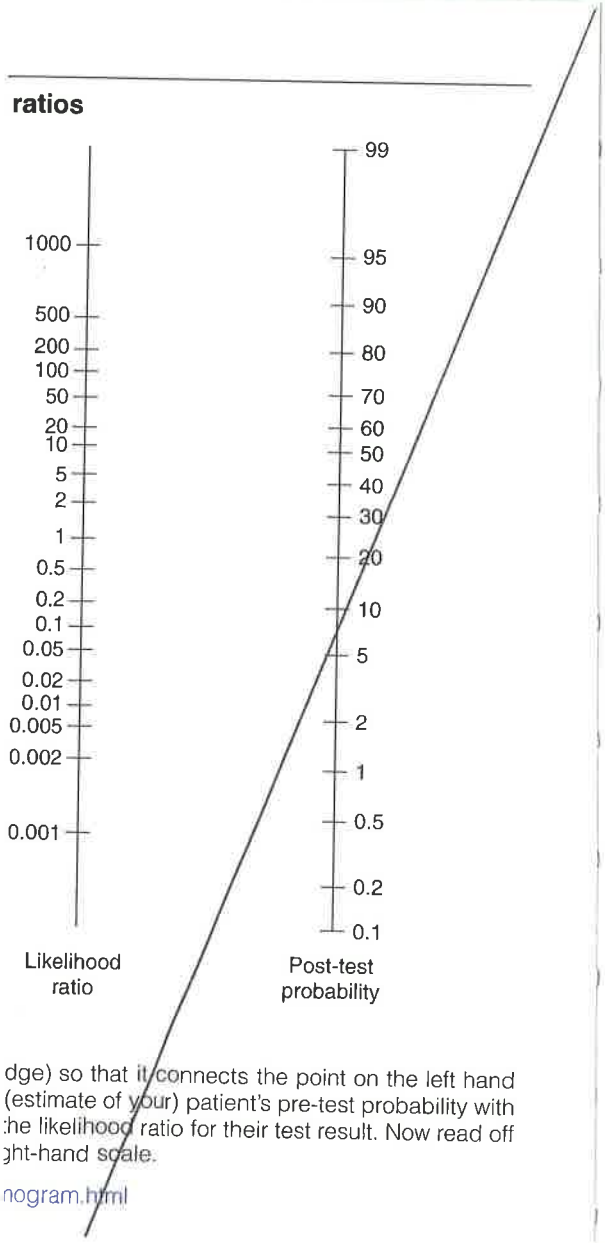
Dear EBM Learners,

The next session will focus on Systematic Review/Meta-analysis. This is the next step up from assessing individual studies.

Please see the following pages for definitions of systematic reviews and meta-analysis how to assess the validity and how to analyze the results.

Most systematic reviews are reviews of therapy, so we will tend to focus on those.



dge) so that it connects the point on the left hand (estimate of your) patient's pre-test probability with :he likelihood ratio for their test result. Now read off ght-hand scale.

nogram.html

Appraising systematic reviews

Is the systematic review valid?

1. Is it a systematic review of high-quality studies which are relevant to your question?
2. Does the methods section adequately describe:
   - a comprehensive search for all the relevant studies?
   - how the reviewers assessed the validity of each study?
3. Are the studies consistent, both clinically and statistically?

Are the results important?

If the review reports odds ratios (ORs), you can generate an NNT if you have an estimate of your patient's expected event rate (PEER).

$$NNT = \frac{1 - \{PEER \times (1 - OR)\}}{(1 - PEER) \times PEER \times (1 - OR)}$$

A systematic review is "a review of a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyse data from studies that are included in the review. Statistical methods may or may not be used to analyse and summarise the results of the included studies" (Cochrane Library 1998, Glossary).

