NCBI Bookshelf. A service of the National Library of Medicine, National Institutes of Health.

Smith PG, Morrow RH, Ross DA, editors. Field Trials of Health Interventions: A Toolbox. 3rd edition. Oxford (UK): OUP Oxford; 2015 Jun 1.

Chapter 3 Reviewing the literature

1. Introduction to reviewing the literature

Systematic reviews are increasingly recognized as an essential step in health care research. They are a method designed to produce an objective, unbiased, up-to-date summary of available evidence. In this chapter, an outline is given of the methods used to systematically review the medical literature and to assess the risk of bias in the identified studies. Results from a systematic review may be summarized as a narrative or a summary estimate produced from a quantitative meta-analysis. In either case, systematic reviews are usually a necessary step in preparing to conduct intervention trials and in setting the results of trials into context.

Before embarking on an intervention trial, it is essential to review what is already known about the questions to be addressed in the trial. The most objective way to do this is to conduct a systematic review of all similar studies that have been published previously on the topic. Such a review should enable an assessment to be made of whether (1) sufficient evidence for the effect of the intervention already exists, or (2) there is a clear scientific rationale for an effect of the intervention, but there is insufficient evidence that the intervention works in practice, or (3) there is an insufficient rationale for an intervention effect. If the review of the published evidence supports (1) or (3), then there may be little justification for conducting a (further) trial. Furthermore, funding agencies may require a systematic review to provide evidence that a new trial is justified, and some journals (including, for example, *the Lancet* (Clark and Horton, 2010)) now require authors to include, in papers reporting the results of a trial, a summary of the findings from a recent systematic review, in order to put their trial into context, or to report their own up-to-date systematic review. For example, before proposing a trial of a new school-based behaviour change intervention to reduce the incidence of HIV infection, it would be essential to review the literature on the effectiveness of previous school-based interventions, and also to review the literature on the rationale underpinning the mechanism by which such an intervention might be expected to be effective.

A proposed trial is worthwhile if the conclusions from a systematic search of the literature provide a strong rationale that the proposed intervention will work, but there is currently insufficient evidence to know how effective, if at all, it is likely to be in the target population for the trial. In addition to wasting time and resources, a trial of an intervention which has already been proven effective may be considered unethical, as participants in the control arm would not receive a beneficial intervention, and conducting a further trial may delay scale-up of the intervention to those who would benefit from it.

In this chapter, we describe methods for conducting systematic reviews of epidemiological studies (including observational studies as well as intervention trials) to judge whether a new intervention trial is justified. We also include sections on assessing the risk of bias in studies and on providing a narrative and quantitative summary of the findings.

Systematic reviews are not trivial undertakings, and not all investigators will have the time or resources to conduct the kind of review that we outline in this chapter. Ideally, other investigators will have conducted a recent review, and it will be possible to utilize their findings. For example, an agency such as the World Health Organization (WHO) might have commissioned a review in order to assist them in setting priorities for disease control or to highlight important areas for research. Those planning to conduct a trial might not need to conduct their own systematic review but could build on the previous work. However, even if an investigator is not going to undertake their own review, it is important that they understand how such reviews are conducted and indeed can assess the quality of published systematic reviews. This chapter should facilitate this.

The insights that a systematic review can give to the reviewers on the effects of an intervention and the quality of previous studies are invaluable. It is highly recommended that all those conducting trials participate in at least one systematic review fairly early in their careers!

2. Systematic reviews

Reviewing the literature can be a daunting task. The volume of information available through published papers, or the Internet, is vast and constantly expanding. Given the volume of literature available, an 'ad hoc' review of the literature is subject to substantial biases if only some studies are included, since the studies that are found this way may well not be representative of all the relevant studies. The best way to ensure an objective and unbiased review of the literature is to conduct a review that follows strict guidelines to minimize bias in selecting and interpreting reported studies.

The basic steps in a systematic review are shown in Box 3.1.

In this chapter, we provide a brief overview of each of these steps. Further details are given in published guidelines, such as the *Cochrane handbook for systematic reviews of interventions* (Higgins and Green, 2008) and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (<http://www.prisma-statement.org>) (Liberati et al., 2009), and books on systematic reviews in health research (Egger et al., 2001, Glasziou, 2001, Khan, 2003).

2.1. Defining the question

The first step in a systematic review is to define the research question. A structured approach for framing the question is useful—the PICOS approach (Population; Interventions (or Exposure); Comparison; Outcomes; Study design) (Higgins and Green, 2008) is used by both Cochrane and PRISMA.

