Standards of Evidence for Research on 'What Works' in the Management of MAM

Patrick Webb, Ph.D.
Professor, Friedman School of Nutrition Science and Policy, Tufts University, Boston, USA

What does this FAQ cover?

This paper presents an overview of classic systematic reviews and statistical meta-analyses which form the backbone of scientific assessments of evidence quality around public health interventions. It explains how these approaches contribute to current knowledge on the effectiveness of MAM interventions. Not covered here are other forms of evidence assessment, such as 'realist reviews'[1] or 'evaluation platform designs'[2] which offer alternative approaches and standards of rigor.

Background

Acknowledgement by the international community of the importance of effective management of Moderate Acute Malnutrition (MAM) has been growing rapidly. The 2008 Lancet series on maternal and child undernutrition, for example, only dealt with Severe Acute Malnutrition (SAM), and did not mention MAM once in the paper on "what works" in terms of evidence-based interventions to tackle undernutrition.[3] Only five years later, the second Lancet series not only included significant attention to MAM as a major problem in its own right, but included the management of MAM among the 10 proven interventions that should be scaled up quickly in countries with the highest burdens of malnutrition.[4] If that be the case, why do we need more research on this topic?

The answer is two-fold: on the one hand, there remain a large number of evidence gaps that require priority attention for policies and programmes to be appropriately designed and implemented. For instance, the International Symposium on Understanding Moderate Malnutrition in Children held in Vienna in 2014 noted that "more evidence on effective programmatic approaches to manage moderate wasting is needed."[5] The Global Nutrition Cluster's MAM Task Force concurs, suggesting that the effectiveness (and cost) of various MAM interventions needs much more attention: "the amount of change that can be achieved and the conditions under which [different approaches are] appropriate to improve nutrition outcomes, requires further research."[6] On the other hand, the quality of evidence used to support policymaking and programming must be improved. While the number of studies dealing with MAM management expands, the rigor of such studies often leaves much to be desired. As the authors of the 2013 Lancet series concluded, "high-quality programmatic research" is needed to help improve the design, implementation and outcome of nutrition interventions where wasting is a prime concern.[4]

Why do we need high standards for research on MAM?

Global norms, standards and protocols for action in the public health and nutrition arenas are typically set by supra-national organizations (such as the World Health Organization) that follow a defined sequence of expert consultations, review of published literature, interaction with key informants, and consolidation of evidence of best practice. Numerous targeted nutrition actions, such as vitamin A supplementation, salt iodization, iron/folate supplementation, and even community-based treatment of SAM, have gained the status of "evidence based interventions" as a result of an accumulation of statistically significant findings from multiple rigorously designed and replicated studies. When carefully reviewed, such intervention studies offer consistent common conclusions which hold up when individual data are pooled (for analyses that rule out biases and confirm results). They form the basis for common scientific and policy agreement on 'what works'. They provide the content for operational guidance based on 'what works' but not yet enough on 'how to' make it work.

We don't yet have that for MAM?

No, not yet. Where empirical evidence on the effectiveness of approaches for the prevention of acute malnutrition and the management of MAM is concerned, such a body of high quality evidence is still lacking. A recent review of the literature on MAM concluded that while the need to act is clear, "the evidence base is sparse. Evidence is particularly lacking for prevention of acute malnutrition."[7] Similarly, the Emergency Nutrition Network argues that where performance of implementation and best practice is concerned, the profession faces a "dearth of published evidence from a range of settings."[8] This extends into the products used to treat MAM, since WHO highlights that "currently there are no evidence-informed recommendations on the composition of supplementary foods specially designed for the management of children with moderate acute malnutrition."[9] As a result, most guidelines and standard setting in relation to MAM has evolved "less through reliance on randomized control trials than on the sharing of lessons learned codified by specially-convened scientific advisory committees and disseminated as operational or technical guidelines by WHO and other UN bodies."[10]

Why isn't that sufficient if 'we know it works'?

