KNOWLEDGE SYNTHESIS:

Literature Searches and Beyond

Ahmed M. Abou-Setta, MD, PhD
Department of Community Health Sciences &
George & Fay Yee Centre for Healthcare Innovation
University of Manitoba
Email: Ahmed_Abou-Setta@umanitoba.ca
Disclosure of Interest

• At the Centre for Healthcare Innovation, we conduct both systematic reviews and rapid reviews, and I have been an author on various funded reviews.
Disclosure of Interest

• At the Centre for Healthcare Innovation, we conduct both systematic reviews and rapid reviews, and I have been an author on various funded reviews

• I have been involved with development of rapid review methodology for the World Health Organization
Ice Breaker!

You tell us...

- What’s your academic background?
- What is your past exposure to Knowledge Synthesis (courses, work experience) ?
- What’s your expectations from this workshop?
Goals

At the end of the workshop, you should have a better idea about:

– What is Knowledge Synthesis
– Why we prefer to be ‘systematic’ in reviewing the evidence
– Different types and approaches to ‘systematically’ review the evidence
– Analysis: when, why, and how
– Different ways of presenting the resulting information to multiple audiences, including key stake-holders, policy-makers, and the public
– Some of the major challenges in conducting knowledge synthesis activities
Archie Cochrane

In 1979 stated that “it is surely a great criticism of our profession that we have not

1) organised a critical summary,

2) by specialty or subspecialty,

3) adapted periodically, of all relevant randomised controlled trials”
Evidence informed decision making

- Meta-Analyses
- Systematic Reviews
- Critically Appraised Topics
- Practice Guidelines
- Critically Appraised Articles
- Randomized Controlled Trials
- Cohort Studies
- Case-Controlled Studies, Case Series/Reports
- Textbooks, Point of Care Resources, Encyclopedic Medical Websites
- Expert Opinion, Personal Experience, Newspapers, Magazines, Blogs

Quality of evidence

Filtered Information
Why Should We Care?

- Primary research is often false and/or biased
Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. In this essay, I discuss the implications of these problems for the conduct and interpretation of research.

Factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a p-value less than 0.05. Research is not most appropriately represented and summarized by p-values, but, unfortunately, there is a widespread notion that medical research articles should be interpreted based only on p-values. Research findings are defined here as any relationship reaching formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. “Negative” research is also very useful.

It can be proven that most claimed research findings are false.

It is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R + 1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, $\alpha$. Assuming that $c$ relationships are being probed in the field, the expected values of the $2 \times 2$ table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Wacholder et al. have called the false positive report probability [10]. According to the 2
Why Should We Care?

- Primary research is often false and/or biased
- Evidence reviews are highly publishable (even with negative findings)
Publications per year

Knowledge Synthesis Publications
Randomized Controlled Trials
Knowledge synthesis Publications

Canada

Manitoba
Why Should We Care?

• Primary research is often false and/or biased
• Evidence reviews are highly publishable (even with negative findings)
• Governments/ funding agencies understand the need for these types of reviews and are willing to fund them
• Many grants will not be accepted without a systematic search showing the need for more primary research
Identifying knowledge gaps

Redundant and unethical
What's Your Question?
Population or problem

• How would you describe the population you are interested in studying?
• What characteristics are important (e.g. age, gender, social attributes, etc.)

Intervention

What intervention or policies (e.g. education, crime prevention, etc.) are you interested in?

Comparison

What alternative or different options do you want to compare against (e.g. status quo)?

Outcome

• What outcome(s) are you interested in?
• How do you measure success?
• How do you measure failure?
• How detrimental outcomes can occur?
Activity #1

* What is your PICO?
What is Knowledge Synthesis?
What is Knowledge Synthesis?

“The contextualization and integration of research findings of individual research studies within the larger body of knowledge on the topic.”

“A synthesis must be reproducible and transparent in its methods, using quantitative and/or qualitative methods.”

Canadian Institutes of Health Research
Example of Knowledge Synthesis

- Systematic review
- Realist syntheses
- Narrative syntheses
- Meta-analyses
- Meta-syntheses
- Practice guidelines
- Consensus conference or expert panel
What is a Literature Review?
What is a Literature Review?

