
Downloaded from: http://researchonline.lshtm.ac.uk/8723/

DOI: 10.1136/bmj.322.7278.98

Usage Guidelines

Please refer to usage guidelines at http://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: Creative Commons Attribution Non-commercial
http://creativecommons.org/licenses/by-nc/3.0/
Systematic reviews from astronomy to zoology: myths and misconceptions

Mark Petticrew

Systematic literature reviews are widely used as an aid to evidence-based decision making. For example, reviews of randomised controlled trials are regularly used to answer questions about the effectiveness of healthcare interventions. The high profile of systematic reviews as a cornerstone of evidence-based medicine, however, has led to several misconceptions about their purpose and methods. Among these is the belief that systematic reviews are applicable only to randomised controlled trials and that they are incapable of dealing with other forms of evidence, such as from non-randomised studies or qualitative research.

The systematic literature review is a method of locating, appraising, and synthesising evidence. The value of regularly updated systematic reviews in the assessment of effectiveness of healthcare interventions was dramatically illustrated by Antman and colleagues, who showed that review articles failed to mention advances in treatment identified by an updated systematic review.1

It is nearly a quarter of a century since Gene Glass coined the term “meta-analysis” to refer to the quantitative synthesis of the results of primary studies.2 The importance of making explicit efforts to limit bias in the review of literature, however, has been emphasised by social scientists at least since the 1960s.3 In recent

disorder and the interplay between them. The psychiatric sciences have sought to convert human misery and pain into technical problems that can be understood in standardised ways and are amenable to technical interventions by experts. But human pain is a slippery thing, if it is a thing at all: how it is registered and measured depends on philosophical and sociomoral considerations that evolve over time and cannot simply be reduced to a technical matter.

Trauma has become a pervasive idiom of distress in Western culture, and day to day usage—as with related terms like “emotional scarring”—is typically metaphorical. But when does it credibly denote a disease akin to physical trauma? The medical discourse on trauma has had heuristic value and some of those diagnosed as having post-traumatic stress disorder do have clinically significant psychiatric dysfunction, however it is labelled (and post-traumatic stress disorder will sometimes do). However, it might be timely for mental health professionals to review our definition of the disorder as a disease and decide whether it has sufficient robustness and explanatory power to apply to the diverse uses to which it is now being put. Society confers on doctors the power to award disease status and the social advantages attached to the sick role. Current practice, which labels people as being mentally ill when they are not, calls this public duty of doctors into question. To conflate normality and pathology devalues the currency of true illness, promotes abnormal illness behaviour, and incurs unnecessary public costs.10

In turn, society might reflect that the medicalisation of life, which has gathered pace in this century, tends to mean that distress is relocated from the social arena to the clinical arena. This is a two-edged sword: there are practical gains for some, but costs may accrue for everyone over time if contributing factors rooted in political and commercial philosophies and practices escape proper scrutiny.

Competing interests: None declared.

however, no logical reason why systematic reviews of most relevant will not "count" as evidence. There is, among some researchers that the studies they consider trials are not common and perhaps reflects a concern prevalent in subjects in which randomised controlled synthesise studies of other designs. This belief is capable of summarising the results only of randomised There is a widespread belief that systematic reviews are randomised controlled trials Systematic reviews include only measures, systematic reviews commonly require more screening thousands of abstracts. As a result of these assess their quality, and even "small" systematic studies, and investigate potential biases in those studies and sources of heterogeneity between study results. They also differ in the measures they typically take to reduce bias, such as using several reviewers working independently to screen papers for inclusion and assess their quality, and even "small" systematic reviews are likely to involve several reviewers screening thousands of abstracts. As a result of these measures, systematic reviews commonly require more time, staff, and money than traditional reviews. Systematic reviews are not simply "bigger," they are qualitatively different.

**Systematic reviews are the same as ordinary reviews, only bigger**

There is a common but erroneous belief that systematic reviews are just the same as traditional reviews, only bigger; in other words, you just search more databases. Systematic reviews are not just big literature reviews, and their main aim is not simply to be "comprehensive" (many biased reviews are "comprehensive") but to answer a specific question, to reduce bias in the selection and inclusion of studies, to appraise the quality of the included studies, and to summarise them objectively. As a result, they may actually be smaller, not bigger, partly because they apply more stringent inclusion criteria to the studies they review. They also differ in the measures they typically take to reduce bias, such as using several reviewers working independently to screen papers for inclusion and assess their quality, and even "small" systematic reviews are likely to involve several reviewers screening thousands of abstracts. As a result of these measures, systematic reviews commonly require more time, staff, and money than traditional reviews. Systematic reviews are not simply "bigger," they are qualitatively different.

