2087-2092.
44. Cray SH, Holtby HM, Kartha VM, Cox PN, Roy WL. Early tracheal extubation after paediatric cardiac surgery: the use of propofol to supplement low-dose opioid anaesthesia. Paediatr Anaesth 2001;11(4):465-471.
45. Cremer OL, Moons KG, Bouman EA, Kruijswijk JE, de Smet AM, Kalkman CJ. Long-term propofol infusion and cardiac failure in adult head-injured patients. Lancet 2001;357(9250):117-118.
46. Culp KE, Augoustides JG, Ochroch AE, Milas BL. Clinical management of cardiogenic shock associated with prolonged propofol infusion. Anesth Analg 2004;99(1):221-226.
47. Da-Silva SS, Wong R, Coquillon P, Gavrilita C, Asuncion A. Partial-exchange blood transfusion: an effective method for preventing mortality in a child with propofol infusion syndrome. Pediatrics 2010;125(6):e1493-e1499.
48. Deng XQ, Wang M, Ji Y. Clinical comparison of propofol and remifentanil anaesthesia with sevoflurane and remifentanil anaesthesia for children with cleft lip and palate repair surgery. Hua Xi Kou Qiang Yi Xue Za Zhi (West China Journal of Stomatology) 2009;27(5):531-534
49. Deramoudt V, Gaudon M, Malledant Y, Chatellier A, Saint-Marc C, Lecallonnec A. Effects of propofol on intraocular pressure in surgery of strabismus in children [Effets du propofol sur la ,pression intraoculairedans la chirurgie du strabisme de l'enfant]. Ann Fr Anesth Reanim 1990;9(1):1-5.
50. Derry S, Loke Y, Aronson J. Incomplete evidence: the inadequacy of database in tracing published adverse drug reaction in clinical trials. BMC Medical Research Methodology 2001,1:7. http://www.biomedcentral.com/1471-2288/1/7
51. Diepstraten J, Chidambaran V, Sadhasivam S, Esslinger HR, Cox SL, Inge TH, Knibbe CA, Vinks AA. Propofol clearance in morbidly obese children and adolescents: influence of age and body size. Clin Pharmacokinet 2012;51(8):543-551. Engelhardt, 2007

52. Engelhardt T, Chan M, McCheyne A, Karsli C, Luginbuehl I, Bissonnette B.. The effect of varying continuous propofol infusions on plasma cyclic guanosine 3',5'-Monophosphate concentrations in anesthetized children. Anesth Analg 2007;105(3):616-619.
53. Engelhardt T, McCheyne AJ, Morton N, Karsli C, Luginbueh I, Adeli K, Walsh W, Bissonnette B. Clinical adaptation of a pharmacokinetic model of Propofol plasma concentrations in children. Pediatr Anaesth 2008;18(3):235-239.
54. Erb TO, Hall JM, Ing RJ, Kanter RJ, Kern FH, Schulman SR, Gan TJ. Postoperative nausea and vomiting in children and adolescents undergoing radiofrequency catheter ablation: a randomized comparison of propofol- and isoflurane-based anesthetics. Anesth Analg 2002;95(6):1577-1581.
55. Erb TO, Kanter RJ, Hall JM, Gan TJ, Kern FH, Schulman SR. Comparison of electrophysiologic effects of propofol and isoflurane-based anesthetics in children undergoing radiofrequency catheter ablation for supraventricular tachycardia. Anesthesiology 2002;96(6):1386-1394.
56. Eriksen J, Povey HM. A case of suspected non-neurosurgical adult fatal propofol infusion syndrome. Acta Anaesthesioll Scand 2006;50(1):117-119. [MEDLINE: 16451160]
57. Ernst E, Pittler M. Assessment of therapeutic safety in systematic reviews: literature review. BMJ 2001;323:546
58. FDA / Astra Zeneca Pharmaceuticals. Dear health care provider letter. www.fda.gov/ medwatch/safety/2001/diprivan deardoc.pdf 2001
59. Festa M, Bowra J, Schell D. Use of propofol infusion in Australian and New Zealand paediatric Intensive care units. Anaesth Intensive Care 2002;30(6):786-793.
60. Fudickar A, Bein B. Propofol infusion syndrome: update of clinical manifestations and pathophysiology. Minerva Anesthesiol 2009;75:339-344.
61. Fudickar A, Smigaj K, Ensenauer R, Fischer G, Dutschke P, Steinfath M, Bein B. Effect of propofol and sevoflurane on acid-base balance during pediatric heart catheterization. Minerva Anestesiol 2013;79(6):626-633.

62. Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. Ann Intern Med 2006; 144(5):364-367
63. Golder S, Loke Y, Bland M. Unpublished data can be of value in systematic reviews of adverse effects: methodological overview. Journal of Clinical Epidemiology 2010; (63) 1071e1081
64. Golder S, Loke Y. Failure or success of electronic search strategies to identify adverse effects data. J Med Libr Assoc. 2012;100(2):130-4.
65. Golder S, Loke Y. Is there evidence for biased reporting of published adverse effects data in pharmaceutical industry-funded studies? Br J Clin Pharmacol 2008;66:767-3.
66. Golder S, Loke Y. The contribution of different information sources for adverse effects data. Int J Technol Assess Health Care 2012;28(2):133-7.
67. Golder S, Loke YK, McIntosh HM. Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. J Clin Epidemiol2008;61:440-8.
68. Golder S, Loke YK, McIntosh HM. Room for improvement? A survey of the methods used in systematic reviews of adverse effects. BMC Med Res Methodol2006;6:3.
69. Golder S, Loke YK, Zorzela L. Some improvements are apparent in identifying adverse effects in systematic reviews from 1994 to 2011. J Clin Epidemiol 2013;66:253-60.
70. Golder S, Loke YK, Zorzela L. Some improvements are apparent in identifying adverse effects in systematic reviews from 1994 to 2011. J Clin Epidemiol2013;66:253-60.
71. Golder S, Loke YK. Is there evidence for biased reporting of published adverse effects data in pharmaceutical industry-funded studies? Br JClinPharmacol 2008; 66(6): 767–773.
72. Golder S, McIntosh HM, Loke Y. Identifying systematic reviews of the adverse effects of health care interventions. BMC Med Res Methodol 2006; 6: 22.
73. Goodman C.M. The Delphi technique: a critique, Journal of Advanced Nursing, 1987, 12, 729-734.
74. GRADEpro [Computer program]. Brozek J, Oxman A, Schünemann H. Version 3.2 for Windows, 2008.

75. Guitton C, Gabillet L, Latour P, Rigal JC, Boutoille D, Al Habash O. Propofol infusion syndrome during refractory status epilepticus in a young adult: successful ECMO resuscitation. Neurocrit Care 2011;15(1):139-145.
76. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ. What is "quality of evidence" and why is it important to clinicians? BMJ 2008;336(7651):995-998.
77. Hammad T, Neyarapally G, Pinheiro S, et al. Reporting of meta-analyses of randomized controlled trials with a focus on drug safety: An empirical assessment. Clin Trials 2013 10: 389 DOI: 10.1177/1740774513479467
78. Hammad T, Neyarapally G, Pinheiro S, Iyasu S, Rochester G, Dal Pan G. Reporting of meta-analyses of randomized controlled trials with a focus on drug safety: An empirical assessment. Clin Trials 2013;10:389-97.
79. Hanna JP, Ramundo ML. Rhabdomyolisis and hypoxia associated with prolonged propofol infusion in children. Neurology 1998;50(1):301-303.
80. Hartley J. Clarifying the abstracts of systematic literature reviews. Bull Med LibrAssoc 2000; 88(4): 332–337.
81. Hasani A, Gecaj-Gashi A, Llullaku S, Jashari H. Postoperative analgesia in children after propofol versus sevoflurane anesthesia. Pain Med 2013;14(3):442-446.
82. Hasani A, Ozgen S, Baftiu N. Emergence agitation in clhildren after propofol versus halothane anesthesia. Med Sci Monit 2009;15(6):CR302-CR306.
83. Hasson F, Keeney S, McKenna. Research guidelines for the Delphi survey technique. journal of Advanced Nursing, 2000, 32(4), 1008-1015
84. Health Canada. Drugs and Health Products. Update on the use of Propofol. http://www.hc-sc.gc.ca/dhp-mps/medeff/advisories-avis/prof/_2002/propofol_pediatric_2_nth-ah-eng.php July10, 2002
85. Hein S, Benken I, Quintel M, Kettler D. Propofol Infusion Syndrome [Das propofol-infusions-syndrom]. Anasth Intensivecmed 2004;45:580-584.
86. Hernandez AV, Walker E, Ioannides-Demos LL, Kattan MW. Challenges in meta-analysis of randomized clinical trials for rare harmful cardiovascular events: the case of rosiglitazone. Am Heart J2008;156:23-30

87. Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011]. Cochrane Collaboration, 2011. www.cochrane-handbook.org.
88. Hill AB. The environment and disease: association or causation? Proc Roy Soc Med 1965;58:295-300.
89. Holzki J, Aring C, Gillor A. Death after re-exposure to propofol in a 3-year-old child: a case report. Paediatr Anaesth 2004;14(3):265-270.
90. Holzki J. Fatal bradyarrhythmia following long term propofol infusion in children [Todliche bradyarrythmien nach propofoldauersedierung im kindesalter]. Journal fur anesthesie und intensivbehandlung 2002;9(1):201-203.
91. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, Schulz KF and the CONSORT Group (2008) CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. PLoS Med 5(1): e20. doi:10.1371/journal. pmed.0050020
92. Hopewell S, Wolfenden L, Clarke M. Reporting of adverse events in systematic reviews can be improved: survey results. J Clin Epidemiol 2007;61:597-602.
93. Howick J, Glasziou P, Aronson JK. The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute? J R Soc Med. May 1, 2009; 102(5): 186–194.
94. Hussein G, Olejniczak P, Carey M, Fisch B, Casey G, Garcha T. Propofoi-induced dystonia: a case report and a review of the literature. Clin Res Regul Aff 1999;16(4):175-181.
95. Ioannidis J, Evans S, Gotzsche P, O'Neill R, Altman D, Schulz K, Moher D. Better Reporting of Harms in Randomized Clinical Trials: an Extension of the CONSORT Statement. Annals of Internal Medicine, 2004; 141 (10):781-788.
96. Ioannidis JP, Hozo I, Djulbegovic B. Optimal type I and type II error pairs when the available sample size is fixed. J ClinEpidemiol. 2013 Aug;66(8):903-910.e2. doi: 10.1016/j.jclinepi.2013.03.002. Epub 2013 May 9.
97. Ioannidis JP, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. JAMA 2001;285:437-43.

