SESSION #3
THERAPY

EVALUATING THE EVIDENCE
VALIDITY AND RESULTS

TODAY WE FOCUS ON VALIDITY

REMINDER: THE EBM PROCESS

- OUR PATIENT
- QUESTION (PICO)
- SEARCHING FOR AN ANSWER (EVIDENCE)
- APPRAISING THE EVIDENCE
  - EVALUATING FOR VALIDITY
  - ANALYZING THE RESULTS
- APPLYING THE RESULTS TO OUR PATIENT

THERAPY

- We have arrived at the next step of the EBM process as noted in the previous slide – Appraisal of the evidence (the articles obtained from a search)
- We will be learning how to appraise a THERAPY ARTICLE
- There are 2 issues to consider when appraising evidence:
  METHODOLOGIC VALIDITY AND RESULTS
- We will focus on VALIDITY for THIS SESSION

VALIDITY FOR A THERAPY ARTICLE

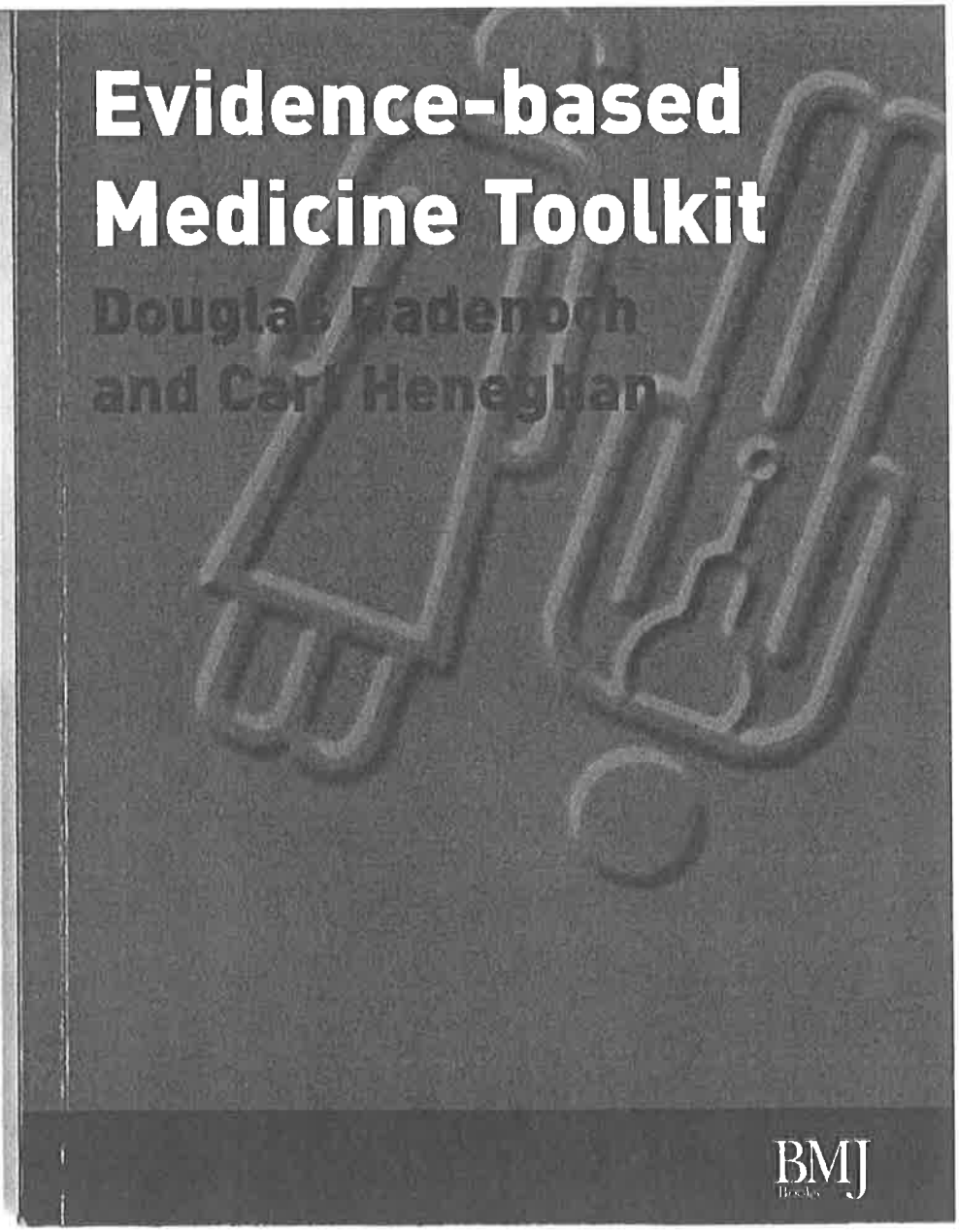
These are the basic validity questions

- Was the assignment of patients to treatment randomized?
- Was follow-up sufficiently long and complete?
- Were all patients analyzed in the groups to which they were randomized (intention to treat)?
- Were patients and clinicians kept blind to treatment?

These, as well as a few other points, will be discussed in the following pages

Evidence-based Medicine Toolkit

Douglas Badenoch
and Carl Heneghan



BMJ Books

mation on the internet, you should be wary
al; ask yourself first: does this website have
ich has been applied to the material?

o.com)

ategories, but there is considerable overlap between
t search, which will list all the Yahoo categories as

/sis, analytic, analytical, analyse, etc.

e adjacent to each other: for example "heart attack"

ildbirth = documents must contain both words

ccur in title of the document (t:natural childbirth) or words
web address (u:uk)

of your search: documents that contain multiple
are ranked highest; those that match your search in
ghest. Other good search engines include Google
no advertising on its simple front-end and a very
n page.

uk/caspfew/: includes introductory exercises, toolkit and

s/searching.html: includes tips on how to target high-quality
(therapy, diagnosis, etc.).
-Based Principles and Practice. Hamilton, ON: BC Decker,

an information skills approach. In M Dawes (ed.), *Evidence-*
ealth care professionals. Edinburgh: Churchill Livingstone,

internet: http://www.shef.ac.uk/~scharr/ir/netting.html.

Appraising therapy articles

Is the study valid?

1. Was there a clearly defined research question?
2. Was the assignment of patients to treatments randomised and was the randomisation list concealed?