Three key features of such a review are:

- a strenuous effort to locate all original reports on the topic of interest
- critical evaluation of the reports
- conclusions are drawn based on a synthesis of studies which meet pre-set quality criteria

When synthesising results, a meta-analysis may be undertaken. This is "the use of statistical techniques in a systematic review to integrate the results of the included studies" (Cochrane Library 1998, Glossary), which means that the authors have attempted to synthesise the different results into one overall statistic. The best source of systematic reviews is the Cochrane Library, available by subscription on CD or via the internet. Many of the systematic reviews so far completed are based on evidence of effectiveness of an intervention from randomised controlled trials (RCTs).

Is the systematic review valid?

1. Is it a systematic review of high-quality studies which are relevant to your question?

This question asks whether the research question in the review is clearly defined and the same as the one you are considering, and whether the studies covered by the review are high quality. Reviews of poor-quality studies simply compound the problems of poor-quality individual studies. Sometimes, reviews combine the results of variable-quality trials (for example randomised and non-randomised trials in therapy); the authors should provide separate information on the subset of randomised trials.

2. Does the methods section describe how all the relevant trials were found and assessed?

The paper should give a comprehensive account of the sources consulted in the search for relevant papers, the search strategy used to find them, **and** the quality and relevance criteria used to decide whether to include them in the review.

*The reviewers' search should aim to minimise **publication bias:** the tendency for negative results to be unequally reported in the literature.*

Search strategy

Some questions you can ask about the search strategy:

- The authors should include hand searching of journals and searching for unpublished literature.
- Were any obvious databases missed?
- Did the authors check the reference lists of articles and of textbooks (citation indexing)?
- Did they contact experts (to get their list of references checked for completeness and to try and find out about ongoing or unpublished research)?
- Did they use an appropriate search strategy: were important subject terms missed?

Did the authors assess the trials' individual validity?

You should look for a statement of how the trials' validity was assessed. Ideally, two or more investigators should have applied these criteria independently and achieved good agreement in their results.

You need to know what criteria were used to select the research. These should include who the study participants were, what outcomes were assessed. A point to consider is th criteria, the less generalisable are the results. Howeve with using very broad inclusion criteria, when heterog an issue.

The importance of a clear statement of inclusion criteria is that studies should be selected on the basis of these criteria (that is, any study that matches these criteria is included) rather than selecting the study on the basis of the results.

3. Are the studies consistent, both clinically and statist

You have to use your clinical knowledge to decide wh patients, interventions, and outcome measures were merit combining their results. If not, this **clinical heterc** review.

Similarly, you would question the review's validity if t each other. Unless this **statistical heterogeneity** ca (such as by differences in patients, dosage, or c should lead you to be very cautious about believing ar review.

Are the results important?

Terms that you will probably come across when lo include vote counting, odds ratios, and relative risks, a

Vote counting

If a systematic review does not contain a meta-anal combining the data from separate trials), the results m count of the number of studies supporting an inter supporting it. This assumes equal weight being given to

Odds ratio (OR)

In measuring the efficacy of a therapy, odds can be us of an event are the probability of it occurring compar occurring.

By dividing the odds of an event in the experimental group by the odds in the control group, we can measure the efficacy of the treatment. ORs are useful because they can be used in a meta-analysis to combine the results of many different trials into one overall measure of efficacy.

If the exp the OR v group has 1; and if t two group

v valid?

high-quality studies which are relevant to

he research question in the review is clearly defined
are considering, and whether the studies covered by
Reviews of poor-quality studies simply compound the
idual studies. Sometimes, reviews combine the results
example randomised and non-randomised trials in
d provide separate information on the subset of

describe how all the relevant trials were

prehensive account
n the search for
rategy used to find
vance criteria used
hem in the review.

*The reviewers' search should aim to minimise **publication bias:** the tendency for negative results to be unequally reported in the literature.*

about the search strategy:

de hand searching of journals and searching for

es missed?
reference lists of articles and of textbooks (citation

o get their list of references checked for completeness
it ongoing or unpublished research)?
iate search strategy: were important subject terms

als' individual validity?

nt of how the trials' validity was assessed. Ideally, two
d have
ndently
ent in

criteria
search.

The importance of a clear statement of inclusion criteria is that studies should be selected on the basis of these criteria (that is, any study that matches these criteria is included) rather than selecting the study on the basis of the results.

