

Critical Reasoning

Lecture-Seminar 8

Evidential Standards in Medicine

Dr. Ioannis Votsis
NCH, Philosophy Faculty

ioannis.votsis@nchlondon.ac.uk



Introduction

What went on before

- The history of medicine is replete with cases where treatments were either ineffective or down-right harmful.

Example:

blood-letting (as a cure all)

Radithor (as a cure all)



NB: Some of these may of course have been good for some things but not others.

The movement and its aims

- Evidence-Based Medicine (EBM) is a relatively recent movement (c. 80s-90s) in the practice of medical research.
- The name seems to suggest that prior to the advent of this movement medicine was not evidence-based.
- This is obviously incorrect. Perhaps a better name would be: ‘Standardised-Evidential Practice in Medicine’.
- EBM emphasises the import of unified standards and the ranking of evidence across medical practice and research.

EBM in Detail

A call to revise and unify standards

- EBM-ers target the overreliance on:
 - (i) clinical experience and judgment
 - (ii) background theory
- They make a number of recommendations, including:
 - * Update the knowledge of medics regularly.
 - * Make results of clinical trials more widely known.
 - * Fix what constitutes the best available evidence.
- The last of these, viz. evidential weight, is the one they focus on the most. We follow suit here.

The naive reading of EBM

- Only evidence sourced from Randomised Controlled Trials (RCTs) establishes anything in medical research.

NB: In short, no RCT → no real evidence.

What is an RCT? *Very roughly*, it's an experiment where half of the people are given a treatment and half a placebo.

- That means traditional sources of evidence like clinical experience and judgment should be dismissed.

Sackett et al. (2000), for example, assert that “If the study was not randomized we’d suggest that you stop reading it and go on to the next article in your search” (p. 108).

Effective non-RCT therapies

- Yet, it's clear that there are many cases of therapies that do not enjoy RCT evidence and seem to work just fine.

Examples:

- * penicillin in cases of (bacterial) pneumonia
- * diuretics in cases of heart failure
- * appendectomy in cases of acute appendicitis
- Also, the question arises whether it is ethical to conduct RCTs when a therapy is already known to work.

The sophisticated reading of EBM

- Other types evidence may be available and *should* be consulted but evidence sourced from RCTs trumps them.
- To be precise, there is a hierarchy of evidence and RCTs sit at the top (or very near it).