3. Were all patients accounted for at its conclusion? Was there an "intention-to-treat" analysis?
4. Were research participants "blinded"?
5. Were the groups treated equally throughout?
6. Did randomisation produce comparable groups at the start of the trial?

Are the results important?

Relative Risk Reduction (RRR) = (CER – EER) / CER
Absolute Risk Reduction (ARR) = CER – EER
Number Needed to Treat (NNT) = 1 / ARR

Is the study valid?

1. Was there a clearly defined research question?

What question has the research been designed to answer? Was the question focused in terms of the population group studied, the intervention received and the outcomes considered?

2. Were the groups randomised?

The most important type of research for answering therapy questions is the randomised controlled trial (RCT). The major reason for randomisation is to create two (or more) comparison groups which are similar. To reduce bias as much as possible, the decision as to which treatment a patient receives should be determined by random allocation.

Concealed randomisation

As a supplementary point, clinicians who are entering patients into a trial may consciously or unconsciously distort the balance between groups if they know the

treatments given to previous patients. For this reason, it is preferable that the randomisation list be concealed from the clinicians.

Why Is this Important?

Randomisation is important because it spreads all confounding variables evenly amongst the study groups, even the ones we don't know about.

Stratified randomisation

True random allocation can result in some differences occurring between the two groups through chance, particularly if the sample size is small. This can lead to difficulty when analysing the results if, for instance, there was an important difference in severity of disease between the two groups. Using stratified randomisation, the researcher identifies the most important factors relevant to that research question; randomisation is then stratified such that these factors are equally distributed in the control and experimental groups.

3. Were all patients accounted for at its conclusion?

There are three major aspects to assessing the follow up of trials:

- Did so many patients drop out of the trial that its results are in doubt?
- Was the study long enough to allow outcomes to become manifest?
- Were patients analysed in the groups to which they were originally assigned (intention-to-treat)?

