

QUALITY IMPROVEMENT RESEARCH

Systematic reviews of the effectiveness of quality improvement strategies and programmes

J Grimshaw, L M McAuley, L A Bero, R Grilli, A D Oxman, C Ramsay, L Vale, M Zwarenstein

Qual Saf Health Care 2003;**12**:298–303

Systematic reviews provide the best evidence on the effectiveness of healthcare interventions including quality improvement strategies. The methods of systematic review of individual patient randomised trials of healthcare interventions are well developed. We discuss methodological and practice issues that need to be considered when undertaking systematic reviews of quality improvement strategies including developing a review protocol, identifying and screening evidence sources, quality assessment and data abstraction, analytical methods, reporting systematic reviews, and appraising systematic reviews. This paper builds on our experiences within the Cochrane Effective Practice and Organisation of Care (EPOC) review group.

that, in comparison with traditional narrative reviews, systematic reviews are an efficient scientific approach to identify and summarise evidence on the effectiveness of interventions that allow the generalisability and consistency of research findings to be assessed and data inconsistencies to be explored. Furthermore, the explicit methods used in systematic reviews should limit bias and improve the reliability and accuracy of conclusions. In this paper we focus on the methods of systematic reviews of the effectiveness of quality improvement strategies and programmes, building on our experiences within the Cochrane Effective Practice and Organisation of Care (EPOC) review group (box 2).^{7–9} (For a more general discussion about the conduct of systematic reviews see the Cochrane Handbook,¹⁰ Egger and colleagues¹¹ and Cooper and Hedges.¹²)

Systematic reviews are “reviews of a clearly formulated question that use explicit methods to identify, select, and critically appraise relevant research and to collect and analyse data from the studies that are included in the review”.¹ Well conducted systematic reviews are increasingly seen as providing the best evidence to guide choice of quality improvement strategies in health care.^{2–4} Furthermore, systematic reviews should be an integral part to the planning of future quality improvement research to ensure that the proposed research is informed by all relevant current research and that the research questions have not already been answered.

Systematic reviews are a generic methodology that can be used to synthesise evidence from a broad range of methods addressing different types of questions (box 1). Mulrow⁶ suggested

FORMING A REVIEW TEAM

When preparing to undertake a systematic review of a quality improvement strategy it is important to assemble a review team with the necessary combination of content and technical expertise. Content expertise may come from consumers, healthcare professionals, and policy makers. Content expertise is necessary to ensure that the review question is sensible and addresses the concerns of key stakeholders and to aid interpretation of the review. Frequently, content experts may not have adequate technical expertise and require additional support during the conduct of reviews. Technical expertise is required to develop search strategies for major databases, hand search

Box 1 Steps involved in undertaking a systematic review

- Stating the objectives of the research.
- Defining eligibility criteria for studies to be included.
- Identifying (all) potentially eligible studies.
- Applying eligibility criteria.
- Assembling the most complete data set feasible
- Analysing this data set, using statistical synthesis and sensitivity analyses, if appropriate and possible.
- Preparing a structured report of the research.

From Chalmers.⁵

Box 2 The Cochrane Effective Practice and Organisation of Care (EPOC) Group

The Cochrane Effective Practice and Organisation of Care (EPOC) group undertakes systematic reviews of the effectiveness of professional, organisational, financial, and regulatory interventions to improve professional practice and the delivery of effective health services.^{6–8} It was established in 1994 and since then has worked with over 180 reviewers worldwide to produce 29 reviews and 22 protocols covering a diverse range of topics including the effectiveness of different continuing medical education strategies, changes in the setting of care and different remuneration systems for primary care physicians.

See end of article for authors' affiliations

Correspondence to:
Dr J Grimshaw, Director of the Clinical Epidemiology Programme, Ottawa Health Research Institute, 1053 Carling Avenue, Ottawa ON K1Y 4E9, Canada;
jgrimshaw@ohri.ca

key journals (when appropriate), screen search results, develop data abstraction forms, appraise quality of primary studies, and statistically pool data (when appropriate).