For example, a systematic review summarized the evidence of the effectiveness of behavioural interventions to prevent HIV infection among young people in sub-Saharan Africa (Napierala Mavedzenge et al., 2011). The review question was structured, using the PICOS approach, as follows:

> *Population*: Among young people aged 10–24 years in sub-Saharan Africa . . .
>
> *Intervention/exposure/comparison*: . . . does exposure to an intervention focusing on reducing HIV risk behaviours, relative to no or minimal intervention, . . .
>
> *Outcomes*: . . . reduce the risk of HIV, STIs, or pregnancy . . .
>
> *Study design*: . . . when evaluated through experimental or quasi-experimental study designs?

A second example, used in this chapter, is a systematic review of the evidence that the use of chewing substances (such as smokeless tobacco or betel nuts) is associated with cardiovascular disease (CVD) in Asia (Zhang et al., 2010). In this case, the question was structured as follows:

> *Population*: Among people in Asian countries . . .
>
> *Intervention/exposure/comparison*: . . . does exposure to chewing substances, relative to not chewing them, . . .
>
> *Outcomes*: . . . increase the risk of CVD . . .
>
> *Study design*: . . . when evaluated through observational epidemiological studies?

Previous systematic reviews had examined this question in the United States of America (USA) and Sweden, but there was no synthesis of the evidence from Asia. If strong evidence for an association was found, this could lead to the development and evaluation of an intervention directed at reducing betel chewing in these populations.

Once the research question is identified, a detailed protocol should be prepared for the review. This will include definition of the search strategy and the planned analyses. There are plans to develop an international register of systematic reviews, led by the Centre for Reviews and Dissemination (<http://www.york.ac.uk/inst/crd/index.htm>), which will enable researchers to register their review protocol. This will extend the register developed by the Cochrane Collaboration (<http://www.cochrane.org>), which was established in 1993 to promote systematic reviews of health care interventions. Researchers undertaking reviews under the Cochrane Collaboration are required to register the protocol for their review in advance, and the review is peer-reviewed before publication. However, many systematic reviews are undertaken outside of the Collaboration and may not currently be registered.

2.2. Identifying relevant literature

The most time-consuming step of a systematic review is to identify studies which address the defined review question. The aim is to have a search strategy which is highly sensitive (i.e. there is a very high probability of including relevant studies), specific (i.e. there is a high probability of excluding non-relevant studies), and precise (i.e. the proportion of studies retrieved which are relevant is high) (Jenkins, 2004).

The first step in defining the search strategy to identify published papers is to set inclusion and exclusion criteria, based on the review question (Table 3.1). Ideally, searches should include papers published in any language (to be fully inclusive and to avoid possible publication bias of those with positive findings being more likely than those with negative findings to be published in English language journals). RCTs are generally regarded as the gold standard for providing evidence of the impact of an intervention, and it is essential to review previous RCTs of similar interventions. However, if there have been few relevant RCTs, non-randomized trials and observational studies should also be reviewed. The initial search may be limited to published papers, but sometimes it is important to include the 'grey' literature (conference abstracts, technical reports, and discussion papers). This is because some completed studies are never published in peer-reviewed journals, and studies are often less likely to be published there if they do not find an effect of the intervention. Inclusion of unpublished studies may therefore reduce bias. However, unpublished studies are difficult to identify and have not undergone peer review, so they may be of poorer quality and insufficient information may be provided to contribute usefully to a review.

2.2.1. Electronic searching

Three commonly used electronic medical databases are MEDLINE (available freely via PubMed at <http://www.ncbi.nlm.nih.gov/PubMed>), Embase (<http://www.embase.com>), and CENTRAL (Cochrane Central Register of Controlled Trials, <http://www.cochrane-handbook.org>). A comprehensive search strategy requires each of these databases to be searched (Higgins and Green, 2008). However, these databases have a North American/European bias, and, for studies in LMICs, it is worth also searching other relevant databases such as LILACS (Latin American Caribbean Health Sciences Literature), African Healthline, GlobalHealth, and Popline. In addition, there are many subject-specific databases, such as PsychInfo (for psychology and related behavioural and social sciences), as well as Internet search engines such as Google Scholar. It may also be useful to search conference databases and trial registries to identify additional papers.

Strategies can be used to identify both free-text words in the database and controlled terms (called MeSH in MEDLINE, i.e. medical subject headings) that are used as keywords. Search strategies need to include the key terms in the review question and use the Boolean operators (such as 'AND', 'OR', 'NOT') to produce a search that is both sensitive and specific to the research question. The search strategy used for the example of chewing substances and CVD in Asia is given in Box 3.2.

Often the reviewers will already know about some key published studies. It is useful to check that all of these have been identified by the electronic database search. If not, a careful review of the search strategy may establish the reason for this, and the search can be amended accordingly.

2.2.2. Reviewing abstracts

The search strategy commonly identifies several thousands of potentially relevant papers. The next step is for two reviewers to independently read through the abstract of each paper and define it as being potentially relevant or not. At this stage, it is recommended to err on the side of caution, i.e. include as 'potentially relevant' if the relevance is unclear from the abstract. The two reviewers should then compare their results and reconcile any differences by discussion, further reference to the abstracts, or a third reviewer independently reading the abstract.