98. Ioannidis JP. Adverse events in randomized trials: neglected, restricted, distorted, and silenced. Arch Intern Med 2009;169:1737-9.

99. Ioannidis JP. Reporting of safety data from randomised trials [Letter]. Lancet. 1998;352:1752-3. [PMID: 9848355]

100. Ivani G, Della Rocca M, De Negri P, Corca D, Cuneo GP, Caffarena PE. Continuous-flow propofol in pediatric oncologic surgery [Le propofol en debit continu en chirurgie oncologique pediatrique]. Cah Anesthesiol 1991;39(6):409-410.

101. Jeleazcov C, Ihmsen H, Schmidt J, Ammon C, Schwilden H, Schuttler J, Fechner J. Pharmacodynamic modelling of the bispectral index response to propofol-based anaesthesia during general surgery in children. Br J Anaesth 2008;100(4):509-516.

102. Jeleazcov C, Schmidt J, Schmitz B, Becke K, Albrecht S. EEG variables as measures of arousal during propofol anaesthesia for general surgery in children: rational selection and age dependence. Br J Anaesth 2007;99(6):845-854.

103. Jorens PG, Van den Eynden GG. Propofol infusion syndrome with arrhythmia, myocardial fat accumulation and cardiac failure. Am J Cardiol 2009;104(8):1160-1162.

104. Kaddu R, Bhattacharya D, MD, Metriyakool K, Thomas R, Tolia V. Propofol compared with general anesthesia for pediatricGI endoscopy: is propofol better? Gastrointest Endosc 2002;55(1):27-32.

105. Kang TM. Propofol infusion syndrome in critically ill patients. Ann Pharmacother 2002;36(9):1453-1456.

106. Karaman Y, Goktay A, Agin H, Karaarslan U.. Propofol infusion syndrome or adrenoleukodystrophy? Paediatr Anaesth 2013;23(4):368-370.

107. Kassebaum N, Hairr J, Goldsmith W, Barwise J, Pandharipande P. Diabetes insipidus associated with propofol anesthesia. J Clin Anesth 2008;20(6):466-468.

108. Kill C, Leonhardt A, Wulf H. Lacticacidosis after short-term infusion of propofol for anaesthesia in a child with osteogenesis imperfecta. Paediatr Anaesth 2003;13(9):823-826.

109. Kiringoda R, Thurm AE, Hirschtritt ME, Koziol D, Wesley R, Swedo SE, O'Grady NP, Quezado ZM. Risks of propofol sedation/anesthesia for imaging studies in pediatric

research: eight years of experience in a clinical research center. Arch Pediatr Adolesc Med 2010;164(6):554-560.

110. Kirkpartrick M, Cole G. Propofol infusion in children: letter to editor. BMJ 1992;305(6863):1223.

111. Klockars JG, Hiller A, Munte S, van Gils M, Taivainen T. Spectral entropy as a measure of hypnosis and hypnotic drug effect of total intravenous anesthesia in children during slow induction and maintenance. Anesthesiology 2012;116(2):340-351.

112. Knibbe CA, Melenhorst-de Jong G, Mestrom M, Rademaker CM, Reijnvaan AF, Zuideveld KP, Kuks PF, van Vught H, Danhof M. Pharmacokinetics and effects of propofol 6% for short-term sedation in paediatric patients following cardiac surgery. Br J Clin Pharmacol 2002;54(4):415-422.

113. Kobacas S, Balcioglu T, Cevik A. The comparison of the effect of inhalation and intravenous anesthesia on the neuromuscular block of rocuronium in children. Turk Anest Rean Cem Meemast 2002;30:127-132.

114. Koch M, De Backer D, Vincent JL. Lactic acidosis: an early marker of propofol infusion syndrome? Intensive Care Med 2004;30(5):522.

115. Kogan A, Efrat R, Katz J, Vidne BA. Propofol-ketamine mixture for anesthesia in pediatric patients undergoing cardiac catheterization. J Cardiothorac Vasc Anesth 2003;17(6):691-693.

116. Koriyama H, Duff JP, Guerra GG, Chan AW, Sedation Withdrawal and Analgesia Team. Is propofol a friend or a foe of the pediatric intensivist? Description of propofol use in a PICU. Pediatr Crit Care Med 2014;15(2):e66-e71.

117. Krause TK, Jansen L, Scholz J, Böttcher H, Wappler F, Burmeister MA, am Esch JS. Propofol anesthesia in children does not induce sister chromatid exchanges in lymphocytes. Mutat Res 2003;542(1-2):59-64.

118. Kruessell MA, Udink ten Cate FE, Kraus AJ, Roth B, Trieschmann U. Use of propofol in pediatric intensive care units: a national survey in Germany. Pediatr Crit Care Med 2012;13(3):e150-e154.

119. Lanigan C, Sury M, Bingham R, Howard R, Mackersie A. Neurological sequelae in children after prolonged propofol infusion. Anaesthesia 1992;47(9):810-811.

120. Laquay N, Pouard P, Silicani MA, Vaccaroni L, Orliaguet G. Early stages of propofol infusion syndrome in paediatric cardiac surgery: two cases in adolescent girls. Br J Anaesth 2008;101(6):880-881.

121. Larsson S, Asgeirsson B, Magnusson J. Propofol-fentanyl anesthesia compared to thiopental-halothane with special reference to recovery and vomiting after pediatric strabismus surgery. Acta Anaesthesiol Scand 1992;36(2):182-186.

122. Lau J, Ioannidis JP, Terrin N, et al. The case of the misleading funnel plot. BMJ. 2006;333(7568):597-600.

123. LeBlanc JG, Blackstock D, Macnab A, Gagnon F, Gagnon R, Russell J, Ring T. Effects of propofol on cerebral oxygenation during cardiopulmonary bypass in children. Can J Anaesth 2000;47(11):1082-1089.

124. Lebovic S, Reich DL, Steinberg LG, Vela FP, Silvay G. Comparison of propofol versus ketarnine for anesthesia in pediatric patients undergoing cardiac catheterization. Anesth Analg 1992;74(4):490-494.

125. Levati A, Colombo N, Arosio EM, Savoia G, Tommasino C, Scialfa G, Boselli L. Propofol anaesthesia in spontaneously breathing paediatric patients during magnetic resonance imaging. Acta Anaesthesiol Scand 1996;40(5):561-565.

126. Liberati A, AltmanD, Tetzlaff J, Mulrow C, Gøtzsche P, Ioannidis J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700

127. Loke Y, Price D, Herxheimer A. Systematic reviews of adverse effects: framework for a structured approach. BMC Medical Research Methodology. July 2007 7:32. http://biomedcentral.com/1471-2288/7/32

128. Loke YK, Aronson JK. Incomplete evidence: the inadequacy of databases in tracing published adverse drug reactions in clinical trials. BMC Med Res Methodol 2001; 1: 7.

129. Loke YK, Derry S. Reporting of adverse drug reactions in randomised controlled trials—a systematic survey. BMC Clin Pharmacol. 2001;1:3. [PMID: 11591227]

130. Loke YK, Golder S, Vandenbroucke J. Comprehensive evaluations of the adverse effects of drugs: importance of appropriate study selection and data sources. Ther Adv Drug Saf 2011;2:59-68.

131. Loke YK, Price D, Herxheimer A. Systematic reviews of adverse effects: framework for a structured approach. BMC Med Res Methodol 2007;7:32.

132. Lynn M.R., Laman E.L. & Englebardt S.P. Nursing administration research priorities: a national Delphi study, Journal of Nursing Administration, 1998, 28(5), 7-11.

133. MacPherson H, Altman DG, Hammerschlag R, Youping L, Taixiang W, White A, Moher D; STRICTA Revision Group. Revised STandards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement. PLoS Med. 2010 Jun 8;7(6):e1000261

134. Mahoney MR, Sargent DJ. Adverse-event rates: journals versus databases. Lancet 2007;369:171e2.

135. Maldini B, Miskulin M. Outpatient arthroscopic knee surgery under combined local and intravenous propofol anesthesia in children and adolescents. Pediatr Anaesth 2006;16(11):1125-1132.

136. Martin P, Murthy V, Petros A. Methabolic, biochemical and hemodynamics effects of infusion of propofol for long term sedation of children undergoing intensive care. British J Anesth 1997;79:276-279.

137. Martin TM, Nicolson SC, Bargas MS. Propofol anesthesia reduces emesis and airway obstruction in pediatric outpatients. Anesth Analg 1993;76(1):144-148.

138. McIntosh HM, Woolacott NF, Bagnall AM. Assessing harmful effects in systematic reviews. BMC Med Res Methodol 2004; 4: 19

139. McKenna H. The Delphi technique: a worthwhile research approach for nursing? Journal of Advanced Nursing, 1994, 19, 1221-1225

140. Mehta N, DeMunter C, Habibi P, Nadel S, Britto J. Short-term propofol infusions in children. Lancet 1999;354(9181):866-867.

141. Meyer P, Langlois C, Soete S, Leydet J, Echenne B, Rivier F, Bonafe A, Roubertie A. Unexpected neurological sequelae following propofol anesthesia in infants: three case reports. Brain Dev 2010;32(10):872-878.

142. Michele TM, Pinheiro S, Iyasu S. The safety of tiotropium–the FDA's conclusions. N Engl J Med 2010; 363: 1097–99.

143. Mierzewska-Schmidt M, Arlukiewicz-Piwowar M, Rawicz M, Woloszczuk-Gebicka B. Propofol infusion syndrome in a 4 year-old girl. Case report. Anestezjologia Intensywna Terapia 2006;38(3):139-142.

144. Mironov PI, Sergeev AV. Anestesic maintenance during cochlear implnatation in children. Anesteziologiya i reannimatologia 2012;1:18-20.

145. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, et al. (1994) Improving the quality of reporting of meta-analysis of randomized controlled trials: The QUOROM statement. Lancet 354: 1896–1900.

146. Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA. 2001;285:1992-5. [PMID: 11308436]

147. Moher D, Liberati A, Tetzlaff J, Altman D. The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PloS Med 2009;6:e1000097.