DEVELOPING A PROTOCOL FOR A SYSTEMATIC REVIEW

Before undertaking a systematic review it is important to develop a formal protocol detailing the background, objectives, inclusion criteria, search methods, and proposed analytical methods to be used in the review. If reviewers do not develop a protocol a priori, there is a danger that the results of the review may be influenced by the data. For example, reviewers may exclude studies with unexpected or undesirable results.¹³ Developing and following a detailed protocol protects against this potential bias. Examples of protocols for reviews of quality improvement strategies are available in The Cochrane Library and from the EPOC website.^{8,9}

INCLUSION CRITERIA FOR SYSTEMATIC REVIEWS

Reviewers need to develop the review question based upon consideration of the types of study (for example, randomised controlled trials), interventions (for example, audit and feedback), study populations (for example, physicians), and outcomes (for example, objective measures of provider behaviour) in which they are interested. In general it is better to choose an estimation approach rather than a hypothesis testing approach in systematic reviews of quality improvement strategies as decision makers want to know something about the size of the expected effects (and the uncertainty around those estimates), and not just whether the null hypothesis can be rejected or not. Moreover, focusing on hypothesis testing tends to focus attention on p values rather than effects.

It is often helpful for reviewers to attempt to frame their research question in terms of the effects of quality improvement strategy *x* on end point *y* in study population *z*. In addition, reviewers should attempt to define a priori any subgroup analyses they wish to undertake to explore effect modifiers (for example, characteristics of the intervention) or other sources of heterogeneity (for example, quality of the included studies).

Design considerations

While cluster randomised trials are the most robust design for quality improvement strategies,¹³ some strategies may not be amenable to randomisation—for example, mass media campaigns. Under these circumstances, reviewers may choose to include other designs including quasi experimental designs.¹⁴ If a review includes quasi experimental studies—for example, interrupted time series designs for evaluating mass media campaigns,¹⁵ the reviewers need to recognise the weaknesses of such designs and be cautious of overinterpreting the results of such studies. Within EPOC, reviewers can include randomised trials, controlled before and after studies, and interrupted time series.⁸

Intervention considerations

Another important issue faced by reviewers is the lack of generally accepted classification of quality improvement strategies; as a result, it is vital that reviewers clearly define the intervention of interest. In our experience it is easier to define interventions based on pragmatic descriptions of the components of an intervention—for example, interactive educational sessions—than theoretical constructs—for example, problem based learning—as the description of interventions in primary studies is commonly poorly reported, especially lacking details of the rationale or theoretical basis for an intervention. Developing the definition of an intervention that can be operationalised within a systematic review frequently requires several iterations, preferably with involvement of content experts outside the review team to ensure that the

Box 3 Examples from the EPOC taxonomy of professional quality improvement strategies¹⁶

- Distribution of educational materials: published or printed recommendations for clinical care including clinical practice guidelines, delivered personally or through mass mailings.
- Educational meetings: healthcare providers who have participated in conferences, lectures, workshops or traineeships.
- Local consensus processes: inclusion of participating providers in discussion to ensure that they agreed that the chosen clinical problem was important and the approach to managing the problem was appropriate.
- Educational outreach visits and academic detailing: use of a trained person who met with providers in their practice settings to give information with the intent of changing the provider's practice. The information given may have included feedback on the performance of the provider(s).
- Local opinion leaders: use of providers nominated by their colleagues as "educationally influential". The investigators must have explicitly stated that their colleagues identified the opinion leaders.
- Patient mediated interventions: new clinical information (not previously available) collected directly from patients and given to the provider e.g. depression scores from an instrument.
- Audit and feedback: any summary of clinical performance of health care over a specified period of time. The summary may also have included recommendations for clinical action. The information may have been obtained from medical records, computerised databases, or observations from patients.
- Reminders: patient or encounter specific information, provided verbally, on paper or on a computer screen, which is designed or intended to prompt a health professional to recall information, including computer aided decision support and drug dosages are included.
- Marketing: a survey of targeted providers to identify barriers to change and subsequent design of an intervention that addresses identified barriers.

resulting definitions are likely to be robust and meaningful. EPOC has developed a taxonomy for quality interventions based on such descriptions that may provide a useful starting point for such discussions (see box 3 for examples).