**Systematic reviews include only randomised controlled trials**

There is a widespread belief that systematic reviews are capable of summarising the results only of randomised controlled trials, and that they cannot be used to synthesise studies of other designs. This belief is prevalent in subjects in which randomised controlled trials are not common and perhaps reflects a concern among some researchers that the studies they consider most relevant will not “count” as evidence. There is, however, no logical reason why systematic reviews of study designs other than randomised controlled trials cannot be carried out. Systematic reviews of non-randomised studies are common, and qualitative studies, for example, can be (and often are) included in systematic reviews. UK guidelines for carrying out systematic reviews do not exclude qualitative research,1 and criteria have been developed to aid in reviewing qualitative studies.2 Even reviews of the effectiveness of interventions do not confine themselves solely to randomised controlled trials; such reviews commonly include other study designs, including non-randomised studies, and case reports.3 In short, there is simply no basis for the belief that systematic reviews can be applied only to randomised controlled trials. The systematic review is simply a methodology that aims to limit bias, and the choice of which study designs to include is a choice that is made by the reviewers. It is not a restriction of the methodology.

**Systematic reviews require the adoption of a biomedical model of health**

This common myth holds that systematic reviews intrinsically adopt a biomedical model that is of relevance only to medicine and that should not be applied to other domains. Related to this is a belief that as health is more than an “absence of illness” other important outcomes of interventions (such as social impacts) need to be considered and that these are somehow inappropriate for inclusion in systematic reviews. Many health and non-health outcomes, however, are regularly defined, measured, and summarised in both qualitative and quantitative primary studies, and these studies can be (and are) included in systematic reviews. Reviews on the Cochrane Database of Systematic Reviews, for example, commonly include "quality of life" as an outcome alongside clinical indicators of the effects of interventions. The argument that it is somehow inappropriate to do systematic reviews of broader health (or non-health) outcomes is simply fallacious. Systematic reviews do not have any preferred “biomedical model,” which is why there are systematic reviews in such diverse topics as advertising, agriculture, archaeology, astronomy, biology, chemistry, criminology, ecology, education, entomology, law, manufacturing, parapsychology, psychology, public policy, and zoology.4-7 A recent paper even adopted systematic review methods to summarise eyewitness accounts of the Indian rope trick.8 In short, the systematic review is an efficient technique for hypothesis testing, for summarising the results of existing studies, and for

<table>
<thead>
<tr>
<th>Table 1 Systematic reviews and traditional narrative reviews compared</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Good quality systematic reviews</strong></td>
</tr>
<tr>
<td><strong>Deciding on review question</strong></td>
</tr>
<tr>
<td><strong>Searching for relevant studies</strong></td>
</tr>
<tr>
<td><strong>Assessing study quality</strong></td>
</tr>
<tr>
<td><strong>Synthesising study results</strong></td>
</tr>
<tr>
<td><strong>Base their conclusions on those studies which are most methodologically sound</strong></td>
</tr>
</tbody>
</table>

---

1. UK guidelines for carrying out systematic reviews do not exclude qualitative research, and criteria have been developed to aid in reviewing qualitative studies.
2. Even reviews of the effectiveness of interventions do not confine themselves solely to randomised controlled trials; such reviews commonly include other study designs, including non-randomised studies, and case reports.
3. In short, there is simply no basis for the belief that systematic reviews can be applied only to randomised controlled trials.
4. Reviews on the Cochrane Database of Systematic Reviews, for example, commonly include "quality of life" as an outcome alongside clinical indicators of the effects of interventions.
5. The argument that it is somehow inappropriate to do systematic reviews of broader health (or non-health) outcomes is simply fallacious.
6. Systematic reviews do not have any preferred “biomedical model,” which is why there are systematic reviews in such diverse topics as advertising, agriculture, archaeology, astronomy, biology, chemistry, criminology, ecology, education, entomology, law, manufacturing, parapsychology, psychology, public policy, and zoology.
7. A recent paper even adopted systematic review methods to summarise eyewitness accounts of the Indian rope trick.
assessing consistency among previous studies; these tasks are clearly not unique to medicine.\textsuperscript{5-16}

