

Systematic Reviews

Simon Gates

8 March 2007

Contents

- Reviewing of research
 - Why we need reviews
 - Traditional narrative reviews
 - Systematic reviews
- Components of systematic reviews
- Conclusions

Key reference

Systematic reviews in health care: meta-analysis in context

M Egger, G Davey Smith, DG Altman
(editors). BMJ Books.

www.systematicreviews.com

Why we need reviews

- Literature develops too fast; 2 million biomedical papers per year (and increasing)
- Impossible for any individual to appraise all literature relevant to a particular field
- Reviews summarise the “state of knowledge” about a particular subject

Why we need reviews

- Each individual study's results affected by chance and its own characteristics
- Most studies too small for firm conclusions
 - Sample size too small to detect clinically important differences
- A study of 1941 trials in schizophrenia found only 58 were large enough to detect an important difference

Why we need reviews

Karl Pearson 1904 (discussing the results of studies of serum inoculation against enteric fever):

“Many of the groups are far too small to allow of any definite opinion being formed at all, having regard to the size of the probable error involved”

Why we need reviews

- By including the results of all relevant studies, we can get a better approximation to the truth than from the results of any individual study

Narrative reviews

- Traditional “narrative” reviews are more like an essay about a subject
- An “expert” describes and summarises results of existing studies
- May be seriously misleading

Narrative reviews

- Subjective
- Studies to be included not assembled in formal and repeatable manner
- Biased set of studies may be included; often selective inclusion of studies supporting author's views

Narrative reviews

- Usually no assessment of methodological quality of studies or strength of evidence they provide
- Obvious that results of more reliable studies are likely to be more accurate
- Well-conducted RCT more likely to give an answer close to the truth than a dodgy case-control study

Narrative reviews

- No formal method of summarising the results of different studies
- Often “vote counting” used
- Add up number of studies “for” and “against” and believe the commonest view
- But takes no account of study design, size and quality
- Common for different narrative reviews to reach opposite conclusions

Beta-blockers after myocardial infarction

Conclusions from two narrative reviews in 1981

“... we still have no clear evidence that beta-blockers improve long-term survival after infarction despite almost 20 years of clinical trials”

Mitchell JRA. Timolol after myocardial infarction. BMJ 1981; 282: 1565-70

“... it seems perfectly reasonable to treat patients who have survived an infarction with timolol.”

Hampton JR. The use of beta-blockers for reduction of mortality after myocardial infarction. Eur Heart J 1981; 2: 259-268

Beta-blockers after myocardial infarction

What was the evidence for effectiveness of beta-blockers in 1981?

Systematic review and meta-analysis show that trials up to 1981 had clearly demonstrated that this treatment was effective.

Systematic reviews

- Treat conduct of a review as a scientific investigation
- Aim to be:
 - Objective
 - Comprehensive
 - Unbiased
 - Repeatable

Systematic reviews

- **There is no alternative**
- Without systematic reviews, every interested person will form their own, personal, biased summary of the research
- This will be strongly influenced by their personal experience and beliefs
- No consensus, strongly-held conflicting beliefs

Systematic reviews

- Systematic reviews may give much more information than narrative reviews
- Estimate of treatment effect and uncertainty around it (using MA)
- Investigation of reasons for variation of treatment effect between studies (e.g. different doses of drug, different settings)

Systematic reviews

- All types of research can be reviewed systematically
- Most familiar (important?) are reviews of randomised trials (RCTs)
- Methodology best developed for reviews of RCTs

Systematic reviews

- Diagnostic accuracy studies
- Genetic studies
- Observational studies
- Prognostic studies

Systematic reviews and meta-analysis

- Meta-analysis (MA) is statistical combination of results of studies included in a review
- Not all systematic reviews contain MA (most frequent in reviews of RCTs)
- Often not appropriate to combine studies
- MA without a systematic review may be misleading

Components of systematic reviews

Overview

- Systematic reviews are a type of research study
- Need to be conducted in accordance with a pre-specified protocol
- As objective as possible
- Reported completely and accurately to ensure they are reproducible

Overview

1. Protocol
2. Search for relevant studies
3. Study selection
4. Assessment of “quality”
5. Data extraction and processing
6. Statistical analyses
7. Reporting

1. Protocol

- Specifies review's question and how review will be conducted
- What types of study design will be included?
- What types of participants?
- What data will be extracted from each study?
- How studies will be located
- How data will be extracted and processed
- Statistical analyses.

1. Protocol - types of study

- Include type of study that are relevant to the question
- E.g. question is whether treatment is effective; include RCTs.
- Other types of study design may be relevant e.g. cohort studies, controlled before/after studies
- Decision about what types of study to include not always straightforward
- If RCT evidence sparse, could include observational studies BUT more likely to be biased and introduces methodological complexity

2. Search for relevant studies

- Aim is to be comprehensive and unbiased
- Include all relevant studies
- Lots of methods that can be used
 - Electronic databases (specify search strategy)
 - Review articles
 - Conference proceedings
 - hand searching of journals
 - Experts, drug companies, etc

3. Selection of studies

- Pre-specified inclusion criteria
 - Participants
 - Interventions
 - Comparisons
 - Study designs
- Often done independently by two people and results compared