The lumping versus splitting debate

A key issue faced by reviewers of quality improvement strategies is deciding how broad the scope of a review should be; this is commonly known as the "lumping" or "splitting" debate.¹⁷ For example, a review team could choose to undertake a review of quality improvement interventions to improve chronic diseases across all healthcare settings and professionals, or a review of quality improvement interventions to improve chronic diseases within primary care, or a review of quality improvement strategies to improve diabetes care within primary care, or a review of audit and feedback to improve all aspects of care across all healthcare settings. The rationale for taking a broad approach ("lumping") is that, because systematic reviews aim to identify the common generalisable features within similar interventions, minor differences in study characteristics may not be crucially important. The rationale for taking a narrower approach ("splitting") is that it is only appropriate to include studies which are very similar in design, study population, intervention characteristics, and outcome recording.

There are good methodological reasons for taking a broad approach. Broad systematic reviews allow the generalisability and consistency of research findings to be assessed across a wider range of different settings, study populations, and behaviours. This reduces the risk of bias or chance results. For

example, Jamtvedt and colleagues undertook a review of audit and feedback to improve all aspects of care across all healthcare settings.¹⁸ They identified 85 studies of which 18 considered the effects of audit and feedback on chronic disease management, 14 considered the effects of audit and feedback on chronic disease management in primary care, and three considered the effects of audit and feedback on diabetes care in primary care settings. By undertaking a broad review they were able to explore whether the effects of audit and feedback were similar across different types of behaviour, different settings, and different types of behaviour within different settings. If they had undertaken a narrow review of audit and feedback on diabetes care in primary care they would have been limited to considering only three studies and may have made erroneous conclusions if these studies suffered from bias or chance results. Very narrowly focused reviews are, in effect, subgroup analyses and suffer all the well recognised potential hazards of such analyses.¹⁹ A more transparent approach is to lump together all similar interventions and then to carry out explicit a priori subgroup analyses.

IDENTIFYING AND SCREENING EVIDENCE SOURCES

Reviewers need to identify what bibliographic databases and other sources the review team will search to identify potentially relevant studies and the proposed search strategies for the different databases. There are a wide range of bibliographic databases available—for example, Medline, EMBASE, Cinahl, Psychlit, ERIC, SIGLE. The review team has to make a judgement about what databases are most relevant to the review question and can be searched within the resources available to them.

The review team has to develop sensitive search strategies for potentially relevant studies. Unfortunately, quality improvement strategies are poorly indexed within bibliographic databases; as a result, broad search strategies using free text and allied MeSH headings often need to be used. Furthermore, while optimal search strategies have been developed for identifying randomised controlled trials,²⁰ efficient search strategies have not been developed for quasi experimental designs. Review teams should include or consult with experienced information scientists to provide technical expertise in this area.

EPOC has developed a highly sensitive search strategy (available at <http://www.epoc.uottawa.ca/register.htm>) for studies within its scope, and has searched Medline, EMBASE, Cinahl and SIGLE retrospectively and prospectively.²¹ We have screened over 200 000 titles and abstracts retrieved by our searches of these databases to identify potentially relevant studies. These are entered onto a database (“pending”) awaiting further assessment of the full text of the paper. Studies which, after this assessment, we believe to be within our scope are then entered onto our database (the “specialised register”) with hard copies kept in our editorial base. We currently have approximately 2500 studies in our specialised register (with a further 3000 potentially relevant studies currently being assessed). In future, reviewers may wish to consider the EPOC specialised register as their main bibliographic source for reviews and only undertake additional searches if the scope of their review is not within EPOC’s scope (see EPOC website for further information about the register).⁹

Preferably two reviewers should independently screen the results of searches and assess potentially relevant studies against the inclusion criteria in the protocol. The reasons for excluding potentially relevant studies should be noted when the review is reported.

