An Introduction to
Systematic Reviews

Susan L. Norris, MD, MPH, MS
Associate Professor
Oregon Health & Science University
Portland, OR
norris@ohsu.edu

Biography, SL Norris

• MD, MS, University of Alberta
• MPH, University of Washington
• Board Certified: general surgery (Canada), family
  medicine (US)
• Clinical practice: Group Health Cooperative,
  Washington state, 1990-1999
• CDC: directed systematic review group with focus on
  diabetes, 1999-2004
• Oregon Health & Science University (current):
  – clinical practice guideline development
  – systematic review methodology (nonrandomized studies,
    sources of bias in reviews)
  – effects of physician-industry relationships on practice
    guidelines and systematic reviews

Disclosures

• Financial: none
• Intellectual
  – Member, GRADE Working Group
  – Methods work: nonrandomized studies, sources of bias in
    reviews
  – Impact of physician-industry relationships on primary research
    and evidence synthesis
• Professional
  – Investigator, Evidence-based Practice Center (AHRQ)
  – Funders: CDC, NIH, AHRQ, American College of Chest
    Physicians, American Urological Association
Ben-Adhem picked up a stone from beside the road. "It had written on it, "Turn me over and read." So he picked it up and looked at the other side.

And there was written, "Why do you seek more knowledge when you pay no heed to what you know already?"

Shah (1968)

Learning Objectives

1. To define the main characteristics of systematic reviews and to differentiate them from narrative reviews.
2. To outline the steps involved in producing a systematic review.
3. To understand potential sources of bias in reviews.
4. To outline five characteristics of a high-quality systematic review and to apply those to a review published in the peer reviewed literature.
5. To learn how to find systematic reviews in Medline or in repositories of systematic reviews.
Part 1.
Introduction

Knowing What Works in Health Care: IOM Report 2008

“...decisions about the care of individual patients should be based on the conscientious, explicit, and judicious use of the current best evidence on the effectiveness of clinical services.”
Part 2. Systematic Reviews: Typology and Terminology
Typology

- Integrative publication
- Systematic review
- Narrative review
- Nonsystematic review
- Review of comparative effectiveness
- Qualitative synthesis
- Quantitative synthesis
- Meta-analysis
- Practice guidelines
- Economic evaluation
- Decision analysis

Integrative Publications

Combine the results of multiple studies
- Reviews
  - Narrative, quantitative
  - Systematic, nonsystematic
- Practice guidelines
- Economic evaluations
- Clinical decision analysis

Types of Reviews

- Nonsystematic (narrative or traditional)
  - Widespread (e.g., NEJM)
  - Typically, 1 or 2 experts write the review based on literature the authors are familiar with
  - Conclusions related more to the experiences and views of the authors than to the evidence
  - Different authors might reach different conclusions on the same topic or even on the same evidence
  - Small but important effects may be missed
- Systematic
Systematic Review

A concise summary of the best available evidence that addresses a sharply defined clinical question (Mulrow 1998)

- Qualitative synthesis = narrative summary
- Quantitative synthesis = meta-analysis

Systematic Review

“In its ideal form, is a review that includes an explicit and detailed description of how it was conducted so that any interested reader would be able to replicate it” (Jadad 1998)

Systematic Reviews: Aims

- Define the extent and limits of the evidence
- Facilitate decision making
  - Do tell you what to do
  - Other factors to consider
    - Equity, judgment, patient preferences, resource constraints
- Clarify what practice/recommendation is based on evidence and what is based on other opinion
- Help to understand inconsistencies
- Help to understand potential applicability to specific situations or settings
- Identify research gaps
Attributes of Systematic Reviews

T  Transparent
R  Reproducible
U  Unbiased
E  Exhaustive
E  Explicit methods

Reviews of Comparative Effectiveness

Comparative effectiveness
= the comparison of effective interventions among patients in typical patient care settings, with decisions tailored to individual patient needs

Sox and Greenfield, 2009

Meta-Analysis

A systematic review that employs statistical methods to combine and summarize the results of several studies (Cook 1995)

Other terms:
- Quantitative synthesis
- Pooling, pooled analysis
When is a Systematic Review Needed?