‘an overview of research on a given topic and answers to related research questions’
What is a Literature Review

• Key characteristics:
  – *Organizes* the literature  
    (make sense of it all... connect the dots)
  – *Evaluates* the literature  
    (high quality to low quality)
  – *Identifies patterns and trends* in the literature
  – *Synthesises* the literature  
    (high quality to low quality)

• But it’s not:
  – *Annotated Bibliography*
  – *One-stop shop* for everything related to a topic
  – *Book review*
What is a Systematic Review?
What is a Systematic Review?

‘an attempt to gather all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research or clinical question through a reproducible, transparent process’
Advantages

- Reduce bias – by gathering ‘all’ evidence
- Transparent
- Replicable
- Resolves controversy between conflicting studies
- Identifies gaps in current research
- Can be basis of cost-effectiveness analyses and knowledge translation projects
- Provides a reliable basis for decision making
# Systematic vs. Literature Review

<table>
<thead>
<tr>
<th>Feature</th>
<th>Literature Review</th>
<th>Systematic Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Question</td>
<td>Broad</td>
<td>Focused</td>
</tr>
<tr>
<td>Search strategy/sources of info</td>
<td>Not usually specified</td>
<td>Comprehensive/explicit search strategy</td>
</tr>
<tr>
<td>Selection criteria</td>
<td></td>
<td>Criterion-based selection; uniformly applied</td>
</tr>
<tr>
<td>Quality assessment</td>
<td></td>
<td>Rigorous critical appraisal</td>
</tr>
<tr>
<td>Synthesis</td>
<td>Qualitative summary</td>
<td>Quantitative summary (usually)</td>
</tr>
<tr>
<td>Conclusions</td>
<td>Based on a sample of the evidence</td>
<td>Based on all available evidence</td>
</tr>
<tr>
<td>Grading</td>
<td>Sometimes performed</td>
<td>Strength of Evidence is Graded</td>
</tr>
</tbody>
</table>
Common Types of Reviews

- Intervention Review (most common)
  - Evidence about the effects of a healthcare intervention
Association Between Omega-3 Fatty Acid Supplementation and Risk of Major Cardiovascular Disease Events
A Systematic Review and Meta-analysis

Evangelos C. Rizos, MD, PhD
Evangelia E. Ntzani, MD, PhD
Eftychia Bika, MD
Michael S. Kostapanos, MD
Moses S. Eliasf, MD, PhD, FASA, FRSH

Treatment with marine-derived omega-3 polyunsaturated fatty acids (PUFAs) for the prevention of major cardiovascular adverse outcomes has been supported by a number of randomized clinical trials (RCTs) and refuted by others. Although their mechanism of action is not clear, their postulated effect on cardiovascular outcomes may be due to their ability to lower triglyceride levels, prevent serious arrhythmias, or even decrease platelet aggregation and lower blood pressure. Current guidelines issued by major societies recommend their use, either as supplements or through dietary counseling, for patients after myocardial infarction (MI), whereas the US Food and Drug Administration has approved their administration.

Context Considerable controversy exists regarding the association of omega-3 polyunsaturated fatty acids (PUFAs) and major cardiovascular end points.

Objective To assess the role of omega-3 supplementation on major cardiovascular outcomes.

Data Sources MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials through August 2012.

Study Selection Randomized clinical trials evaluating the effect of omega-3 on all-cause mortality, cardiac death, sudden death, myocardial infarction, and stroke.

Data Extraction Descriptive and quantitative information was extracted; absolute and relative risk (RR) estimates were synthesized under a random-effects model. Heterogeneity was assessed using the Q statistic and I². Subgroup analyses were performed for the presence of blinding, the prevention setting, and patients with implantable cardioverter-defibrillators, and meta-regression analyses were performed for the omega-3 dose. A statistical significance threshold of .0063 was assumed after adjustment for multiple comparisons.

Data Synthesis Of the 3635 citations retrieved, 20 studies of 68,680 patients were included, reporting 7044 deaths, 3993 cardiac deaths, 1150 sudden deaths, 1837 myocardial infarctions, and 1490 strokes. No statistically significant association was observed with all-cause mortality (RR, 0.96; 95% CI, 0.91 to 1.02; risk reduction [RD] −0.004; 95% CI, −0.01 to 0.002), cardiac death (RR, 0.91; 95% CI, 0.85 to 0.98; RD, −0.01; 95% CI, −0.02 to 0.00), sudden death (RR, 0.87; 95% CI, 0.75 to 1.01; RD, −0.003; 95% CI, −0.01 to 0.006), myocardial infarction (RR, 0.89; 95% CI, 0.76 to 1.04; RD, −0.002; 95% CI, −0.007 to 0.002), and stroke (RR, 1.05; 95% CI, 0.93 to 1.18; RD, 0.001; 95% CI, −0.002 to 0.004) when all supplement studies were considered.