QUALITY ASSESSMENT AND DATA ABSTRACTION

Studies meeting the inclusion criteria should be assessed against quality criteria. While there is growing empirical evi-

dence about sources of bias in individual patient randomised trials of healthcare interventions,²² quality criteria for cluster randomised trials and quasi experimental are less developed. EPOC has developed quality appraisal criteria for such studies based upon threats to validity of such studies identified by Cook and Campbell²³ (available from EPOC website).⁹ Reviewers should develop a data abstraction checklist to ensure a common approach is applied across all studies. Box 4 provides examples of data abstraction checklist items that reviewers may wish to collect. Data abstraction should preferably be undertaken independently by two reviewers. The review team should identify the methods that will be used to resolve disagreements.

PREPARING FOR DATA ANALYSIS

The methodological quality of primary studies of quality improvement strategies is often poor. Reviewers frequently need to make decisions about which outcomes to include within data analyses and may need to undertake re-analysis of some studies. In this section we highlight methods for addressing two common problems encountered in systematic reviews of quality improvement strategies—namely, reporting of multiple end points and handling unit of analysis errors in cluster randomised studies.

Reporting multiple outcomes

Commonly, quality improvement studies report multiple end points, for example—changes in practice for 10 different preventive services or diagnostic tests. While reviewers may choose to report all end points, this is problematic both for the analysis and for readers who may be overwhelmed with data. The review team should decide which end points it will report and include in the analysis. For example, a review team could choose to use the main end points specified by the investigators when this is done, and the median end point when the main end points are not specified.²¹

Handling unit of analysis errors in primary studies

Many cluster randomised trials have potential unit of analysis errors; practitioners or healthcare organisations are randomised but during the statistical analyses the individual patient data are analysed as if there was no clustering within practitioner or healthcare organisation.^{14, 24} In a recent systematic review of guideline dissemination and implementation strategies over 50% of included cluster randomised trials had

Box 4 Examples of data abstraction checklist items

- Inclusion criteria
- Type of targeted behaviour
- Participants
 - Characteristics of participating providers
 - Characteristics of participating patients
- Study setting
 - Location of care
 - Country
- Study methods
 - Unit of allocation/analysis
 - Quality criteria
- Prospective identification by investigators of barriers to change
- Type and characteristics of interventions
 - Nature of desired change
 - Format/sources/recipient/method of delivery/timing
- Type of control intervention (if any)
- Outcomes
 - Description of the main outcome measure(s)
- Results

Derived from EPOC data abstraction checklist.^{8, 16}

such unit of analysis errors.²¹ Potential unit of analysis errors result in artificially low p values and overly narrow confidence intervals.²⁵ It is possible to re-analyse the results of cluster randomised trials if a study reports event rates for each of the clusters in the intervention and control groups using a *t* test, or if a study reports data on the extent of statistical clustering.^{25–26} In our experience it is rare for studies with unit of analysis errors to report sufficient data to allow re-analysis. The point estimate is not affected by unit of analysis errors, so it is possible to consider the size of the effects reported in these studies even though the statistical significance of the results cannot be ascertained (see Donner and Klar²⁷ for further discussion on systematic reviews of clustered data and Grimshaw and colleagues²⁰ and Ramsay and colleagues²⁸ for further discussion of other common methodological problems in primary studies of quality improvement strategies).

METHODS OF ANALYSIS/SYNTHESIS

Meta-analysis

When undertaking systematic reviews it is often possible to undertake meta-analyses that use “statistical techniques within a systematic review to integrate the results of individual studies”.¹ Meta-analyses combine data from multiple studies and summarise all the reviewed evidence by a single statistic, typically a pooled relative risk of an adverse outcome with confidence intervals. Meta-analysis assumes that different studies addressing the same issue will tend to have findings in the same direction.²⁹ In other words, the *real* effect of an intervention may vary in magnitude but will be in the same direction. Systematic reviews of quality improvement strategies typically include studies that exhibit greater variability or *heterogeneity* of estimates of effectiveness of such interventions due to differences in how interventions were operationalised, targeted behaviours, targeted professionals, and study contexts. As a result, the real effect on an intervention may vary both in magnitude and direction, depending on the modifying effect of such factors. Under these circumstances, meta-analysis may result in an artificial result which is potentially misleading and of limited value to decision makers. Further reports of primary studies frequently have common methodological problems—for example, unit of analysis errors—or do not report data necessary for meta-analysis. Given these considerations, many existing reviews of quality improvement strategies have used qualitative synthesis methods rather than meta-analysis.