1. When there is a substantive clinical or public health question
2. Uncertainty exists
   - Effects of an intervention, specific population or setting
   - Balance of benefits and harms
   - Small effects in individual studies
   - Need to explain inconsistency in results across studies
3. Need to define knowledge gaps, develop a research agenda
4. Several primary studies exist
   – Often with disparate findings
   – There may be situations where there is a need to establish the lack of evidence

Steps in a Systematic Review

• Develop the review question
• Develop inclusion/exclusion criteria
• Search for literature
• Quality assess individual studies
• Data abstraction and analysis
• Synthesis of findings
• Grading the strength of evidence

Part 3. Formulating an Answerable Systematic Review Question
Formulating an Answerable Question

- Most important step in the review
- Depends on perspective
- Results have impact on patient care
- Not dependent on available data
- Considerations:
  - Common conditions
  - Areas of uncertainty
  - Resource-intensive
  - New treatments
  - Heterogeneity of effects

Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials

GCS Smith, J Pell
BMJ 2003
Levels of questions

- Can it work?
- Does it work?
- When/in whom does it work?
- Is it worth it?
- What research is needed?

Brian Haynes, EBM

What is the Purpose of Different Questions?

- Can it work?
  - Useful for: Coverage decisions
- How well does it work overall?
  - Useful for: Quality measurement and improvement
- When/in whom does it work?
  - Useful for: Patient/clinician decisions; health systems
- Is it worth it?
  - Useful for: Informing patient choice
- What kind of research is needed?
  - Researchers, funding agencies

What is Needed to Answer Those Questions?

- Can it work?
  - Studies of high internal validity + quantitative synthesis
- How well does it work overall?
  - Studies with external & internal validity
- When/in whom does it work?
  - External validity & internal validity
- Is it worth it (benefits vs. harms)?
  - Quantitative synthesis + decision models
- What kind of research would be most useful?
  - Qualitative synthesis, broad inclusion criteria
Formulating a Review Question

P  Person, population  
I  Intervention, exposure  
C  Comparator  
O  Outcome  
(T)  Timing  
(S)  Setting, study design

Example: The Well-Formulated Question

Intervention  
Outcomes  
Population  
Condition of Interest

Does anticoagulation therapy improve outcomes in early treatment of patients with ischemic stroke?

- Heparin
- Dicumarol
- CY 222sc
- Phenindione
- Kabi 2165sc

- Recurrent Stroke
- Extra/Intracranial haemorrhage
- Pulmonary Embolism
- Venous Thrombosis
- Death

- ≥18 yrs.
- Hospitalized ≤48h from onset

Abrupt impairment of brain function caused by a variety of pathologic changes involving cranial blood vessels.
Setting inclusion and Exclusion Criteria

What Evidence is Admissible?
- Consider topic by topic
- One size does not fit all
- Importance of protocol
- Educated decisions *a priori*
- Be explicit
- Document and defend changes
- Starts with a clear question

What Evidence is Admissible?
- **P** Person, population
- **I** Intervention, exposure
- **C** Comparator
- **O** Outcome
- **T** Timing
- **S** Setting, study design
Inclusion Criteria for a Drug Class Review (TZDs)

Populations:
- Adults with type 2 diabetes:
  ADA or WHO definition
- Adults with prediabetes:
  Defined as...
- Adults with metabolic syndrome:
  ATP definition...
  - ↑ TG level
  - ↓ HDL-cholesterol
  - ↑ BP
  - ↑ FPG

Part 4. Searching the Literature

Image from cartoonbank.com removed.
Image description: Picture of sheriffs in forest with 3 bloodhound dogs; Dogs are using a laptop on the ground. Caption reads: First, they do an online search.
Literature Searches for a Systematic Review

- **Purpose:**
  - To locate all relevant studies
  - Comprehensive, unbiased collection of studies:
    - Failure to search and identify all relevant sources can lead to a biased review
- **Challenges**
  - Locating literature indexed in electronic databases
  - Locating literature not-indexed
  - Identifying, locating unpublished studies
  - Balance resource allocation and return

Methods of Searching for Relevant Studies

- **Electronic Databases**
  - Medline, Cochrane, PsychINFO, Embase, etc.
- **Trial registries**
- **Hand searching**
- **Grey literature searches**
- **Experts**
- **Reference lists of reviews and included studies**

Rationale for Comprehensive Searching

- **Limitations of the search methods**
  - Only a small proportion of citations that are identified by comprehensive electronic searching (eg, MEDLINE) are relevant
  - The typical MEDLINE search retrieves 50%-80% of relevant studies
  - Ability to search more precisely depends on database
Rationale for Comprehensive Searching

• Non-English language articles
  – not easily accessible on MEDLINE
  – Non-English language articles are of similar quality to English language articles and results are negative more often

• Publication Bias
  – 25-50% of studies are never published, even as abstracts
  – Studies with negative findings less likely to be published

Rationale for Comprehensive Searching

Failure to search and identify all relevant sources can lead to a biased review

Part 5. Assessing the Risk of Bias of Individual Studies
Bias

"a systematic error, or deviation from the truth, in results or inferences"