2.2.3. Reviewing full articles

Full copies of all papers, the abstracts of which were considered to be potentially relevant, should be obtained (electronically, from libraries, or by emailing the author). They should be reviewed by the two reviewers who independently assess whether or not each paper meets each of the inclusion/exclusion criteria. Discrepancies should be resolved as for the abstracts.

2.2.4. Hand searching

The next step in the search strategy is usually to review the reference lists of all the eligible studies identified from the electronic database search, to identify any studies that were missed by that search but have been referenced in the

eligible papers.

Previous review papers should also be read to check that no known papers have been omitted. Finally, it is legitimate, though sometimes time-consuming, to include unpublished studies which can be identified through colleagues or contact with the investigators of unpublished studies, for example, identified through Internet searches or trial registers. It is also important to identify ongoing studies, where possible, as these may be included in updates of the review.

2.2.5. Flow chart of search strategy

The template for a flow chart summarizing the search results is given in Figure 3.1. In the example of behavioural interventions among young people in sub-Saharan Africa, a total of 1173 papers were identified from the electronic databases, of which 137 were deemed potentially relevant after review of their titles and abstracts, and full-text articles were obtained. After excluding those not meeting the inclusion criteria, the final review included 40 papers, representing 23 studies (as sometimes the results of one study were reported in more than one paper) (Napierala Mavedzenge et al., 2011). For the example of chewing substances and CVD in Asia, 1756 publications were identified from electronic databases, of which only six were eligible for inclusion in the analysis of CVD (Zhang et al., 2010).

2.3. Descriptive synthesis of studies

When the eligible papers have been identified, a data extraction form should be completed for each study, which contains fields enabling a detailed description of the study design and of the results. For example, descriptive elements would include the PICOS components, as discussed in Section 2.1. The results should focus on the pre-specified outcomes in the review protocol and would include outcome measures, definition of exposures/interventions, measures of effect, and 95% confidence intervals (CIs). The form should be pilot-tested on a few sample papers and revised, as appropriate. Two reviewers then read each paper in detail independently, summarize the paper on to the data extraction form, and appraise the risk of biases. A common shortcut, which is permissible, is that one reviewer completes the data extraction form and the other then checks and edits it, with the final version based on a discussion of any discrepancies.

The next step is to begin to summarize the evidence from the eligible studies as a whole. All reviews should include a descriptive table of the included studies, which summarize the study population, intervention, comparison, outcome, and study design. One of the 23 studies that were identified in the review of behavioural interventions among young people is summarized in Table 3.2.

In the table that summarizes the results of each study, all the primary and secondary outcome measures should be included. For a binary outcome, this would include the proportion with the outcome among the exposed and unexposed groups, the appropriate measure of effect (e.g. risk ratio (RR), rate ratio (RR), or odds ratio (OR)), and 95% CI. For continuous outcomes, the mean, standard deviation in the exposed and unexposed, plus the effect measure (e.g. standardized mean difference) should be given.

2.4. Assessing risk of bias in the studies

Once the description of each study is completed, an evaluation should be conducted of the extent of potential bias and error that may have arisen, either from the design or the analysis of each of the original studies. The main aim of this is to guide interpretation of the findings of the review. In some cases, it may be decided to exclude a study which is flawed to the extent that the results are considered likely not to be valid. Alternatively, a sensitivity analysis might be conducted to evaluate how the summary results differ if results from more flawed studies are included or excluded.

There are several methods for assessing the risk of bias, including checklists or ‘quality score’ scales. The recommendation of the Cochrane Collaboration and the PRISMA guidelines is to use a ‘domain-based evaluation’, in which critical assessments are made for domains such as blinding of participants and generation of the random sequence (for randomized studies) (Higgins and Green, 2008). For observational studies, there are additional possible sources of bias. For example, in case-control studies, check should be made on the external validity of case selection, the choice of control group, and adjustment for confounding factors.

Table 3.3 summarizes some of the sources of potential bias in RCTs and observational studies.

The assessment of potential biases should be tailored to the research question. For each review, there should be consideration of whether one potential bias is more important to the interpretation of findings than others. For example, if an outcome is measured objectively (for example, mortality), then blinding of those evaluating the outcome is not going to be very important. In contrast, if loss to follow-up is high and associated with the outcome, then this could cause substantial bias.

A table summarizing the risk of bias in each study should be completed independently by two reviewers, and any differences reconciled by discussion or reference to a third reviewer. Summarizing the results can be done in different ways—some authors rank the studies in order of quality; others divide them into those with low, medium, or high risk of bias. These decisions should be taken independently of the results of the studies, if possible, before examining the results, and the reviewers need to decide which studies (if any) will be taken forward to a quantitative meta-analysis of findings.