Conclusion Overall, omega-3 PUFA supplementation was not associated with a lower risk of all-cause mortality, cardiac death, sudden death, myocardial infarction, or stroke based on relative and absolute measures of association.

JAMA. 2012;308(10):1024-1033

www.jama.com
Common Types of Reviews

• Intervention Review (most common)
  – Evidence about the effects of a healthcare intervention

• Diagnostic Accuracy Review
  – Evidence around diagnostic accuracy of different screening tests
Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review

Ann Van den Bruel, Tonya Haj-Hassan, Matthew Thompson, Frank Buntinx, David Mant, for the European Research Network on Recognising Serious Infection Investigators

Summary
Background Our aim was to identify which clinical features have value in confirming or excluding the possibility of serious infection in children presenting to ambulatory care settings in developed countries.

Methods In this systematic review, we searched electronic databases (Medline, Embase, DARE, CINAHL), reference lists of relevant studies, and contacted experts to identify articles assessing clinical features of serious infection in children. 1939 potentially relevant studies were identified. Studies were selected on the basis of six criteria: design (studies of diagnostic accuracy or prediction rules), participants (otherwise healthy children aged 1 month to 18 years), setting (ambulatory care), outcome (serious infection), features assessed (assessable in ambulatory care setting), and sufficient data reported. Quality assessment was based on the Quality Assessment of Diagnostic Accuracy Studies criteria. We calculated likelihood ratios for the presence (positive likelihood ratio) or absence (negative likelihood ratio) of each clinical feature and pre-test and post-test probabilities of the outcome. Clinical features with a positive likelihood ratio of more than 5.0 were deemed red flags (ie, warning signs for serious infection); features with a negative likelihood ratio of less than 0.2 were deemed rule-out signs.

Findings 30 studies were included in the analysis. Cyanosis (positive likelihood ratio range 2.66-52.20), rapid breathing (1.26-9.78), poor peripheral perfusion (2.39-38.80), and petechial rash (6.18-83.70) were identified as red flags in several studies. Parental concern (positive likelihood ratio 14.40, 95% CI 9.30-22.10) and clinician instinct (positive likelihood ratio 23.50, 95% CI 16.80-32.70) were identified as strong red flags in one primary care study. Temperature of 40°C or more has value as a red flag in settings with a low prevalence of serious infection. No single clinical feature has rule-out value but some combinations can be used to exclude the possibility of serious infection—for example, pneumonia is very unlikely (negative likelihood ratio 0.07, 95% CI 0.01-0.46) if the child is not short of breath and there is no parental concern. The Yale Observation Scale had little value in confirming (positive likelihood ratio range 1.10-6.70) or excluding (negative likelihood ratio range 0.16-0.97) the possibility of serious infection.

Interpretation The red flags for serious infection that we identified should be used routinely, but serious illness will still be missed without effective use of precautionary measures. We now need to identify the level of risk at which clinical action should be taken.

Funding Health Technology Assessment and National Institute for Health Research National School for Primary Care Research.
Common Types of Reviews

• Intervention Review (most common)
  – Evidence about the effects of a healthcare intervention

• Diagnostic Accuracy Review
  – Evidence around diagnostic accuracy of different screening tests

• Prognostic Review
  – Evidence of models or predictors of patient outcomes
Influence of smoking cessation after diagnosis of early stage lung cancer on prognosis: systematic review of observational studies with meta-analysis

A Parsons, research fellow,1 A Daley, senior lecturer, NIHR career scientist,2 R Begh, research associate,1 P Aveyard, clinical reader, NIHR career scientist1

ABSTRACT

Objective To systematically review the evidence that smoking cessation after diagnosis of a primary lung tumour affects prognosis.

Design Systematic review with meta-analysis.


Study selection Randomised controlled trials or observational longitudinal studies that measured the effect of quitting smoking after diagnosis of lung cancer on prognostic outcomes, regardless of stage at presentation or tumour histology, were included.

Data extraction Two researchers independently identified studies for inclusion and extracted data. Estimates were combined by using a random effects model, and the I² statistic was used to examine heterogeneity. Life tables were used to model five year survival for early stage non-small cell lung cancer and limited stage small cell lung cancer, using death rates for continuing smokers and quitters obtained from this review.

