Critical Appraisal

Refresher Day

Paul Stevenson

April 2018

Study Day **Schedule**

...approximate.

10:00 – 10:30	Introduction and exercise 1
10:30 – 11:00	Thinking about facilitating a tutorial
11:00 – 11:30	Coffee and read paper for apprasal
11:30 – 13:00	Appraisal of an RCT
13:00 – 13:30	Lunch
13:30 – 14:00	Thinking about teaching statistics
14:00 - 15:00	Appraisal of a Systematic Review
15:00 – 15:30	Coffee and Break
15:30 – 16:00	Appraisal of Qualitative Research

Study Day **Format**

Aims:

– Refresh and develop critical appraisal skills & knowledge

- Increase confidence to deliver critical appraisal training to others.

1.Practical Exercises – Critical Appraisal of articles

2. Statistics – explore how to interpret common statistical methods

3. Delivering your own courses – Examples and things to consider

4. Learner led – Ask questions, support each other, adapt pace and focus to our outcomes.

Realistic **Expectations**!

We will cover a lot in one day.....but we won't cover everything

MSc in EBHC Medical Statistics

This is a course for health professionals who wish to strengthen their statistical skills and ability to analyse data. Students will gain the confidence in carrying out the methods that are widely used in medical statistics, and interpreting the results for the practice of evidence-based health care.

This is a joint programme between the Nuffield Department of Primary Care Health Sciences and the Department for Continuing Education's Continuing Professional Development Centre. The Programme works in collaboration with the renowned Centre for Evidence-Based Medicine in Oxford.

Programme details

The MSc in EBHC Medical Statistics is a part-time course.

There are two compulsory modules, four option modules (two from group 1 and two more either from group 1 or 2) and a dissertation.

Compulsory Modules

- Essential Medical Statistics
- Statistics for Clinical Trials

Optional Modules - 1

- Meta-analysis
- Big Data Epidemiology
- Statistical Computing with R and Stata (online)

Optional Modules - 2

- Introduction to Study Design and Research Methods
- Systematic Reviews
- Evidence-based Diagnosis and Screening

Key facts

Part-time: 2-4 years

Start Date: October 2018

Course status: open

Deadlines: 12 noon UK time (midday) on:

- Friday 19 January 2018
- Friday 9 March 2018

Later applications may be considered if places are available

Fee rates for the academic year 2018/19

Annual Award Fee: £6,085

Module fee: each £1,850 (per taught module, 6 required)

Dissertation fee: £5,550 (equivalent to 3 module fees)

- Not a Research/Statistics Methods Tutorial

Critical Appraisal

Introduction

Paul Stevenson

April 2018

How do we make decisions?

…. Pragmatically and under time constraints

Belief – “it worked for me”

Anecdote – “I heard it worked for someone”

Tradition – “We’ve always done it this way”

Instinct - “I feel confident that this will work”

Marketing- “ The company that makes this drug gave me great free stuff”

Experience- “This has worked before”

Example Introductory Activity

Are the 3 vehicles the same size?

Theories of **learning**

Moon's Levels of Learning -Effective learning is seen when a person progresses through a cycle of five stages

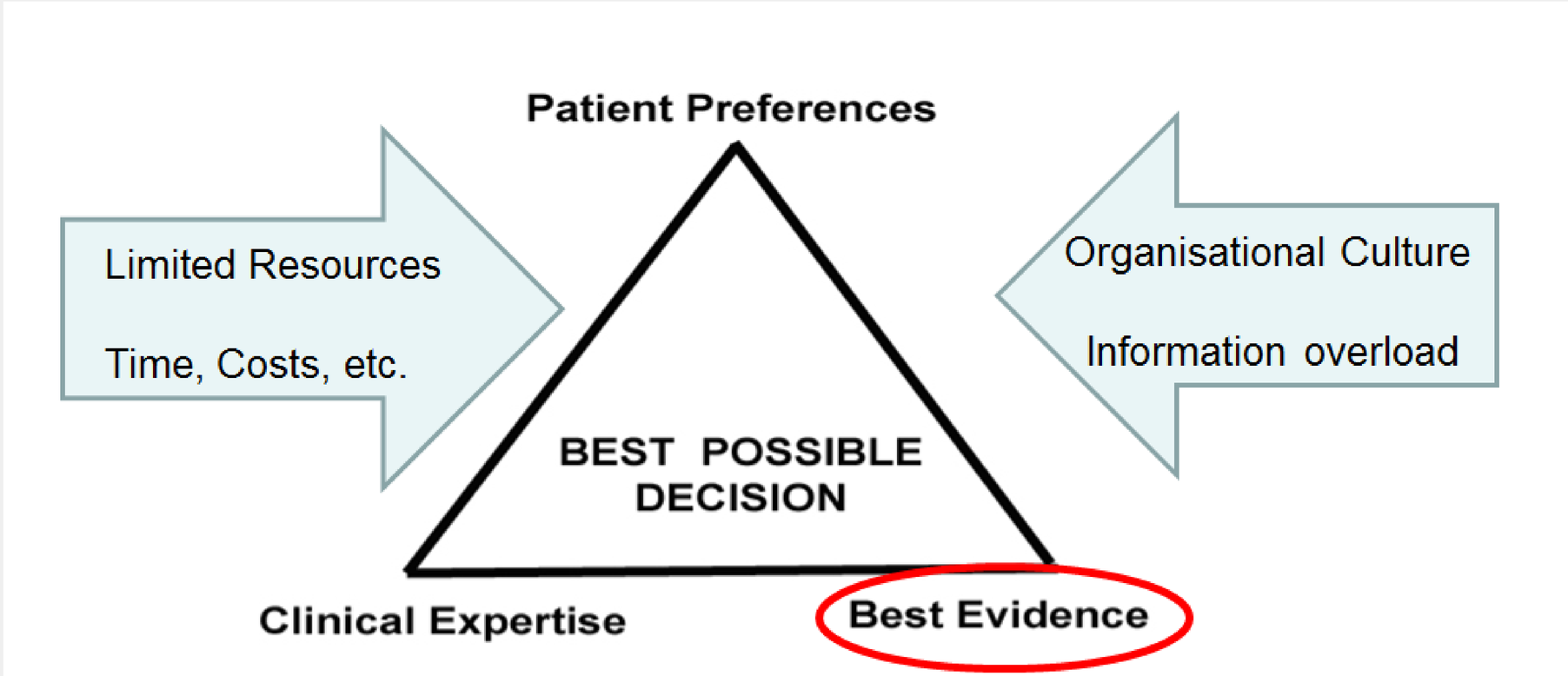
Stage 1: Noticing	The student has to register the topic, event or incident as being interesting or important in some way.
Stage 2: Making sense	The student thinks more about what they have noticed and tries to understand it better.
Stage 3: Making meaning	The student starts to ask questions and to connect ideas together.
Stage 4: Working with meaning	The student makes links with other ideas and events. They would probably refer to literature and other research. At this point, reflection on the learning is likely to be taking place.
Stage 5: Transformative learning	The student has reached the point where they can formulate new ideas of their own. They know what they would do if a similar situation arose in the future.

Felder-Silverman Learning Style

Sensing learners Concrete, practical, oriented towards facts and procedures.	**Intuitive learners** Conceptual, innovative, oriented toward theories and meanings.
Visual learners Prefer visual representations of presented material – pictures, diagrams, flow charts.	**Verbal learners** Prefer written and spoken explanations.
Inductive learners Prefer presentations that proceed from the specific to the general.	**Deductive learners** Prefer presentations that go from the general to the specific.
Active learners Learn by trying things out, working with others.	**Reflective learners** Learn by thinking things through, working alone.
Sequential learners Linear, orderly, learn in small incremental steps.	**Global learners** Holistic, systems thinkers, learn in large leaps.

What is **Critical Appraisal**

Critical appraisal is the **process** of carefully and **systematically** examining **research** to judge its **trustworthiness**, and its **value** and **relevance** in a particular context.



Critical Appraisal

Three points of focus...

- **The Message** – what are the findings of this paper?
- **Validity** – can you trust the results?
- **Applicability** – can the results be generalised to your own group of patients.

- Stage 1- **Methods** - How was data collected?
- Stage 2 – **Analysis** - How was data analysed?
- Stage 3- **Conclusion** – What does it mean?

“Shocking research shows woman who use mobile phones are more likely to have children with behavioural problems”

The Daily Telegraph

Mail Online health

Mums-to-be warned exposing babies in the womb to mobile phones 'could give them behaviour problems', report

THE Sun

Login | Register — Thu, 26 Apr 2012 ▶ Salford | 11C

HOME
WOMAN
Fabulous Mag
Mums & Dads
Beauty
Fashion
Health
Real Life
Betty Brisk
VIDEO
NEWS
Politics
Hold Ye Front Page
Captain Crunch
Sun City
SPORT
Football
Dream Team
Cricket

health — EDITED BY EMMA MORTON

Mobiles in ‘brat’ risk to mothers

By EMMA MORTON, Health and Science Editor

MUMS-TO-BE run a greater risk of having a naughty child if they regularly use a mobile when pregnant, a study says today.

Medics believe that microwave radiation emitted by the handsets could wreak unseen damage on an unborn baby's brain, leading to behavioural problems.

In the biggest study of its kind, 28,000 Danish seven-year-olds and their mothers were questioned.

Researchers found that children were 30 per cent more likely to display problems if they had been exposed regularly to the radiation while in the womb — compared with those whose mothers did not use a mobile.

The Telegraph

Mobile phones could damage unborn babies, researchers claim

Radiation from mobile phones may affect the brain development of unborn babies, the lead author of a controversial animal study has claimed.

A pregnant patient of yours has read the above newspaper articles and wants your advice.

Should she stop using or reduce the use of her mobile phone until after her child is born?
Read the newspaper article. What advice would you give her?
Can you justify your opinion?

5 minutes to come to a decision.

Critical Appraisal

Why we should be using scientific research to inform decisions

How was **data collected** (Methods)?

- Research relied on mothers to *recall* their mobile use during pregnancy.
- Behaviour assessed by mothers completing likert scale
- Exposure/Dose – what is ‘regular’ mobile-phone use?

What does the **data analysis** tell us (Analysis)?

- Small Difference: of 28,000 children, only 4% v 3%

  i.e over 95% of children showed no adverse effects.
- Newspaper states increased risk of 30%

Could the **results** be due to any other factor or bias (Conclusion)?

- Confounding Factors – Parenting style, socio-economic

Why did I give you the newspaper article?

1. **Reduce Anxiety** - familiar format, easy to read.

2. **Assess Group** - Creates an opportunity to estimate skills and personality of individuals and group.

3. **Introduces basic concepts** that we will look at in more depth during the rest of the tutorial.

Think about the structure of your tutorial

Facilitate an effective learning environment.

Bias in Research

Bias is pervasive because we want to confirm *our own* beliefs

- Selection Bias – difference between groups (age, health status, socio-economic status, regression effect)
- Performance Bias – difference in results due to care provided differently to groups
- Recall Bias – incorrect recall of past exposure
- Interviewer Bias – weighted questions etc.
- Measurement Bias – difference in result due to method / time results measured, Surrogate measures

The Catalogue of Bias - https://catalogofbias.org

- Bjork-Shiley Mechanical Heart Valves

- Many thousands implanted in the early 80s
- Serious flaw led to fatal valve malfunction

Why wasn't this shown in the initial research?

- Data collected at hospital discharge stage
  - **Follow up period** not long enough

Why do we need to be **critical**

Research isn't easy and limitations and errors occur frequently

CBT for Smoking Cessation .

What outcome to measure?

- Number of Cigarettes smoked?
- Number of people who quit?
- At what point(s) will data be collected?
- One month ? Six months? 3years?

How will data be measured?

- Participant diary? Questionnaire?
- Biochemical check, carbon-monoxide in breath?

Selecting a Research paper for Tutorial Work !

Do you send the paper prior to tutorial?

I find there are always some who don't read it in advance, so I always factor in reading time into the session.

What Topic should you look at?

Something familiar, you want participants to focus on the critical appraisal process rather than the content of the paper.

What Content should the paper include ? -

Find out what the participants want to learn and then identify a paper that contains a good example that can be used for learning.

You don't want 'perfect' research- there needs to be flaws and uncertainty to generate discussion.

Do you have to use an entire paper?

No, think about using specific sections of a paper rather than all of it: Especially for illustrating statistics.