148. Moher D, Liberati A, Tetzlaff J, Altman D. The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PloS Med2009;6:e1000097

149. Moher D, Schulz K, Simera I, Altman D. Guidance for Developers of Health Research Reporting Guidelines. PloS Medicine, 2010; 7 (2) e1000217.

150. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. PloS Med 2007;4:e78.

151. Moseley AM, Elkins MR, Herbert RD, Maher CG, Sherrington C. Cochrane reviews used more rigorous methods than non-cochrane reviews: survey of systematic reviews in physiotherapy. J Clin Epidemiol 2009;62:1021-30.

152. Muir Gray JA. Evidence-based in Healthcare. How to make health policy and management decisions. Churchill Livingstone,2001.

153. Murphy M, Black N, Lamping D, McKee C, Sanderson C, et al. Consensus Development Methods and their use in clinical guideline development. Health Technol Assess 1998; 2: 1-88.

154. Murray DM, Thorne GC, Rigby-Jones AE, Tonucci D, Grimes S, Tooley MA, Wolf AR. Electroencephalograph variables, drug concentrations and sedation scores in children emerging from propofol infusion anaesthesia. Paediatr Anaesth 2004;14(2):143-151.

155. Papaioannou V, Dragoumanis C, Theodorou V, Pneumatikos I. The propofol infusion 'syndrome' in intensive care unit:from pathophysiology to prophylaxis and treatment.. Acta Anesth Belg 2008;59:79-86.

156. Papanikolau PN, Ioannidis JP. Availability of large-scale evidence on specific harms from systematic reviews of randomized trials. Am J Med 2004;117:582-9.

157. Park HJ, Kim YL, Kim CS, Kim SD, Kim HS. Changes of bispectral index during recovery from general anesthesia with 2% propofol and remifentanil in children. Pediatr Anesth 2007;17(4):353-357.

158. Parke TJ, Stevens JE, Rice AS, Greenaway CL, Bray RJ, Smith PJ, Waldmann CS, Verghese C. Metabolic acidosis and fatal myocardial failure after propofol infusion in children: five case reports. BMJ 1992;305(6854):613-616.

159. Patel KN, Simon HK, Stockwell CA, Stockwell JA, DeGuzman MA, Roerig PL, Rigby MR. Pedaitric procedural sedation by a dedicated nonanesthesiology pedaitric sedation service using propofol. Pediatr Emerg Care 2009;25(3):133-138.

160. Patermann B, Buzello S, Duck M, Paul M, Kampe S. Accidental tenfold overdose of propofol in a 6-month oldi nfant undergoing elective craniosynostosis repair. Anaesthesia 2004;59(9):912-914.

161. Pedersen N, Jensen A, Kilmose L, Olsen K.. Propofol-remifentanil or sevoflurane for children undergoing magnetic resonance imaging? A randomised study. Acta Anaesthesiol Scand 2013;57(8):988-995.

162. Peeters MY, Prins SA, Knibbe CA, DeJongh J, van Schaik RH, van Dijk M, van der Heiden IP, Tibboel D, Danhof M. Propofol pharmacokinetics and pharmacodynamics for depth of sedation in nonventilated infants after major craniofacial surgery. Anesthesiology 2006;104(3):466-47

163. Pepperman M, Macrae D. A comparison of propofol and other sedative use in paediatric intensive care in the United Kingdom. Paediatr Anaesth 1997;7(2):143-153.

164. Perrier ND, Baerga-Varela Y, Murray MJ. Death related to propofol use in an adult patient. Crit Care Med 2000;28(8):3071-3074

165. Pershad J, Wan J, Anghelescu DL. Comparison of propofol with pentobarbital/midazolam/fentanyl sedation for magnetic resonance imaging of the brain in children. Pediatrics 2007;120(3):e629-e636.

166. Phillips B, Ball C, Badenoch D,Straus S, Haynes B ,Dawes M. Oxford Centre for Evidence-based Medicine. Levels of Evidence Model. 1998. Updated by Jeremy Howick March 2009. Oxford Centre for Evidence-based Medicine - Levels of Evidence (March 2009) http://www.cebm.net/index.aspx?o=1025

167. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJW. Reporting of non inferiority and equivalence randomized trials: An extension of the CONSORT statement. JAMA. 2006; 295:1152-1160.

168. Pilkington K, Boshnakova A. Complementary medicine and safety: A systematic investigation of design and reporting of systematic reviews. Complement Ther Med(2011), doi:10.1016/j.ctim.2011.10.002

169. Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. Arch Intern Med 2009;169:1756-61.

170. Playfor SD, Venkatesh K. Current patterns of propofol use in PICU in the United Kingdom and North America. Paediatr Anaesth 2004;14(6):501-504.

171. Powell C. The Delphi technique: myths and realities. J Adv Nursing, 2003, 41: 376–382.

172. Prins SA, Peeters MY, Houmes RJ, van Dijk M, Knibbe CA, Danhof M, Tibboel D. Propofol 6% as sedative in children under 2 years of age following major craniofacial surgery. Br J Anaesth 2005;94(5):630-635.

173. Reed MD, Yamashita TS, Marx CM, Myers CM, Blumer JL. A pharmacokinetically based propofol dosing strategy for sedation of the critically ill, mechanically ventilated pediatric patient. Critl Care Med 1996;24(9):1473-1481.

174. Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011. [Other:]

175. Rigby-Jones AE, Nolan JA, Priston MJ, Wright PM, Sneyd JR, Wolf AR. Pharmacokinetics of propofol infusions in critically ill neonates, infants, and children in an intensive care unit. Anesthesiology 2002;97(6):1393-1400.

176. Roberts RJ, Barletta JF, Fong JJ, Schumaker G, Kuper PJ, Papadopoulos S, Yogaratnam D, Kendall E, Xamplas R, Gerlach AT, Szumita PM, Anger KE, Arpino PA, Voils SA, Grgurich P, Ruthazer R, Devlin JW. Incidence of propofol-related infusion syndrome in critically ill adults: a prospective, multicenter study. Crit Care 2009;13(5):R169. [MEDLINE: 19874582]

177. Robinson JD, Melman Y, Walsh EP. Cardiac conduction disturbances and ventricular tachycardia after prolonged propofol infusion in an infant. Pacing Clin Electrophysiol 2008;31(8):1070-1073.

178. Rosen DJ, Nicoara A, Koshy N, Wedderburn RV. Too much of a good thing? Tracing the history of the propofol infusion syndrome. J Trauma 2007;63(2):443-447.

179. Sabsovich I, Rehman Z, Yunen J, Coritsidis G. Propofol infusion syndrome: a case of increasing morbidity with traumatic brain injury. Am J Crit Care 2007;16(1):82-85.

180. Sackett DL, Wennberg JE. Choosing the best research design for each question. BMJ1997; 315(7123):1636.

181. Salvitano BM, Diaz MVP, Flores MdCL, Aldana LR. Intravenous general anesthesia: Propofol-Fentanyl vs Midazolan-Fentanyl in pediatric patient [Anestesia general endovenosa: propofol-fentanyl vs. midazolam-fentanyl en el paciente pediatrico]. Rev Anest Mex 1997;9(6):204-211.

182. Santaguida PL, Raina P. The development of the McHarm quality assessment scale for adverse events: Delphi consensus on important criteria for evaluating harms. Available at http://hiru.mcmaster.ca/epc/mcharm.pdf2008;Accessed on May 14, 2008

183. Scharf O, Colevas AD. Adverse event reporting in publications compared with sponsor database for cancer clinical trials. J ClinOncol 2006;24:3933e8.

184. Schmidt J, Fechner J, Fritsch B, Schmitz B, Carbon R, Rosch W, Albrecht S. Propofol-remifentanil versus sevoflurane-remifentanil for anesthesia for pediatric procedures in infants, children and adolescents. Anaesthesist 2001;50(10):757-766.

185. Sepulveda P, Cortınez LI, Saez C, Penna A, Solari S, Guerra I, Absalom AR. Performance evaluation of paediatric propofol pharmacokinetic models in healthy young children. Br J Anaesth 2011;107(4):593-600.

186. Shamseer L, Galipeau J, Turner L, Moher D. Improving the reporting and usability of research studies. Can J Anesth 2013; 60:337–344.

187. Shankar V. Severe propofol infusion syndrome in a child treated with continous veno-venous hemofiltration (CVVH). Crit Care Med 2009;37(12 sup):A540.

188. Sheridan RL, Keaney T, Stoddard F, Enfanto R, Kadillack P, Breault L. Short- term propofol infusion as an adjunct to extubation in burned children. J Burn Care Rehabil 2003;24(6):356-360.

189. Shuster JJ, Jones LS, Salmon DA. Fixed vs random effects meta-analysis in rare event studies: the rosiglitazone link with myocardial infarction and cardiac death. Stat Med. 2007 Oct 30;26(24):4375-85.

190. Shuster JJ. Empirical vs natural weighting in random effects meta-analysis. Stat Med. 2010 May 30;29(12):1259-65. doi: 10.1002/sim.3607.

191. Sigh S, Loke YK. Drug safety assessment in clinical trials: methodological challenges and opportunities. Trials2012;13:138

192. Srinivasan M, Turmelle M, Depalma LM, Mao J, Carlson DW. Procedural sedation for diagnostic imaging in children by pedaitric hospitalists using propofol: analysis of nature, frequency, and predictors of adverse events and interventions. J Pediatric 2012;160(5):801-806.

193. Steinmetz J, Holm-Knudsen R, Sorensen MK, Eriksen K, Rasmussen LS. Hemodynamic differences between propofol–remifentanil and sevoflurane anesthesia for repair of cleft lip and palate in infants. Pediatr Anaesth 2007;17(1):32-37.

194. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, et al.Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011;343:d4002

195. Strickland RA, Murray MJ. Fatal metabolic acidosis in a pediatric patient receiving an infusion of propofol in the intensive care unit: is there a relationship? Crit Care Med 1995;23(2):405-409.

196. Svensson ML, Lindberg L. The use of propofol sedation in a paediatric intensive care unit. Nurs Crit Care 2012;17(4):198-203.

197. Teng SN, Kaufman J, Czaja AS, Friesen RH, da Cruz EM. Propofol as a bridge to extubation for high-risk children with congenital cardiac disease. Cardiol Young 2011;21(1):46-51.

198. The PLoS Medicine Editors (2011) Best Practice in Systematic Reviews: The Importance of Protocols and Registration. PLoS Med 8(2): e1001009. doi:10.1371/journal.pmed.1001009

199. Ting TW, Lee JH. Acute pancreatitis after propofol infusion in a teenage patient. Anaesth Intensive Care 2012;40(3):561-562.