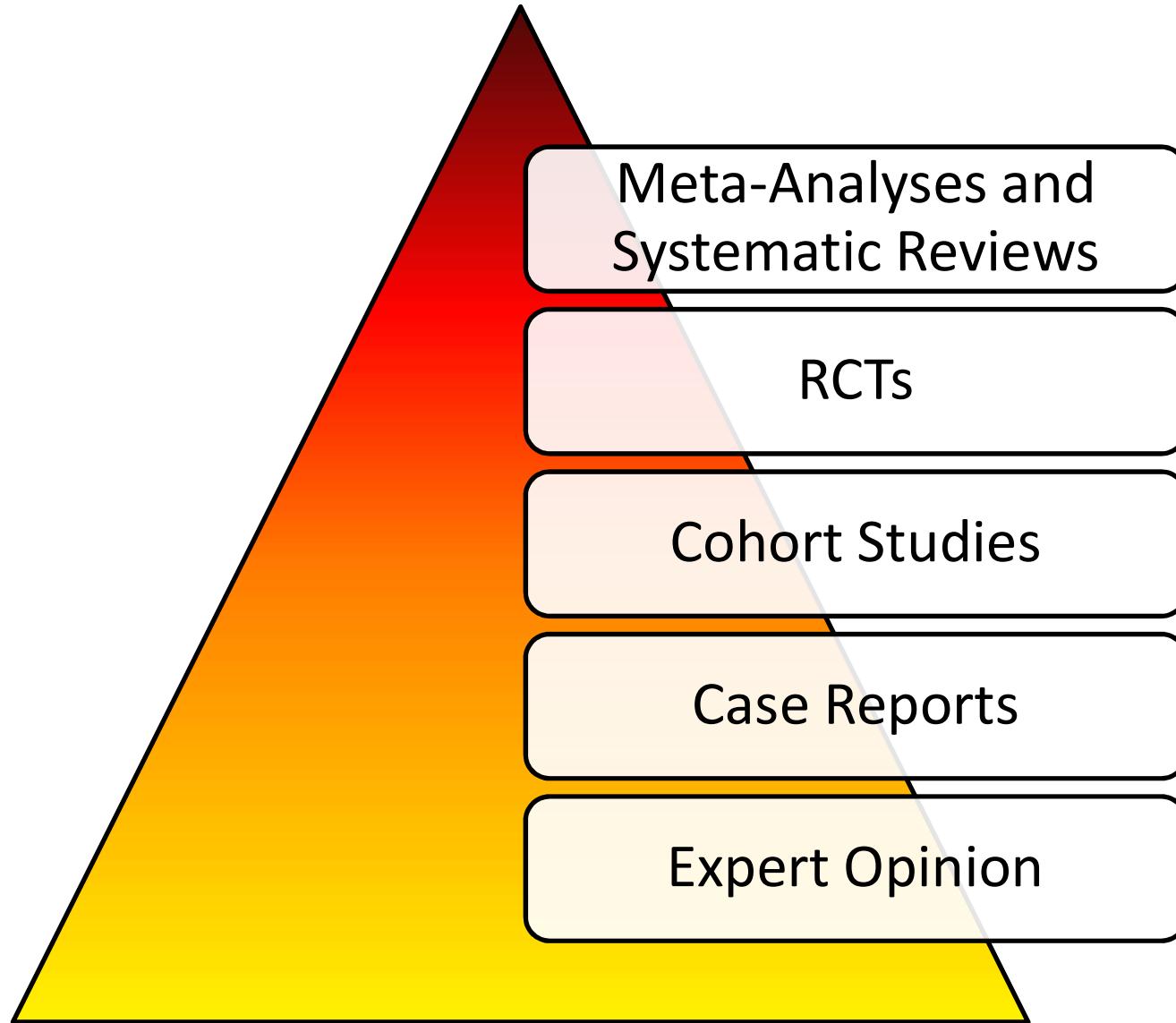
“EBM is not restricted to randomized trials and meta-analyses. . . . some questions about therapy do not require randomized trials (successful interventions for otherwise fatal conditions) or cannot wait for the trials to be conducted. And if no randomized trial has been carried out for our patient’s predicament, we must follow the trail to the next best external evidence and work from there” (Sackett et al. 2000, p. 72).

Evidence hierarchies

- Some sources of evidence are meant to be epistemically privileged to others.
- They are thus are ranked higher than them. What determines which ranks higher?
- Generally speaking, but by no means always, this is fixed strictly by imposing further constraints on the source.

Example: RCTs are CTs with randomisation.

Evidence hierarchies: An example



Weighing evidence

- At this point, the following question arises:

How exactly are we to make a combined judgment when our evidence comes from different sources?

- Ranking sources provides an order of preference but does not tell us how to weigh these.
- For example, ranking makes clear that RCTs are more desirable than CTs but does not tell us how much more.
- Assigning weights to evidence allows us to do that. The trouble is to determine weights in a *non-arbitrary* way.



Rank the sources of evidence (1 best... 5 worst) according to EBM hierarchies.

What does the naive reading of EBM say?

RCTs are the
best evidence

CTs are the
only evidence

RCTs are the
only evidence

CTs are the
best evidence

Dissecting RCTs

Why (R)CTs?

- Why RCTs? Before we answer this question, we first need to fathom why CTs?
- In any study, we want to ensure that any effect is attributed to a specific action (in medicine: a treatment).
- **Problem:** We cannot know what would happen to *the same* objects or subjects *in the absence* of that action.
- To overcome this problem we use a comparison group which is not subjected to that specific action.