Governments around the world, including in countries carrying a high burden of MAM or where humanitarian emergencies are frequent occurrences, look to global standards when they approve national protocols for intervention in public health within their frontiers. All agencies working in such countries should abide by national rules and laws. While there does exist a growing body of information on MAM programming experiences (anecdotal, based on project-reporting, or derived from small studies), the quality is not yet sufficient to pass the bar on scientific rigor, and hence feed into standard-setting. Many existing studies have a lot to offer, but if excluded from a formal systematic review (which does not only include randomized control trials) their potentially valuable contributions on MAM evidence may be largely hidden to the operational community because their findings are not being reported as acceptable evidence. In this case, 'not acceptable' often means that study findings were not included in a systematic review (due to incomparable outcome measures, study design, population, etc. as outlined below), which serves as the foundation for most critiques of evidence quality.

What is a systematic review?

The term systematic review has a formal definition and protocol for implementation; in other words, it is not simply a search through the literature resulting in a narrative summary. Originating

in the medical and life sciences disciplines to help add rigor to assessments of the clinical efficacy of various interventions or drugs, systematic reviews seek to a) identify potential biases in the selection of publications/data/findings explicit and transparent, and b) minimize such biases to the extent possible in order to make conclusions and policy recommendations on what works based on evidence that is deemed to be credible and 'the best there is'. The approach involves a comprehensive search strategy (with carefully defined 'search terms') which has the goal of identifying, appraising, and synthesizing all relevant studies on a narrowly-formulated topic. The systematic search throws up large numbers of paper abstracts and titles that are screened for eligibility by more than one researcher. Full texts of apparently eligible studies are similarly assessed, with only those meeting all inclusion criteria retained for full review. Often, the authors of retained studies are contacted by the reviewers with a request for any missing information and for access to the data (for later meta-analysis). The individual studies are then graded by more than one researcher for 'quality' characteristics; that is, the risk of bias inherent in the study design and implementation of each piece of research, as well as adequacy of sample size, appropriate choice of indicators, appropriate subject inclusion criteria, and more. In many (but not all) cases, systematic reviews also include a meta-analysis, which involves pooling data or computed results from included studies to synthesize outcomes into a single quantitative estimate or summary effect size.

What is involved in a meta-analysis?

Meta-analysis is a way to combine and then compare the results of separate (independent) studies and synthesizing conclusions that evaluate the overall effects being researched. Meta-regression is a tool used in meta-analysis to examine the impact of various variables on each study's effect size using regression-based techniques. To allow a study inclusion in meta-analysis requires good documentation of all elements of the database (the metadata) so that requests for sharing of data can be responded to favourably.

Why is this approach 'better' than any other to determine quality of evidence?

There are many standards of evidence used in science, and different disciplines often use a variety of methodological approaches, analytical techniques, and their own types of evidence when formulating guidance and policy. In the public health domain, which plays a key role in defining the appropriate products and protocols to use in saving lives and managing malnutrition, the standards of evidence commonly deemed acceptable by ministries of health, the World Health Organization, and non-governmental organizations such as Doctors Without Borders (MSF) lie at the higher end of the rigor scale. While the so-called gold standard for experimental design in public health is the double blind, placebo controlled randomized control trial (RCTs for short), this type of study design is often not possible or even appropriate for answering questions about interventions in humanitarian settings or non-emergency contexts where, say, child wasting is widely prevalent. As such, a range of research designs are published relating to the management of MAM that offer varying degrees of experimental rigor. For example, the 2014 paper by Langendorf et al.[11] on uses of cash and/or food in management of wasting in Niger, was reviewed by Kerac and Seal [7] who noted that despite the large number of children included in the study it was not a randomized trial, and since it was "more akin to an observational study… findings could equally be due to inter-site differences resulting in bias or unmeasured confounding". That said, the reviewers noted that the study still had many methodological strengths that should not be ignored, including "rigorous and detailed reporting and analysis." In other words, a range of designs can (and should) be considered in any review of evidence, and their inclusion or exclusion should be framed by a clear set of criteria.

Who sets those criteria for inclusion or exclusion of evidence?