Conclusions This review provides preliminary evidence that smoking cessation after diagnosis of early stage lung cancer improves prognostic outcomes. From life table modelling, the estimated number of deaths prevented is larger than would be expected from reduction of cardiorespiratory deaths after smoking cessation, so most of the mortality gain is likely to be due to reduced cancer progression. These findings indicate that offering smoking cessation treatment to patients presenting with early stage lung cancer may be beneficial.

INTRODUCTION

Worldwide, lung cancer is the most commonly diagnosed form of cancer.1 In the United Kingdom, its annual incidence is second only to that of breast cancer, accounting for around 39 000 new cancer diagnoses annually.2 In countries that have seen a high prevalence of smoking, around 90% of diagnoses of lung cancer are attributable to cigarette smoking.3 The increased incidence from smoking is proportional to the length and intensity of smoking history.4 On average, a lifetime smoker has a 20-fold increase in the risk of developing lung cancer compared with a
Other Types of Systematic Reviews

• Overviews of reviews
  – Evidence from already published *systematic reviews* for a given topic or disciplinary area
Alma-Ata: Rebirth and Revision 2

Supporting the delivery of cost-effective interventions in primary health-care systems in low-income and middle-income countries: an overview of systematic reviews

Simon Lewin, John N Lavis, Andrew D Oxman, Gabriel Bastias, Mickey Chopra, Agustín Ciapponi, Signe Flottorp, Sebastian García Martí, Tomas Pantoja, Gabriel Roda, Nathan Souza, Shaun Treweek, Charles S Wiysonge, Andy Haines

Strengthening health systems is a key challenge to improving the delivery of cost-effective interventions in primary health care and achieving the vision of the Alma-Ata Declaration. Effective governance, financial and delivery arrangements within health systems, and effective implementation strategies are needed urgently in low-income and middle-income countries. This overview summarises the evidence from systematic reviews of health systems arrangements and implementation strategies, with a particular focus on evidence relevant to primary health care in such settings. Although evidence is sparse, there are several promising health systems arrangements and implementation strategies for strengthening primary health care. However, their introduction must be accompanied by rigorous evaluations. The evidence base needs urgently to be strengthened, synthesised, and taken into account in policy and practice, particularly for the benefit of those who have been excluded from the health care advances of recent decades.

Introduction

In 1978, representatives from 134 countries gathered in Alma-Ata in the former USSR and declared that primary health care, “based on practical, scientifically sound and socially acceptable methods and technology made universally accessible through people’s full participation”, was key to delivering health for all by the year 2000. Recent years have seen a renewed interest in primary resources needed to improve delivery of cost-effective interventions; and the fragmented and weakened state of health systems in many countries.

More generally, there have been calls to redress the balance between the now dominant vertical, disease-focused programmes and the horizontal, systems-focused perspective that underpins most approaches for
Other Types of Systematic Reviews

• Overviews of reviews
  – Evidence from already published systematic reviews for a given topic or disciplinary area

• Scoping reviews
  – Map out previous primary research and systematic reviews for a given topic or disciplinary area

• Rapid reviews
  – Rapidly assess (usually ≤ 6 weeks) the evidence about the effects of a healthcare intervention
Other Types of Systematic Reviews

• Realist reviews
  – Deal with complex ‘system’ issues and attempt to provide explanation rather than judgment (e.g. answer questions like ‘how’, ‘why’, and for ‘whom’)

• Provincial reviews
  – Provides overview on what is happening in other jurisdictions/regions/states/provinces/countries (e.g. answer questions like ‘what are others doing’)

The Effectiveness and Safety of Preschool Hearing Screening Programs

November 2012
Learning from success
Questions
Activity #2

* What was their question?
* Was the chosen review type the best suited to answer this question?
Activity #3

* What is your question?
* What review type is best suited to answer this question?
Main steps involved in the conduct of a ‘traditional’ systematic review
Define a focused 4-part review question (Patient, Intervention, Comparison and Outcome)

PubMed, Embase, Web of Science, Cochrane CENTRAL and subject specific databases; Contact authors, experts, comments, citation tracking

Use filters for specific study designs (e.g. PubMed Clinical Queries filters, and Cochrane filters for RCTs)

Run searches on all relevant databases and sources

Save all citations (titles/abstracts) in a reference manager Document search strategies that were employed These citations are ready for first screen (N₀)