4. Assessment of quality

- Studies vary in quality (i.e. their susceptibility to bias)
- Need to have criteria to judge how likely it is that each study suffers from bias
- Likelihood of bias will affect the conclusions that can be drawn from the review
- Criteria will vary depending on type of study
- Example, for RCTs:
 - Randomisation and allocation concealment
 - Blinding
 - Completeness of follow-up and intention to treat analysis

4. Assessment of quality

- Often poor reporting means that quality is hard to assess
- Checklists for items that should be reported
 - CONSORT (RCTs)
 - STROBE (observational studies)
 - STARD (diagnostic studies)
- Checklists and scales for study quality have also been proposed
- Many scales exist for RCTs (none really recommended!)
- QUADAS for diagnostic studies (Whiting et al, BMC Medical Research Methodology 2006 Mar 6;6:9)

4. Assessment of quality

- Quality assessment for RCTs [see Chapter 5 in Egger, Davey Smith & Altman]
- Much empirical research on quality scores
- Use of scores not recommended
- Better to assess different aspects of quality separately

5. Data extraction

- Analogous to data collection in a trial
- Needs to be done in a way that avoids bias
- Often done independently by two reviewers
- Major problem of missing information in studies
- Often some subjectivity involved

6. Statistical analysis

[More on this in tomorrow's session]

- Need to judge whether MA is appropriate
- Often will not be, e.g.
 - studies used different interventions
 - quality too poor
 - settings or populations too different
- Subgroup analyses
- Sensitivity analyses
- Investigations of publication bias etc

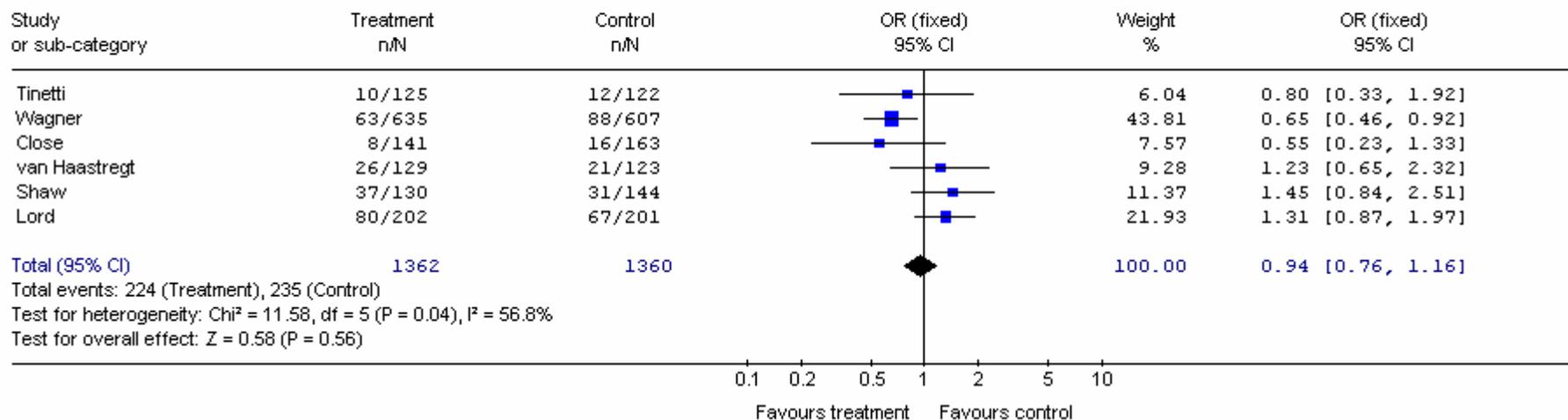
7. Reporting

- Needs to record accurately what was done so that study can be repeated
- Guidelines for reporting of systematic reviews:
 - QUORUM (RCTs) [Moher D et al. Lancet. 1999 Nov 27;354(9193):1896-1900]
 - MOOSE (Observational studies) [Stroup DF et al JAMA 2000;283:2008–12]

7. Graphical display of results

- “Forest plot” most commonly used to display results
- Shows
 - the effect estimate for each study
 - the uncertainty around each estimate (confidence interval)
 - If meta-analysis used, the overall (pooled) effect estimate and uncertainty around this

Review: Falls clinic effectiveness
 Comparison: 01 All studies
 Outcome: 04 Fall-related injury



Problems and Limitations

Problems and limitations

- The reliability of systematic reviews is dependent on the studies that they include
- Several ways in which bias may occur
- Reviews of RCTs are best studied; problems generally much worse in other reviews

1. Inclusion bias

- If the set of studies is not comprehensive, results of the review may be biased
- Publication bias: studies showing a positive effect more likely to be published and hence included in reviews
- May bias results
- Trial registration helps to prevent this
- Other related biases; language bias, grey literature bias etc.

2. Quality of included studies

- If primary studies are biased, review will also be biased (garbage in, garbage out)
- Danger of a precise but wrong result
- Combining results of a large high quality study with small low quality studies may give worse estimate

2. Quality of included studies

- Major problem with poor reporting
- Often impossible to tell whether conduct of study was of adequate quality
- E.g. randomisation: often procedures are not described

Conclusions

- Traditional narrative reviews are outdated and should no longer be conducted (but still are!)
- Systematic review methodology offers major advantages in the strength of conclusions that can be drawn
- It should always be used for reviewing of research