2.5. Quantitative synthesis of results

2.5.1. Forest plots

Following the descriptive analysis and assessment of risk of bias, it may or may not be appropriate to conduct a formal meta-analysis that quantifies the overall effect of the intervention. If, for example, the study populations, interventions, and reported outcomes differed substantially, the authors may decide to focus on describing the studies, their results, applicability, and limitations in a narrative review, rather than produce a quantitative summary. This was the case for the systematic review of interventions in young people in sub-Saharan Africa (Napierala Mavedzenge et al., 2011).

In other cases, it might be useful to summarize the data quantitatively. A first step for this is to produce a graph, called a forest plot, which displays the measure of effect (e.g. OR) for each study, together with a horizontal line denoting the CI. Before constructing such a graph, it is important to consider whether the results from the different studies are indeed measuring the same effect and are comparable to each other. For example, a smoking cessation intervention may have a different effect in pregnant women than among teenage girls. In such cases, it would be beneficial to present results stratified by subgroups, in whom effects might be expected to differ. As with all analyses, these subgroups should be defined in advance and included in the review protocol. For example, in the review of chewing substances in Asia, it was decided a priori to stratify by geographical region, to minimize confounding due to the presence or absence of tobacco in chewing substances, as this was thought to differ between regions.

In this example, the six eligible studies included five cohort studies and one case-control study. The forest plot is shown in Figure 3.2. The solid vertical line indicates a relative risk (RR) of one, representing no association between the exposure and outcome. In this example, all six studies had a RR greater than one, indicating an increased risk of CVD among individuals who used chewing substances, and the 95% CI did not include one for four of these studies, indicating strong evidence of an association. The forest plot also includes an overall (summary) estimate of the RR. This is a weighted average of the effects from each of the studies.

There are two main methods of obtaining the summary measure of an intervention effect. In a ‘fixed-effects’ model, it is assumed that the true effect of exposure (or the intervention) is the same in each study, any variation between studies being solely due to chance. In contrast, a ‘random-effects’ model may be used, in which the true effect of exposure for the individual studies are assumed to inherently vary (e.g. due to differences in the populations or residual confounding factors). In a random-effects model, the weights allow for this between-study variation, as well as the random variation.

In Figure 3.2, a random-effects model was used, and the weights for each study are given on the right-hand side of the forest plot. The overall (summary) estimate is RR = 1.26, with a 95% CI of 1.12–1.40. Note that this summary estimate is more precise (i.e. has a narrower CI) than any one of the individual studies. By undertaking a systematic review and meta-analysis, the reviewers can now report that there is strong evidence that, in these populations, exposure to chewing substances was associated with an increased risk of CVD of around 26%, compared with non-users.

2.5.2. Examining heterogeneity

The effect sizes of individual studies will inevitably be different from each other, but it is important to assess whether this difference is likely to be due to random variation (i.e. the true underlying effect will be the same) or to real

differences in underlying effect sizes in the individual studies. It is therefore essential to examine the consistency of the effects and to quantify the heterogeneity (or difference) in effect sizes between studies. Several measures are available for this, one of which is the I^2 statistic (Higgins et al., 2003). This statistic is the percentage of total variation across studies that is due to heterogeneity, rather than chance. A value of I^2 of 0% indicates no observed heterogeneity, and larger values indicate increasing heterogeneity. The principal advantage of the I^2 statistic is that it does not depend on the number of studies included in the meta-analysis and so can be used even for meta-analyses containing relatively few studies, which typically have low power to detect heterogeneity using other measures.

In our example, the value of I^2 is 35.9%, with a p-value of 0.17, indicating little evidence of heterogeneity. The reviewers were therefore justified in presenting the summary estimate. If, in contrast, the I^2 statistic suggests evidence of heterogeneity, for example if I^2 was 70%, further exploration of the causes of heterogeneity would be needed, for example by undertaking (pre-specified) subgroup analyses. If there was no longer evidence of heterogeneity within subgroups, this would indicate that the stratifying characteristics were an important source of heterogeneity, and results should be presented within subgroups, rather than overall.

3. Software available for systematic reviews and meta-analyses

Systematic reviews involve managing large quantities of information. There are various software packages available which can be used to prepare systematic reviews. For example, the Cochrane Collaboration produces a freely available program called RevMan which is a Windows-based software package designed to enter reviews in the Cochrane format. This includes an analysis module (MetaView) for quantitative summaries.

Results of searches from electronic databases can also be automatically downloaded into a reference manager software package, such as EndNote, and, from there, exported into database packages, such as Excel, for review and assessment of abstracts. Standard statistical packages, such as Stata, include modules for meta-analyses.

4. Reporting findings from systematic reviews

There are several guidelines for reporting results of a systematic review. The most recent are the PRISMA guidelines (<http://www.prisma-statement.org>) which are given in Table 3.4 (Moher et al., 2009). These include a full description of the rationale for the review, the research question, methods used, and analyses. Reviewers will then need to summarize their main findings, including the strengths and limitations of the review, the strength of the evidence for each main outcome, and the relevance to different population groups.