Sources of Example Critical Appraisals

Home | Login

BestBETs

BEST EVIDENCE TOPICS

About
Databases
Submit
Browse BETs
Browse Critical Appraisals
Browse all by Topic
Accumulator BETs
Background
Teaching
News
Resources
Contact

Search Critical Appraisals

Filter Help

Filter Options

BET No:
BET Title:
Publication Type:
Linked to BET:
Appraisal Title:
Author:
Uncategorised:

Case-control checklist (including harm)
Cohort
Decision rule
Diagnosis
Economic
Educational intervention
Guideline
Prognosis
Qualitative
Randomised control trial
Review or meta-analysis
Screening
Survey (including pre-test probabilities)

-- Filter --

Sort by Group GO

Show key to symb

Appraisal D

Case-control

Kostis et al

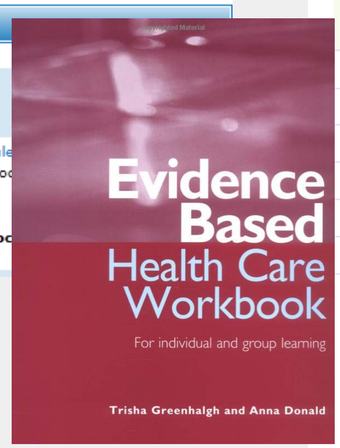
Weekend versus Weekday Admission and Mortality from Myocardial Infarction

Case-control checklist (including harm)

ens U, Scheule

perative myoc

erative myoc



NHS choices Your health, your choices

Enter a search term

Health A-Z | Live Well | Care and support | Health news | Services near you

Home › Behind the Headlines

Behind the Headlines

Your guide to the science that makes the news

Categories

- Cancer
- Genetics and stem cells
- Food and diet
- Obesity
- Neurology
- Lifestyle and exercise
- Older people
- Heart and lungs
- Medication
- Pregnancy and child
- Mental health
- Medical practice
- Diabetes

'Apple-shaped' women may have increased heart attack risk

New drug for advanced stage of multiple sclerosis
Friday March 23 2018

Non-hormonal alternative to HRT shows promise in treating hot flushes
Wednesday March 14 2018

Up to 1 in 5 antibiotics may be prescribed inappropriately
Tuesday February 27 2018

Big new study confirms antidepressants work better than placebo
Thursday February 22 2018

'Painkillers best option for sore throats' say new NHS guidelines
Friday January 26 2018

University of South Australia

STUDY RESEARCH PARTNER NEWS & EVENTS LOGIN

Home > Research > Sansom Institute for Health Research > International Centre for Allied Health Evidence > Resources > iCAHE Journal Club Critical Appraisals

International Centre for Allied Health Evidence

Resources
- Critical Appraisal Tools
- Glossary of terms
- Guideline Clearinghouse
- iCAHE Journal Club Critical Appraisals
- Assisting Implementation Starter Pack
- Research Hub
- iCAHE Outcome Calculators
- iCAHE textbooks
- Useful websites
- iCAHE's Learning Hub
- WCPT Clinical Practice Guidelines Symposium

Services
Publications
Projects
International Projects
Partners and Collaborators
iCAHE Allied Health Conference

iCAHE Journal Club Critical Appraisals

As part of the Journal Clubs that are run at iCAHE, a large volume of Critical Appraisal Summaries (CA Summaries) of published literature are created. These are useful resources to all stakeholders and as such we provide a library of all summaries.

For each summary, iCAHE provides the citation details and methodological quality of the study identified to address the clinical question developed by the journal club. Copyright issues preclude iCAHE from putting the full text of the critically appraised papers on the website. The citations are provided so that clinicians can access the article from their own library sources. These pages are regularly updated.

Please choose a category:
Aged Care
Cardiac Rehabilitation
Chronic Disease Management
Clinical Supervision
General Topics
Hand Rehabilitation
Mental Health
Neurological Rehabilitation
Nutrition & Dietetics
Orthopaedic Rehabilitation
Paediatric Care
Palliative Care
Speech Therapy
Women's Health
DECD Journal Club

Search for CA summaries — Search tips

Enter a keyword

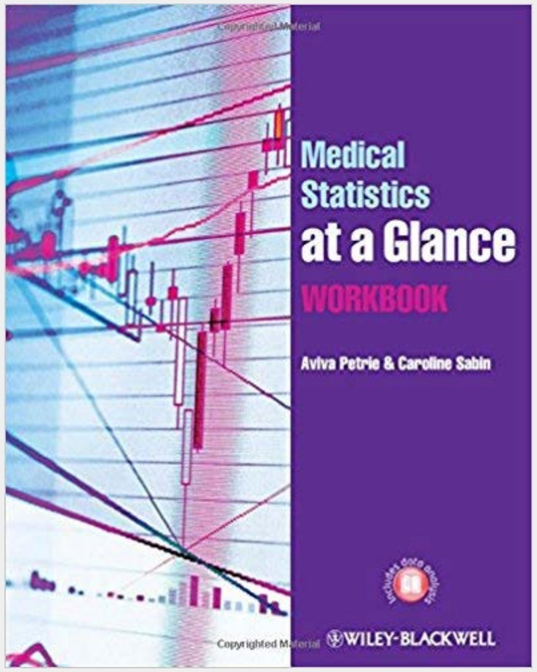
Search

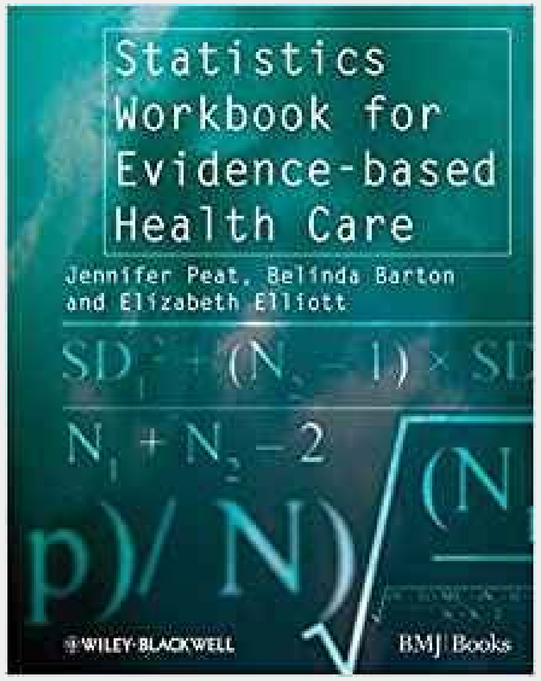
iCAHE will continually update this compendium of critically appraised research publications as a resource for health services interested in implementing journal clubs within their health services.

Internet search for **Journal Club Critical Appraisals** can be useful for specific disciplines

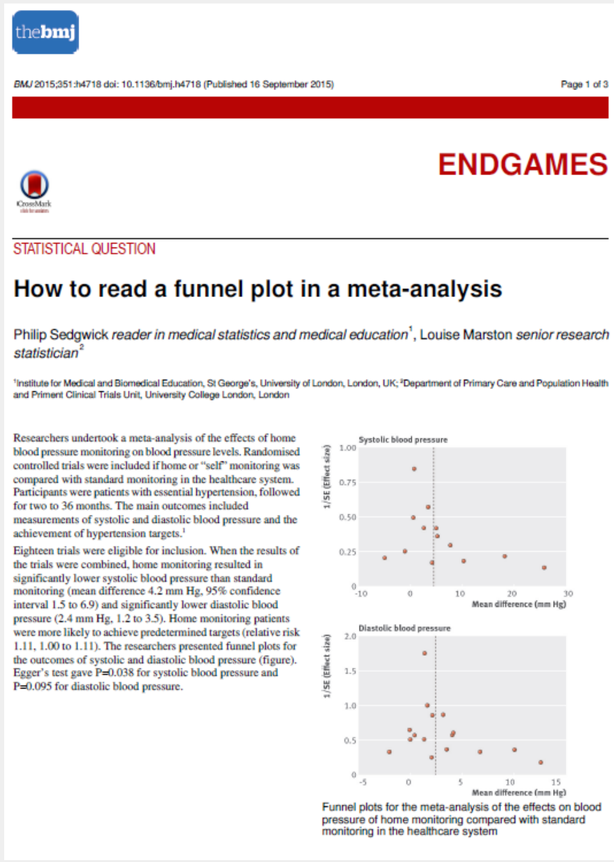
Where to get Example Statistics Exercises

Statistics Workbooks





BMJ Endgames series



thebmj

BMJ 2015;351:h4718 doi: 10.1136/bmj.h4718 (Published 16 September 2015) Page 1 of 3

ENDGAMES

STATISTICAL QUESTION

How to read a funnel plot in a meta-analysis

Philip Sedgwick *reader in medical statistics and medical education*[1], Louise Marston *senior research statistician*[2]

[1]Institute for Medical and Biomedical Education, St George's, University of London, London, UK; [2]Department of Primary Care and Population Health and Priment Clinical Trials Unit, University College London, London

Researchers undertook a meta-analysis of the effects of home blood pressure monitoring on blood pressure levels. Randomised controlled trials were included if home or "self" monitoring was compared with standard monitoring in the healthcare system. Participants were patients with essential hypertension, followed for two to 36 months. The main outcomes included measurements of systolic and diastolic blood pressure and the achievement of hypertension targets.[1]

Eighteen trials were eligible for inclusion. When the results of the trials were combined, home monitoring resulted in significantly lower systolic blood pressure than standard monitoring (mean difference 4.2 mm Hg, 95% confidence interval 1.5 to 6.9) and significantly lower diastolic blood pressure (2.4 mm Hg, 1.2 to 3.5). Home monitoring patients were more likely to achieve predetermined targets (relative risk 1.11, 1.00 to 1.11). The researchers presented funnel plots for the outcomes of systolic and diastolic blood pressure (figure). Egger's test gave P=0.038 for systolic blood pressure and P=0.095 for diastolic blood pressure.

Funnel plots for the meta-analysis of the effects on blood pressure of home monitoring compared with standard monitoring in the healthcare system

a) Failure to include in the meta-analysis all of the relevant trials that have been conducted may have been due to reporting bias

b) A funnel plot can suggest whether relevant trials were not included in the meta-analysis only as a result of publication bias

c) The funnel plots for systolic and diastolic blood pressure indicate that not all of the relevant trials that have been conducted were identified

d) The result of Egger's test indicates that asymmetry exists in the funnel plot for the outcome of systolic blood pressure

Answers

Statements *a*, *c*, and *d* are true, whereas *b* is false.

The aim of the meta-analysis was to investigate the effects on blood pressure of home monitoring compared with standard monitoring in the healthcare system. The outcomes included systolic and diastolic blood pressure. The purpose of the meta-analysis was to combine the sample estimates of the treatment effects (difference between home and standard monitoring in outcome) to give a total overall estimate of the population parameter for each outcome, thereby reducing a large amount of information to a manageable quantity. The population parameter is the difference in outcome (treatment effect) between home monitoring and standard monitoring that would be observed in the population if both methods were applied to all members. However, the researchers may not have identified all the relevant trials that had been conducted. If so, the total overall estimates produced by the meta-analysis would probably overestimate the population parameters. Failure to include all relevant trials in a meta-analysis can be due to reporting bias (*a* is true), a collective term for various types of bias.[2] Reporting bias occurs when the reporting of research findings is influenced by the nature and direction of trial results, and it includes publication bias, language bias, citation bias, and time lag bias. Failure to include all of the relevant studies that have been conducted in a meta-analysis is often wrongly attributed solely to publication bias.

Failure to include in the meta-analysis all of the relevant trials that have been conducted can be shown graphically using the funnel plot. The funnel plot may show bias resulting from various sources, including all types of reporting bias, but it is not possible to identify which of the reporting biases may be present. Researchers often incorrectly indicate that the purpose of the plot is to detect whether trials were not included in the meta-analysis solely because of publication bias (*b* is false).

In the above meta-analysis a separate funnel plot was presented for each of the outcomes of systolic and diastolic blood pressure. The funnel plot was a scatter plot of the estimated effect size (mean difference between home monitoring and standard monitoring) of blood pressure plotted on the horizontal axis against the reciprocal of standard error of the estimated effect on the vertical axis for the trials identified. The standard error provides a measure of the precision of the effect size as an estimate of the population parameter.[3] Typically, trials with smaller sample sizes produce less precise estimated effects. As sample size increases the precision of the estimated effect increases and the size of the standard error decreases, and therefore the reciprocal of the standard error increases in size. Hence, trials with less precise estimated effects scatter more widely at the bottom of the plot. If the samples for the trials were selected from the population at random, the estimated treatment effects would be expected to scatter around the total overall estimate of the meta-analysis (represented by the vertical line on the plot). As sample size increases, because the precision of the estimated effects increases, the spread of points would be expected to narrow and the scatter plot would resemble a funnel.