200. Ture H, Mercan A, Koner O, Aykac B, Ture U. The effects of propofol infusion on hepatic and pancreatic function and acid-base status in children undergoing craniotomy and receiving phenytoin. Anesth Analg 2009;109(2):366-371.

201. Uezono S, Goto T, Terui K, Ichinose F, Ishguro Y, Nakata Y, Morita S. Emergence agitation after sevoflurane versus propofol in pediatric patients. Anesth Analg 2000;91(3):563-566

202. van Gestel JP, Blusse van Oud-Alblas JH, Malingre M, Ververs FF, Braun KP, van Nieuwenhuizen O. Propofol and thiopental for refractory status epilepticus in children. Neurology 2005;65(4):591-592.

203. van Meurs M, van der Starre PJA. Green Urine [Groene Urine]. Nederlands Tijdschrift voor Anesthesiologie 2000;13(3):87-89.

204. van Straaten E, Hendriks JJ, Ramsey G, Vos GD. Rhabdomyolysis and pulmonary hypertension in a child, possibly due to long-term high-dose propofol infusion. Intensive Care Med 1996;22(9):997-1001.

205. Vandenbroucke JP; von Elm E; Altman D; Gøtzsche P; Mulrow C ; Pocock S; et al. InitiativeStrengthening the Reporting of Observational Studies in Epidemiology (STROBE). Guidelines for reporting observation studies. PLoS Med. 2007 DOI: 10.1371/journal.pmed.0040297

206. Vandermeer B, Bialy L, Johnston B, Hooton N, Hartling L, Klassen TP, et al. Meta-analyses of safety data: a comparison of exact versus asymptotic methods. Stat Methods Med Res 2009;18:421-32.

207. Veldhoen ES, Hartman BJ, van Gestel JP. Monitoring biochemical parameters as an early sign of propofol infusion syndrome: false feeling of security. Pediatr Crit Care Med 2009;10(2):e19-e21.

208. Waalkens H, Plotz F, Verkade H, Strengers J, Knoester H, Mandema J. Continuous propofol infusion contraindicated in children [Langdurige sedatie met propofol gecontraindiceerd bij kindern]. Nederlands Tijdschirift Geneeskunde 1997;141(9):440-442.

209. Warren FC, Abrams KR, Golder S, Sutton AJ. Systematic review of methods used in meta-analyses where a primary outcome is an adverse or unintended event. BMC Med Res Methodol 2012;12:64.

210. Weir PM, Munro HM, Reynolds PI, Lewis IH, Wilton NC. Propofol infusion and the incidence of emesis in pediatric outpatient strabismus surgery. Anesth Analg 1993;76(4):760-764.

211. Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, et al. (2012) PRISMA-Equity 2012 Extension: Reporting Guidelines for Systematic Reviews with a Focus on Health Equity. PLoS Med. 9(10):e1001333.

212. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm.

213. Welzing L, Vierzig A, Junghaenel S, Eifinger F, Oberthuer A, Trieschmann U, Roth B. Remifentanil and propofol for weaning of mechanically ventilated pediatric intensive care patients. Eur J Pedaitr 2011;170(4):477-481.

214. Westhout FD, Muhonen MG, Nwagwu CI. Early propofol infusion syndrome following cerebral angiographic embolization for giant aneurysm repair. Case report. J Neurosurg 2007;106(2 Suppl):139-142.

215. Williams GD, Jones TK, Hanson KA, Morray JP. The hemodynamic effects of propofol in children with congenital heart disease. Anesth Analg 1999;89(6):1411-1416.

216. Withington DE, Decell MK, Al Ayed T. A case of propofol toxicity: further evidence fora causal mechanism. Pediatric Anaesth 2004;14(6):505-508.

217. Wolf A, Weir P, Segar P, Stone J, Shield J. Impaired fatty acid oxidation in propofol infusion syndrome. Lancet 2001;357(9256):606-607.

218. Wong J. Propofol Infusion Syndrome. Am J Therapeut 2010;17:487-491.

219. Woodward WM, Barker I, John RE, Peacock JE. Propofol infusion vs thiopentone/isoflurane anaesthesia for prominent ear correction in children. Paediatr Anaesth 1997;7(5):379-383.

220. Xia WF, Liu Y, Zhou QS, Tang QZ, Zou HD. Comparison of the effects of propofol and midazolam on inflammation and oxidase stress in children with congenital heart disease undergoing cardiac surgery. Yonsei Med J 2011;52(2):326-332.

221. Xia WF, Liu Y, Zhou QS, Tang QZ, Zou HD. Protective effect of propofol and its relation to postoperation recovery in children undergoing cardiac surgery with cardiopulmonary bypass. Pediatr Cardiol 2011;32(7):940-946.

222. Yang W, Gu H, Chen B, Li J, Fan Q, Yuan Y, Wang X. Evaluation of SLOG/TCI-III pediatric system on target control infusion of propofol. J Transl Med 2011;9(187):1-6.

223. Yazici Y. Some concerns about adverse event reporting in randomized clinical trials. Bull NYU Hosp Jt Dis 2008;66:143-5

224. Zorzela L, Golder S, Liu Y, Pilkington K, Hartling L, et all. Quality of reporting in systematic reviews of adverse events: systematic review. BMJ 2014;348:f7668 doi: 10.1136/bmj.f7668.

225. Zorzela L, Hartling L, Loke Y, Joffe A, Ioannidis JP, Golder S, Santaguida PL, Altman D, Vohra S. The Development of PRISMA Harms using a Modified Delphi Technique. Thesis Chapter 3. Unpublished.

226. Zorzela LM, Punja S, Joffe A, Hartling L, Vandermeer B, Loke Y, Vohra S. Propofol infusion for paediatric sedation. Cochrane Database of Systematic Reviews 2012, Issue 4. Art. No.: CD009813. DOI: 10.1002/14651858.CD009813

227. Zou H, Zhou Q, Xia W, Wu L, Cheng B.. Cardio protective effects of propofol and midazolam in children with congenital heart disease undergoing open heart surgery. Chinese Medical J 2007;87(33):2309-2312.

228. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D for the CONSORT and Pragmatic Trials in Healthcare (Practihc) group. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008; 337;a2390

Appendix 1: Cochrane Protocol: Propofol infusion for paediatric sedation

Propofol infusion for paediatric sedation (Protocol)

Zorzela LM, Punja S, Joffe A, Hartling L, Vandermeer B, Loke Y, Vohra S

THE COCHRANE COLLABORATION®

This is a reprint of a Cochrane protocol, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2012, Issue 4

http://www.thecochranelibrary.com

WILEY
Publishers Since 1807

TABLE OF CONTENTS

[Intervention Protocol]

Propofol infusion for paediatric sedation

Liliane M Zorzela[1], Salima Punja[2], Ari Joffe[3], Lisa Hartling[1], Ben Vandermeer[4], Yoon Loke[5], Sunita Vohra[1]

[1]Department of Pediatrics, University of Alberta, Edmonton, Canada. [2]Department of Medicine, University of Alberta, Edmonton, Canada. [3]Department of Pediatrics, Division of Pediatric Intensive Care, University of Alberta and Stollery Children's Hospital, Edmonton, Canada. [4]Department of Pediatrics, Alberta Research Centre for Child Health Evidence & University of Alberta Evidence-based Practice Centre, Edmonton, Canada. [5]School of Medicine, University of East Anglia, Norwich, UK

Contact address: Liliane M Zorzela, Department of Pediatrics, University of Alberta, 8727-118 street, Edmonton, Alberta, T6G1T4, Canada. lilizorzela@hotmail.com.

Editorial group: Cochrane Anaesthesia Group.
Publication status and date: New, published in Issue 4, 2012.

Citation: Zorzela LM, Punja S, Joffe A, Hartling L, Vandermeer B, Loke Y, Vohra S. Propofol infusion for paediatric sedation. *Cochrane Database of Systematic Reviews* 2012, Issue 4. Art. No.: CD009813. DOI: 10.1002/14651858.CD009813.

ABSTRACT

This is the protocol for a review and there is no abstract. The objectives are as follows:

1. To identify the propofol dose and duration of infusion associated with the primary outcome (PRIS and cardiac arrest)

2. To describe the patient characteristics associated with the primary outcome

3. To identify a possible association between serious adverse events related to propofol use and the patient care setting (paediatric intensive care, operating room, paediatric emergency or other)

4. To describe the length of time from initiation and ending of propofol infusion to the development of the primary outcome

5. To identify the mode of surveillance for AE in different study designs (active surveillance, passive or specific monitoring procedures for collecting AE) and the relationship with identified AE

BACKGROUND

Propofol is an anaesthetic agent frequently used in paediatric sedation. Some of the valuable features of this drug include its quick onset of action and its short half life such that the patient is fully awake a few minutes after the infusion is stopped. Propofol plays an important role in allowing neurologic assessment of a sedated patient and also during the weaning of mechanical ventilation.

Since 1992, there have been several case reports of metabolic acidosis, rhabdomyolysis, liver dysfunction, cardiac arrhythmias and cardiac arrest associated with a prolonged propofol infusion in high doses. This phenomenon became referred to as 'propofol infusion syndrome' (PRIS). It is a poorly understood syndrome, usually causing death. PRIS was originally only reported in paediatric patients (Bray 1998; Hanna 1998; Parke 1992; Wolf 2001) but several more recent case reports of adult patients (Cremer 2001; Eriksen 2006; Kang 2002) with the same signs and symptoms suggest it is not exclusive to the paediatric population.

As PRIS is a rare disorder, it is unlikely to be identified in randomized clinical trials. These studies are often only powered to detect the efficacy of interventions, and therefore do not have sufficient data to confirm their safety (Loke 2011). Although a multicentre adult clinical trial found a 1.1% incidence of PRIS in critically ill adults receiving propofol (Roberts 2009), the only comparable paediatric study has not yet been published. Partial data reported a comparison of the brand name of propofol Diprivan 2%, Diprivan 1% versus standard sedative agents (for example, lorazepam, chloral hydrate, fentanyl, ketamine, morphine, or phenobarbital) in paediatric critical care patients and identified a significant increase in mortality in patients receiving propofol (the first group had 11% mortality; second group had 9% mortality; and the standard sedation group had only 4% mortality) (Health Canada 2002). Unfortunately, this study was never published in a peer reviewed journal and many uncertainties regarding the incidence of PRIS in paediatric patients remain. While available reports discuss an association between PRIS and propofol infusion, causation has not been clearly determined; nor has the effect of dose, duration, or setting should these relationships exist.