Finally, the results of the systematic review need to be assessed for their implications for policy and future research. One system to assist with interpreting results of systematic reviews is the GRADE system (Grading of Recommendations Assessment, Development, and Evaluation) (Guyatt et al., 2008). This gives guidelines as to whether results from a systematic review provide ‘strong’ or ‘weak’ evidence. This includes not only results of a systematic review, but also an evaluation of the balance between desirable and undesirable effects, and whether the intervention represents a wise use of resources.

References

Clark, S. and Horton, R. 2010. Putting research into context—revisited. *Lancet*, **376**, 10–11.10.1016/S0140-6736(10)61001-X [PubMed: 20609979] [CrossRef]

Egger, M., Smith, G. D., and Altman, D. G. 2001. *Systematic reviews in health care: meta-analysis in context*. London: BMJ Books.10.1002/9780470693926 [CrossRef]

Glasziou, P. 2001. *Systematic reviews in health care: a practical guide*. Cambridge, New York: Cambridge University Press.10.1017/CBO9780511543500 [CrossRef]

Guyatt, G. H., Oxman, A. D., Vist, G. E., et al. 2008. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, **336**, 924–6.10.1136/bmj.39489.470347.AD [PMC free article: PMC2335261] [PubMed: 18436948] [CrossRef]

Higgins, J. and Green, S. P. 2008. *Cochrane handbook for systematic reviews of interventions* [Online]. Oxford: Wiley-Blackwell. Available at: <http://tectutorials.com/Resources/AHRQ%20Modules/UoCTrainingMaterials/CochraneHB/booktext.pdf>.

Higgins, J. P., Thompson, S. G., Deeks, J. J., and Altman, D. G. 2003. Measuring inconsistency in meta-analyses. *BMJ*, **327**, 557–60.10.1136/bmj.327.7414.557 [PMC free article: PMC192859] [PubMed: 12958120] [CrossRef]

Jenkins, M. 2004. Evaluation of methodological search filters—a review. *Health Information & Libraries Journal*, **21**, 148–63.10.1111/j.1471-1842.2004.00511.x [PubMed: 15318913] [CrossRef]

Khan, K. S. 2003. *Systematic reviews to support evidence-based medicine: how to review and apply findings of healthcare research*. London: Royal Society of Medicine Press.

Liberati, A., Altman, D. G., Tetzlaff, J., et al. 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine*, **6**, e1000100.10.1371/journal.pmed.1000100 [PMC free article: PMC2707010] [PubMed: 19621070] [CrossRef]

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and Group, P. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, **6**, e1000097.10.1371/journal.pmed.1000097 [PMC free article: PMC2707599] [PubMed: 19621072] [CrossRef]

Napierala Mavedzenge, S. M., Doyle, A. M., and Ross, D. A. 2011. HIV prevention in young people in sub-Saharan Africa: a systematic review. *Journal of Adolescent Health*, **49**, 568–86.10.1016/j.jadohealth.2011.02.007 [PubMed: 22098767] [CrossRef]

Zhang, L. N., Yang, Y. M., Xu, Z. R., Gui, Q. F., and Hu, Q. Q. 2010. Chewing substances with or without tobacco and risk of cardiovascular disease in Asia: a meta-analysis. *Journal of Zhejiang University SCIENCE B*, **11**, 681–9.10.1631/jzus.B1000132 [PMC free article: PMC2932878] [PubMed: 20803772] [CrossRef]