Sometimes the standard error, rather than the reciprocal of the standard error, is plotted on the vertical axis. Because trials with larger sample sizes produce more precise estimated effects and therefore smaller standard errors, the vertical axis may be inverted—with zero at the top—so that the scatter of points resembles a funnel. Measures of precision of the estimated effects other than the standard error are sometimes used, including the reciprocal of the sample size or variance of the estimated effect. Sometimes lines are superimposed on a funnel plot to resemble the limits of the predicted funnel shape in the estimated effects, thereby aiding visual interpretation.

If all of the relevant trials that have been conducted were included in a meta-analysis, a funnel plot would be expected to be symmetrical in shape—that is, the points would be scattered in the shape of a funnel centrally around the total overall estimated effect. If not all of the relevant trials were included then the plot would be asymmetrical. Assessment of symmetry in a funnel plot is typically subjective. Any assessment is particularly difficult when the number of trials is small; funnel plots are thought to be unreliable methods of investigating potential bias if the number of studies is less than 10.

Visual inspection of the funnel plots (figure) in the above meta-analysis suggests asymmetry for both systolic and diastolic blood pressure. It therefore seems that not all of the relevant trials that had been conducted were included in the meta-analysis (*c* is true). For both systolic and diastolic blood pressure, studies seem to be missing at the bottom of the plot towards the left hand side. Such studies would probably be trials with large standard errors and small sample sizes, with a mean blood pressure for standard monitoring that was lower than for home monitoring. However, it is only an assumption that these studies were ever undertaken.

Formal statistical tests exist for assessing asymmetry in a funnel plot, including Egger's test. The null hypothesis for Egger's test is that symmetry exists in the funnel plot, with the alternative indicating that asymmetry is present. The P value for Egger's test was 0.038 for systolic blood pressure and 0.095 for diastolic blood pressure. Hence there was evidence of asymmetry at the 5% level of significance in the funnel plot for systolic blood pressure (*d* is true) but not for diastolic blood pressure. Although there was a discrepancy between the visual inspection of the funnel plot and Egger's test result for diastolic blood pressure, the test result should be interpreted in the context of visual inspection of the funnel plot. Sometimes statistical tests for detecting asymmetry in a funnel plot have low statistical power.

As described above, asymmetry in a funnel plot may be caused by reporting bias. However, it can also be the result of poor methodological design in the trials identified—for example, the lack of blinding to treatment allocation, which makes the measurements prone to ascertainment bias.[4] Typically, poor methodological design results in estimated treatment effects being spuriously inflated. Poor methodological design is a common problem in trials with small sample sizes and it leads to an absence of studies on one side at the base of the funnel, resulting in asymmetry in the funnel plot. This might explain the asymmetry of the funnel plots in the above meta-analysis.

If, based on the funnel plot, it is suspected that not all relevant trials have been included in a meta-analysis, the effect sizes and standard errors for those studies thought to be missing can be predicted using a method called "trim and fill." The researchers

Why many tutorials appraise an RCT

Source of bias	Effect on treatment efficacy	Size of the effect	References
Randomisation	Increase	Non-randomised studies overestimate treatment effect by 41% with inadequate method, 30% with unclear method	KF Schultz, I Chalmers, RJ Hayes, DG Altman. Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. Journal of the American Medical Association 1995 273: 408-12.
Randomisation	Increase	Completely different result between randomised and non-randomised studies	Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Randomization is important in studies with pain outcomes: systematic review of transcutaneous electrical nerve stimulation in acute postoperative pain. British Journal of Anaesthesia 1996; 77: 798-803.
Blinding	Increase	17%	KF Schultz, I Chalmers, RJ Hayes, DG Altman. Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. Journal of the American Medical Association 1995 273: 408-12.
Blinding	Increase	Completely different result between blind and non-blind studies	Ernst E, White AR. Acupuncture for back pain: A meta-analysis of randomised controlled trials. Arch Int Med 1998, 158: 2235-2241.

- Looking at the reasons why an RCT is more robust illustrates potential shortcomings in other methodologies.

Critical Appraisal Checklists

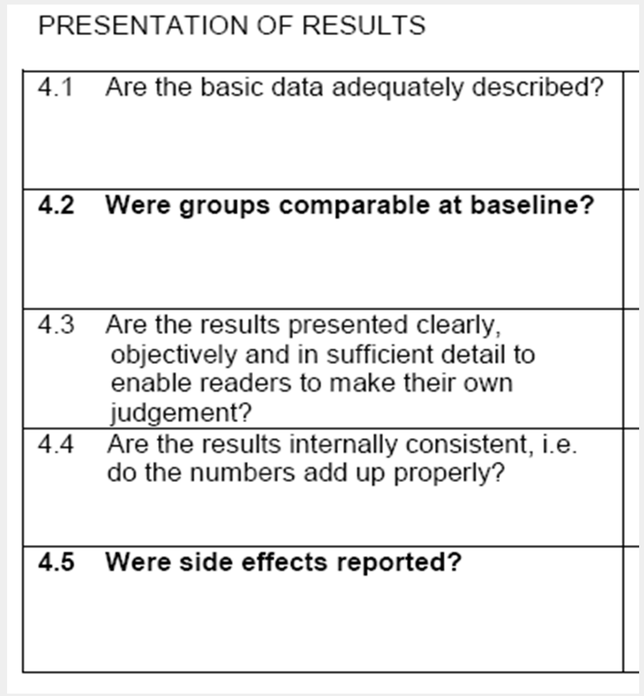
www.delfini.org/index_Resources.htm

www.cebm.net/2014/06/critical-appraisal/

https://casp-uk.net/casp-tools-checklists/

http://joannabriggs.org/research/critical-appraisal-tools.html

www.bestbets.org/links/BET-CA-worksheets.php



PRESENTATION OF RESULTS

4.1 Are the basic data adequately described?	
4.2 Were groups comparable at baseline?	
4.3 Are the results presented clearly, objectively and in sufficient detail to enable readers to make their own judgement?	
4.4 Are the results internally consistent, i.e. do the numbers add up properly?	
4.5 Were side effects reported?	

Using a checklist :

- Ensures you are **consistent** in how you evaluate multiple papers.
- **Simplifies** the process by breaking it into smaller ‘chunks’ of information

Group Exercise

Appraisal 1
RCT – Smoking Cessation

Paul Stevenson

April 2018

Class Exercise: Appraisal 1

- Read the research paper you have been provided with
- Using the CASP RCT appraisal tool
- Work as a group
- Reach a consensus and answer each of the **10 questions** in the CASP tool.
- You have approx **30 minutes** to read the paper and answer

Intention to Treat Analysis

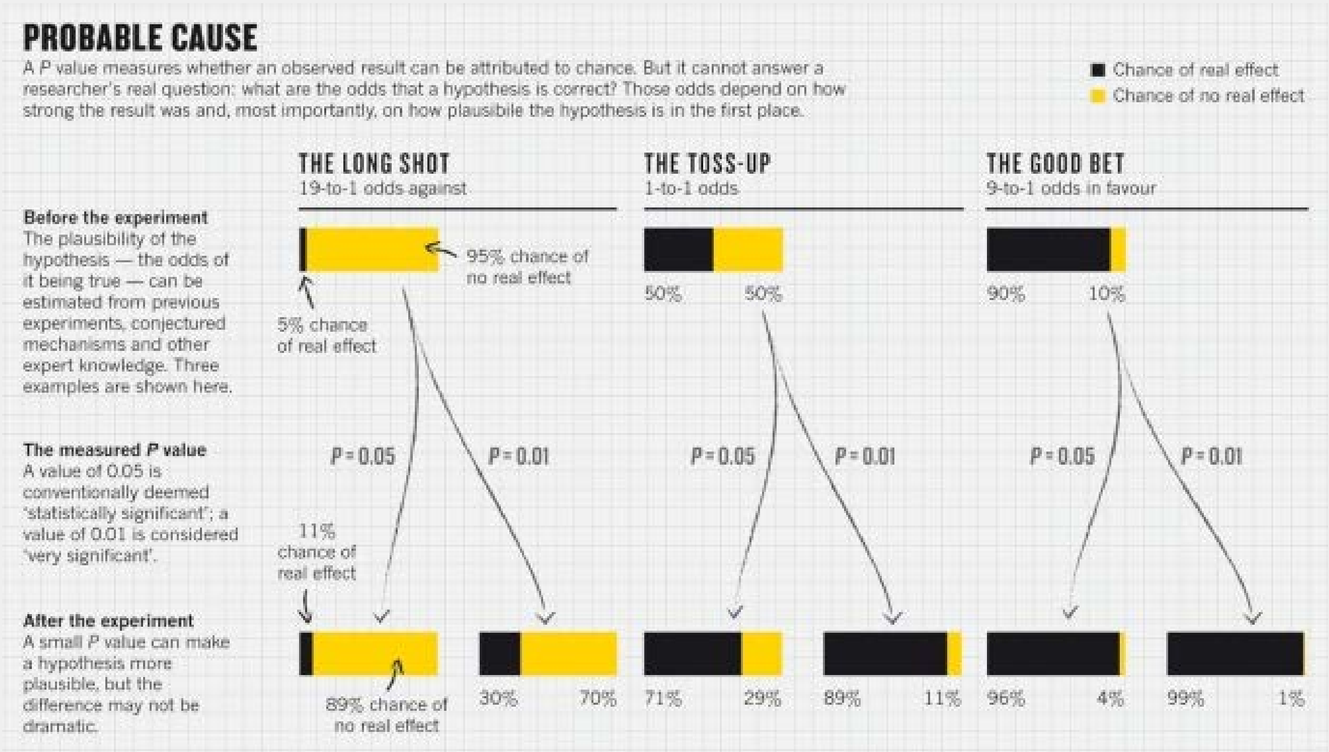
- How are drop-outs dealt with in the research.

Group A : Cohort of 20	Group B : Cohort of 20
5 people drop out due to side effects	5 people drop out due to side effects
10 people have a successful event	10 people have a successful event
5 people are unaffected	5 people are unaffected
10/15 = 0.67	10/20 = 0.50

P value

(statistical significance)

- A p-value is calculated to assess whether trial results are likely to have occurred simply through chance.
- A p-value of 0.05 or less is considered ‘statistically significant’

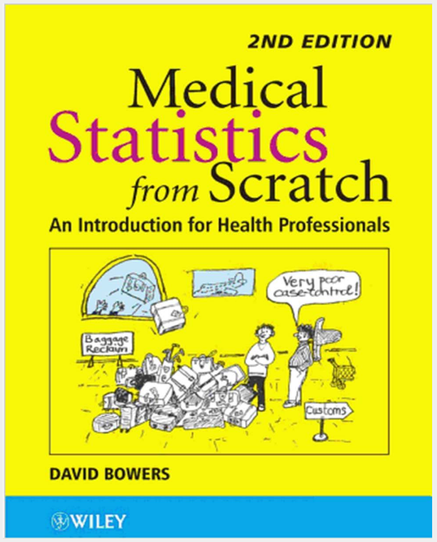


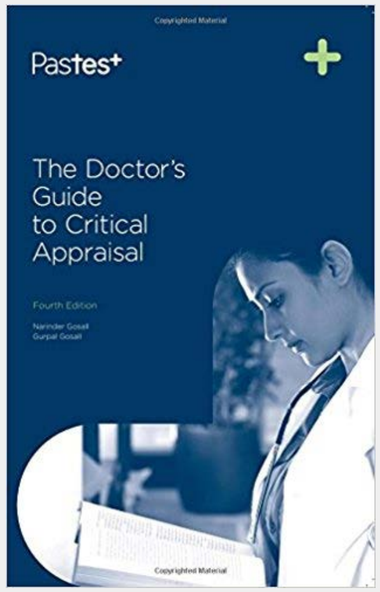
- confidence intervals are a potentially more useful approach to assessing the role of chance.