Although deriving an average effect across a heterogeneous group of studies is unlikely to be helpful, quantitative analyses can be useful for describing the range and distribution of effects across studies and to explore probable explanations for the variation that is found. Generally, a combination of quantitative analysis, including visual analyses, and qualitative analysis should be used.

Qualitative synthesis methods

Previous qualitative systematic reviews of quality improvement strategies have largely used vote counting methods that add up the number of positive and negative comparisons and conclude whether the interventions were effective on this basis.^{2–30} Vote counting can count either the number of comparisons with a positive direction of effect (irrespective of statistical significance) or the number of comparisons with statistically significant effects. These approaches suffer from a number of weaknesses. Vote counting comparisons with a positive direction fail to provide an estimate of the effect size of an intervention (giving equal weight to comparisons that show a 1% change or a 50% change) and ignore the precision of the estimate from the primary comparisons (giving equal weight to comparisons with 100 or 1000 participants). Vote counting comparisons with statistically significant effects suffer similar problems; in addition, comparisons with potential

unit of analysis errors need to be excluded because of the uncertainty about their statistical significance and underpowered comparisons observing clinically significant but statistically insignificant effects would be counted as “no effect comparisons”.

To overcome some of these problems, we have been exploring more explicit analytical approaches reporting:

- the number of comparisons showing a positive direction of effect;
- the median effect size across all comparisons;
- the median effect size across comparisons without unit of analysis errors; and
- the number of comparisons showing statistically significant effects.²¹

This allows the reader to assess the likely effect size and consistency of effects across all included studies and whether these effects differ between studies, with and without unit of analysis errors. By using these more explicit methods we are able to include information from all studies, but do not have the same statistical certainty of the effects as we would using a vote counting approach. An example of the impact of this approach is shown in box 5.

Exploring heterogeneity

When faced with heterogeneity in both quantitative and qualitative systematic reviews, it is important to explore the potential causes of this in a narrative and statistical manner (where appropriate).³² Ideally, the review team should have identified potential effect modifiers a priori within the review protocol. It is possible to explore heterogeneity using tables, bubble plots, and whisker plots (displaying medians, interquartile ranges, and ranges) to compare the size of the observed effects in relationship to each of these modifying variables.¹⁸ Meta-regression is a multivariate statistical technique that can be used to examine how the observed effect sizes are related to potential explanatory variables. However, the small number of included studies common in systematic review of quality improvement strategies may lead to overfitting and spurious claims of association. Furthermore, it is important to recognise that these associations are observational and may be confounded by other factors.³³ As a result,

Box 5 Impact of using an explicit analytical approach

Freemantle *et al*³¹ used a vote counting approach in a review of the effects of disseminating printed educational materials. None of the studies using appropriate statistical analyses found statistically significant improvements in practice. The authors concluded: “This approach has led researchers and quality improvement professionals to discount printed educational materials as possible interventions to improve care”.

In contrast, Grimshaw *et al*³¹ used an explicit analytical approach in a review of the effects of guideline dissemination and implementation strategies. Across four cluster randomised controlled trials they observed a median absolute improvement of +8.1% (range +3.6% to +17%) compliance with guidelines. Two studies had potential unit of analyses, the remaining two studies observed no statistically significant effects. They concluded: “These results suggest that educational materials may have a modest effect on guideline implementation . . . However the evidence base is sparse and of poor quality”. This approach, by capturing more information, led to the recognition that printed educational materials may result in modest but important improvements in care and required further evaluation.

such analyses should be seen as exploratory. Graphical presentation of such analyses often facilitates understanding as it allows several levels of information to be conveyed concurrently (fig 1).