Users of evidence should determine what is credible and acceptable according to the use to which it will be applied. Even small qualitative studies or focus group interviews can generate valuable information to a project implementer about local context, cultural constraints or political concerns. However, when seeking to generalize findings across settings and promote a consensus position among scientists on evidence-based recommendations, minimum criteria are needed to establish credible norms. The principles of systematic reviews have evolved over time, but most professionals refer for guidance to the PRISMA standards[12], the protocols defined by the Cochrane Collaboration[12], or the Campbell Collaboration.[13] Each of these has carefully defined the approach to evaluating research quality and the rationale for criteria used when excluding studies from a review or meta-analysis of data.

When might my studies on MAM be excluded from a systematic review?

There are several types of criteria used to screen papers into or out of a systematic review: first, the study design. Most reviews will automatically include randomised controlled trials (RCTs), including cluster-randomised controlled trials (cRCTs), and non-randomised controlled clinical trials (CCTs). Some, such as the Cochrane review by Lazzarini et al. of specially formulated foods for treating children with MAM, include controlled before-and-after studies (CBAs), but only if they met the additional criteria of a) contemporaneous data collection (were data in the experimental and control sites collected in the same time frame) and b) inclusion of appropriate control sites (that are comparable to the intervention site in terms of setting and population characteristics).[15] However, even if the research design appears to be sound, studies can be excluded if they do not offer sufficient detail on the interventions considered, on potential mitigating or confounding factors, if they do not have an appropriate counterfactual as part of the study, or if the sample size was too small to allow for inference of statistical significance. For example, a review of nutrition surveys carried out in Ethiopia between 2003 and 2008 (more than 340 surveys), concluded that too few surveys provided sufficient detail on the representativeness of the study population, contextual concerns, baseline conditions, trend data, or potential confounders (such as previous or contemporaneous interventions), nor were there always plausible links made between programme components and expected outcomes.[10] In other words, many studies are excluded because insufficient information was provided by the authors even when the study design was of a high quality.

Why would a lack of information disqualify a good study?

If the reviewers cannot obtain full answers about potential sources of bias (and hence the quality of the study) from the paper itself or from the authors, they cannot have full confidence in the interpretation of findings reported. Thus, information gaps downgrade the perception of study quality. A bias can be defined as a systematic error or deviation from the truth in results or inferences.[13] There are several key categories of bias risk that are commonly used in screening study quality:

i) Selection bias, which refers to systematic differences that may exist between baseline characteristics of groups that are compared, but not taken into account in the analysis.

ii) Avoiding bias in participant access to interventions. This requires researchers to share adequate information for reviewers to assess any bias in treatment (i.e. those who were supposed to get it did get it, there was no systematic exclusion of certain categories of population, etc.).

iii) Complete outcome data provided (no missing findings). Many studies report positive impacts without presenting full data on, for example, default rates, people lost to follow-up, mortality or relapse rates. Even if provided, information is needed on the distribution of such outcomes across intervention groups to be able to assess any introduced bias by sub-group.
iv) Selective reporting. It is important to avoid 'cherry-picking' results to the exclusion of findings that do not support the study's hypotheses. Reviewers consider if pre-specified primary outcomes are all reported and if analysis methods proposed in a research protocol were actually applied.
v) Additional sources of risk include ensuring no or minimal leakage of intervention effects into control groups (called protection against contamination); baseline confounders (systematic differences between intervention and control groups not specified); and blinding of outcome assessment (if the study involved an intervention trial, reviewers assess whether the outcomes could have been influenced by lack of blinding, that is, the researchers themselves do not know which group receive the intervention and which does not). The latter is not often relevant for studies of MAM treatment, but could play a role in a randomized trial for prevention.