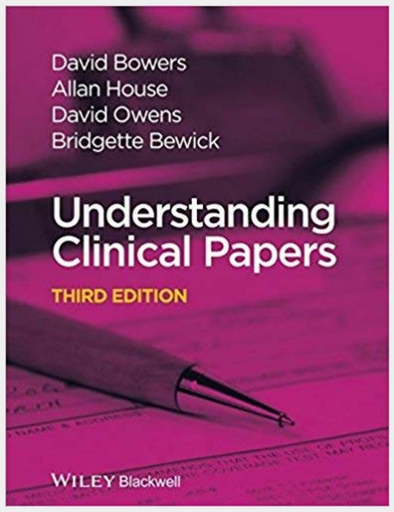
Recommended Reading
…or listening

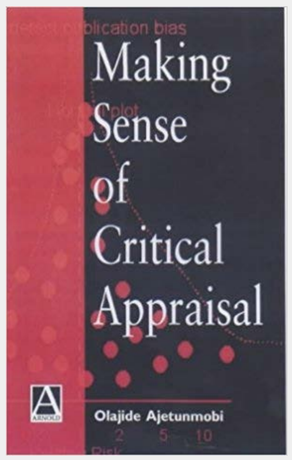
- Testing Treatments:

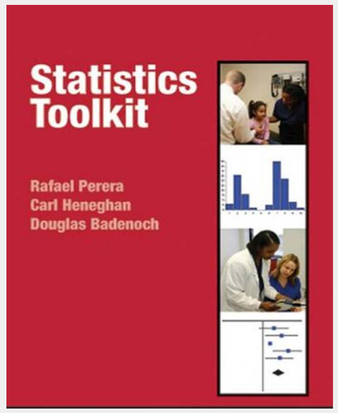
https://en.testingtreatments.org/book/

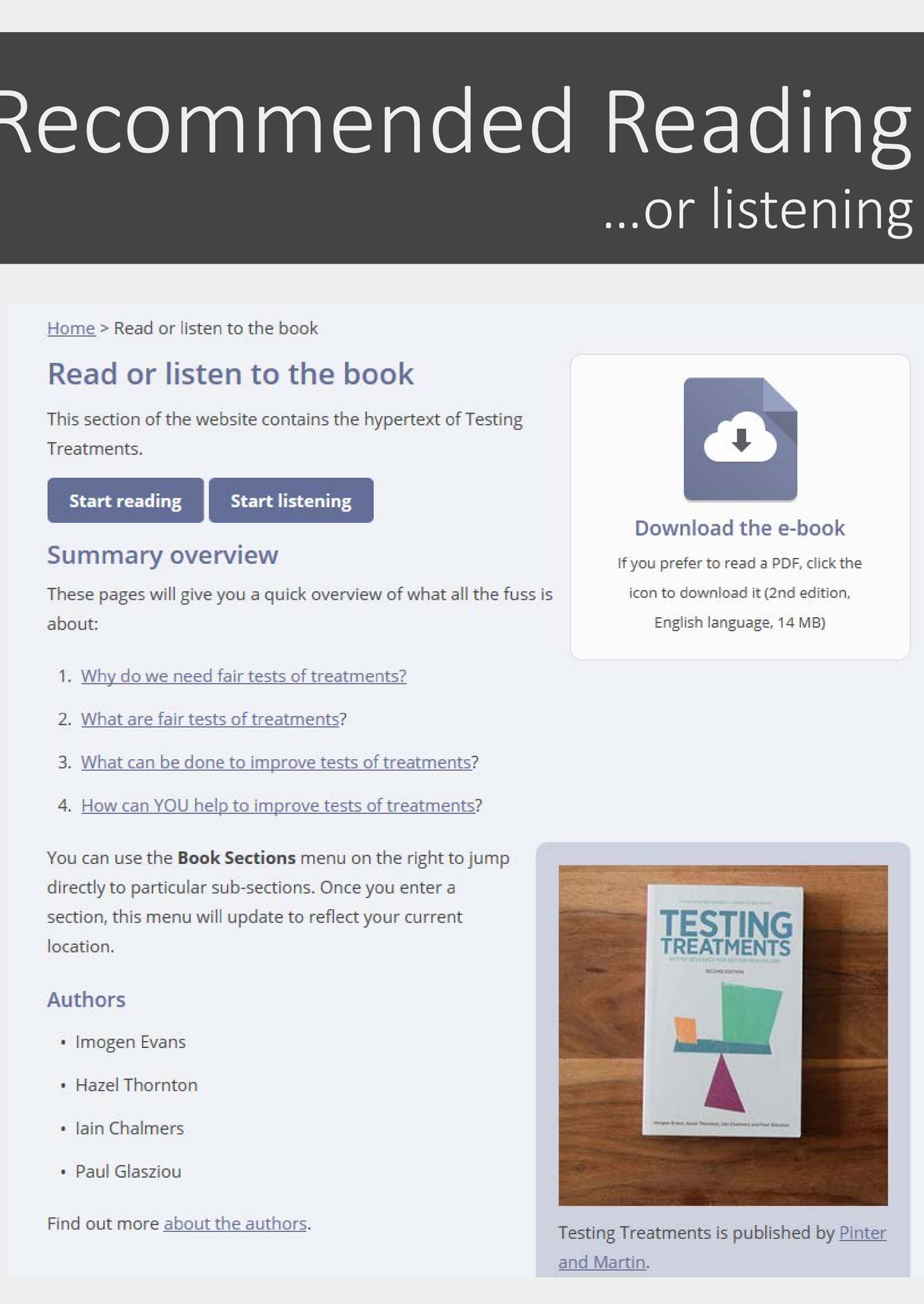












Facilitating Critical Appraisal

- Observe the room and conversations

- Give people space but ensure everyone is able to contribute
  - Tactics for Extroverts
  - Tactics for Introverts
  - Tactics for the disengaged
  - Tactics for groups with mixed abilities and experiences

- Tricky Questions – What to do when you don’t know the answer.

- Facilitation vs Teaching

Critical Appraisal

Systematic Reviews

Paul Stevenson

April 2018

Introduction

We will appraise a systematic review together:

1. First half of appraisal focus on review process

1. Second half of appraisal focus on Meta-analysis and Stats

Selecting a paper for tutorials

Detailed systematic reviews tend to be several pages long

- Typical Cochrane Review
  – **70 pages**

- Typical Review published in BMJ
  – **12 pages**

You might want to use excerpts from multiple reviews

- Allows you to focus attention on key learning points
- Reduces the amount of time spent reading papers.

High quality systematic reviews seek to:

... These are the things you should be considering when you carry out a critical appraisal

1. **Identify all** relevant published and unpublished evidence

2. **Select** studies for **inclusion**

3. **Assess the quality** of each study

4. **Synthesize the findings** from individual studies in an unbiased way

5. **Interpret the findings** and present an **impartial summary** of the findings with due consideration of any flaws in the evidence.

Stage 1 Example

What is the objective of this systematic review

Cochrane Database of Systematic Reviews

Vaccines for preventing influenza in healthy adults

New search | Review | Intervention

Vittorio Demicheli ✉, Tom Jefferson, Eliana Ferroni, Alessandro Rivetti, Carlo Di Pietrantonj

First published: 1 February 2018

Objectives

To assess the effects (efficacy, effectiveness, and harm) of vaccines against influenza in healthy adults, including pregnant women.

PICO

Patients – Healthy Adults including pregnant women

(Age range? Any existing conditions excluded?)

Types of participants

Healthy individuals aged 16 to 65 years, irrespective of influenza immune status. We excluded studies considering more than 25% of individuals outside this age range. We also included pregnant women together with their newborns.

Intervention – Vaccines against influenza

Types of interventions

Live, attenuated, or killed vaccines, or fractions thereof, administered by any route, irrespective of antigenic configuration.

Comparison – no action

Outcome – efficacy and effectiveness in reduction of influenza cases / Assessment of potential harms from vaccination.

1. Numbers and seriousness (complications and working days lost) of symptomatic influenza and influenza-like illness (ILI) cases occurring in vaccine and placebo groups.

Stage 2 – Literature Search

Has the process successfully identified **all** of the relevant research on this topic?

- Search all relevant database sources of information
- Search highly relevant publications
- Obtain unpublished studies (check trial registers)
- ‘Pearl Grow’ using reference lists from appropriate papers.

Stage 2 – Example

Has the process successfully identified **all** of the relevant research on this topic?

Electronic searches

We searched the Cochrane Central Register of Controlled Trials (CENTRAL; 2016, Issue 12) searched 31 December 2016 via the Cochrane Library), which contains the Cochrane Acute Respiratory Infections Group's Specialised Register; MEDLINE (PubMed) (January 1966 to 31 December 2016); Embase (Elsevier) (1990 to 31 December 2016); WHO International Clinical Trials Registry Platform (ICTRP; www.who.int/ictrp/en, 1 July 2017); and ClinicalTrials.gov (www.clinicaltrials.gov, 1 July 2017). See Appendix 2 for the search strategies used to identify trials.

Not included CINAHL , is that important?

Searching other resources

In order to identify further trials, we read the bibliographies of retrieved articles and handsearched the journal *Vaccine* from its first issue to the end of 2009. The results of the handsearches are included in CENTRAL. In order to locate unpublished trials for the first edition of this review, we wrote to manufacturers and first or corresponding trial authors of studies in the review.

Embase (Elsevier)

#1 'influenza vaccine'/de
#2 'influenza'/exp
#3 'influenza virus a'/exp OR 'influenza virus b'/exp
#4 flu:ab,ti OR influenza*:ab,ti
#5 #2 OR #3 OR #4
#6 'vaccine'/de OR 'acellular vaccine'/de OR 'dna vaccine'/de OR 'inactivated vaccine'/de OR 'live vaccine'/de OR 'subunit vaccine'/de OR 'virus vaccine'/de OR 'virosome vaccine'/de OR 'recombinant vaccine'/de
#7 'immunization'/de OR 'vaccination'/de OR 'active immunization'/de OR 'immunoprophylaxis'/de OR 'mass immunization'/de
#8 vaccin*:ab,ti OR immuni*:ab,ti OR inocul*:ab,ti
#9 #6 OR #7 OR #8
#10 #5 AND #9
#11 #1 OR #10
#12 'randomized controlled trial'/exp OR 'single blind procedure'/exp OR 'double blind procedure'/exp OR 'crossover procedure'/exp
#13 random*:ab,ti OR placebo*:ab,ti OR factorial*:ab,ti OR crossover*:ab,ti OR 'cross-over':ab,ti OR 'cross over':ab,ti OR assign*:ab,ti OR allocat*:ab,ti OR volunteer*:ab,ti OR ((singl* OR doubl*) NEAR/3 (blind* OR mask*)):ab,ti
#14 #12 OR #13
#15 #11 AND #14

No search for generic and commercial names of vaccines (ie Fluzone, Agrippal, Fluenz etc.)

Stage 3a – Assess for Inclusion

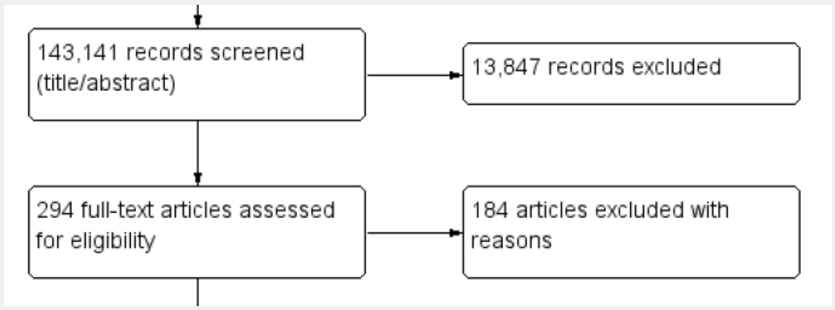
- Which trials are relevant to the review?

- Inclusion & Exclusion Criteria

- Recommendation is to have two independent researchers carry out the screen and select process.

Stage 3a – Example

Which studies are included in the review? Is there any bias that could result from the exclusion/inclusion criteria .



Selection of studies

Two review authors (AR, CDP) independently excluded all initially identified and retrieved articles not fulfilling the inclusion criteria. In the case of disagreement, one review author (VD) acted as arbitrator. We recorded the selection process in sufficient detail to complete a PRISMA flow diagram and 'Characteristics of excluded studies' table (Moher 2009).