APPRAISING SYSTEMATIC REVIEWS OF QUALITY IMPROVEMENT STRATEGIES

Systematic reviews of quality improvement strategies are of varying quality and potential users of such reviews should appraise their quality carefully. Fortunately, Oxman and colleagues have developed and validated a checklist for appraising systematic reviews including nine criteria scored as “done”, “partially done”, and “not done”, and one summary criterion scored on a 1–7 scale where 1 indicates “major risk of bias” and 7 indicates “minor risk of bias” and one summary

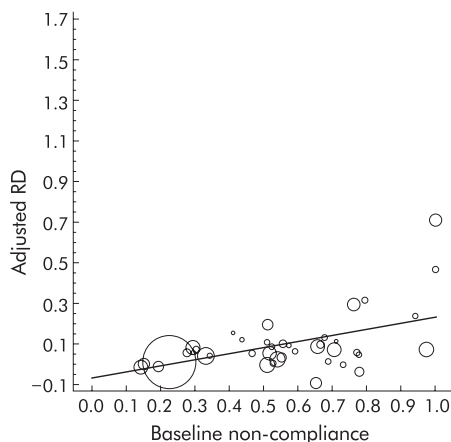


Figure 1 Graphical presentation of results using a bubble plot. This bubble plot, from a review on the effects of audit and feedback,¹⁸ shows the relationship between the adjusted risk difference and baseline non-compliance. The adjusted risk difference represents the difference in non-compliance before the intervention from the difference observed after the intervention. Each bubble represents a study, and the size of the bubble reflects the number of healthcare providers in the study. The regression line shows a trend towards increased compliance with audit and feedback with increasing baseline non-compliance.

Box 6 Checklist for appraising systematic reviews

- Were the search methods used to find evidence (primary studies) on the primary question(s) stated?
- Was the search for evidence reasonably comprehensive?
- Were the criteria used for deciding which studies to include in the review reported?
- Was bias in the selection of articles avoided?
- Were the criteria used for assessing the validity of the studies that were reviewed reported?
- Was the validity of all the studies referred to in the text assessed using appropriate criteria (either in selecting studies for inclusion or in analysing the studies that are cited)?
- Were the methods used to combine the findings of the relevant studies (to reach a conclusion) reported?
- Were the findings of the relevant studies combined appropriately relative to the primary question addressed by the review?
- Were the conclusions made by the author(s) supported by the data and/or the analysis reported in the review?
- Overall, how would you rate the scientific quality of this review?

Items 1–9 scored as done, not clear, not done. Item 10 scored on a scale of 1 (major risk of bias) to 7 (minimal risk of bias).

Adapted from Oxman and Oxman and Guyatt.^{33 34}

Key messages

- Systematic reviews provide the best evidence on the effectiveness of healthcare interventions including quality improvement strategies.
- Systematic reviews allow the generalisability and consistency of research findings to be assessed and data inconsistencies to be explored across studies.
- The conduct of systematic reviews requires content and technical expertise.
- The Cochrane Effective Practice and Organisation of Care (EPOC) group has developed methods and tools to support reviews of quality improvement strategies.

criterion (box 6).^{35 36} Grimshaw and colleagues used this scale to appraise 41 systematic reviews of quality improvement strategies published by 1998; the median summary quality score was 4, indicating that they had some methodological flaws.³ Common methodological problems within these reviews included failure to adequately report inclusion criteria, to avoid bias in the selection of studies, to report criteria to assess the validity of included studies; and failure to apply criteria to assess the validity of selected studies. Unit of analysis errors were rarely addressed in these reviews.

CONCLUSION

Systematic reviews are increasingly recognised as the best evidence source on the effectiveness of different quality improvement strategies. In this paper we have discussed issues that reviewers face when conducting reviews of quality improvement strategies based on our experiences within the Cochrane Effective Practice and Organisation of Care group. The main limitation of current systematic reviews (and the main challenge confronting reviewers) is the quality of evaluations of quality improvement strategies. Fortunately, well done systematic reviews provide guidance for future studies. Indeed, at present the main contribution of systematic reviews in this area may be to highlight the need for more rigorous evaluations, but there are indications that the quality of evaluations is improving.²⁰ Those planning and reporting evaluations of quality improvement should do so in the context of a systematic review. Similarly, those planning quality improvement activities should consider the results of systematic reviews when doing so.