Cochrane Manual, 9/08

Quality

"extent to which the authors conducted their research to the highest possible standards"

Cochrane Manual, 9/08

Internal validity =
The extent to which the results of a study can be reliably attributed to the intervention under evaluation

vs. external validity = applicability of results to other populations and settings
vs. precision = measure of likelihood of random errors
Why Assessment of Risk of Bias is Important

– Assess and minimize bias in the review
– Guide interpretation of findings

Hierarchy of Evidence

STUDY DESIGN

- Randomized Controlled Trials
- Cohort Studies and Case Control Studies
- Case Reports and Case Series
- Expert opinion

Sources of bias in the production and dissemination of evidence

Methodological quality

- 1. Selection bias
- 2. Information bias
- 3. Measurement bias
- 4. Attrition bias
- 5. Performance bias
- 6. Detection bias
- 7. Reporting bias
- 8. External validity bias
- 9. Random error bias
- 10. Bias due to index tests
- 11. Bias due to reference standards
- 12. Bias due to selection of the筛
- 13. Information bias
- 14. Measurement bias
- 15. Performance bias
- 16. Selection bias
- 17. Reference standard bias
- 18. Index test bias
- 19. Reference standard bias
- 20. Index test bias
- 21. Reference standard bias
- 22. Index test bias
- 23. Reference standard bias
- 24. Index test bias

Swedell J Clin A 1979
Classification Scheme for Bias

- Selection bias
- Performance bias
- Attrition bias
- Detection bias
- Reporting bias

Bias in RCTs

Sample

Randomisation

Treatment

Outcome assessment

Analysis
Risk of Bias Tool

- For assessing RCTs only
- Developed by international group (largely Cochrane based)
- Described in detail in Cochrane manual (9/08) www.cochrane.org

Components of RoB Tool

- Sequence generation
- Allocation concealment
- Blinding (by outcome)
- Incomplete outcome data (by outcome)
- Selective outcome reporting
- Other

Cochrane Risk of Bias Tool

<table>
<thead>
<tr>
<th>Domain</th>
<th>Domain description</th>
<th>Sequence generation</th>
<th>Allocation concealment</th>
<th>Blinding of participants, personnel and outcome assessors</th>
<th>Incomplete outcome data</th>
<th>Selective outcome reporting</th>
<th>Other sources of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No information</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
</tr>
<tr>
<td>Yes</td>
<td>No information</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
</tr>
<tr>
<td>Yes</td>
<td>No information</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
<td>Yes / No / Unclear</td>
</tr>
</tbody>
</table>
What do You do with the RoB Assessment?

- Stratify studies by RoB
- Formal comparisons of effects by RoB with meta-regression or test for differences across subgroups
- Present all studies and provide a narrative discussion
- Restrict primary analysis of all studies and then of low, or low and unclear RoB
- Exclude high RoB studies in synthesis
- Sensitivity analysis: how are conclusions affected if high RoB studies were included?

Part 6.
Evidence Synthesis

Synthesis (noun)

1 a : the composition or combination of parts or elements so as to form a whole; c : the combining of often diverse conceptions into a coherent whole;

Merriam Webster Online Dictionary, 2009
Synthesis in Systematic Reviews

- Explore relationships in the data
  - Describe patterns across included studies
    - Direction of effects
    - Size of effects
  - Draw conclusions about the likely size and direction of effect
- Assess robustness of the data and conclusions
  - Strength of the evidence
  - Applicability of conclusions on effect size to different populations and/or contexts


Additional Elements of Synthesis in Systematic Reviews

- Identify gaps in the evidence
  - Little or no *direct* evidence
  - Poor quality evidence
  - Variation that is not explainable
  - Future research recommendation
    - Based on findings
    - Specific enough for funders

Additional Elements of Synthesis in Systematic Reviews

- Limitations of the review
  - Scope
    - Any important elements not addressed
  - Search strategies
    - Databases searched
    - Language restrictions
  - Study selection and abstraction
    - Two reviewers vs one
    - Indirect evidence
  - Limitations of analyses
    - Low power due to few studies
Quantitative Synthesis: Meta-analysis

- Pool results of a specific outcome across multiple studies
- Studies are weighted by a measure of dispersion from each study
- Provides an overall point estimate of effect
  - Narrower confidence interval
- May be appropriate when:
  - limited sample sizes prevent finding a difference where one may exist
  - studies have conflicting results: assess heterogeneity
- Designed to address direct evidence for a single outcome of interest
- Data obtained (ideally) from a systematic review

Meta-analysis: Enoxaparin versus Unfractionated Heparin


Part 7.
Grading the Strength of Evidence Across Studies
Grading a Body of Evidence

- Subjective summary
  - Assessor applies a descriptive word to evidence
    - e.g., good, fair, poor
  - Imprecise and subjective
- Formal systems for grading a body of evidence
  - Many systems exist, most flawed
  - Gives greater weight to better designs
  - Difficult to apply when contradictory evidence exists
  - Different systems consider different study features: quality, design, consistency, number of studies, etc.