Types of studies

Any randomised controlled trial (RCT) or quasi-RCT comparing influenza vaccines in humans with placebo or no intervention, or comparing types, doses, or schedules of influenza vaccine. We only

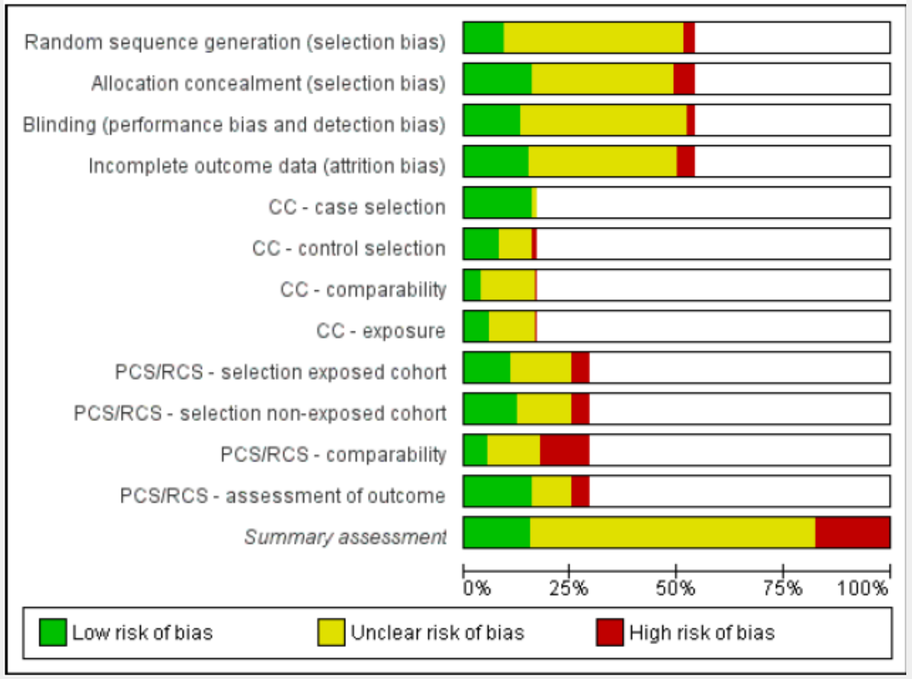
Baxter 2012	No design: controlled case series
Baxter 2013	Self controlled time series study
Belongia 2009	Case-control study, no harm assessment
Belshe 2001	No original data
Benke 2004	Questionnaire survey; non-comparative analysis
Beran 2013	Absence of an adequate control group (quadrivalent versus trivalent inactivated vaccine; low versus normal adjuvant content)
Betts 1977b	Trial with swine vaccine (Hsw1N1, A/New Jersey/76)

Stage 3b – Assess Methods

Two review authors (CDP, AR) independently assessed the methodological quality of the included studies using criteria from the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). In case of disagreement, one review author (VD) acted as arbitrator in assigning quality judgements.
We classified studies according to the following key domains for assessing risk of bias (Higgins 2011).

The overall quality of the retrieved studies was poor and was affected by poor reporting or limited descriptions of the studies' designs. A detailed description is provided in the Risk of bias in included studies section of the review.
The main problems with influenza vaccine studies are their poor quality and discrepancies between the data presented, their conclusions, and the authors' recommendations.

Figure 2. 'Risk of bias' graph: review authors' judgements about each risk of bias item presented a percentages across all included studies.



Group Exercise

Critical Appraisal of a Systematic Reviews

Part A (**Questions 1 -9**)

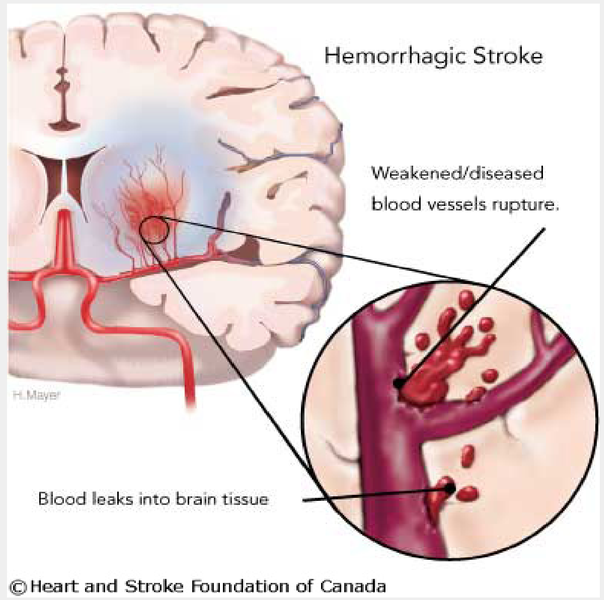
Paul Stevenson

April 2018

Group Exercise

- Read the review article:

Effects of vitamin E on Stroke



- Using the supplied appraisal tool:
- As a group reach a consensus to answer the **first 9 questions** of the tool.

Meta-Analysis

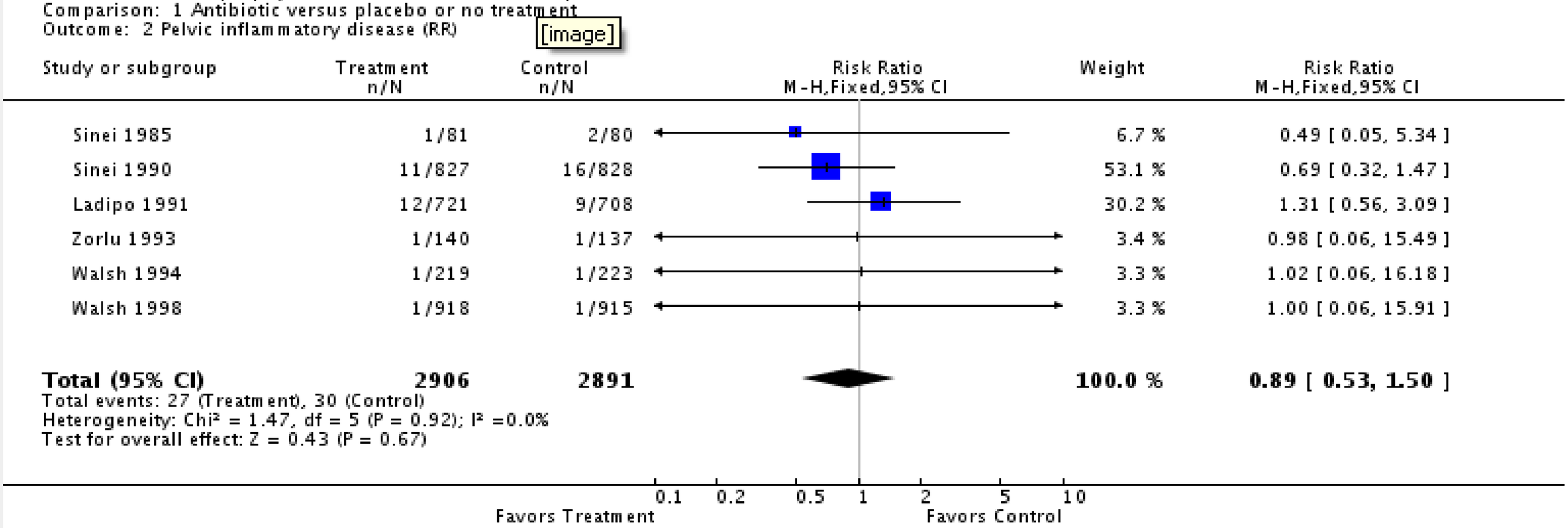
- A "**meta-analysis**" is a statistical approach to combine the data derived from a systematic-review.

Meta-Analysis

Antibiotic prophylaxis for intrauterine contraceptive device insertion

Analysis 1.2. Comparison 1 Antibiotic versus placebo or no treatment, Outcome 2 Pelvic inflammatory disease (RR).

Review: Antibiotic prophylaxis for intrauterine contraceptive device insertion
Comparison: 1 Antibiotic versus placebo or no treatment
Outcome: 2 Pelvic inflammatory disease (RR)



- Forest Plot , Odds ratio , Confidence Interval , Weighting , Heterogeneity , P and I^2

Heterogeneity

The diversity between studies

- Are the studies sufficiently similar to justify being amalgamated into a single review?

- Ideally, the studies being combined should all be undertaken in the same way and to the same protocols.

Sources of heterogeneity in systematic reviews:

- statistical (variation in point estimates between trials)
- methodological (variation in study methods: e.g. blinding)
- clinical (variation in intervention, participants, outcome measurement, setting)

Heterogeneity – The I^2 statistic

- One measure of heterogeneity is I^2, it indicates the percentage of variance in a meta-analysis that is attributable to study heterogeneity

- The I^2 statistic quantifies % of variation which is not due to chance.
  - scores heterogeneity between 0% and 100%.
  - The higher the score the greater the heterogeneity

    (low is good; A value of <25% is considered low.)

- Be Careful! Meta-analyses with 2-4 studies are often not adequate to accurately estimate heterogeneity. This results in an incorrect zero between study variance estimate, leading to a false homogeneity assumption.

Heterogeneity

Analysis 2.3. Comparison 2 Danazol versus progestagens, Outcome 3 Number of women reporting adverse events.

Review: Danazol for heavy menstrual bleeding

Comparison: 2 Danazol versus progestagens

Outcome: 3 Number of women reporting adverse events

Study or subgroup	Danazol n/N	Progestagen n/N	Odds Ratio M-H,Fixed,95% CI	Odds Ratio M-H,Fixed,95% CI
Bonduelle 1991	10/10	14/14		0.0 [0.0, 0.0]
Buyru 1995	11/20	7/20		2.27 [0.64, 8.11]
Dunphy 1998	8/9	2/7		20.00 [1.42, 282.45]
Higham 1993	17/19	11/18		5.41 [0.94, 30.97]
Total (95% CI)	**58**	**59**		**4.05 [1.61, 10.21]**

Total events: 46 (Danazol), 34 (Progestagen)

Heterogeneity: $Chi^2 = 2.30$, $df = 2$ ($P = 0.32$); $I^2 = 13\%$

Test for overall effect: $Z = 2.96$ ($P = 0.0030$)

Heterogeneity: $Chi^2 = 2.30$, $df = 2$ ($P = 0.32$); $I^2 = 13\%$

I squared = 13% indicating there is a good degree of similarity between the results.

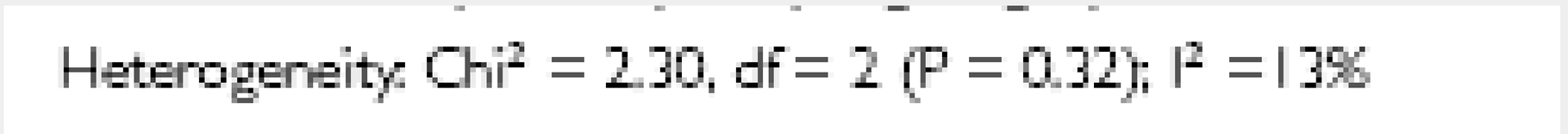
Heterogeneity - Chi^2

- Chi-squared test of heterogeneity, which is often shown at the bottom of forest plots.

- A significant result ($P < 0.05$) implies significant differences between the trials (heterogeneity) and draws in to question the wisdom of combining the studies.

- When the chi^2 stat is greater than the df stat this indicates heterogeneity

- However, this is a weak test and so heterogeneity can be present even with a non-significant result.



Weighting

Antibiotic prophylaxis for intrauterine contraceptive device insertion

Analysis 1.2. Comparison 1 Antibiotic versus placebo or no treatment, Outcome 2 Pelvic inflammatory disease (RR).