Each of these elements is graded as having 'low risk of bias' (hence a study element of high quality), 'high risk', or 'unclear risk' (with researchers often referring to the GRADE working group recommendations on how to determine the quality of evidence).[16] GRADE stands for Grading, Recommendation, Assessment, Development and Evaluation – a process commonly used by the WHO in their internal review process that leads to formal protocols or guidelines on practices. Reviewers will also consider the risks of bias inherent in the study design used (relative to primary questions posed), whether an assessment of sample population heterogeneity was conducted, if sensitivity analysis was pursued on analysis specifications and results, and whether appropriate analytical methods were applied to the data.[17]

This seems like a high bar for inclusion; are many studies excluded from systematic reviews for one or other of the above reasons?

Most systematic reviews present a flow diagram that describes how many titles/abstracts were captured in the database and literature searches; these usually number in the tens of thousands. The number of full texts reviewed and excluded is also presented, resulting in a much small sub-set of studies which did meet all or most of the inclusion and quality criteria laid out above. In other words, the majority of studies identified as being potentially useful for the review are not in the end included in that review. For example, Lazzerini et al. conducted their literature search for studies dealing with specially-formulated foods used in the treatment of MAM.[15] They searched 13 electronic databases (each clearly identified) as well as consulting with key informants who are familiar with the grey (unpublished) literature on the topic. Their search generated 8,900 references for possible inclusion in the review, but of these only 298 represented studies appropriate to the analysis. The full text of these articles was read by multiple researchers who found that 287 had to be excluded for a variety of reasons (which are usually tabulated as an appendix to ensure transparency of reporting on the review process). Thus, only 11 studies (representing 8 trials) were ultimately used for the systematic review analysis and final Cochrane report. Publication bias is also a major concern of systematic reviews. It is possible that studies finding 'no effect' are omitted from publication (do not pass peer review for academic journals), and therefore do not show up as part of systematic reviews. This would clearly bias conclusions in favour of positive results where they exist.

What do existing systematic reviews on MAM tell us?

Between 2010 and 2014, there were roughly a dozen systematic reviews (some as yet unpublished) relating to the management of wasting. They did not all use the same exclusion criteria, databases, cut-offs for range of years of publication considered, or focus on the same types of research designs or interventions. That said, there are a lot of common conclusions across those studies, including:

1. There exists moderate-to-high quality evidence that food supplements of various kinds are effective in the treatment of SAM and MAM (where effective is defined as meeting minimum SPHERE standards for exit from treatment). In other words, existing products and protocols used for treating wasting are 'known to work'.
2. Lipid-based ready-to-use foods (RUFs) tend to generate faster and higher weight gain than grain-based fortified blended foods (FBFs). The clinical significance of this remains to be determined in terms of sustainability of recovery (relapse rate post-exit from treatment) and type of weight gain (lean vs fat).
   Differences in height gain are not usually significant (studies are often too short and not powered to detect statistical differences in height).
3. There is little evidence so far of a statistically significant difference between types of foods used in treatment regimes in terms of mortality outcomes, default rates, or progression from MAM to SAM.

In other words, current treatments that follow existing protocols for ensuring adequate programme coverage, inclusion of those in need of treatment, and necessary quality of implementation do work in diverse settings where wasting is a serious threat to child mortality.

What do existing systematic reviews relating to MAM *not* tell us?

What the current state of evidence does not allow us to conclude with any confidence is the cost-effectiveness of a range of approaches, the potential contribution of home-based diets to improving outcomes, the effectiveness of existing products and approaches to the prevention of MAM (i.e. containing mild wasting such that it does not evolve into MAM), incidence rates versus prevalence of MAM, seasonality effects of wasting and how that affects recovery rates and relapses, sub-group analyses to explore effectiveness controlling for initial conditions, health history, co-morbidities, etc., the role of intensive behavior change communication and/or provision of cash/vouchers with or without food in the management of MAM, cognitive versus physiological outcomes of treatment and recovery (i.e. going beyond anthropometry as an outcome metric), or the relative dose-response of food treatments containing various levels of animal source protein, specific amino acids, different forms of micronutrients, or probiotics. There is also very little about the nature of 'recovery' being children meeting a weight gained per day criterion that allows them to exit a treatment regimen when they have crossed a physiological threshold. The extent to which any immune system damage, metabolism or any cognitive impacts have been repaired at that point remains unknown, but of concern given that up to one third of 'recovered' children may relapse.