In several countries, including Canada and the United States, regulatory agencies (Health Canada 2002; US FDA 2001) have contraindicated the use of propofol infusion in critically ill paediatric patients, however propofol is used freely as a continuous infusion in paediatric operating rooms and emergency rooms. There is an urgent need to review the current knowledge of adverse effects of propofol infusion in paediatric patients to support clinicians and regulatory bodies to promote evidence-based decision-making when evaluating the best drug choices for continuous anaesthesia in paediatric patients.

Description of the condition

Propofol infusion syndrome (PRIS) has been described mostly by case reports or case series (Bray 1998; Cremer 2001; Eriksen 2006; Hanna 1998; Kang 2002; Parke 1992; Wolf 2001) of patients receiving prolonged infusions of propofol (more than 48 hours) and in high doses (more than 4 mg/kg/hour). Patients with PRIS develop metabolic acidosis, rhabdomyolysis, liver dysfunction, and arrhythmias, often followed by cardiac arrest and death.

Description of the intervention

The intervention being evaluated in this systematic review is the use of propofol as a continuous infusion for 60 minutes or longer in paediatric patients (aged more than 28 days and less then 19 years old).

The length of propofol infusion necessary to cause PRIS is unclear. We will assume in this review that induction doses of propofol do not cause PRIS. There is an association of PRIS with longer duration and higher doses of propofol in paediatric studies (Bray 1998; Hanna 1998; Parke 1992). We will further evaluate the association of duration of propofol infusion and dosage with PRIS as a subgroup analysis.

Why it is important to do this review

Propofol is an anaesthetic agent with short onset of action and short half life. These two characteristics make the drug clinically useful. However, there have been case reports of an association between propofol used in prolonged infusions and in high dosages with metabolic acidosis, liver dysfunction, arrhythmias and death. Regulatory bodies, such as FDA and Health Canada (Health Canada 2002; US FDA 2001) contraindicate the use of propofol as a continuous infusion in critically ill paediatric patients, mostly based on case reports (Bray 1998; Hanna 1998; Parke 1992) but propofol is freely used in paediatric emergency rooms and operating rooms. We understand that PRIS in not exclusive to the paediatric population (Cremer 2001; Eriksen 2006; Kang 2002) but, due to limitations of propofol use in critically ill children and no restriction of its use in other paediatric populations, we believe it is important to systematically review the use of propofol infusion in children in an attempt to identify risk factors and address its safety concerns.

The goal of this review is to identity the association of PRIS and other serious adverse events (AEs) associated with use of propofol as a continuous infusion in paediatric patients, and therefore identify, if possible, any risk factors associated with these events, such as patient setting, professional experience, dosage and duration of infusion. By doing a systematic review of all study designs we will be able to gather all the evidence available in an attempt answer those questions.

OBJECTIVES

1. To identify the propofol dose and duration of infusion associated with the primary outcome (PRIS and cardiac arrest)

2. To describe the patient characteristics associated with the primary outcome

3. To identify a possible association between serious adverse events related to propofol use and the patient care setting (paediatric intensive care, operating room, paediatric emergency or other)

4. To describe the length of time from initiation and ending of propofol infusion to the development of the primary outcome

5. To identify the mode of surveillance for AE in different study designs (active surveillance, passive or specific monitoring procedures for collecting AE) and the relationship with identified AE

METHODS

Criteria for considering studies for this review

Types of studies

We will not limit our inclusion criteria by study design. The Cochrane Adverse Effects Methods Group recommends the use of non-randomized studies for reviews of adverse events (Higgins 2011).

We will include randomized or quasi-randomized controlled trials of parallel group or cross-over design, which use individual or cluster randomization. In the case of cross-over trials, we will include only the first arm (we will consider events that happened after the cross-over in a sensitive analysis). Since our primary outcome is a rare event, we will also include cohort studies (prospective and retrospective), historically controlled studies, cross-sectional studies, case series and case reports.

Types of participants

The target population consists of paediatric patients (age ranging from 28 days to 19 years) receiving sedation using propofol for more than 60 minutes in a hospital (intensive care unit, operating room, emergency room, or any other location within the hospital) or other medical setting (for example dentist offices). The indication for sedation will not be a restriction at this time; instead, it will be the subject of a subgroup analysis. This is a review of adverse events of propofol infusion in paediatric patients, so it has broad inclusion criteria (any children receiving propofol infusion for sedation).

Types of interventions

The intervention being studied is the use of propofol as an anaesthetic drug given by continuous infusion. It is not clear the length of infusion necessary to cause PRIS. The case reports are usually associated with prolonged use (more than 48 hours) and in high doses (more than 4 mg/kg/hour). We will include patients receiving infusions lasting 60 minutes or longer, as an arbitrary number. Further, we will exclude studies in which propofol bolus or infusion was given in more than one study arm, in order to have a propofol free comparison group and to limit the possible risk factor to one single arm. As propofol does not provide analgesia, the group receiving propofol can also receive other sedative, analgesic drugs concomitantly.

The comparison group, when applicable, will be any sedative or analgesic agent different from propofol, alone or in combination.

Types of outcome measures

Primary outcomes

The primary outcome is the report of PRIS or any other serious adverse events resulting in cardiac arrest (with return to spontaneous circulation or use of extra-corporeal life support) or death associated with propofol use as a continuous infusion in paediatric patients. The follow-up time to measure the primary outcome for this study will be the development of the outcome of interest (PRIS or cardiac arrest) or time of hospital discharge for patients who did not develop the primary outcome.

The primary outcomes will be defined as the following.

Cardiac arrest, a reduction of cardiac output requiring any of the following: chest compressions; defibrillation, epinephrine boluses or cardiac mechanical support (extra-corporeal life support (ECLS); or ventricular assist device (VAD)).

Propofol infusion syndrome (PRIS) will be defined as metabolic acidosis, arterial pH $\leq$ 7.3 along with a serum bicarbonate $\leq$ 18 mg/dL; plus the presence of any signs in any the below categories (adapted from Roberts 2009):

1. rhabdomyolysis (the breakdown of muscle fibres resulting in the release of muscle fibre contents (myoglobin) into the bloodstream), defined as creatine phosphokinase (CPK) $\geq$ 10,000 IU/L or positive serum or urine myoglobin test or positive urinary casts for haemoglobin;
2. hypotension (initiation of a vasopressor agent or increase of $\geq$ 20% from baseline;
3. hepatic transaminitis (increase in aspartate aminotransferase or alanine aminotransferase, or both, $\geq$ 3 times above baseline);
4. hypertriglyceridaemia (serum triglyceride concentration $\geq$ 400 mg/dL);
5. hypoxia (partial pressure of arterial oxygen $\leq$ 60 mm Hg);
6. hyperthermia (temperature $\geq$ 38.3 °C);
7. cardiac dysfunction that includes asystole, pulseless electrical activity, ventricular fibrillation, sustained ventricular

tachycardia of 30 seconds or longer, myocardial failure (ejection fraction ≤ 40%), or bradycardia (heart rate ≤ 60 bpm);

8. renal failure that includes oliguria (urine output ≤ 0.5 mL/kg/hr for ≥ 6 hours), anuria (urine output ≤ 10 mL/hr for ≥ 6 hours), elevation in serum creatinine (increase of ≥ 1 mg/dL from baseline), or hyperkalaemia (serum K^+ ≥ 6 mg/dL (excluding other known causes or haemolyzed specimens).

We will also capture events that do not clearly fulfil the above inclusion criteria. These events will be reported as an unclear outcome definition and will be the subject of a sensitivity analysis.

Search methods for identification of studies

Electronic searches

We will search the Cochrane Central Register of Controlled Trials (CENTRAL) (*The Cochrane Library*, latest version); Ovid MEDLINE; and Ovid EMBASE (1988 to date).

The proposed search strategy is given in Appendix 1 (MEDLINE), Appendix 2 (EMBASE) and Appendix 3 (CENTRAL).

We will not impose a language restriction.

Searching other resources

We will search the websites of regulatory pharmaceutical authorities such as Current Problems in Pharmacology (www.mhra.gov.uk); the Australian Adverse Drug Reaction Bulletin (www.tga.gov.au/adr/aadrb.htm); the European Public Assessment Reports from the European Medicine Evaluation Agency (www.emca.eu) and the Food and Drug Administration FDA Medwatch (www.fda.gov/medwatch).

We will screen references of all retrieved articles to identify additional publications. No conference or meeting abstracts will be searched.

The search will not be limited by study type. The search date will start in 1977, the year when propofol was initially commercialized, to the present.

Data collection and analysis

Selection of studies

All studies meeting the following criteria will be included.

1. Paediatric population: age ranging from 28 days to 19 years, exclusively; if studies have mixed population we will contact authors for individual data or data specific to children.
2. Use of propofol as a continuous infusion for more than 60 minutes.
3. The comparison group (when applicable) will be any other sedative or analgesic agents (alone or in combination) different from propofol.

There will not be any study design restrictions. One author (LZ) will carry out the searches. Two review authors (LZ and SP) will independently screen all titles and abstracts and classify them as either:

- relevant (meeting all the inclusion criteria);
- possible (meeting some of the inclusion criteria or unclear whether meets inclusion criteria);
- rejected (failing to meet the inclusion criteria).

The studies classified as relevant or possible will have their full text reviewed by both review authors and any disagreement will be resolved by a third person (SV). For studies not published in English, a translator will provide the translation of titles and abstracts for screening and the full text if the study meets the inclusion criteria.

Data extraction and management

Two authors will independently extract data using standardized data extraction forms (Appendix 4) developed a priori by the authors. Data will be entered into RevMan Version 5.1. We will use two different data extraction forms, one for randomized or quasi-randomized clinical trials and a second for the non-randomized studies. The data extraction forms were piloted using two different studies of different designs.

We will extract the following data:

• study information, including type of study, inclusion and exclusion criteria;

• study population characteristics, baseline data such as age and gender, reason for hospital admission, setting, history of organ dysfunction prior to propofol initiation;

• medication details, including name of drug, dose, route, frequency and duration of treatment;

• definitions given for the reported adverse events, and methods used to monitor and report these.

Assessment of risk of bias in included studies

Clinical trials

Two review authors (LZ and SP) will independently assess the methodological quality of each trial using the Cochrane 'Risk of bias' tool (Higgins 2011). We will assess the following sources of bias for all study designs.