Reviewer 1 screens all titles/abstracts and makes selections for second screen

Reviewer 2 screens all titles/abstracts and makes selections for second screen

Reviewers meet and resolve disagreements on citations they do not agree on. The final number (N) selected after this process is ready for second screen (review of full-text articles)

Get full texts of all articles identified for second screen (N)

Excluded after second screen

Articles considered eligible after full text review (by two reviewers) is the final set of studies for inclusion (n₀)

Studies included in the final analysis (n₁)

Each article gets a unique ID number

Reviewer 1 extracts data (including quality assessment) from the final selected article

Reviewer 2 extracts data (including quality assessment) from the final selected articles

Collect outcomes as cell values of a 2x2 table, if possible

Contact authors for missing data; email authors short, structured questionnaires; reminders help!

Software suggestions: Access, Excel

Exploration of heterogeneity: graphical methods (e.g. Galbraith plots), subgroup analyses, and meta-regression

Use QUORIM or MOOSE as guides for report writing

Import data and analyse using software Tabulate study characteristics Generate forest plots of effect measures Check for heterogeneity Pool effect measures if heterogeneity is not a concern. If heterogeneity is found, identity sources of heterogeneity Consider subgroup and sensitivity analyses Explore possibility of publication bias

Interpret, discuss results and write the report. Discuss applicability of results and limitations of the review Make recommendations for practice or policy, and research

Check for heterogeneity: Chi-squared or I-squared tests; these tests have low power; consider a conservative p value of <0.10 for significance

You made it! Celebrate!!!

Team

Core team:
- Project coordinator
- Content expert(s)
- Methodologist(s)
- Trial search coordinator(s)
- Analyst(s)
- Researcher assistant(s)

Supportive team:
- Medical editor(s)
- Business manager
- Secretary
- Database manager/ IT support
## Budget

<table>
<thead>
<tr>
<th>Project phase</th>
<th>Project member</th>
<th>Hourly cost</th>
<th>No. of hours</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol preparation/ registration</td>
<td>All team members</td>
<td></td>
<td>2 – 5 days</td>
<td>$3,000 – $9,000</td>
</tr>
<tr>
<td>Literature search</td>
<td>Librarian</td>
<td>$200</td>
<td>2 – 5 days</td>
<td>$3,200 – $8,000</td>
</tr>
<tr>
<td><strong>Screening 5,000 – 10,000 titles/abstracts</strong></td>
<td>2 Research assistants</td>
<td>$200</td>
<td>5 – 10 days</td>
<td>$8,000 – $16,000</td>
</tr>
<tr>
<td><strong>Screening 75 – 100 full-text reports</strong></td>
<td>2 Research assistants</td>
<td>$200</td>
<td>3 – 5 days</td>
<td>$4,800 – $8,000</td>
</tr>
<tr>
<td><strong>Data extraction/ quality assessment for 20 – 40 full-text reports</strong></td>
<td>2 Research assistants</td>
<td>$200</td>
<td>8 – 16 days</td>
<td>$12,800 – $25,600</td>
</tr>
<tr>
<td>Data analysis</td>
<td>Statistician</td>
<td>$200</td>
<td>3 – 5 days</td>
<td>$4,800 – $8,000</td>
</tr>
<tr>
<td>Preliminary presentation of results/discussion with team members &amp; content experts</td>
<td>All team members</td>
<td></td>
<td>$1000</td>
<td>$1000</td>
</tr>
<tr>
<td>Report preparation</td>
<td>Medical writer</td>
<td>$100</td>
<td>5 – 10 days</td>
<td>$4,000 – $8,000</td>
</tr>
<tr>
<td></td>
<td>Review manager</td>
<td>$200</td>
<td>5 – 10 days</td>
<td>$8,000 – $16,000</td>
</tr>
</tbody>
</table>

*$50,000 – $100,000+*
# Timeline

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literature search/ snowballing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level I Screening</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level II Screening</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data extraction/ quality assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preliminary presentation of results</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Report preparation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Considerations:

- If we started today… Probably finish sometime in Dec 2015 (earliest)
- Holidays, sick days, other projects interfere with timeline
- Delays happen… need to add a ‘cushion’ to timeline
- Realistic end date… Mar 2016
- Some reviews can last up to 2 years or longer
What if the decision can’t wait...