ACKNOWLEDGEMENTS

The Cochrane Effective Practice and Organisation of Care (EPOC) group is funded by the UK Department of Health. Jeremy Grimshaw holds a Canada Research Chair in Health Knowledge Transfer and Uptake. The Health Services Research Unit is funded by the Chief Scientist Office of the Scottish Office Department of Health. The views expressed are those of the authors and not necessarily of the funding bodies. Phil Alderson became an EPOC editor in September 2002; the development of many of these methods predated his arrival and Phil did not consider his contribution sufficient for authorship. We would like to acknowledge his ongoing contribution to EPOC.

We would like to thank all the reviewers who have undertaken quality improvement reviews and helped us to develop many of the ideas within this paper. This paper reflects our experiences of undertaking reviews of quality improvement strategies over the last decade. It promotes some of the tools developed by EPOC over this time period, most of which are available freely from our website. We hope that this paper will increase the number and quality of systematic reviews of quality improvement strategies and that some of these will be done in collaboration with EPOC.

REFERENCES

- 1 **Cochrane Collaboration.** Glossary. In: Cochrane Collaboration. *Cochrane Library*. Issue 2. Oxford: Update Software, 2003.
- 2 **NHS Centre for Reviews and Dissemination.** *Effect Health Care* 1999;5:1–16.