Group composition

- We call the group subjected to the specific action the *treatment group* and the other one the *control group*.
- Objects or subjects in the two groups must:
 - * possess the same condition under study (e.g. same illness)
 - * be differentially subjected to actions (e.g. treatment vs. no-treatment)
- But even this is not enough!



Confounding factors

- Suppose the two groups contain subjects with a given illness and the specific action is a treatment (vs. a placebo).
- If those in the treatment group recover or at least recover faster, have we proven that the treatment is successful?
- The answer is NO! All sorts of *confounding* factors, other than the treatment, could have led to that positive effect.

Examples:

- * Individuals in the control group may have been suffering from a harsher form of that medical problem.
- * Individuals in the control group may be more frail.

Other things must be made equal!

- Only if the two groups are *equal in all relevant respects*, do we know that the treatment is the only difference.
- More realistically:

The closer the two groups are to being equal, the more likely that we'll find the true contribution of the treatment.

- Various ways to mitigate this problem. For example:
 - * Remove subjects with confounding factors from the study.
 - * Equally distribute these factors by intentional selection.

The R in RCT

- One major obstacle in the quest to make *other things equal* is the fact that there may be *unknown confounding factors*.
- Randomisation is meant to solve this problem since it is an unbiased method of distributing subjects into the groups.
- The hope is that any known or unknown confounding factors are *evenly* distributed between the two groups.

NB: After a random assignment we can check whether the known confounders have been evenly distributed.

Randomisation: Quotations

“In a randomized trial, the only difference between the two groups being compared is that of most interest: the intervention under investigation” (M. Clarke, Director, Cochrane Centre).

“[Presumably t]he strongest argument for the epistemic superiority of randomized trials... is precisely that RCTs are alleged to solve the problem of ‘unknown factors’: a randomized trial is – allegedly – controlled for all factors known and unknown” (Worrall 2007, p. 1004).

A confounding factor

undermines our confidence in the effectiveness of a treatment

is always known

overdetermines the effectiveness of a treatment

is always unknown

What is randomisation meant to achieve?

Start the presentation to see live content. Still no live content? Install the app or get help at PollEv.com/app

Internal Validity Problems

Internal validity

- A study is internally valid when the treatment's impact is accurately measured in the study population.

Questioning the epistemic power of RCTs

- Worrall (2002; 2007a; 2007b) argues that there are issues concerning the internal validity of RCTs.
- He doesn't deny their usefulness but dismisses their claimed superiority over other kinds of evidence sources.
- One of the arguments he criticises is that randomisation guards against all (i.e. known and unknown) confounders.

Randomisation offers no guarantees

- Worrall uses a well-known ‘study’ to promote his view.

Leibovici et al. (2001):

Records of 3393 inpatients with bloodstream infections at Rabin Medical Centre admitted and released 1990-1996.

In 2000, random number generator assigns patients into 2 groups: control group (nothing), ‘treatment’ group (prayer).

Hospital stay length and duration of fever much lower in the ‘treatment’ group. *Conclusion:* ‘Treatment’ is effective.

- **Upshot:** Clearly, the ‘treatment’ is *not* effective. Thus, Worrall claims, the R in RCT is no guarantee of effectiveness.

Indefinitely many confounders?

- The underlying reason why Worrall is doubtful of the epistemic power of RCTs is his belief that:

“... the list of possible ‘unknown’ factors is indefinitely long...” (1007-8).
- Suppose, for the sake of the argument, that unknown confounders are indeed numerous in a given study.
- Then the probability that any randomisation will evenly distribute all of them goes down with the size of the study.
- **Warning:** We should not assume their numerosity a-priori!
Invariant results over multiple RCTs count against this.

Confounder strength

- Beyond not being able to guarantee the even distribution of confounders there is also a problem relating to strength.
- Suppose:
 - * a study tracks 100 patients
 - * 20 of them possess the same unknown confounder
 - * no other known or unknown confounders
 - * randomisation distributes these evenly
- Does this mean that any positive result can safely be attributed to the intervention (i.e. treatment)?
- Alas, no! The same confounder may manifest itself in different strengths, e.g. variable immune response.



The End