1. Selection bias, including randomization: describes the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups. Describes the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.

2. Performance bias: describes all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provides any information relating to whether the intended blinding was effective.

3. Detection bias: describes all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provides any information related to whether the intended blinding was effective.

4. Attrition bias: describes the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. States whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition or exclusions where reported, and any re-inclusions in analyses performed by the review authors.

5. Selective reporting bias: states how the possibility of selective outcome reporting was examined by the review authors, and what was found. We are aware that the adverse events related to a therapy are poorly reported in clinical trials.

6. Other source of bias: states any important concerns about bias not addressed in the other domains in the tool.

If particular questions or entries were pre-specified in the study's protocol, responses should be provided for each question or entry.

Two authors will assess the risk of bias of each trial, following the domain-based evaluation as described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). The authors will assess the domains for risk of bias; in the first three domains as 'Low risk of bias'; 'High risk of bias' and 'Unclear', which means there is an uncertain risk of bias.

Non-randomized studies

We will assess the risk of bias using the Newcastle-Ottawa Quality Assessment scale (Wells 2008) for cohorts and case controls. We will rate studies based on selection of participants, comparability between groups and exposure to the intervention (for case controls) and outcome (for cohort studies). The authors will assess risk of bias in studies, using a 'star' if the option (a) or sometimes (b) are selected. The more stars a study receives, the lower the risk of bias. Studies with seven or more stars will be considered to have a low risk of bias.

Each domain can only receive one star, with the exception of the field comparability, which can receive two stars.

Case series and case reports will be clinically assessed by the validity of the report, if in accordance with the PRIS definition adopted for this review.

The authors will assess the risk of bias of case control studies based on the following domains.

Selection

1. Is the case definition adequate?

a) Yes, with independent validation. * (It requires some independent validation (e.g. > 1 person, record, time, process to extract information, or reference to primary record source such as x-rays or medical or hospital records)).

b) Yes, if present some record linkage (e.g. ICD codes in database) or self-report with no reference to primary record.

c) No description.

2. Representativeness of the cases

a) Consecutive or obviously representative series of cases. * (It includes all eligible cases with outcome of interest over a defined period of time, all cases in a defined catchment area, all cases in a defined hospital or clinic, group of hospitals, health maintenance organization, or an appropriate sample of those cases (e.g. random sample).

b) Not satisfying requirements in part (a), or not stated.

3. Selection of controls

This item assesses whether the control series used in the study is derived from the same population as the cases and essentially would have been cases had the outcome been present.

a) Community controls * (i.e. same community as cases and would be cases if had outcome).

b) Hospital controls, within same community as cases (i.e. not another city) but derived from a hospitalised population.

c) No description.

4. Definition of controls

a) No history of disease * (endpoint). If cases are first occurrence of outcome, then it must explicitly state that controls have no history of this outcome. If cases have new (not necessarily first) occurrence of outcome, then controls with previous occurrences of outcome of interest should not be excluded.

b) No description of source of controls. No mention of history of outcome.

Comparability

1. Comparability of cases and controls on the basis of the design or analysis

A maximum of 2 stars can be allotted in this category.

a) Either cases and controls must be matched in the design or confounders must be adjusted for in the analysis for the most important factor, or both *.

b) Study controls are adjusted for any other additional factor *.

Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability.

If the odds ratio for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment.

Exposure

1. Ascertainment of exposure
a) Secure record * (e.g. surgical records)
b) Structured interview where blind to case control status *
c) Interview not blind to case control status
d) Written self-report or medical record only
e) No description
2. Same method of ascertainment for cases and controls
a) Yes *
b) No
3. Non-response rate
a) Same rate for both groups *
b) Non-respondents described
c) Rate different and no designation
The authors will assess risk of bias on cohort studies, based on the following domains.

Selection

1. Representativeness of the exposed cohort (this Item assesses the representativeness of exposed individuals in the community, not the representativeness of the sample of children from some general population.
a) Truly represents of the average of …….. in the community *.
b) Somewhat representative of the average ……in the community *.
c) Selected group of users
d) No description of the derivation of the cohort.
2. Selection of the non-exposed cohort
a) Drawn from the same community as the exposed cohort *
b) Drawn from a different source
c) Written self-report
d) No description
3. Ascertainment of exposure
a) A secure record * (e.g. surgical records)
b) Structured interview *
c) Written self-report
d) No description
4. Demonstration that outcome of interest was not present at start of study
In the case of mortality studies, outcome of interest is still the presence of a disease or incident, rather than death. That is to say that a statement of no history of disease or incident earns a star.
a) Yes*.
B) No.

Comparability

1. Comparability of cohorts on the basis of the design or analysis
A maximum of 2 stars can be allotted in this category.
a) Study controls for …….. (most important factor) *.
b) Study controls for any additional factor *.
Either exposed and non-exposed individuals must be matched in the design or confounders must be adjusted for in the analysis, or both. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability. Note: If the relative risk for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment.

Outcome

1. Assessment of outcome
For some outcomes (for example fractured hip), reference to the medical record is sufficient to satisfy the requirement for confirmation of the fracture. This would not be adequate for vertebral fracture outcomes, where reference to x-rays would be required.
a) Independent or blind assessment stated in the paper, or confirmation of the outcome by reference to secure records (x-rays, medical records, etc.) *.
b) Record linkage (e.g. identified through ICD codes on database records) *.
c) Self-report (i.e. no reference to original medical records or x-rays to confirm the outcome).
d) No description.
2. Was follow-up long enough for outcomes to occur (an acceptable length of time should be decided before quality assessment begins)?
a) Yes *
b) No
3. Adequacy of follow-up of cohorts
This item assesses the follow-up of the exposed and non-exposed cohorts to ensure that losses are not related to either the exposure or the outcome.
a) Completed follow-up in all participants accounted for *.
b) Participants lost to follow-up unlikely to introduce bias.
c) Describes the percentage lost to follow-up, but no description of whom lost to follow-up.
d) No statement
We will report the 'Risk of bias' table as part of the table 'Characteristics of included studies' and present a 'Risk of bias summary' figure, which will detail all of the judgements made for all included studies in the review.

Measures of treatment effect

Different study designs will provide different information about the adverse events being studied. Case reports or case series are likely to report more detailed information on duration, dose and other medications prior to the adverse event. RCTs will provide a quantitative estimate (relative risk, incidence of AE and cardiac

arrest). The combination of all these data will help to build up a complete PRIS picture.

As the primary outcome is the report of PRIS or cardiac arrest, these events will be reported as number of events for all study designs. In case series or case reports, the events will be qualitatively reported and will be the subject of a sensitive analysis.

We will report odds ratio (OR) and 95% confidence interval (CIs). Adverse events will be reported as 'events' and also as number of patients; we will report events using risk difference when a comparison group is present. For an events -rate below 1%, the Peto one step odds ratio will be used as it is the least biased and most powerful method, also providing the best confidence interval coverage.

Unit of analysis issues

We will address any issues related to the unit of analysis as recommended in the Cochrane Handbook (Higgins 2011). In the case of cross-over trials, we will include only the first arm (we will not consider any event that happened after the cross-over).

Dealing with missing data

We will attempt to obtain any relevant missing data from the study authors. We will contact authors, through the contact e-mail provided, to seek individual patient data (for those who developed the primary outcome).

We will perform intention-to-treat (ITT) analyses where trials have not included all randomized participants in their analyses. ITT analysis is a strategy for analysing data in which all participants are included in the group to which they were assigned, regardless of whether they completed the intervention given to the group. The missing participants will be counted as developing the outcome of interest, and not, and those two scenarios will be reported. Performing an ITT analysis is important as it prevents bias caused by loss of participants.

Assessment of heterogeneity

Heterogeneity in this review may arise from differences in comparison groups across trials (that is standard of care may involve different drugs in different trials). In the analysis phase they will be treated as one single 'control' group. A subgroup analysis will then be done if more than two trials are found using the same comparison drug (for example ketamine).

As we will include non-randomized trials in this review, heterogeneity will be greater than if RCTs alone were included. We will analyse studies grouped by design (RCTs, cohort, case control and case series). A meta-analysis will be of RCT data only; the analysis of other study designs will be descriptive. Heterogeneity will be addressed using forest plots, and statistical heterogeneity will be assessed using the I^2 statistic and the following as a rough guide for interpretation: 0% to 40% might not be important, 30% to 60% may represent moderate heterogeneity, 50% to 90% may represent substantial heterogeneity, and 75% to 100% considerable heterogeneity (Higgins 2011).

Assessment of reporting biases

If 10 or more studies are included, we will assess for publication bias using a funnel plot as described by Light and Pillemer (Light 1984). A funnel plot is a graph that plots the measurement effect for each study against the standard error (SE) of the measurement effect precision (1/SE). Asymmetry of the funnel plot will also help in identifying other possible biases, such as reporting bias (publication bias, selective outcome reporting, selective analysis reporting), poor methodological designs or inadequate analysis; but the asymmetry could be due to true heterogeneity, by chance or artefactual.

We will compare the fixed-effect model estimate against the random-effects model to assess the possible presence of small sample bias in the published literature (that is the intervention effect is more beneficial in the smaller studies). In the presence of small sample bias, the random-effects model estimate of the intervention effect is more beneficial than the fixed-effect model estimate (Sterne 2011).

The selective reporting bias will be assessed using the Cochrane 'Risk of bias' tool and it will be reported as described in the section Assessment of risk of bias in included studies.

Data synthesis

The main comparison will be propofol infusion and control. In this review, the control group will be called standard care and it will include the most common sedative agents used for paediatric sedation.

All analyses will be performed using the Review Manager software. For rare events such as PRIS, we will use the Peto one-step odds ratio method. It is the least biased and most powerful method and provides the best confidence interval coverage provided there is no substantial imbalance between treatment and control group sizes within studies, and treatment effects are not exceptionally large (Higgins 2011).

We will pool data from studies that are sufficiently homogenous and with the same study design in order to perform a meta-analysis. Data across different study designs will therefore not be pooled. Analysis will be performed using RevMan 5.1 and forest plots will be produced. Cross-over trials will be analysed as parallel trials, using only the first phase.

Data from case series and case reports will be purely descriptive.

Subgroup analysis and investigation of heterogeneity

We will perform subgroups for the primary outcome (PRIS and cardiac arrest) based on the following.