Review: Antibiotic prophylaxis for intrauterine contraceptive device insertion
Comparison: 1 Antibiotic versus placebo or no treatment
Outcome: 2 Pelvic inflammatory disease (RR)

Study or subgroup	Treatment n/N	Control n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
Sinei 1985	1/81	2/80		6.7 %	0.49 [0.05, 5.34]
Sinei 1990	11/827	16/828		53.1 %	0.69 [0.32, 1.47]
Ladipo 1991	12/721	9/708		30.2 %	1.31 [0.56, 3.09]
Zorlu 1993	1/140	1/137		3.4 %	0.98 [0.06, 15.49]
Walsh 1994	1/219	1/223		3.3 %	1.02 [0.06, 16.18]
Walsh 1998	1/918	1/915		3.3 %	1.00 [0.06, 15.91]
Total (95% CI)	**2906**	**2891**		**100.0 %**	**0.89 [0.53, 1.50]**

Total events: 27 (Treatment), 30 (Control)
Heterogeneity: Chi² = 1.47, df = 5 (P = 0.92); I² =0.0%
Test for overall effect: Z = 0.43 (P = 0.67)

0.1 0.2 0.5 1 2 5 10
Favors Treatment
Favors Control

Weighting

...all studies are treated equally

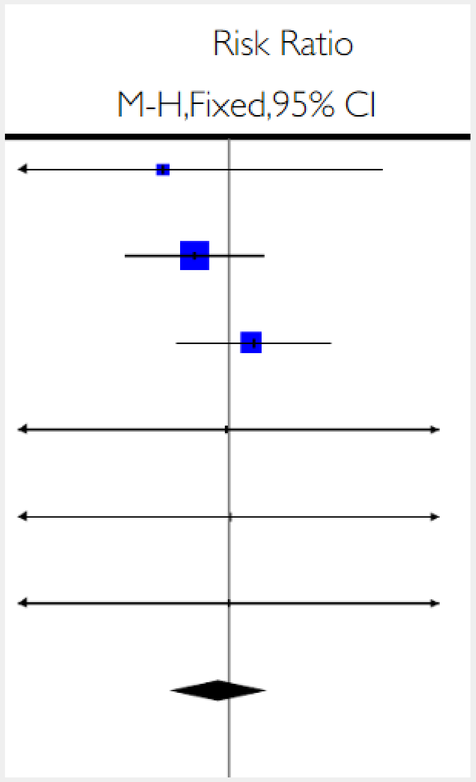
- Meta-analysis gives more weight to studies with more precise estimates.

- Greater weight given to:
  - Larger study sample size
  - Smaller confidence Intervals (consistency of results)
  - Higher methodological quality

Forest Plot : OR and CI

Antibiotic prophylaxis with ICD to prevent pelvic inflammatory disease

- Visual ⟷ Numerical



Risk Ratio
M-H,Fixed,95% CI

0.49 [0.05, 5.34]

0.69 [0.32, 1.47]

1.31 [0.56, 3.09]

0.98 [0.06, 15.49]

1.02 [0.06, 16.18]

1.00 [0.06, 15.91]

0.89 [0.53, 1.50]

Shows a 11% reduction in infection from antibiotics

But.. CI shows expected range of results to be between a 47% reduction and a 50% increase.

Interpretation of Results

- What does an 11% risk reduction imply ?

- Is it clinically significant?
  - How likely is an event
  - What are the consequences of an event
  - How easy is it to implement the intervention

Relative vs Absolute Risk, & NNT

Total (95% CI)	**2906**	**2891**		**100.0 %**	**0.89 [0.53, 1.50]**

disease (RR).

Review: Antibiotic prophylaxis for intrauterine contraceptive device insertion

Comparison: 1 Antibiotic versus placebo or no treatment

Outcome: 2 Pelvic inflammatory disease (RR)

Study or subgroup	Treatment n/N	Control n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
Sinei 1985	1/81	2/80		6.7 %	0.49 [0.05, 5.34]
Sinei 1990	11/827	16/828		53.1 %	0.69 [0.32, 1.47]
Ladipo 1991	12/721	9/708		30.2 %	1.31 [0.56, 3.09]
Zorlu 1993	1/140	1/137		3.4 %	0.98 [0.06, 15.49]
Walsh 1994	1/219	1/223		3.3 %	1.02 [0.06, 16.18]
Walsh 1998	1/918	1/915		3.3 %	1.00 [0.06, 15.91]
Total (95% CI)	**2906**	**2891**		**100.0 %**	**0.89 [0.53, 1.50]**

Total events: 27 (Treatment), 30 (Control)

Heterogeneity: $Chi^2 = 1.47$, df = 5 (P = 0.92); $I^2 = 0.0\%$

Test for overall effect: Z = 0.43 (P = 0.67)

Test for subgroup differences: Not applicable

0.1 0.2 0.5 1 2 5 10

Relative vs Absolute Risk, & NNT

- Experimental event rate – 27/2906 =0.009
- Controlevent rate – 30/2891 =0.010

- Relative Risk Reduction (CER-EER/CER)

  (0.010 – 0.009)/0.010 = 0.1 **(10% relative risk reduction)**

- Absolute Risk Reduction (CER-EER)

  0.010 – 0.009 = 0.001 **(less than 1% absolute risk reduction)**

- Numbers Needed To Treat (1/ARR)

  1/0.001 = 1,000

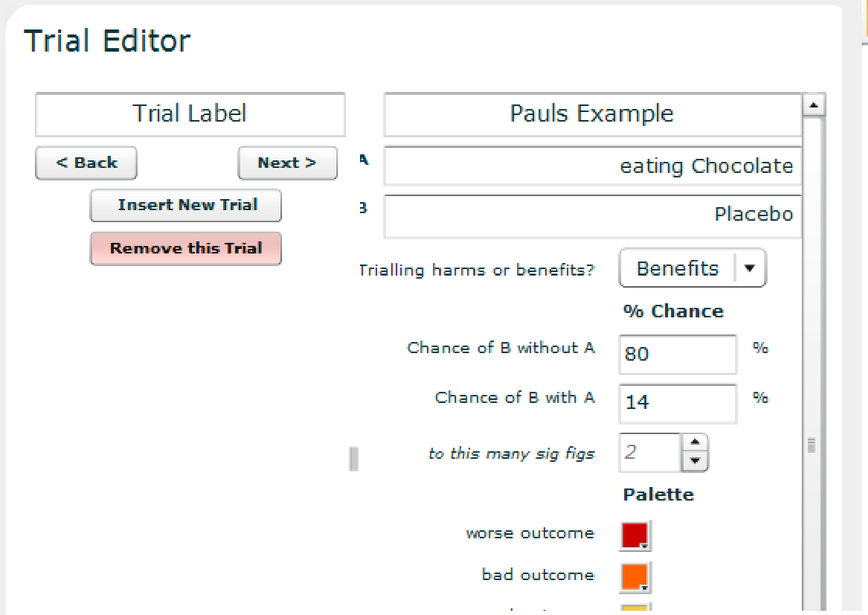
  (need to give 1,000 patients antibiotics to prevent one additional infection)

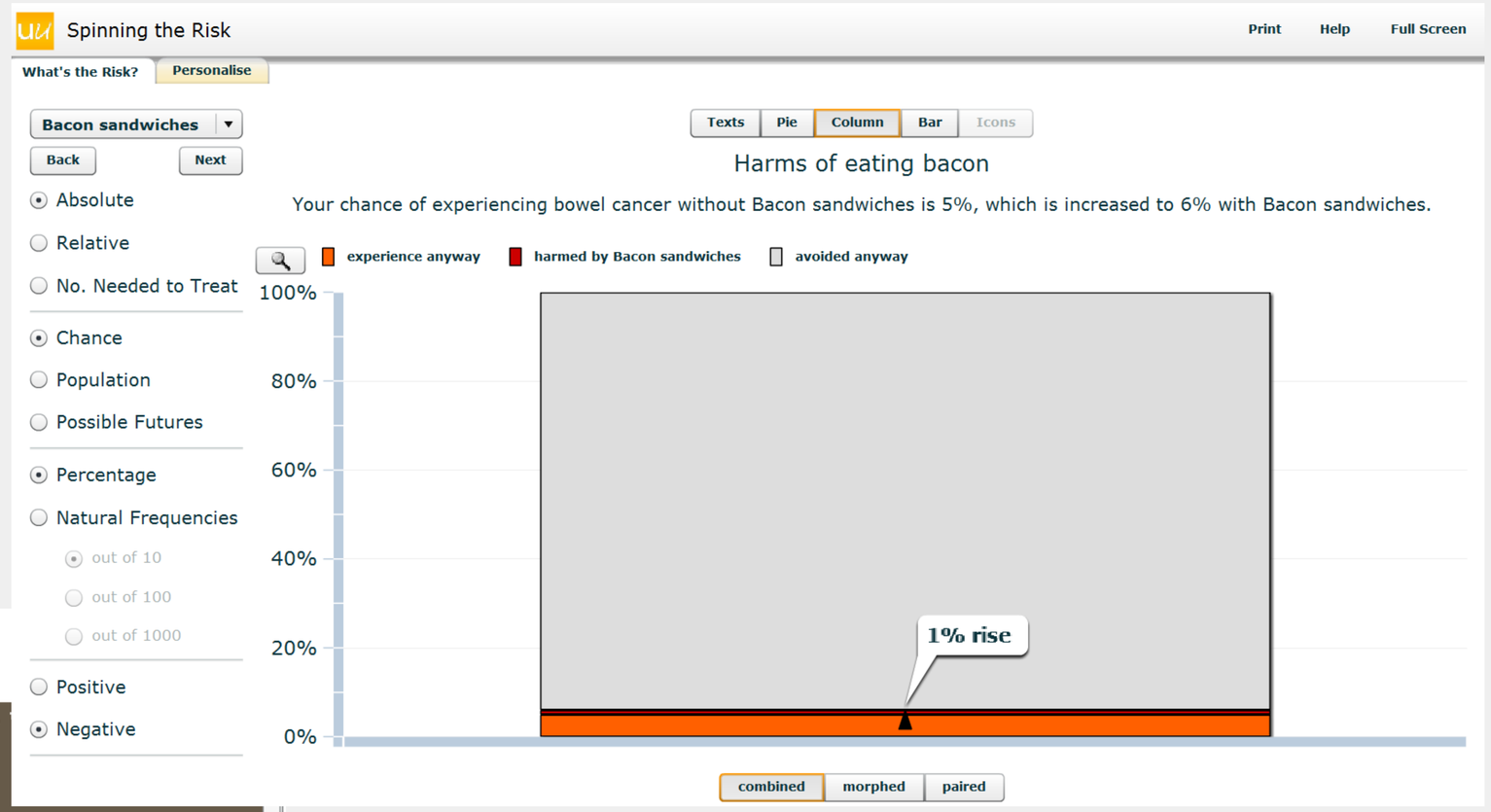
  Odds Ratio (EER/CER) – 0.009/0.010 = 0.9

  - OR of 1 indicates no difference. If CI *do not* include 1 there is statistical significance

Interpretation of Results

https://understandinguncertainty.org/files/RiskDisplay11.swf





Group Exercise

Critical Appraisal of a Systematic Reviews

Part B (**Questions 10 -18**)

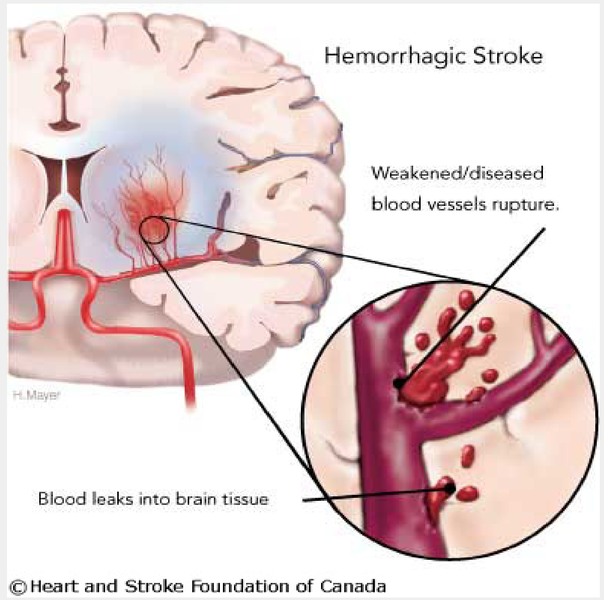
Paul Stevenson

April 2018

Group Exercise

- Read the review article:

Effects of vitamin E on Stroke



- Using the CASP appraisal tool:
- As a group reach a consensus to answer **questions 10-18** of the tool.