Figures

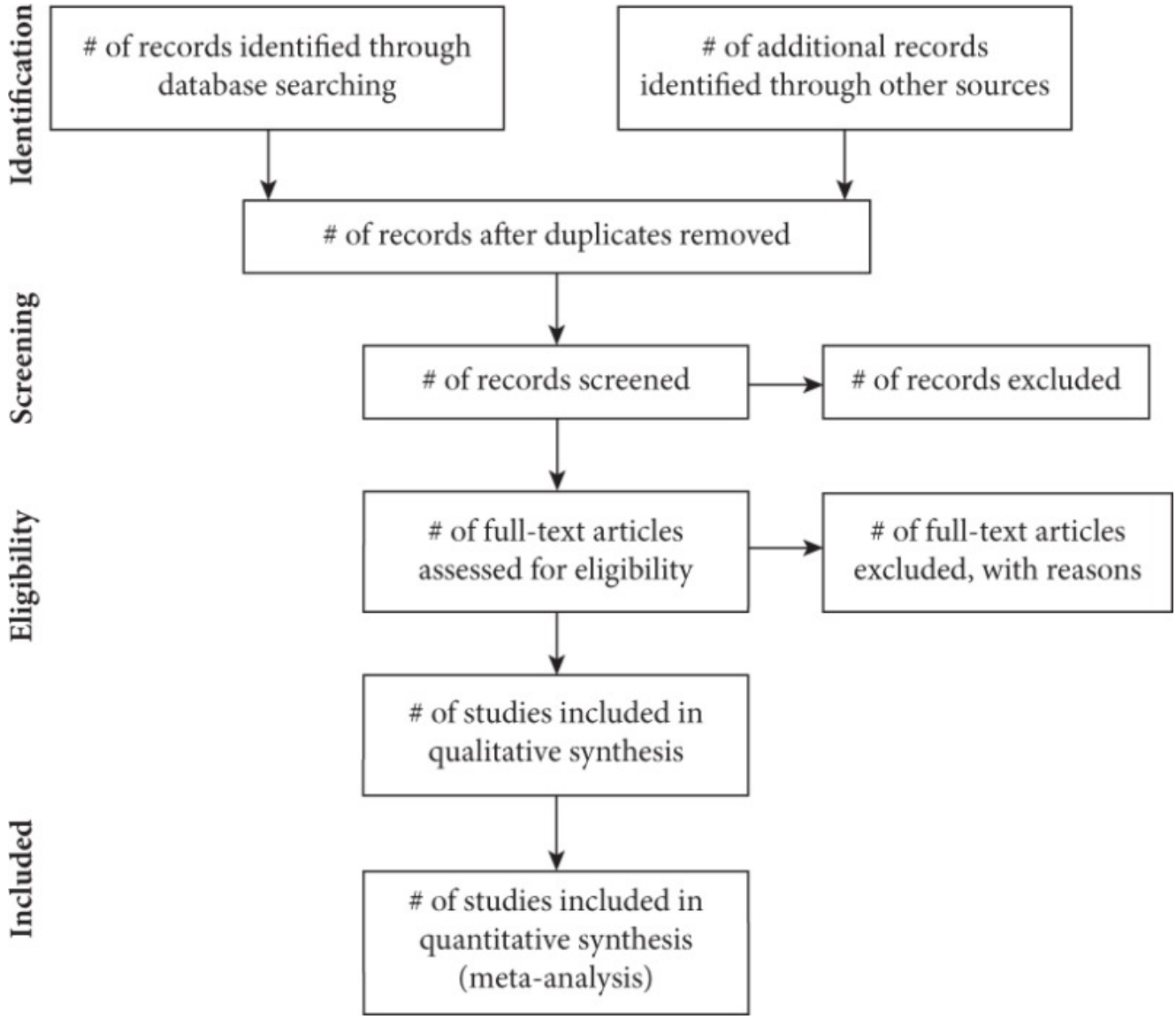

Figure 3.1

Flow diagram of study selection process.

From Moher et al., Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, *PLoS Medicine*, Volume 6, Issue 7, e1000097, Copyright © Moher et al. 2009. This figure is reproduced from an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