Drop-out rates

The undertaking of a clinical trial is usually time-consuming and difficult to complete properly. If less than 80% of patients are adequately followed up then the results may be invalid. The American College of Physicians has decided to use 80% as its threshold for inclusion of papers into the *ACP Journal* and *Evidence-Based Medicine*.

Length of study

Studies must allow enough time for outcomes to become manifest. You should use your clinical judgement to decide whether this was true for the study you are appraising, and whether the length of follow up was appropriate to the outcomes you are interested in.

Intention-to-treat

Sometimes, patients may change treatment aims during the course of a study, for all sorts of reasons. If we analysed the patients on the basis of what treatment they got rather than what they were allocated (intention-to-treat), we have altered the even distribution of confounders produced by randomisation. So, all

patients should be analysed in the groups to which they were originally randomised, even if this is not the treatment they actually got.

4. Were the research participants "blinded"?

Ideally, patients and clinicians should not know whether they are receiving the treatment. The assessors may unconsciously bias their assessment of outcomes if they are aware of the treatment. This is known as observer bias.

So, the ideal trial would blind patients, carers, assessors and analysts alike. The terms single, double and triple blind are sometimes used to describe these permutations. However, there is some variation in their usage and you should check to see exactly who was blinded in a trial. Of course, it may have been impossible to blind certain groups of participants, depending on the type of intervention. Note also that concealment of randomisation, which happens before patients are enrolled, is different from blinding, which happens afterwards.

Placebo control

Patients do better if they think they are receiving a treatment than if they do not; the placebo effect is a widely accepted potential bias in trials.

So, the ideal trial would perform **"double-blind" randomisation** (where both the patient and the clinician do not know whether they are receiving active or placebo treatment), and where the randomisation list is concealed from the clinician allocating treatment (see above). In some cases, it would not be possible to blind either or both of the participants (depending on the type of intervention and outcome), but researchers should endeavour to carry out blind allocation and assessment of outcomes wherever possible.

5. Equal treatment

It should be clear from the article that, for example, there were no co-interventions which were applied to one group but not the other and that the groups were followed similarly with similar check-ups.

6. Did randomisation produce comparable groups at the start of the trial?

The purpose of randomisation is to generate two (or more) groups of patients who are similar in all important ways. The authors should allow you to check this by displaying important characteristics of the groups in tabular form.

Outcome measures

An outcome measure is any feature that is recorded to determine the progression of the disease or problem being studied. Outcomes should be objectively defined and

measured wherever possible. Often, outcomes are expressed as mean values of measures rather than numbers of individuals having a particular outcome. The use of means can hide important information about the characteristics of patients who have improved and, perhaps more importantly, those who have got worse.

Are the results important?

Two things you need to consider are how large is the treatment effect and how precise is the finding from the trial.

In any clinical therapeutic study there are three explanations for the observed effect:

1. Bias.
2. Chance variation between the two groups.
3. The effect of the treatment.

Once bias has been excluded (by asking if the study is valid), we must consider the possibility that the results are a chance effect.

p Values

Alongside the results, the paper should report a measure of the likelihood that this result could have occurred if the treatment was no better than the control. The *p* value is a commonly used measure of this probability.

For example, a p value of <0.01 means that there is a less than 1 in 100 (1%) probability of the result occurring by chance; p<0.05 means this is less than 1 in 20 probability.

Quantifying the risk of benefit and harm

Once chance and bias have been ruled out, we must examine the difference in event rates between the control and experimental groups to see if there is a significant difference. These **event rates** can be calculated as shown below:

	Control	Experimental	
Event	a	b	Control event rate (CER) = a / (a+c)
No event	c	d	Experimental event rate (EER) = b / (b + d)

Relative risk reductic

Relative risk reducti event rate (EER) cor

Absolute risk reducti

Absolute risk redu experimental group.

ARR is a more clinic "factors out" the bas when compared to a below, where the sa greater absolute ber

	CER	EE
1	0.36 (36%)	0.3 (34
2	0.036% (3.6%)	0.0 (3.4

Number needed to tr

Number needed to t absolute number of p is the inverse of the a