1. Comparison group: if two or more trials are found using the same comparison group (for example ketamine), these studies will be subgrouped in an attempt to identify any risk increase or reduction for developing the primary outcome.

2. Dosage of propofol: we will subgroup studies using propofol infusion at equal to or less than 4mg/kg/h or more than 4mg/kg/h as this seems to be the dosage cut-off for reporting PRIS in children, but it is not clear what dose of propofol is necessary to cause PRIS. This subgroup analysis will be done as an attempt to measure a dose-effect relationship between propofol and the development of PRIS.

3. Duration of propofol infusion: we will subgroup studies with propofol infusion lasting less than or equal to 12 hours, between 12 and 24 hours and more than 24 hours. This subgroup analysis was chosen as the duration of exposure of propofol required to cause PRIS is unclear. It will be done in an attempt to determine if there is any relationship between duration of infusion of propofol and the development of PRIS.

4. Indication for sedation: if enough trials are found under similar settings, for example sedation for mechanical ventilation in intensive care unit or sedation procedures in the operating room, these studies will be subgrouped in an attempt to identify patient setting (location within the healthcare facility) and indication in an association with the primary outcome.

Sensitivity analysis

We will also do a sensitivity analysis of patients who were lost to follow-up as patients developing the outcome of interest and not developing it as a way to evaluate how robust our findings are.
We will compare the fixed-effect model estimate against the random-effects model to assess the possible presence of small sample bias in the published literature (that is the intervention effect is more beneficial in the smaller studies). In the presence of small sample bias, a sensitivity analysis will be done to identify the possible source of this difference, for example comparing pharmaceutical company and academic generated reports, and we will also analyse the potential positive outcome bias report for small studies.

Summary of findings table

We will summarize the evidence in a 'Summary of findings' table, as recommended by the Cochrane Anaesthesia Group (CARG) and described by Guyatt (Guyatt 2008).
Guyatt uses the GRADE approach. It appraises the quality of a body of evidence based on the extent to which one can be confident that an estimate of effect or association reflects the item being assessed. The quality of a body of evidence considers within study risk of bias (methodologic quality), the directness of the evidence, heterogeneity of the data, precision of effect estimates and risk of publication bias.
We will use GRADEPro system (GRADEpro 2008) to assess the quality of the body of evidence associated with the specific outcomes below.

1. Mortality (or any other serious adverse event as previously defined by WHO).
2. Any adverse event requiring intervention (change in heart rate, blood pressure, respiratory drive or hypoxia).
3. Propofol infusion syndrome.

ACKNOWLEDGEMENTS

We would like to thank John Carlisle (content editor), Nathan Pace (statistical editor), William B. McIlvaine, Mike Sury (peer reviewers) and Janet Wale (consumer editor) for their help and editorial advice during the preparation of this protocol for the systematic review.

REFERENCES

Additional references

Bray 1998
Bray RJ. Propofol infusion syndrome in children. *Paediatric Anaesthesia* 1998;**8**(6):491–9. [MEDLINE: 9836214]

Cremer 2001
Cremer OL, Moons KG, Bouman EA, Kruijswijk JE, Smet AM, Kalkman CJ. Long-term propofol infusion and cardiac failure in adult head-injured patients. *Lancet* 2001;**357**(9250):117–8. [MEDLINE: 11197401]

Eriksen 2006
Eriksen J, Povey HM. A case of suspected non-neurosurgical adult fatal propofol infusion syndrome. *Acta Anaesthesiologica Scandinavica* 2006;**50**(1):117–9. [MEDLINE: 16451160]

GRADEpro 2008
Brozek J, Oxman A, Schünemann H. GRADEpro. Version 3.2 for Windowns, 2008.

Guyatt 2008
Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HFJ. What is "quality of evidence" and why is it important to clinicians. *BMJ* 2008;**336**:995–8.

Hanna 1998
Hanna JP, Ramundo ML. Rhabdomyolysis and hypoxia associated with prolonged propofol infusion in children. *Neurology* 1998;**50**(1):301–3. [MEDLINE: 9443502]

Health Canada 2002
Health Canada. Drugs and Health Products. Update on the use of Propofol. http://www.hc-sc.gc.ca/dhp-mps/medeff/

advisories-avis/prof/_2002/propofol_pediatric_2_nth-ah-eng.php July10, 2002.

Higgins 2011
Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011] . The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Kang 2002
Kang TM. Propofol infusion syndrome in critically ill patients. *The Annals of Pharmacotherapy* 2002;**36**(9): 1453–6. [MEDLINE: 12196066]

Light 1984
Light R, Pillemer D. *Summing Up: The Science of Reviewing Research*. Cambridge: Harvard University Press, 1984.

Loke 2011
Loke Y, Golder S, Vandenbroucke J. Comprehensive evaluations of the adverse effects of drugs: importance of appropriate study selection and data sources. *Therapeutic Advances in Drug Safety* 2011;**2**(2):59–68. [: Loke 2011]

Parke 1992
Parke TJ, Stevens JE, Rice AS, Greenaway CL, Bray RJ, Smith PJ, et al.Metabolic acidosis and fatal myocardial failure after propofol infusion in children: five case reports. *BMJ* 1992;**305**(6854):613–6. [MEDLINE: 1393073]

RevMan 5.1
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). 5.1. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011.

Roberts 2009
Roberts R, Barletta J, Fong J, Schumaker G, Kuper P. Incidence of propofol-related infusion syndrome in critically ill adults: a prospective, multicenter study. *Critical Care* 2009;**13**(5):R169. [MEDLINE: 19874582]

Sterne 2011
Sterne J, Sutton A, Ioannidis J, Terrin N, Jones D, Lau J, et al.Recomendations for examining and interpreting funnel plot asymmetry in meta-analysis of randomised controlled trials. *BMJ* 2011;**342**:343:d4002 doi: 10.1136/. [MEDLINE: 21784880]

US FDA 2001
FDA /Astra Zeneca Pharmaceuticals. Dear health care provider letter. www.fda.gov/ medwatch/safety/2001/ diprivan deardoc.pdf 2001.

Wells 2008
Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_ epidemiology/oxford.htm.

Wolf 2001
Wolf A, Weir P, Segar P, Stone J, Shield J. Impaired fatty acid oxidation in propofol infusion syndrome. *Lancet* 2001; **357**(9256):606–7. [MEDLINE: 11558490]

* *Indicates the major publication for the study*

APPENDICES

Appendix 1. MEDLINE search strategy

1. (propofol* or 2,6-diisopropylphenol* or disoprofol or disoprivan or Diprivan or fresofol or recofol or ivofol or aquafol or rapinovet).mp. or exp Propofol/
2. (child* or p?ediatric* or infant* or teen* or youth*).mp. or exp child/ or infant/ or exp Pediatrics/ or exp Minors/ or Adolescent/
3. 1 and 2

Appendix 2. Ovid EMBASE search strategy

1. (propofol* or 2,6-diisopropylphenol* or disoprofol or disoprivan or diprivan or fresofol or recofol or ivofol or aquafol or rapinovet).mp. or exp PROPOFOL/
2. (child* or p?ediatric* or infant* or teen* or youth*).mp. or exp child/ or infant/ or exp PEDIATRICS/ or exp juvenile/ or **adolescent/**
3. 1 and 2

Appendix 3. CENTRAL search strategy

#1 MeSH descriptor Propofol explode all trees
#2 propofol* or diisopropylphenol* or disoprofol or disoprivan or diprivan or fresofol or recofol or ivofol or aquafol or rapinovet
#3 (#1 OR #2)
#4 MeSH descriptor Child explode all trees
#5 MeSH descriptor Infant explode all trees
#6 MeSH descriptor Minors explode all trees
#7 MeSH descriptor Pediatrics explode all trees
#8 MeSH descriptor Adolescent explode all trees
#9 child* or p?ediatric* or infant* or teen* or youth*
#10 (#4 OR #5 OR #6 OR #7 OR #8 OR #9)
#11 (#3 AND #10)

Appendix 4. Data extraction form

Propofol Infusion for paediatric sedation
CARG 239
Clinical trials
Reviewer: ……..
Study ID (last name/year) …… Date……..
Title……
Language:…….. Country……
Inclusion criteria (must fulfil all brackets)
( ) only paediatric patients (1 month to 19 years old)
( ) used propofol as a continuous infusions (mg/kg/h or mcg/kg/min) for more than 60min
( ) propofol used in only one study arm
( ) it is a controlled clinical trial (there is a comparison group)
Definitions:
Cardiac arrest will be defined as significant reduction on cardiac output requiring any of the following: chest compressions; defibrillation, epinephrine boluses or cardiac mechanical support (Extra corporeal life support -ECLS; or ventricular assist device-VAD).
Propofol infusion syndrome (PRIS) will be defined as metabolic acidosis, defined as arterial pH $\leq$ 7.30 along with a serum bicarbonate $\leq$ 18 mg/dL; plus the presence of any signs in any the below categories. A patient was deemed to have experienced a particular manifestation category if they experienced any manifestation within the category (adapted from Roberts 2009)
1) rhabdomyolysis (CPK $\geq$ 10,000 IU/L) or positive serum or urine myoglobin test or positives urinary casts for haemoglobin;
2) hypotension (initiation of a vasopressor agent or increase of $\geq$20% from baseline vasopressor infusion prior to propofol initiation);
3) hepatic transaminitis (increase in the aspartate aminotransferase and/or alanine aminotransferase $\geq$ 3 times above baseline);
4) hypertriglyceridaemia (serum triglyceride concentration $\geq$ 400 mg/dL);
5) hypoxia (partial pressure of arterial oxygen $\leq$ 60 mmHg);
6) hyperthermia (temperature $\geq$ 38.3°C);
7) cardiac dysfunction that included asystole, pulseless electrical activity, ventricular fibrillation, sustained ventricular tachycardia of 30 seconds or longer, myocardial failure (ejection fraction $\leq$40%), or bradycardia (heart rate $\leq$ 60 bpm)
8) renal failure that included oliguria (urine output $\leq$ 0.5 mL/kg/hr for $\geq$ 6 hours), anuria (urine output $\leq$ 10 mL/hr for $\geq$ 6 hours), elevation in serum creatinine (increase of $\geq$ 1 mg/dL from baseline), or hyperkalaemia (serum K+ $\geq$ 6 mg/dL (excluding other known causes or haemolyzed specimens)