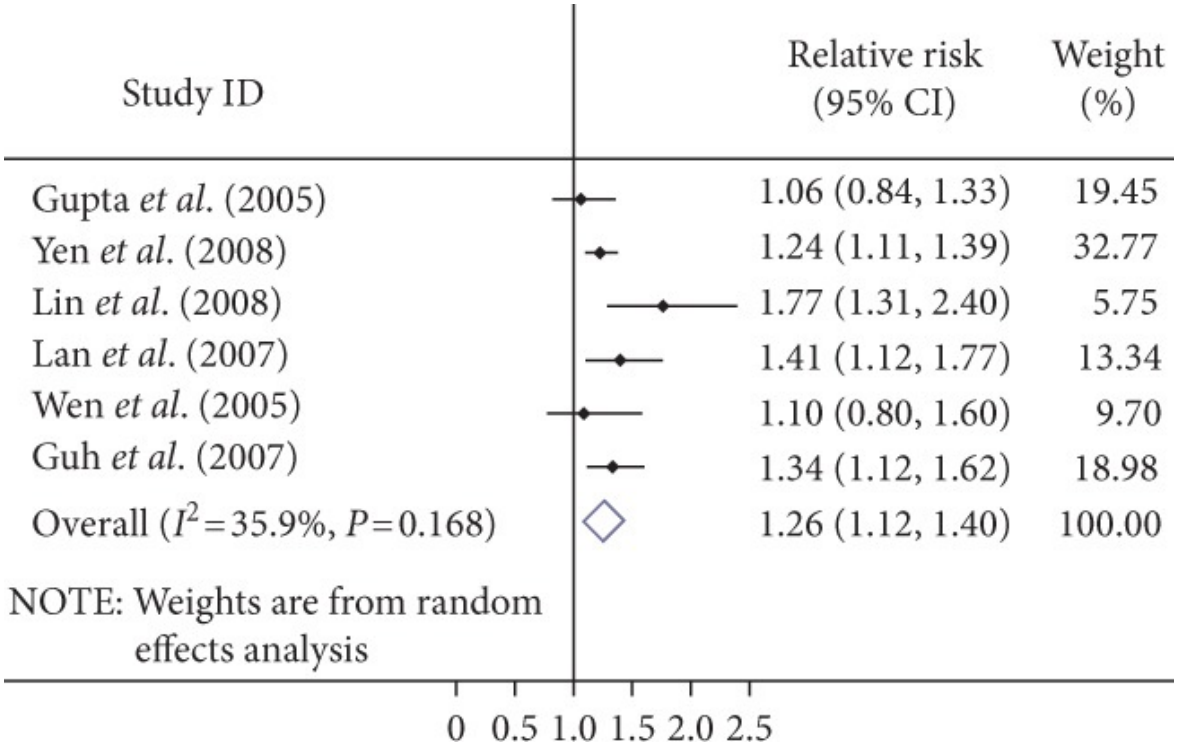

Figure 3.2

Forest plot for the association of exposure to chewing substances and risk of CVD in Asia.

Reproduced from Zhang, L. N. et al., Chewing substances with or without tobacco and risk of cardiovascular disease in Asia: a meta-analysis, *Journal of Zhejiang University Science B*, Volume 11, Issue 9, pp. 681–9, Copyright © Zhejiang University and Springer-Verlag Berlin Heidelberg 2010, with permission from Springer and Springer Science and Business Media. This image is not covered by the Creative Commons licence terms of this publication. For permission to reuse please contact the rights holder.

Tables

Table 3.1 Inclusion criteria: example for the systematic review of behavioural interventions to prevent HIV infection among young people in sub-Saharan Africa

PICOS component (see text)	Inclusion criteria	Exclusion criteria
*P*opulation	Young people aged 10–24 years. In studies with a wider age range, there must be an analysis of the impact of the intervention in young people (10–24 years) or, at least, in part of that age range. In sub-Saharan Africa. Based in a school, and/or health facility, and/or geographically defined community.	Study population not representative of a general population of young people (for example, young sex workers). Fewer than 100 people in the study.
*I*ntervention/exposure	Behavioural intervention focused on one or more of the following: (i) improving sexual and reproductive health skills and behaviour (ii) reducing the risk of sexually transmitted diseases (STDs) (iii) reducing unintended pregnancies (iv) increasing utilization of health services for treatment of STIs and/or behaviours related to more appropriate service utilization.	
*C*omparison	No or minimal behavioural intervention.	No suitable comparison group (for example, non-randomized study with post-intervention data only). No adjustment for differences between groups that might bias the findings.
*O*utcome	At least one of the following measured: (i) prevalence or incidence of HIV infection (ii) prevalence or incidence of another STI (iii) prevalence or incidence of pregnancy (measured by laboratory test or clinically observed) (iv) reported sexual and reproductive health behaviour (including treatment-seeking behaviour).	Measured less than 3 months after the intervention starts.
*S*tudy design	Published in 2005–2008 (because an earlier systematic review had covered the period up to the end of 2004). Randomized and non-randomized epidemiological studies which included a contemporaneous comparison group or a before–after/time series analysis in the intervention group only.	

Table 3.2 Description of one of the studies included in the systematic review of youth interventions against HIV infection in sub-Saharan Africa

Study, location, and programme	Type of intervention and setting	Target population, primary objectives, comparison, and study outcomes	Intervention description	Study design
United Republic of Tanzania, MEMA kwa Vijana	**Schools:** Teacher-led. Curriculum-based sexual and reproductive health education. **Health facility:** Interventions to facilitate youth friendliness of service providers, linked to interventions in the community and in other sectors (schools), to promote acceptance and utilization	**Target population**: Persons aged 12–19 years in rural areas. **Primary objectives:** Delayed sexual initiation, increased condom use, decreased number of sexual partners, and increased use of health services, especially for sexual and reproductive health services. **Comparison arm:** Current (very limited) sexual and reproductive health education in schools, and no additional interventions within health facilities or in the wider community. **Study outcomes:** **Primary:** HIV incidence; HSV2 prevalence. **Secondary:** pregnancy (by test and self-reported); prevalence of other STIs (by test and self-reported); knowledge and attitudes related to sexual and reproductive health issues; self-reported sexual risk behaviours, including sexual debut during trial follow-up, use of condoms, number of sexual partners, use of health services if reported a potential STI.	In-school teacher-led and peer-assisted programme. Covered refusal, self-efficacy, self-esteem, STI/HIV, sexuality, contraception, social values, respect, gender. Used drama, stories, and games. Also included interventions to make government health services more youth-friendly, youth condom promotion and distribution, and limited community-wide interventions. Ten to 15 lessons per year over 3 years.	Cluster randomized trial. Ten intervention clusters, ten control clusters.