Indeed, the state of evidence is generally weak and has many gaps that urgently need to be filled. Lazzerini et al. point to "limitations in the completeness of evidence and its generalizability."[15] Similarly, Lenters et al., in their systematic review and meta-analysis of treatments of SAM and MAM in low- and middle-income countries concluded that evidence-based generalizations must still be treated with caution due to "gaps in our ability to estimate effectiveness."[18] As a result, the WHO's Briefing Paper relating to the wasting goal of the World Health Assembly *Global Nutrition*

Targets 2025, which is to reduce and then maintain levels of child wasting to below 5%, states that the world still needs to "develop better understanding of the major causal factors of wasting", and strengthen the methods used "to assess population levels of wasting for accurately assessing burden of acute malnutrition for service planning, design and monitoring"- in other words, strengthen the evidence base for action at all levels.[19]

What should studies include if they are to be incorporated into future systematic reviews?

First, consult the (high quality) literature and research experts to determine an appropriate study design that will convincingly answer the question being posed, and that does not unknowingly simply duplicate other ongoing studies.[20] Second, while every study faces trade-offs in terms of feasibility, cost, desirability and rigor, poorly conceptualized or badly conducted studies should be anathema to the professional nutrition community. The idea that 'some data must be better than no data' is false where the data collected are either not used or useless and when scare resources could have been used in other productive ways. Third, carefully consider (and avoid) the list of design criteria that would make a study ineligible for inclusion in a formal systematic review. Fourth, ensure that you understand the list of potential biases or information gaps that would result in an 'unclear' or 'high risk' of bias grading for your study; aim for 'low risk' assessment and provide as much detail not just on study design as possible, but on the characteristics of the population, setting and interventions considered. Less is *not* more when it comes to convincing reviewers that you tried to take possible biases into account. Fifth, be transparent in acknowledging study limitations and do not over-interpret the findings that can be supported by the data. Sixth, ensure that if a study seeks to compare a health intervention or products that they are compare on equal terms; that is, apples are not compared with oranges. Too many studies have sought to argue that one approach or treatment is superior (or not inferior) to another, when the amount of product used is different, packaging and delivery are different, messaging and programme interaction with beneficiaries is different, duration of treatment is different, etc., which does not provide an appropriate basis for like-with-like comparison. Finally, make strong recommendations of operational and policy significance where they are warranted.

Conclusions

The bar is set high where research quality is concerned. Nevertheless, more and more organizations are investing in generating stronger evidence of what works (and what doesn't) where the management of MAM is concerned. As Kerac and Seal put it, the publication of increasingly high quality studies on MAM treatment and prevention feeds directly into improved practice on the ground; as such researchers and practitioners alike should continue to reach for ever higher standards of research rigor so that future dialogue on best practice "can be more scientific and focused on evidence rather than ideologies."[7]

Acknowledgements

The author acknowledges the valuable comments of Robert Houser, Devika Suri, Hedwig DeConinck, Henrik Friis, Christian Fabiansen, Bernardette Cichon and Nicky Dent on earlier drafts of this paper. Some of the work on which this paper is based was generously supported by USAID's Office of Food for Peace (OFFP) under its Food Aid Quality Review grant to Tufts University. The

insightful contributions of Dina Esposito, Judy Canahuati and Rufino Perez at OFFP are acknowledged in that context.