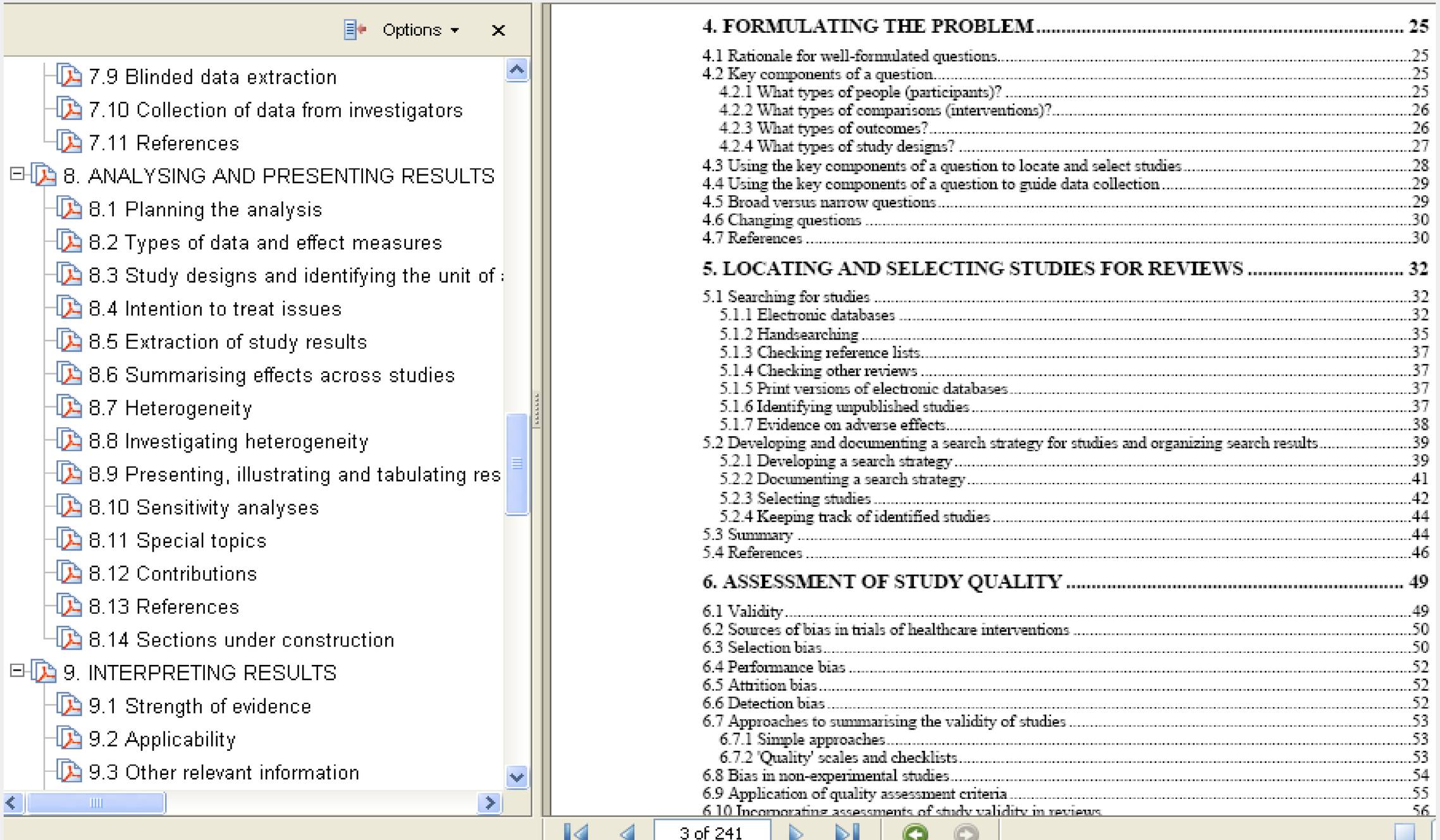
Useful Resources...

Paul Stevenson

April 2018

Cochrane Handbook

http://training.cochrane.org/handbook



CRD Guide
www.york.ac.uk/media/crd/Systematic_Reviews.pdf

CHAPTER 1

CORE PRINCIPLES AND METHODS FOR CONDUCTING A SYSTEMATIC REVIEW OF HEALTH INTERVENTIONS

Assessing the robustness of the synthesis

Towards the end of the synthesis process, the analysis of relationships as described above should lead into an overall assessment of the strength of the evidence. This is essential when drawing conclusions based on the narrative synthesis.

Robustness can relate to the methodological quality of the included studies (such as risk of bias), and/or the credibility of the product of the synthesis process. Obviously, these are related. The credibility of a synthesis will depend on both the quality and the quantity of the evidence base it is built on, and the method of synthesis and the clarity/transparency of its description. If primary studies of poor methodological quality are included in the review in an uncritical manner then this will affect the integrity of the synthesis. Attempts to minimize the introduction of bias might include 'weighting' the findings of studies according to technical quality (i.e. giving greater credence to the findings of more methodologically sound studies) and providing a clear justification for this. Similarly, a clear description of the potential sources of bias within the synthesis process itself helps establish credibility with the reader.

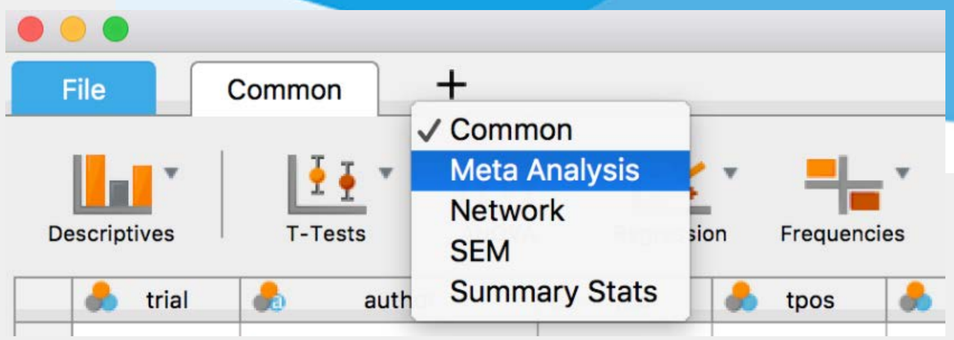
Table 1.4 describes the tools and techniques that might be employed at this stage of the synthesis.

Table 1.4: Assessing the robustness of the synthesis

Use of validity assessment	Use of specific rules to define weak, moderate or good evidence. An example is the approach used by the US Centers for Disease Control and Prevention[131] although there are many other evidence grading systems available. Decisions about the strength of evidence are explicit although the criteria used are often debated.
Reflecting critically on the synthesis process	Use of a critical discussion to address methodology of the synthesis used[132] (especially focusing on its limitations and their potential influence on the results); evidence used (quality, validity, generalisability) – with emphasis on the possible sources of bias and their potential influence on results of the synthesis; assumptions made; discrepancies and uncertainties identified; expected changes in technology or evidence (e.g. identified ongoing studies); aspects that may have an influence on implementation and effectiveness in real settings. Such a discussion would provide information on both the robustness and generalisability of the synthesis.
Checking the synthesis with authors of primary studies	It is possible to consult with the authors of included primary studies in order to test the validity of the interpretations developed during the synthesis and the extent to which they are supported by the primary data.[133] The authors of the primary studies may have useful insights into the possible accuracy and generalisability of the synthesis; this is most likely to be useful when the number of primary studies is small. This is a technique that has been used with qualitative evidence.

Statistical Software - JASP

https://jasp-stats.org/



Qualitative Research

April 2018

What is Qualitative Research

- **Narrative Data** – text, spoken word etc.

- Collecting the data:
  - Questionnaires
  - Interviews
  - Discussion / Focus Groups
  - Logs / Diaries / Documents
  - Observation / Reflection

Qualitative Approach

- Individuals construct their own **versions of reality**

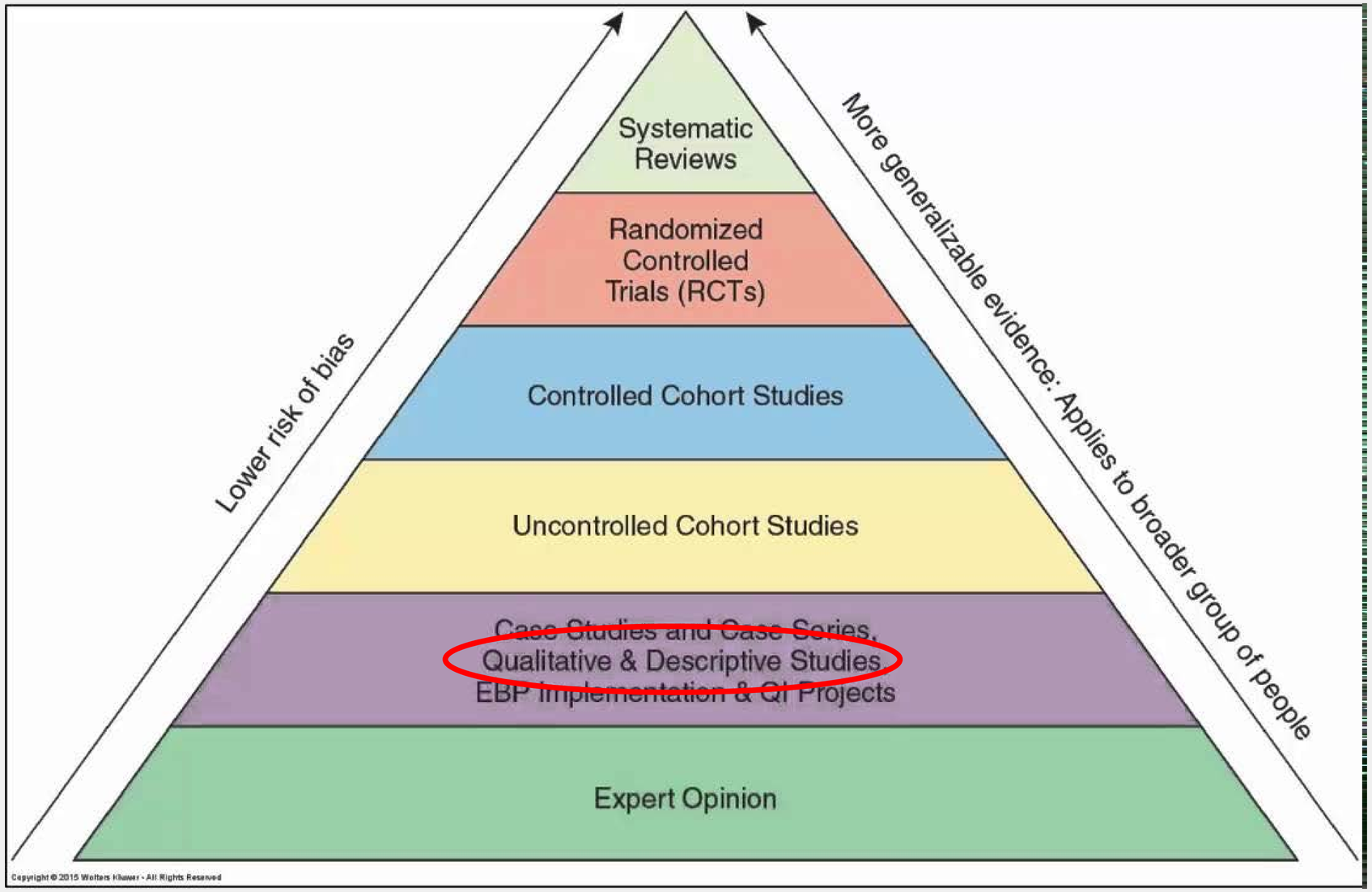
- Qualitative research aims to understand and interpret how **individuals experience** and sense a specific environment.

- **Inductive** – generates new theory ..rather than tests existing theory.

- Concerned with opinions, experiences and feelings of individuals. **Subjective** data.

Main Criticisms of Qualitative Methods

- Why is qualitative research so low in the pyramid of evidence?



Main Criticisms of Qualitative Methods

- The data collection methods (interviews, questionnaires, and observation) may be more at risk of observer effect and bias than quantitative methods.

- Authors may be ‘selective’ in their choice of quotes so that they can promote a specific predetermined conclusion.

- Results may not be generalisable due to small sample sizes, or because subjects were not chosen at random.
  (There can be multiple, simultaneous valid truths)

Why Use Qualitative Data ?

Quantitative research can show us that 65% of diabetic patients do not attend self-management education sessions.

Qualitative research can provide possible explanations why those patients are not engaging with education sessions.