### Systematic reviews are of no relevance to the real world

Systematic reviews have been portrayed as being obsessed solely with disease outcomes and with randomised controlled clinical trials carried out in simple, closed healthcare systems, which are of no relevance to the complex social world outside evidence based medicine. In fact researchers have been carrying out systematic reviews of policy and other social interventions since the 1970s. For example, there have been at least a dozen systematic reviews investigating the effectiveness of delinquency and correctional programmes for the treatment of offenders, one of which reviewed 400 studies to detect a 10% reduction in delinquency, when previous (non-systematic) reviews had been unable to discern any positive effect of correctional treatments.\textsuperscript{17}

Systematic reviews have also been widely used to examine an array of contemporary and often contentious “real world” issues. These range from reviews of the effectiveness of policy and other interventions to systematic reviews of social issues. Complex “real world” issues are not beyond the remit of systematic reviews. This is highlighted by a recent report that summarised systematic reviews of both randomised and non-randomised studies of issues such as prevention of vandalism, crime deterrence, drug misuse, domestic violence, child abuse, and many others.\textsuperscript{18} These and many other examples show that systematic reviews can provide a credible evidence base to support policymaking.

### Systematic reviews necessarily involve statistical synthesis

This myth derives from a misunderstanding about the different methods used by systematic reviews. Some reviews summarise the primary studies by narratively describing their methods and results. Other reviews take a statistical approach (meta-analysis) by converting the data from each study into a common measurement scale and combining the studies statistically. The above myth assumes that such reviews can only be done this way. Many systematic reviews, however, do not use meta-analytic methods. Some of those which do, probably shouldn’t; for example, it is common practice to pool studies without taking into account variations in study quality, which can bias the review’s conclusions. It has been pointed out that one of the allures of meta-analysis is that it gives an answer, no matter whether studies are being combined meaningfully or not.\textsuperscript{19} Systematic reviews should not therefore be seen as automatically involving statistical pooling as narrative synthesis of the included studies is often more appropriate and sometimes all that is possible. A recent methodological review provides clear guidance on when and how to carry out meta-analyses of randomised and non-randomised studies.\textsuperscript{19}

### Systematic reviews have to be done by experts

Although expert practitioners are often involved in systematic reviews, most systematic reviewers are not expert practitioners. Even among those carrying out reviews of healthcare interventions, clinical experts are often in the minority. This is not to suggest that clinical input is irrelevant in systematic reviews of clinical interventions. Clearly, such input is invaluable in the location and interpretation of the evidence, and expert opinion is particularly valuable when evidence is sparse.\textsuperscript{20} Systematic reviews, however, are not the sole provenance of expert practitioners (such as clinical experts). For example, potential users of systematic reviews, such as consumers and policymakers, can be involved in the process. This can help to ensure that reviews are well focused, ask relevant questions, and are disseminated effectively to appropriate audiences.\textsuperscript{21}

### Systematic reviews can be done without experienced information/library support

Systematic reviews can indeed be carried out without proper information or library support, though