- 3 **Grimshaw JM**, Shirran L, Thomas RE, *et al*. Changing provider behaviour: an overview of systematic reviews of interventions. *Med Care* 2001;**39**(Suppl 2):112–45.
- 4 **Grimshaw JM**, Eccles MP, Walker AE, *et al*. Changing physician behavior. What works and thoughts on getting more things to work. *J Cont Educ Health Professions* 2002;**22**:237–43.
- 5 **Chalmers I**. Trying to do more good than harm in policy and practice: the role of rigorous, transparent, up-to-date, replicable evaluations. *Ann Am Acad Political Soc Sci* 2003 (in press).
- 6 **Mulrow CD**. Systematic reviews: rationale for systematic reviews. *BMJ* 1994;**309**:597–9.
- 7 **Mowatt G**, Grimshaw JM, Davis D, *et al*. Getting evidence into practice: the work of the Cochrane Effective Practice and Organisation of Care Group (EPOC). *J Cont Educ Health Professions* 2001;**21**:55–60.
- 8 **Alderson P**, Bero L, Grilli R, *et al*, eds. Cochrane Effective Professional and Organisation of Care Group. In: Cochrane Collaboration. *Cochrane Library*. Issue 2. Oxford: Update Software, 2003.
- 9 **Effective Practice and Organisation of Care Group (EPOC)**. 2003. <http://www.epoc.uottawa.ca> (accessed 27 April 2003).
- 10 **Clarke M**, Oxman AD, eds. Cochrane Reviewers' Handbook 4.1.6. In: Cochrane Collaboration. *Cochrane Library*. Issue 2. Oxford: Update Software, 2003. Updated quarterly.
- 11 **Egger M**, Davey Smith G, Altman DG, eds. *Systematic reviews in health care. Meta-analysis in context*. London: BMJ Publishing, 2001.
- 12 **Cooper H**, Hedges LV, eds. *The handbook of research synthesis*. New York: Russell Sage Foundation, 1994.
- 13 **Egger M**, Davey Smith G. Principles and procedures for systematic reviews. In: Egger M, Davey Smith G, Altman DG, eds. *Systematic reviews in health care. Meta-analysis in context*. London: BMJ Publishing, 2001.
- 14 **Eccles M**, Grimshaw J, Campbell M, *et al*. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care* 2003;**12**:47–52.
- 15 **Grilli R**, Ramsay CR, Minozi S. Mass media interventions: effects on health services utilisation. In: Cochrane Collaboration. *Cochrane Library*. Issue 2. Oxford: Update Software, 2003.
- 16 **Effective Practice and Organisation of Care Group (EPOC)**. *The data collection checklist, section 2.1.1*. <http://www.epoc.uottawa.ca/checklist2002.doc> (accessed 27 April 2003).
- 17 **Goetzsche PC**. Why we need a broad perspective on meta-analysis. *BMJ* 2000;**321**:585–6.
- 18 **Jamtvedt G**, Young JM, Kristoffersen DT, *et al*. Audit and feedback: effects on professional practice and health care outcomes. In: Cochrane Collaboration. *Cochrane Library*. Issue 2. Oxford: Update Software, 2003 (submitted for publication).
- 19 **Oxman AD**, Guyatt GH. A consumer's guide to subgroup analyses. *Ann Intern Med* 1992;**116**:78–84.
- 20 **Lefebvre C**, Clarke MJ. Identifying randomised trials. In: Eggar M, Davey Smith G, Altman DG, eds. *Systematic reviews in health care. Meta-analysis in context*. London: BMJ Publishing, 2001.
- 21 **Grimshaw JM**, Thomas RE, MacLennan G, *et al*. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2003 (in press).
- 22 **Juni P**, Altman DG, Eggar M. Assessing the quality of randomised controlled trials. In: Eggar M, Davey Smith G, Altman DG, eds. *Systematic reviews in health care. Meta-analysis in context*. London: BMJ Publishing, 2001.
- 23 **Cook TD**, Campbell DT. *Quasi-experimentation: design and analysis issues for field settings*. Chicago: Rand McNally, 1979.
- 24 **Whiting-O'Keefe QE**, Henke C, Simborg DW. Choosing the correct unit of analysis in medical care experiments. *Med Care* 1984;**22**:1101–14.
- 25 **Ukoumunne OC**, Gulliford MC, Chinn S, *et al*. Methods for evaluating area-wide and organisation based interventions in health and health care: a systematic review. *Health Technol Assess* 1999;**3**(5).
- 26 **Rao JNK**, Scott AJ. A simple method for the analysis of clustered binary data. *Biometrics* 1992;**48**:577–85.
- 27 **Donner A**, Klar N. Issues in meta-analysis of cluster randomised trials. *Stat Med* 2002;**21**:2971–80.
- 28 **Ramsay C**, Matowe L, Grilli R, *et al*. Interrupted time series designs in health technology assessment: lessons from two systematic reviews of behaviour change strategies. *Int J Technol Assess Health Care* 2003 (in press).
- 29 **Peto R**. Why do we need systematic overviews of randomised trials? *Stat Med* 1987;**6**:233–41.
- 30 **Bushman BJ**. Vote counting methods in meta-analysis. In: Cooper H, Hedges L, eds. *The handbook of research synthesis*. New York: Russell Sage Foundation, 1994.
- 31 **Freemantle N**, Harvey EL, Wolf F, *et al*. Printed educational materials: effects on professional practice and health care outcomes (Cochrane Review). In: *Cochrane Library*. Issue 1. Oxford: Update Software, 1997.
- 32 **Thompson SG**. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ* 1994;**309**:1351–5.
- 33 **Sterne JAC**, Egger M, Davey Smith G. Investigating and dealing with publication and other biases. In: Eggar M, Davey Smith G, Altman DG, eds. *Systematic reviews in health care. Meta-analysis in context*. London: BMJ Publishing, 2001.
- 34 **Oxman AD**. Checklists for review articles. *BMJ* 1994;**309**:648–51.
- 35 **Oxman AD**, Guyatt GH. The science of reviewing research. *Ann NY Acad Sci* 1993;**703**:123–31.

Want full text but don't have
a subscription?

Pay per view

For just US\$8 you can purchase the full text of individual articles using our secure online ordering service. You will have access to the full text of the relevant article for 48 hours during which time you may download and print the pdf file for personal use.

www.qshc.com