Qualitative Research

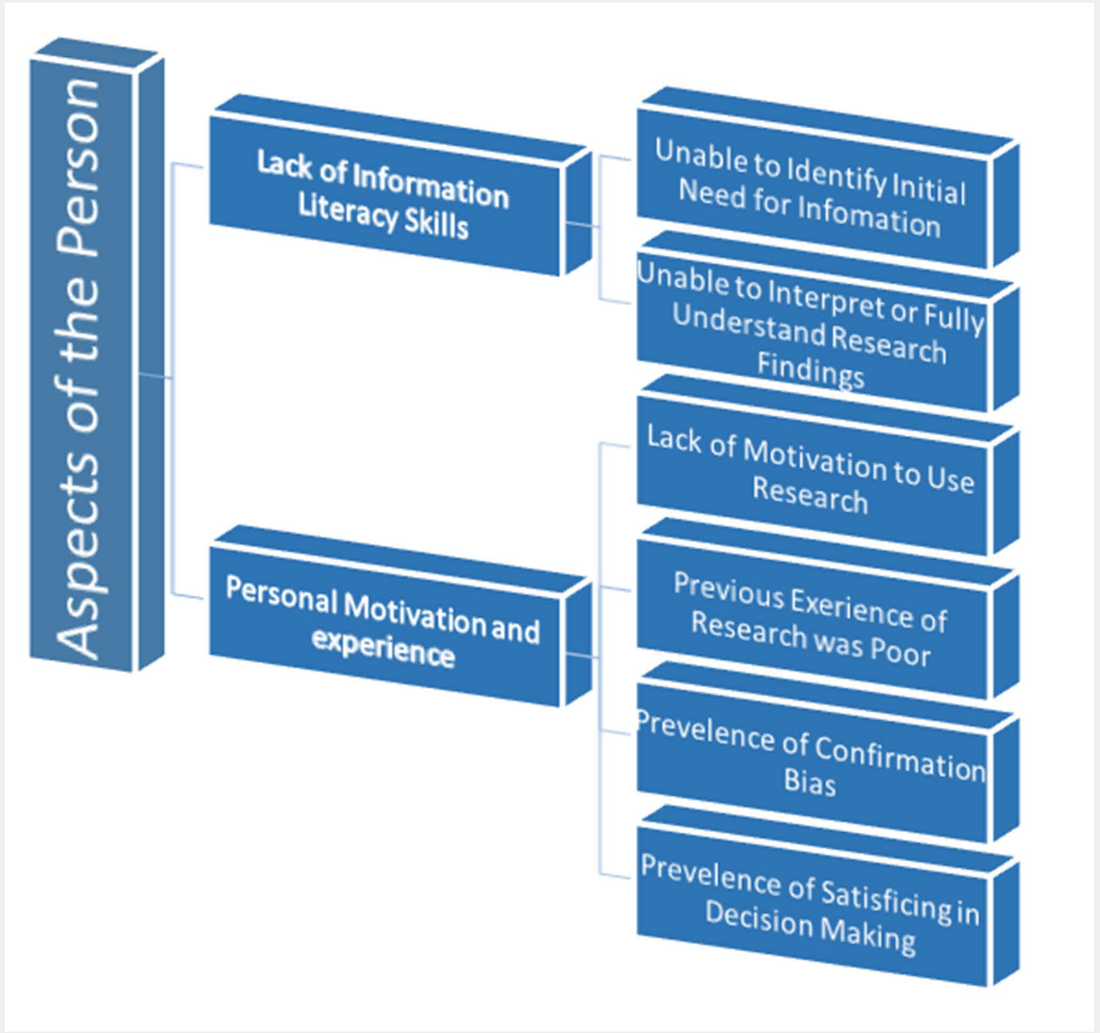
In your group:

Take **5 minutes**

Write down various **descriptions** of what is inside your bag of Randoms...

Thematic Analysis

- Analysing interviews individually
- Identify themes
- List of themes grows quickly to begin with then becomes stable (theoretical saturation).
- Systematic: important to attribute all statements to a theme



Further Information Recommendation: *Coding Manual for Qualitative Researchers*
www.sagepub.com/sites/default/files/upm-binaries/24614_01_Saldana_Ch_01.pdf

Qualitative Research

In your group:

Take **5 minutes**

Identify **themes and sub-themes** that you can see in the data/descriptions we have...

Qualitative Research

Possible Themes

- Colour – red , green, etc
- Taste – sweet, fruity, lime.
- Category – vehicle, sport equipment
  - Sub- Category
    - Sport Equipment
      - Footwear – Roller-skate, Tennis Shoes

Qualitative Research

Why

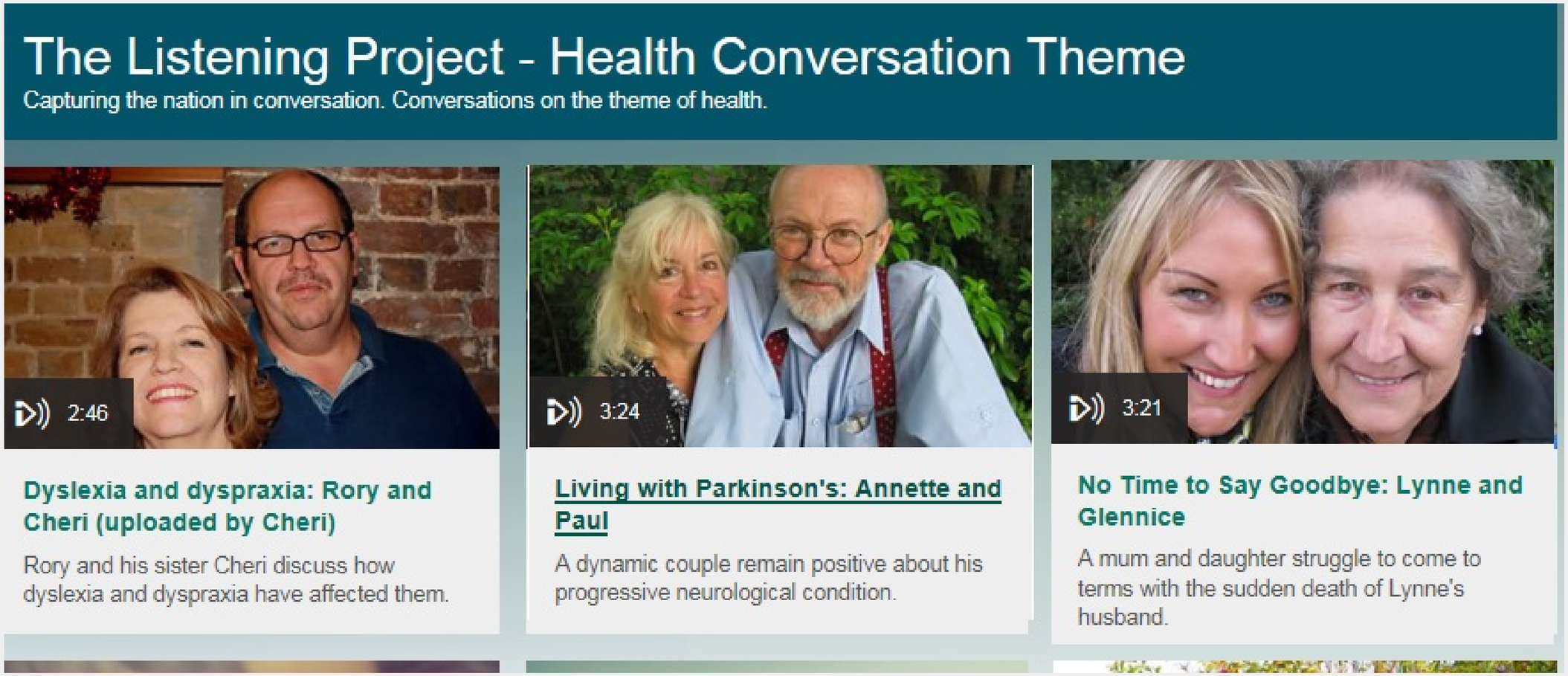
Quick way to (hopefully) show that :

- Different interpretations / Focus can be made by different researchers.
- Illustrate that results will vary dependent on the sample.

Alternative Classroom Exercise

www.bbc.co.uk/programmes/p00p1h38

- As an exercise to understand coding and thematic analysis the Radio 4 Listening Project has hundreds of short interviews that you could utilise .



Critical Appraisal: Qualitative Research

Improving rigour in qualitative research

1 Saturation

2 Multiple coders

3 Triangulation

What size sample?

Is it correct to state that ‘more mentions equates to greater importance’ ?

- Not statistical research
- Not looking for probabilities or measuring outcomes.

- One persons view can be as valid as 100 people
- Often what isn’t mentioned is equally interesting

However…

- **Saturation**: data collection from additional cases no longer elicits new additional information.

Multiple Coders

- The data is coded by more than one researcher.

  The researchers compare their coding to identify if any bias or difference in interpretation exists.

- It is common to limit the intercoder reliability test to a sample of the body of content.

Participant Validation

- Data or results are returned to participants to check for accuracy and resonance with their experiences..

Triangulation & Multiple Coders

- By including **multiple sources** of information we can increase the robustness of the research findings.

- Method Triangulation =
  - questionnaire + interview + observation
- Space Triangulation =
  - acute hospital + community setting + home
- Person Triangulation =
  - patients + nurses + pharmacists + social workers

Transcripts: Data Interpretation

- Not always easy to interpret transcribed data:
- He was ALL RIGHT
    - (He was alright, I liked him)
- HE was all right
    - (He was alright but I wasn’t keen on the others)
- He WAS all right
    - (He used to be alright but isn’t any more)
- He was all right?
    - (You might think he’s alright but I don’t)

Critical Appraisal - Tools

Context II: Setting	10. Within what geographical and care setting is the study carried out? 11. What is the rationale for choosing this setting? 12. Is the setting appropriate and/or sufficiently specific for examination of the research question? 13. Is sufficient detail given about the setting? 14. Over what time period is the study conducted?
Context III: Sample (events, persons, times and settings)	15. How is the sample (events, persons, times and settings) selected? (For example, theoretically informed, purposive, convenience, chosen to explore contrasts) 16. Is the sample (informants, settings and events) appropriate to the aims of the study? 17. Is the sample appropriate in terms of depth (intensity of data collection - individuals, settings and events) and width across time, settings and events (For example, to capture key persons and events, and to explore the detail of inter-relationships)? 18. What are the key characteristics of the sample (events, persons, times and settings)?
Context IV: Outcomes	19. What outcome criteria are used in the study? 20. Whose perspectives are addressed (professional, service, user, carer)? 21. Is there sufficient breadth (e.g. contrast of two or more perspective) and depth (e.g. insight into a single perspective)?

Best Bets - http://bestbets.org/ca/pdf/qualitative.pdf

McMaster University - www.unisa.edu.au/Global/Health/Sansom/Documents/iCAHE/CATs/McMasters_qualreview_version2%200.pdf

Critical Appraisal

- **Who was studied** – are they the ‘right’ cohort?
  - Also , who wasn’t included in the study.

- **What context** did the research take place in?
  - Setting, timeframe (i.e. winter / summer)

- **How was data collected** – is there possibility of bias?
  - i.e. Researchers role and relationship with subject

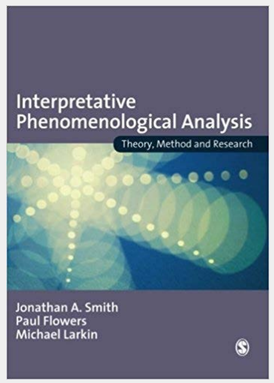
- **How was data analysed** – is it systematic?

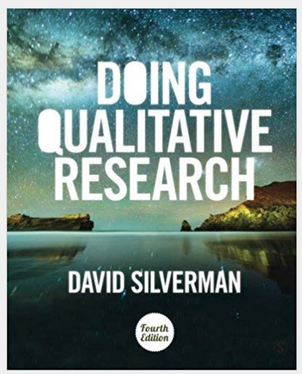
Class Exercise – Appraisal of Paper

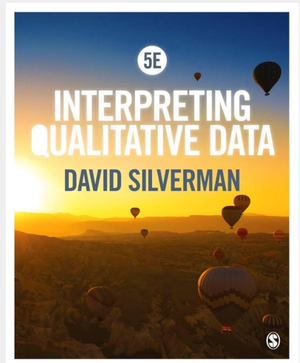
- Use the CASP appraisal tool to assess the quality of the paper you have been given.
- Answer the 10 questions in the tool
- You have approximately **30 minutes** to do this.

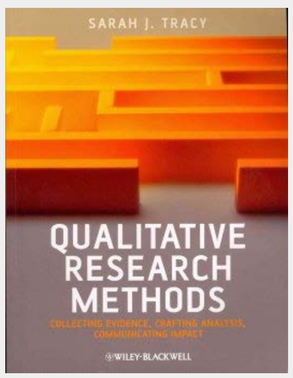
Continued Learning

- Read some Qualitative Textbooks to get a greater understanding of Methods and Methodology.









- Critically Appraise a qualitative LIS paper or healthcare paper with your colleagues for practice.

THANK YOU for your participation .

Any Questions?

Paul.Stevenson@anhst.nhs.uk.