---

**Table 2** Examples of systematic reviews in the “real world”

<table>
<thead>
<tr>
<th>Review question</th>
<th>Methods</th>
<th>Authors’ conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does spending more money on schools improve educational outcomes?</td>
<td>Meta-analysis of effect sizes from 38 publications\textsuperscript{17}</td>
<td>Systematic positive relation between resources and student outcomes</td>
</tr>
<tr>
<td>Do women or men make better leaders?</td>
<td>Review of organisational and laboratory experimental studies of relative effectiveness of women and men in leadership and managerial roles\textsuperscript{18}</td>
<td>Aggregated over organisational and laboratory experimental studies in sample, male and female leaders were equally effective</td>
</tr>
<tr>
<td>Does sexual orientation of the parent matter?</td>
<td>Review investigating impact having homosexual as opposed to heterosexual parents has on emotional wellbeing and sexual orientation of child\textsuperscript{19}</td>
<td>Results show no differences between heterosexual and homosexual parents in terms of parenting styles, emotional adjustment, and sexual orientation of child(ren)</td>
</tr>
<tr>
<td>Are fathers more likely than mothers to treat their sons and daughters differently?</td>
<td>Review of 39 published studies\textsuperscript{20}</td>
<td>Fathers’ treatment of boys and girls differed most in areas of discipline and physical involvement and least in affection or everyday speech. Few differences for mothers</td>
</tr>
<tr>
<td>Is job absenteeism an indicator of job dissatisfaction?</td>
<td>Review of 23 research studies\textsuperscript{21}</td>
<td>Yes; stronger association was observed between job satisfaction and frequency of absence than between satisfaction and duration of absence</td>
</tr>
<tr>
<td>Are jurors influenced by defendants’ race?</td>
<td>Meta-analytic review of experimental studies\textsuperscript{22}</td>
<td>Results are consistent in finding that race influences sentencing decisions</td>
</tr>
<tr>
<td>Is there a relation between poverty, income inequality, and violence?</td>
<td>Review of 34 studies reporting on violent crime, poverty, and income inequality\textsuperscript{23}</td>
<td>Results suggest that homicide and assault may be more closely associated with poverty or income inequality than rape or robbery</td>
</tr>
</tbody>
</table>

References in this table are given on the BMJ’s website.
researchers are not typically experienced in information retrieval and their searches are likely to be less sensitive, less specific, and slower than those done by information professionals. Improvements to information technology are likely to facilitate the retrieval and filtering of information from electronic databases, but currently this remains a challenging task. Producing a good systematic review requires skill in the design of search strategies and benefits from professional advice on the selection of sources of published and unpublished studies.

Systematic reviews are a substitute for doing good quality individual studies

It would be comforting to think that systematic reviews were a sort of panacea, producing final definitive answers and precluding the need for further primary studies. Yet they do not always provide definitive answers and are not intended to be a substitute for primary research. Rather, they often identify the need for additional primary studies as they are an efficient method of identifying where research is currently lacking. Systematic reviews can therefore lead to more, not less, primary research. They can also prevent unnecessary new primary studies being carried out—for example, when meta-analyses show the effectiveness of an intervention by pooling many primary studies.

Conclusion

I have covered a selection of some of the more common myths and misunderstandings about systematic reviews. There are others (such as the myth that systematic reviews are not research but are something that researchers should be expected to do anyway without particular skills, training, or funding). Awareness of the non-clinical applications of systematic reviews is increasing, and the establishment of the Campbell Collaboration, a sibling of the Cochrane Collaboration, will contribute to this by preparing, maintaining, and disseminating systematic reviews of the effects of social and educational policies and practices. There are undoubtedly many methodological challenges to be faced in the application of systematic reviews outside clinical specialties. For example, there may be difficulties in incorporating appropriate contextual information and in incorporating the results of relevant qualitative research; and there may be problems of implementation and dissemination. There may also be considerable problems relating to the identification of unpublished studies and “grey” literature. These problems are also common in reviews of healthcare interventions and do not themselves preclude the use of systematic review methods.

In conclusion, I suggest that many criticisms of systematic reviews are ill founded. In particular, systematic reviews are commonly and erroneously perceived solely to be aids to clinical decision making, and this underestimates their wider uses. Despite methodological and other challenges, systematic reviews are already helping to identify “what works” beyond the world of evidence based medicine, and their potential role is more wide ranging than is often realised.

I thank Iain Chalmers, Sally Macintyre, and Trevor Sheldon for comments and for suggesting myths.

Funding: Chief Scientist Office of the Scottish Executive Department of Health

Competing interests: None declared.

3 Chalmers I, Hedges L. A brief history of research synthesis. Eval Health Prof (in press).
5 Popay J, Rogers A, Williams G. Rationale and standards for the systematic review of qualitative literature in health services research. Qual Health Res 1998;8:341-5.
18 Contributors to the Cochrane Collaboration and the Campbell Collaboration. Evidence from systematic reviews of research relevant to implementing the ‘wider public health agenda’. NHS Centre for Reviews and Dissemination [http://www.york.ac.uk/inst/crd/sph.htm August 2000].
38 Contributors to the Cochrane Collaboration and the Campbell Collaboration. Evidence from systematic reviews of research relevant to implementing the ‘wider public health agenda’. NHS Centre for Reviews and Dissemination [http://www.york.ac.uk/inst/crd/sph.htm August 2000].