1 day

1 week

1 month
Barriers

• Time:
  – Full review: 6 to 8 months
  – Rapid review: 4 to 6 weeks
  – Other review types: variable

• Cost:
  – Full review: $25,000 to over $500,000
  – Rapid review: $5,000 to $25,000
  – Other review types: variable
Limitations

- Empty reviews: no studies/evidence available
- Negative findings: no conclusive results
- Biased results: Studies are often poor quality or at unclear to high risk of bias
- Limited answer to complex questions: results limited to specific clinical question
- Results not directly linked to practice change
Questions
Activity #4

* What is your timeline?
* What is your budget?
* What are your barriers?
Theoretical case study

- A friend’s grandmother fell and broke her hip
- She is fine but had to have immediate surgery
- When you visit her in the hospital, she is drowsy, confused and seems to be in some pain
- The nurse tells you that this is ‘normal’ for her age… but you are not convinced
- So… you (after whipping out your smartphone) decide to search for information on managing pain in hip fracture patients
Theoretical case study

Questions we need to answer…

What’s the question (again)???

What are we looking for???

Where should we search???

Can we ‘trust’ what is being said???

What if we find lots of ‘reports’???

What if these ‘reports’ are conflicting???

How do we ‘tell’ others what we found???
What’s the question (again)???
What’s the question (again)???

- The question will determine the inclusion and exclusion criteria (PICOTS format):
  - **P**opulation of interest
  - **I**nterventions and **C**omparators
  - **O**utcomes of interest

What's the question (again)???
What’s the question (again)???

• The question will determine the inclusion and exclusion criteria (PICOTS format):
  – **P**opulation of interest
  – **I**nterventions and **C**omparators
  – **O**utcomes of interest
  – **T**iming (duration)
  – Appropriate **S**ettings
  – Appropriate **S**tudy designs
What’s the question (again)???

• Poorly formulated question:
  What drugs best manage pain?

• Well formulated question:
  In older adults (≥50 years) admitted to the hospital following acute hip fracture, what is the effectiveness of pharmacologic and/or non-pharmacologic pain management interventions for controlling acute (up to 30 days post-fracture) and chronic pain (up to 1 year post-fracture) compared with usual care or other interventions in all settings?
What are we looking for???
## Inclusion/Exclusion criteria

- **Must be defined a priori**
- **Allows inclusion/exclusion of studies based on objective criteria**

<table>
<thead>
<tr>
<th>Study design</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Randomized controlled trials, nonrandomized controlled trials (e.g. quasi-randomized trials), cohort studies (prospective or retrospective), case-control studies</td>
<td>Observational study designs with no comparison group (case reports, case series, cross-sectional studies)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participants</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Older adults (≥50 years old) of either sex admitted to hospital with acute hip fracture due to low energy trauma</td>
<td>Majority (&gt;80%) of participants &lt;50 years; acute hip fractures due to high energy trauma</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interventions</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pharmacological and/or nonpharmacological pain management monotherapy or combination therapy, regardless of mode of administration or time point during the care pathway</td>
<td>Interventions directly related to surgical/nonsurgical treatment of the hip fracture and not a pain management intervention</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comparator</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Usual care (as defined by study authors) or another intervention(s) for pain management, administered as monotherapy or combination therapy</td>
<td>Initial care for patients is substantially different than the current practices in North America (e.g., based on time to discharge from acute care to subacute care)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primary: Acute and chronic pain Secondary: Mortality, functional status, pain med. Use Adverse events: mental status, quality of life, length of stay</td>
<td>None of the aforementioned outcomes were available from the trial report or through communication with the study’s corresponding author</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Timing</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>From time of trauma leading to acute hip fracture and thereafter</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Setting</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All settings</td>
<td></td>
</tr>
</tbody>
</table>
Where should we search???
Where should we search???

- Common general medical citation databases:
  - Medline
  - PubMed
  - Embase
  - Scopus
  - Web of Knowledge

- Specialty citations also available:
  - CINAHL: nursing literature
  - PsycInfo: psychology and psychiatry literature
  - LILACS: Spanish and Portuguese literature
Figure 2. Overlap of active titles (source type “Journal” only) between the citation databases and the medical databases.

Sources for pre-synthesized evidence

- **Cochrane Library**  [http://www.thecochranelibrary.com](http://www.thecochranelibrary.com)
- **Centre for Reviews and Dissemination**  [http://www.york.ac.uk/inst/crd/](http://www.york.ac.uk/inst/crd/)
  - DARE - Database of Abstracts and reviews of Effects
  - NHS EED – NHS Economic Evaluation Database
  - HTA – Heath Technology Assessment Database
- **Trip database**  [http://www.tripdatabase.com](http://www.tripdatabase.com)
- **MEDLINE/EMBASE**
  - Use appropriate filters or MeSH/EMTREE headings
- **Up-to-date**
- **SR review protocol registries:**  PROSPERO & Cochrane
Can we ‘trust’ what is being said???
Would You Say You “Had Sex” If . . . ?