Table 3.3 Methods for assessing risk of bias in RCTs and observational studies

Source of bias	Definition	Assessment for RCTs	Assessment for observational studies
Selection bias	Systematic differences between the comparison groups	Generation of random allocation Allocation concealment	Selection of exposed/unexposed Selection of cases/controls
Performance bias	Systematic differences in the care provided (apart from intervention)	Blinding of participant and provider Misclassification of exposure	Systematic differences in those exposed and unexposed Misclassification of exposure
Attrition bias	Systematic differences between the comparison groups in withdrawals from the study	Intention-to-treat analysis Outcome data not available for all participants	Differing follow-up rates between exposed and unexposed (or participation rates in cases and controls)
Detection bias	Systematic difference in outcome assessment	Blinding of those evaluating outcome	

Table 3.4 PRISMA guidelines for systematic reviews and meta-analyses

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 (for example, Web address), and, if available, provide registration information, including registration number	
Eligibility criteria	6	Specify study characteristics (for example, PICOS, length of follow-up) and report characteristics (for example, years considered, language, publication status) used as criteria for eligibility, giving rationale	
Information sources	7	Describe all information sources (for example, 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 (for example, 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 (for example, 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 (for example, 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 (for example, I^2) for each meta-analysis	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (for example, publication bias, selective reporting within studies)	
Additional analyses	16	Describe methods of additional analyses (for example, sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified	
RESULTS			

Section/topic	#	Checklist item	Reported on page #
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 (for example, 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 (for example, 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 (for example, health care providers, users, and policy makers)	
Limitations	25	Discuss limitations at study and outcome level (for example, risk of bias) and at review level (for example, 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 (for example, supply of data); role of funders for the systematic review	

From Moher et al., Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, *PLoS Medicine*, Volume 6, Issue 7, e1000097, Copyright © Moher et al. 2009. This table is reproduced from an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Boxes

Box 3.1 The five basic steps in a systematic review

1. Defining the question.
2. Identifying relevant studies in a predefined, systematic way.
3. Assessing the quality of each relevant study.
4. Summarizing the evidence.
5. Interpreting the findings.

Box 3.2 Example of a search strategy for evidence of an association between chewing substances and CVD, ischaemic heart disease, or cerebrovascular disease in Asia

We searched PubMed (up to July 2010), using the terms: (‘cardiovascular diseases’ [MeSH] OR (‘cardiovascular’ [All Fields] AND ‘diseases’ [All Fields]) OR ‘cardiovascular diseases’ [All Fields] OR ‘cerebrovascular disorders’ [MeSH] OR (‘cerebrovascular’ [All Fields] AND ‘disorders’ [All Fields]) OR ‘cerebrovascular disorders' [All Fields] OR ‘stroke’ [MeSH] OR ‘stroke’ [All Fields] OR 'mortality' OR death*) AND (‘betel quid’ OR ‘betel-quid’ OR ‘betel nut’ OR ‘betel nuts’ OR ‘areca nut’ OR ‘areca nuts’ OR ‘paan’ OR ‘pan’ OR 'snuff' OR 'snus' OR ‘gul’ OR ‘gutka’ OR ‘khaini’ OR ‘loose leaf’ OR ‘maras’ OR ‘mawa’ OR ‘mishri’ OR ‘naswar’ OR ‘*Areca catechu*’ OR ‘tooth powder’ OR ‘shammah’ OR ‘tobacco chewing gum’ OR ‘zarda’ OR ‘tobacco, smokeless’ [MeSH] OR ‘smokeless tobacco’ OR ‘chewing tobacco’ OR ‘non-smoking tobacco’) AND (‘cohort studies’ [MeSH] OR ‘cross-sectional studies’ [MeSH] OR ‘case control studies’ [MeSH] OR (‘cohort’ [TI] AND stud* [TI]) OR (case* [TI] AND control* [TI]) OR 'prospective' OR 'retrospective' OR 'cross-sectional' OR ‘cross sectional’), which yielded 1006 potentially relevant references. We adapted the searching strategy for a second search in ISI Web of Science (updated 19 July 2010) and found another 739 references. We identified all observational studies, including cohorts, case-control studies, and cross-sectional studies, provided that they explored the association between ever using chewing substances and the occurrence (incidence or mortality) of CVD and reported the strength of the associations with a quantitative risk estimate. There was no limitation on the language, study year, or publication status.

Text extract reproduced from Zhang, L. N. et al., Chewing substances with or without tobacco and risk of cardiovascular disease in Asia: a meta-analysis, *Journal of Zhejiang University Science B*, Volume 11, Issue 9, pp.681–9, Copyright © Zhejiang University and Springer-Verlag Berlin Heidelberg 2010. This box is not covered by the Creative Commons licence terms of this publication. For permission to reuse please contact the rights holder.

Bookshelf ID: NBK305502