Study participants inclusion criteria
……

Study participant exclusion criteria
……

1) Patient demographics:
1.1) Number of patients enrolled…… (……. Females)
1.2) Age (average) …… from…….. (months/years) to …….. (months/years)
1.3) Patient location: ( ) intensive care
( ) diagnostic imaging (MRI/ CT/other ……..)
( ) operating room due ……
( ) other ……..
1.4) Patients background (reason for hospital admission)
……
1.5) primary outcome being measured……...
2) Study drug: PROPOFOL # patients on this group: ……
Age (average/SD+ min/max) ……(years/months)
Males/Females……
Are any signs of organ dysfunction prior to propofol infusion? ( ) Yes ( ) No
( ) cardiac : on inotropic support; previous arrhythmia; previous reduced ejection fraction by ECHO;
( ) hepatic: elevated liver enzymes
( ) hypoxia (pao2 <60)
( ) renal: oliguria, anuria or elevation of serum creatinine
2.1) Propofol used as an infusion at average dose of ……. mg/kg/h (or) mcg/kg/min minimum of …… and maximum of ……. mg/kg/h (or) mcg/kg/min
2.2) Duration of propofol infusion …….. hours (min ……./ max…….)
2.3) Other drugs used in association with propofol (name and dose)
a) prior to propofol infusion……
b) during propofol infusion ……
c) after propofol infusion was stopped ……
2.4) Any report of PRIS or cardiac arrest (events of interest) on this group?
( ) yes ( ) no
How were these reports collected? ……
( ) spontaneous report ( ) active surveillance ( ) not clear
2.5) How many events of interest in total? …… In how many unique patients ……
2.6) Report the AE found (kind of event, methods of monitoring and diagnosing the AE, treatment and outcome)
……

……

……

2.7) Possible confounding factors or biases as per:
Authors:
……

SR extractor:
……

……

2.8) If PRIS or cardiac arrest were identified, what is it relationship with propofol infusion? (time from initiation/ending of infusion to the event) ……
……
2.9) How long was the follow-up for?
……
2.10) Number of patients that lost follow up and reasons for it
Propofol group:
……..

……

2.11) Any other events of interest reported, not fulfilling the PRIS inclusion criteria?
……

2.12) Were any patients excluded from the adverse effect analysis? Why?
……

3) Comparison group: # patients on this group……
Age (average/SD+ min/max) ……(years/months)
Males/Females:……..
Are any signs of organ dysfunction prior to initiation of sedation infusion? ( ) Yes ( ) No
( ) cardiac : on inotropic support; previous arrhythmia; previous reduced ejection fraction by ECHO;
( ) hepatic: elevated liver enzymes
( ) hypoxia (pao2 <60)
( ) renal: oliguria, anuria or elevation of serum creatinine
3.1) drug (s) dosage
drug (1) ……dose…… duration……
drug (2) ……dose……..duration ……
drug (3) …….. dose…….. duration ……
drug (4) ……dose …….. duration ……
3.2) Any serious adverse event (PRIS or cardiac arrest) reported on this group?
( ) yes ( ) no
How were these reports collected? ……
( ) spontaneous report ( ) active surveillance ( ) not clear
3.4) How many AEs in total? …… In how many unique patients ……
3.5) Report the AE found (kind of event, methods of monitoring and diagnosing the AE, treatment and outcome)
……

3.6) Possible confounding factors or biases as per:
Authors:
……

SR extractor:
……

……

3.7) If PRIS or cardiac arrest was identified, what is it relationship the drugs been used? (time from initiation/ending of drugs been used to the event)
……

3.8) How long was the follow-up for?
……
3.9) Number of patients that lost follow up and reasons for it
Comparison group:
……..
……

3.10) Any other events of interest reported, not fulfilling the PRIS inclusion criteria?
……

3.11) Were any patients excluded from the adverse effect analysis? Why?
……

4) Study bias:
4.1) Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence.? Low ( ) High ( ) Unclear ( )
4.2) Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment. Low ( ) High ( ) Unclear ( )
4.3) Performance bias due to knowledge of the allocated interventions by participants and personnel during the study. Low ( ) High ( ) Unclear ( )
4.4) Detection bias due to knowledge of the allocated interventions by outcome assessors
Low ( ) High ( ) Unclear ( )
4.5) Attrition bias due to amount, nature or handling of incomplete outcome data.
Low ( ) High ( ) Unclear ( )
4.6) Reporting bias due to selective outcome reporting.
Low ( ) High ( ) Unclear ( )
Other biases:
4.7) Bias due to problems not covered elsewhere in the table. Low ( ) High ( ) Unclear ( )
Why? ……
Overall risk of bias: Low ( ) High ( ) Unclear ( )
Propofol Infusion for paediatric sedation
CARG 239
No- randomized studies
Reviewer: ……
Study ID (last name/year) …… Date……..
Title……
Language:…….. Country……..
Kind of study
( ) case report (number of patients……..)
( ) case-control
( ) cohort study
( ) other ……..
Propofol:
Patient's characteristics:
Age (average/SD)..... years/months
Male/Female.....
Reason for hospital admission……
Patient location: ( ) intensive care
( ) diagnostic imaging (MRI/ CT/other ……)

( ) operating room due ……
( ) other ……..
Are any signs of organ dysfunction prior to propofol infusion? ( ) Yes ( ) No
( ) cardiac : on inotropic support; previous arrhythmia; previous reduced ejection fraction by ECHO;
( ) hepatic: elevated liver enzymes
( ) hypoxia (pao2 <60)
( ) renal: oliguria, anuria or elevation of serum creatinine
Propofol dose and duration of infusion:
……
Other drugs used (name and dosage)
- Prior to propofol infusion
……..
……

- During propofol infusion:
……

……

- After propofol infusion was stopped
……

……

Adverse events reported (kind of event, duration, treatment used and outcome)
……

……

……

……

……

……

3.6) Possible confounding factors or biases:
By the study authors:
……

……

By data extractor:
……

……

3.7) If PRIS or cardiac arrest was identified, what is it relationship the drugs been used? (time from initiation/ending of drugs been used to the event)
……

Comparison group (if applicable)

Drug(s) used for comparison (name and dosage)
……

……

……

Number of patients……
Male/female……
Reason for hospital admission……
Patient location: ( ) intensive care
( ) diagnostic imaging (MRI/ CT/other ……)
( ) operating room due ……
( ) other ……..
Personal delivering the intervention: Pediatric trained: YES ( ) NO ( ) ( ) Not clear
( ) anaesthesiologist
( ) intensivists
( ) emergency physician
( ) other……
How was the control selected?
……..
Were the patients similar between intervention and control? ……
……

Adverse events reported (kind of event, duration, treatment used and outcome)
……

……

……

……

……

……

Possible confounding factors:
By the study authors:
……

……

By the reviewer:
……

If PRIS or cardiac arrest was identified, what is it relationship the drugs been used? (time from initiation/ending of drugs been used to the event)
……

……

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE

CASE CONTROL STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

1) Is the case definition adequate?

a) yes, with independent validation

b) yes, e.g. record linkage or based on self reports

c) no description

2) Representativeness of the cases

a) consecutive or obviously representative series of cases

b) potential for selection biases or not stated

3) Selection of Controls

a) community controls

b) hospital controls

c) no description

4) Definition of Controls

a) no history of disease (endpoint)

b) no description of source

Comparability

1) Comparability of cases and controls on the basis of the design or analysis

a) study controls for _____ (Select the most important factor.)

b) study controls for any additional factor

(This criteria could be modified to indicate specific control for a second important factor.)

Exposure

1) Ascertainment of exposure

a) secure record (e.g. surgical records)

b) structured interview where blind to case/control status

c) interview not blinded to case/control status

d) written self report or medical record only

e) no description

2) Same method of ascertainment for cases and controls

a) yes

b) no

3) Non-Response rate

a) same rate for both groups

b) non respondents described

c) rate different and no designation

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE

COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort

a) truly representative of the average _____ (describe) in the community

b) somewhat representative of the average _____ in the community

c) selected group of users e.g. nurses, volunteers

d) no description of the derivation of the cohort

2) Selection of the non exposed cohort

a) drawn from the same community as the exposed cohort

b) drawn from a different source

c) no description of the derivation of the non exposed cohort

3) Ascertainment of exposure

a) secure record (e.g. surgical records)
b) structured interview
c) written self report
d) no description
4) Demonstration that outcome of interest was not present at start of study
a) yes
b) no
Comparability
1) Comparability of cohorts on the basis of the design or analysis
a) study controls for _____ (select the most important factor)
b) study controls for any additional factor (This criteria could be modified to indicate specific control for a second important factor.)
Outcome
1) Assessment of outcome
a) independent blind assessment
b) record linkage
c) self report
d) no description
2) Was follow-up long enough for outcomes to occur
a) yes (select an adequate follow up period for outcome of interest)
b) no
3) Adequacy of follow up of cohorts
a) complete follow up - all subjects accounted for
b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost)
c) follow up rate < ____% (select an adequate %) and no description of those lost
d) no statement

HISTORY

Protocol first published: Issue 4, 2012

CONTRIBUTIONS OF AUTHORS

Conceiving the review: Liliane Zorzela (LZ), Salima Punja (SP), Lisa Hartling (LH), Ari Joffe (AJ), Yoon Loke, (YL) Sunita Vohra (SV)

Co-ordinating the review: LZ, SP, LH, AJ, YL, SV

Undertaking manual searches: LZ

Screening search results: LZ, SP

Organizing retrieval of papers: LZ

Screening retrieved papers against inclusion criteria: LZ, SP

Appraising quality of papers: LZ, SP

Abstracting data from papers: LZ, SP

Writing to authors of papers for additional information: LZ

Providing additional data about papers: LZ and Ben Vandermeer

Obtaining and screening data on unpublished studies: LZ, SP

Data management for the review: LZ

Entering data into Review Manager (RevMan 5.1): LZ

RevMan statistical data: LZ and Ben Vandermeer (BV)

Other statistical analysis not using RevMan: BV, LZ

Interpretation of data: LZ, SP, LH, AJ, YL, SV

Statistical inferences: LZ, SP, LH, AJ, YL, SV

Writing the review: LZ, SP, LH, v, YL, SV

Securing funding for the review: none

Performing previous work that was the foundation of the present study: LH, YL, SV

Guarantor for the review (one author): SV

Person responsible for reading and checking review before submission: LZ, SP, LH, AJ, YL, SV

DECLARATIONS OF INTEREST

None known

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- New Source of support, Not specified.