Stephanie A. Sanders, PhD
June Machover Reinisch, PhD

The degree to which individuals vary with respect to the behavioral criteria involved in labeling an interaction as having “had sex” has implications for both clinical and research purposes. Recent public discourse regarding whether oral-genital contact constitutes having “had sex” highlights the importance of explicit criteria in contrast with implicit assumptions in this area. Unfortunately, a review of the literature demonstrates that empirical exploration of what is included in definitions of having “had sex” for the general public in the United States remains scant. Social and legal definitions of “sex,” “sex act,”

Context The current public debate regarding whether oral sex constitutes having “had sex” or sexual relations has reflected a lack of empirical data on how Americans as a population define these terms.

Objective To determine which interactions individuals would consider as having “had sex.”

Methods A question was included in a survey conducted in 1991 that explored sexual behaviors and attitudes among a random stratified sample of 599 students representative of the undergraduate population of a state university in the Midwest.

Participants The participants originated from 29 states, including all 4 US Census Bureau geographic regions. Approximately 79% classified themselves as politically moderate to conservative.

Main Outcome Measure Percentage of respondents who believed the interaction described constituted having “had sex.”

Results Individual attitudes varied regarding behaviors defined as having “had sex”: 59% (95% confidence interval, 54%-63%) of respondents indicated that oral-genital contact did not constitute having “had sex” with a partner. Nineteen percent responded similarly regarding penile-anal intercourse.

Conclusions The findings support the view that Americans hold widely divergent opinions about what behaviors do and do not constitute having “had sex.”

JAMA. 1999;281:275-277
University Students’ Definitions of Sexual Abstinence and Having Sex

E. Sandra Byers · Joel Henderson · Kristina M. Hobson

Received: 23 March 2007 / Revised: 25 September 2007 / Accepted: 25 September 2007 / Published online: 19 January 2008
© Springer Science+Business Media, LLC 2008

Abstract We asked 298 heterosexual Canadian university students about their definitions of the terms abstinence and having sex. For both terms, students were provided with a list of 17 sexual behaviors and indicated whether they would include each in their definition. The majority of both male and female students included activities that did not involve genital stimulation in their definition of sexual abstinence and did not include these activities in their definition of having sex. Conversely, most students did not include bidirectional sexual stimulation (penile–vaginal intercourse or penile–anal intercourse) in their definitions of sexual abstinence but did include them in the definition of having sex. Students were quite mixed in whether activities involving unidirectional genital stimulation (e.g., oral sex, genital fondling) constituted abstinence, having sex, or neither abstinence nor having sex. However, they were more likely to see these behaviors as abstant than as having sex. Students were more likely to rate a behavior as abstinence if orgasm did not occur. A canonical correlation analysis was used to examine the patterns of association between a number of predictors and inclusions of behaviors involving no genital stimulation, unidirectional stimulation, and bidirectional genital stimulation in abstinence definitions. The results indicated that male participants who were more involved with their religion and sexually conservative, less sexually experienced, and who had not received sexual health education at home were more likely to define bidirectional genital stimulation and less likely to define no genital stimulation and unidirectional sexual stimulation as sexual abstinence. The research and health promotion implications of these results are discussed.

Keywords Sexual abstinence · Having sex · Sexual definitions · Sexual health

Introduction

Sexuality researchers and policy makers invariably rely on self-report data to gather information on the frequency and prevalence of sexual behavior. The accuracy of these data rests, in part, on a common understanding of the questions being asked and the sexual terms being used. Clear definitions and unambiguous wording are especially important to sex education programs and public health initiatives for youth (Smith, Steen, Spaulding-Givens, & Schwendinger, 2003). However, recipients of these programs may rely on their own idiosyncratic understanding of sexual terms such as having sex and abstinence. In fact, a series of studies conducted in the United States, United Kingdom, Australia, and Canada have shown that there is considerable variability in how university students define having sex (Pitts & Rahman, 2001; Randall & Byers, 2003; Richter & Song, 1999; Sanders & Reinsch, 1999; see also Bogart, Cecil, Wagstaff, Pinkerton, & Abramson, 2000; Carpenter, 2001). Although the vast majority of participants (more than 94%) in these four studies included penile–vaginal intercourse in their definitions of sex, fewer participants (between 70% and
Can we ‘trust’ what is being said???