Grades of Recommendation Assessment, Development, and Evaluation (GRADE)

- Since 2000
- Researchers/guideline developers with interest in methodology
- Aim: to achieve a common, transparent and sensible system for grading the quality of evidence and the strength of recommendations

GRADE Uptake

- World Health Organization
- National Institute Clinical Excellence (NICE)
- Agency for Healthcare Research and Quality (AHRQ)
- Canadian Agency for Drugs and Technology in Health
- Allergic Rhinitis in Asthma Guidelines (ARIA)
- American Thoracic Society
- American College of Chest Physicians
- UpToDate
- British Medical Journal
- American College of Physicians
- Cochrane Collaboration
- European Society of Thoracic Surgeons
- Clinical Evidence
- American Urological Association
- Many other organizations
Grading Body of Evidence: GRADE method

- Directness of evidence
  - Health outcomes vs intermediate outcomes, efficacy outcomes
  - Direct comparisons vs indirect comparisons
- Precision of estimates
  - Degree of certainty surrounding an effect estimate with respect to a given outcome
- Consistency of results
  - Effect sizes have the same sign
  - The range of effect sizes is narrow
- Risk of bias
  - Study design
  - Aggregate risk of bias of the studies under consideration

*GRADE Working Group: wwwGRADEWorkingGroup.org

Rate Quality of the Evidence

= Confidence in an estimate of effect

- High: considerable confidence in estimate of effect
- Moderate: further research may change estimate
- Low: further research is very likely to impact the estimate
- Very low: estimate of effect is very uncertain
Part 8.
Resources for Identifying or Performing High Quality Systematic Reviews
Archie Cochrane

Image of Archive Cochrane removed.

The Cochrane Collaboration

Image of Cochrane logo removed.
Recommendations on reporting of systematic reviews

- QUORUM, 1999
  - Quality of Reporting of Meta-analyses

- MOOSE, 2000
  - Meta-analysis of observational studies in epidemiology (Stroup et al. JAMA April 2000)

- PRISMA, 2009

Quality Assessment of Systematic Reviews: AMSTAR

11-Items answered as:
- Yes
- No
- Can’t answer
- Not Applicable
  - Addresses review planning, searching, quality assessment, synthesis, conflicts of interest

Shea et al. BMC Medical Research Methodology 2007

Locating Systematic Reviews

- Search Strategies in Medline
  - Study type: systematic review
  - EBM reviews
  - “systematic review” or “Medline” as text term in title or abstract
- Key things to look for:
  - Year of last search
  - Inclusion criteria: PICOTS
  - Quality of the review
Organizations that Produce High Quality Systematic Reviews

- Agency for Healthcare Research and Quality
  - Effective Healthcare Program (effectivehealthcare.ahrq.gov)
  - Evidence-based Practice Center program (www.ahrq.gov)
- US Preventive Services Task Force
  - Supported by EPC program (www.ahrq.gov/clinic/uspstfix.htm)
- Cochrane Collaboration reviews (www.cochrane.org)
- Health Technology Assessment programs
  - E.g. UK NCCHTA (www.ncchta.org)
- National Institute for Clinical Excellence (www.nice.org.uk)
- Drug Effectiveness Review Project (www.ohsu.edu/drugeffectiveness)

Resources for Performing Systematic Reviews

- Organizations
  - Cochrane Collaboration: www.cochrane.org
  - NICE, SIGN
- Books
  - Systematic Reviews in Health Care: Meta-Analysis in Context: Eggers et al. BMJ Books

Caveat Emptor: Systematic Reviews

- SRs should give the best estimate of any true effect, but
  - MUST be well done
- Findings must be taken in concert with other evidence and factors
  - Evidence outside the scope of the review may be relevant
  - Cost and implementation implications
  - Generalizability of findings
- Readers sometimes have unrealistic expectations
  - Limitations of the available evidence
    - Lack of evidence:
      - No large and/or good quality trials
      - No trials at all
      - Interventions or populations of interest not specifically studied
  - Limitations of a review process
    - The review does not make decisions for you
    - An individual review may not answer all questions related to a given decision
    - While a systematic approach reduces bias, it does not eliminate it
- Reviews become outdated
Suggested Readings


References & Resources


References & Resources


References & Resources