These should include who the study participants were, what was done to them, and what outcomes were assessed. A point to consider is that the narrower the inclusion criteria, the less generalisable are the results. However, this needs to be balanced with using very broad inclusion criteria, when heterogeneity (see below) becomes an issue.

3. Are the studies consistent, both clinically and statistically?

You have to use your clinical knowledge to decide whether the groups of patients, interventions, and outcome measures were similar enough to merit combining their results. If not, this **clinical heterogeneity** would invalidate the review.

Similarly, you would question the review's validity if the trials' results contradicted each other. Unless this **statistical heterogeneity** can be explained satisfactorily (such as by differences in patients, dosage, or durations of treatment), this should lead you to be very cautious about believing any overall conclusion from the review.

Are the results important?

Terms that you will probably come across when looking at systematic reviews include vote counting, odds ratios, and relative risks, amongst others.

Vote counting

If a systematic review does not contain a meta-analysis (a statistical method for combining the data from separate trials), the results may be presented as a simple count of the number of studies supporting an intervention and the number not supporting it. This assumes equal weight being given to each study, regardless of size.

Odds ratio (OR)

In measuring the efficacy of a therapy, odds can be used to describe risk. The odds of an event are the probability of it occurring compared to the probability of it not occurring.

By dividing the odds of an event in the experimental group by the odds in the control group, we can measure the efficacy of the treatment. ORs are useful because they can be used in a meta-analysis to combine the results of many different trials into one overall measure of efficacy.

If the experimental group has lower odds, the OR will be less than 1; if the control group has lower odds, the OR will be above 1; and if there is no difference between the two groups, the OR will be exactly 1.

To calculate the NNT for any OR and PEER:

$$NNT = \frac{1 - [PEER \times (1 - OR)]}{(1 - PEER) \times PEER \times (1 - OR)}$$

Logarithmic odds

Odds ratios are usually plotted on a log scale to give an equal line length on either side of the line of "no difference". If odds ratios are plotted on a log scale, then a log odds ratio of 0 means no effect, and whether or not the 95% confidence interval crosses a vertical line through zero will lead to a decision about its significance.

Binary or continuous data

Binary data (an event rate: something that either happens or not, such as numbers of patients improved or not) is usually combined using odds ratios. Continuous data (such as numbers of days, peak expiratory flow rate) is combined using differences in mean values for treatment and control groups (weighted mean differences or WMD) when units of measurement are the same, or standardised mean differences when units of measurement differ. Here the difference in means is divided by the pooled standard deviation.

How precise are the results?

The statistical significance of the results will depend on the extent of any confidence limits around the result (see p17). The review should include confidence intervals for all results, both of individual studies and any meta-analysis.

Further reading

Altman D. *Practical Statistics for Medical Research*. Edinburgh: Churchill Livingstone, 1991.

Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomised control trials and recommendations of clinical experts. *JAMA* 1992;**268**:240–8.

Cochrane Library: http://www.update-software.com

NHS Centre for Reviews and Dissemination: http://www.york.ac.uk/inst/crd/

Oxman AD *et al*. Users' Guides to the Medical Literature VI: How to use an overview. *JAMA* 1994;**272**(17):1367–71.

Sackett DC, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine: How to practice and teach EBM*. Churchill Livingstone, 2000.

Seers K. Systematic review. In M Dawes *et al*. (eds) *Evidence-Based Practice: a primer for health care professionals*. Edinburgh: Churchill Livingstone, 1999, pp85–100.

Appraising articles on harm/aetiology

Is the study valid?

1. Was there a clearly defined research question?
2. Were there clearly defined, similar groups of patient
3. Were exposures and clinical outcomes measured th groups?
4. Was the follow up complete and long enough?
5. Does the suggested causative link make sense?

Are the valid results from this study import

		Adverse outco	
		Present (case)	Abs (con
Exposure	Yes (Cohort)	a	b
	No (Cohort)	c	d
Totals		a + c	b +

In a randomised trial or cohort study: Relative risk
In a case–control study: Odds ratio =