• Quality assessment:

systematic review validity = validity of primary studies

  – Quality of included studies need to be assessed
  – Common tools include:
    • Cochrane Risk of bias tool (randomized trials)
    • New-Castle – Ottawa Scale (cohorts/ case controls)
    • AMSTAR tool (Systematic reviews)
Can we ‘trust’ what is being said???

• Quality assessment:
  – Randomized trials are considered to be more rigorous than observational studies
  – Systematic reviews based on well-designed trials will likely be more valid and accurate
What if we find lots of ‘reports’???
What if we find lots of ‘reports’???

• Meta-analysis:
  – Combining data from different trials to get a summary effect estimate
  – Assumes clinical homogeneity between individual trial PICOTS
  – Assumes a ‘reasonable’ amount of statistical homogeneity between individual trials
# Forest Plot

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Events HES</th>
<th>Total HES</th>
<th>Events Control</th>
<th>Total Control</th>
<th>Weight</th>
<th>Risk Ratio (M-H, Random, 95% CI)</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berard 1995</td>
<td>5</td>
<td>155</td>
<td>4</td>
<td>152</td>
<td>1.0%</td>
<td>1.23 [0.34, 4.48]</td>
<td>1995</td>
</tr>
<tr>
<td>Schortgen 2001</td>
<td>13</td>
<td>65</td>
<td>11</td>
<td>64</td>
<td>3.3%</td>
<td>1.16 [0.56, 2.40]</td>
<td>2001</td>
</tr>
<tr>
<td>Brunkhorst 2008</td>
<td>81</td>
<td>297</td>
<td>51</td>
<td>303</td>
<td>18.0%</td>
<td>1.62 [1.19, 2.21]</td>
<td>2008</td>
</tr>
<tr>
<td>McIntyre 2008</td>
<td>3</td>
<td>21</td>
<td>1</td>
<td>19</td>
<td>0.4%</td>
<td>2.71 [0.31, 23.93]</td>
<td>2008</td>
</tr>
<tr>
<td>James 2011</td>
<td>2</td>
<td>58</td>
<td>3</td>
<td>57</td>
<td>0.6%</td>
<td>0.66 [0.11, 3.78]</td>
<td>2010</td>
</tr>
<tr>
<td>Du 2011</td>
<td>1</td>
<td>21</td>
<td>0</td>
<td>21</td>
<td>0.2%</td>
<td>3.00 [0.13, 69.70]</td>
<td>2010</td>
</tr>
<tr>
<td>Vlachou 2010</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>17</td>
<td></td>
<td>Not estimable</td>
<td>2011</td>
</tr>
<tr>
<td>Ferner 2012</td>
<td>87</td>
<td>400</td>
<td>65</td>
<td>400</td>
<td>20.8%</td>
<td>1.34 [1.00, 1.79]</td>
<td>2012</td>
</tr>
<tr>
<td>Guidet 2012</td>
<td>21</td>
<td>100</td>
<td>11</td>
<td>96</td>
<td>3.9%</td>
<td>1.83 [0.93, 3.59]</td>
<td>2012</td>
</tr>
<tr>
<td>Myburgh 2012</td>
<td>235</td>
<td>3500</td>
<td>196</td>
<td>3500</td>
<td>51.8%</td>
<td>1.20 [1.00, 1.44]</td>
<td>2012</td>
</tr>
</tbody>
</table>

**Total (95% CI)**

<table>
<thead>
<tr>
<th>Events HES</th>
<th>Total HES</th>
<th>Events Control</th>
<th>Total Control</th>
<th>Weight</th>
<th>Risk Ratio (M-H, Random, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4629</td>
<td>4629</td>
<td>100.0%</td>
<td>448</td>
<td>342</td>
<td>1.32 [1.15, 1.50]</td>
</tr>
</tbody>
</table>

Total events: 448, 342

Heterogeneity: Tau² = 0.00; Chi² = 5.07, df = 8 (P = 0.75); I² = 0%

Test for overall effect: Z = 4.08 (P < 0.0001)
What if we find conflicting ‘reports’???
What if we find conflicting ‘reports’???
How do you present the results??
Questions
Where can I search for evidence
Thank You