References

1. Pawson R, Tilley N. 2004. 'Realist evaluation', in Matthieson S (ed.) *Encyclopaedia of Evaluation* Newbury Park, UK: Sage. At: http://www.communitymatters.com.au/RE_chapter.pdf.
2. Victora C, Black R, Boerma J, Bryce J. 2011. Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations. *Lancet*, 377: 85–95. DOI:10.1016/S0140-6736(10)60810-0.
3. Bhutta Z, Ahmed T, Black R, Cousens S, Dewey K, Giugliani E, Haider B, Kirkwood B, Morris S, Sachdev H, Shekar M for the Maternal and Child Undernutrition Study Group. 2008, *Lancet* 371: 417–40, DOI:10.1016/S0140-6736(07)61693-6.
4. Bhutta Z, Das J, Rizvi A, Gaffey M, Walker N, Horton S, Webb P, Lartey A, Black R. for the Nutrition Interventions Review Group and the Maternal and Child Nutrition Study Group. 2013. Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost? *Lancet,* http://dx.doi.org/10.1016/S0140-6736(13)60996-4.
5. IAEA (International Atomic Energy Agency). 2014. International Symposium on Understanding Moderate Malnutrition in Children for Effective Interventions. Vienna, Austria, 26-29 May, 2014.
6. GCN/MTF (Global Nutrition Cluster/MAM Task Force). 2014. *Moderate Acute Malnutrition: A Decision Tool For Emergencies*. New York, NY.
7. Kerac M, Seal A. 2014. Preventing Acute Malnutrition in Young Children: Improving the Evidence for Current and Future Practice. *PLOS Medicine*, 11 (9): e1001715.
8. Emergency Nutrition Network. 2012. *Mid Upper Arm Circumference and Weight-for-Height Z-score as Indicators of Severe Acute Malnutrition: a Consultation of Operational Agencies and Academic Specialists to Understand the Evidence, Identify Knowledge Gaps and to Inform Operational Guidance*, London, UK.
9. WHO (World Health Organization). 2012. *Technical note: Supplementary foods for the management of moderate acute malnutrition in infants and children 6–59 months of age*. Geneva, Switzerland.
10. Webb P, Boyd E, de Pee S, Lenters L, Bloem M, Schultink W. 2014. Nutrition in emergencies: Do we know what works? *Food Policy* 49: 33–40.
11. Langendorf C, Roederer T, de Pee S, Brown D, Doyon S, et al. (2014) Preventing Acute Malnutrition among Young Children in Crises: A Prospective Intervention Study in Niger. *PLoS Med* 11(9): e1001714. doi:10.1371/journal.pmed.1001714.
12. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). 2014. http://www.prisma-statement.org/(last accessed Sept 10, 2014).
13. The Cochrane Collaboration. 2014. http://www.cochrane.org/cochrane-reviews (last accessed October 24, 2014).
14. The Campbell Collaboration. 2014. http://www.campbellcollaboration.org/ (last accessed September 10, 2014).
15. Lazzerini M, Rubert L, Pani P. Specially formulated foods for treating children with moderate acute malnutrition in low- and middle-income countries. Cochrane Database of Systematic Reviews 2013, Issue 6. Art. No.: CD009584. DOI: 10.1002/14651858.CD009584.pub2.
16. GRADE http://www.gradeworkinggroup.org/
17. Guyatt G, Oxman A, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schunemann H. GRADE guidelines 2. Framing the question and deciding on important outcomes. *Jou Clin Epidemiol* 2011; **64**:395–400.
18. Lenters, L., K. Wazny, P. Webb, T. Ahmed and Z. Bhutta. 2013. Treatment of Severe and Moderate Acute Malnutrition in Low- and Middle-Income Settings: A Systematic Review, Meta-analysis and Delphi Process. *BMC Public Health* 13 (Suppl. 3): S23. http://www.biomedcentral.com/1471-2458/13/S3/S23
19. WHO. 2014. *Global Nutrition Targets 2025: reduce and maintain childhood wasting to less than 5%*. Technical Brief on WHA Targets No. 6. Geneva, Switzerland. (September 2014 draft). Mimeo.
20. REFINE (Research Engagement on Food Innovation for Nutritional Effectiveness). http://refinenutrition.org/about-us.htm (last accessed September 15, 2014).

Appendix 1: PRISMA Checklist (dated 2009)

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	
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.	

Section/topic	#	Checklist item	page #
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.	
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.	
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.	

Source: Moher D, Liberati A, Tetzlaff J, Altman DG, the PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(6): e1000097. doi:10.1371/journal.pmed1000097

.