PH 458/231
Evidence & Policy
Lecture 10

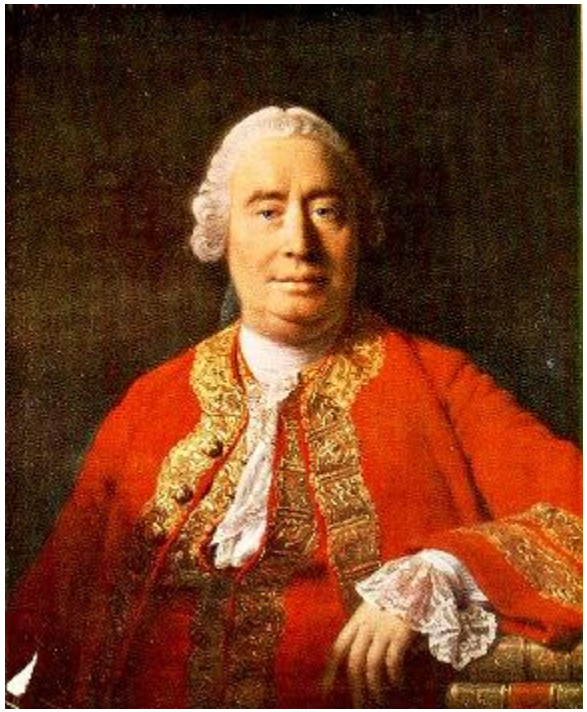
John Worrall
LAK 3.02
Office Hours:
Monday 13.30-14.30
Thursday 12.30 – 13.30

“Evidence-Based Medicine”

- What else?

Evidence based everything!

- “A wise man proportions his belief to the evidence.” David Hume *Enquiry Concerning Human Understanding*



“Evidence-Based Medicine”

- H. W. Haggard *The Doctor in History*:
- The King [Charles II] was bled to the extent of a pint from his right arm. Next [his doctor] drew 8 ounces of blood from his left shoulder… gave an emetic to make the King vomit, two physics, and an enema containing antimony, rock salt, marshmallow leaves , violets, beetroot, camomile flowers, fennel seed, linseed, cardamom seed, saffron, cochineal, and aloes. The King’s head was then shaved and a blister raised on his scalp. A sneezing powder of hellebore root was given to purge his brain, and a powder of cowslip administered to strengthen it, for it was taken as known in those days that the nasal secretion came from the brain.

The king died.

“Evidence-Based Medicine”

- Story of death of George Washington
- Bloodletters thought that they had evidence that their therapy was effective
- But they didn’t *really*
- A whole series of post hoc ergo propter hoc fallacies
- ‘Physics’ and procedures placebos (at best!)

“Evidence-Based Medicine”

- Medieval French surgeon Henri de Mondeville:
- ‘Keep up your patient’s spirits by music of viols and ten-stringed psaltery, or by forged letters describing the death of his enemies, or by telling him that he has been elected to a bishopric, if a churchman.’
- Thomas Jefferson (in his diary 1807) wrote that one of the most successful physicians he had ever known told him that he
- ‘used more bread pills, drops of coloured water and powders of hickory ash than all other medicines put together’

“Evidence-Based Medicine”

- EBM began as a movement at McMaster in late 80s/early 90s
- Its big concern: perhaps not just a historical phenomenon?
- By no means an ‘idle’ worry
- Internal mammary artery ligation for angina (1958/9)
- So EBM set out to subject medical interventions to the test of *real* evidence of efficacy
- Inspired by Archie Cochrane
- Given that the bloodletters (and the ‘old style consultants’ with their clinical expertise) believed they had evidence for the efficacy of their treatments
- The obvious question for EBM: what is “real” evidence?

“Evidence-Based Medicine”

- Sackett: “EBM is a new paradigm that de-emphasises intuition, unsystematic clinical expertise, and pathophysiologic rationale as sufficient grounds for clinical decision making and stresses the examination of evidence from clinical research …”
- Where ‘evidence from clinical research’ meant evidence from RCTs

“Evidence-Based Medicine”

- Initial impression: EBM says *only* evidence from RCTs really counts as valid scientific evidence

“Evidence-Based Medicine”

- EBM → RCT → ‘Gold standard’ (double blind)

“Evidence-Based Medicine”

- *BMJ*, 2001:
- ‘Britain has given the world Shakespeare, Newtonian physics, the theory of evolution, parliamentary democracy and the randomized controlled trial.’

“Evidence-Based Medicine”

- The idea that only RCTs provide really telling evidence is not sustainable
- Misrepresentation of EBM?
- Maybe – but
- ‘[if] the study was not randomized we’d suggest that you stop reading it and go on to the next article in your search’ (Sackett et al *Evidence Based Medicine,* 3rd edition, p.108)

“Evidence-Based Medicine”

- *Some* pro-EBM cases (grommets for glue ear; suppression of ventricular ectopic beats,...)
- But more measured voices immediately pointed out:
- Lots of contrary cases:

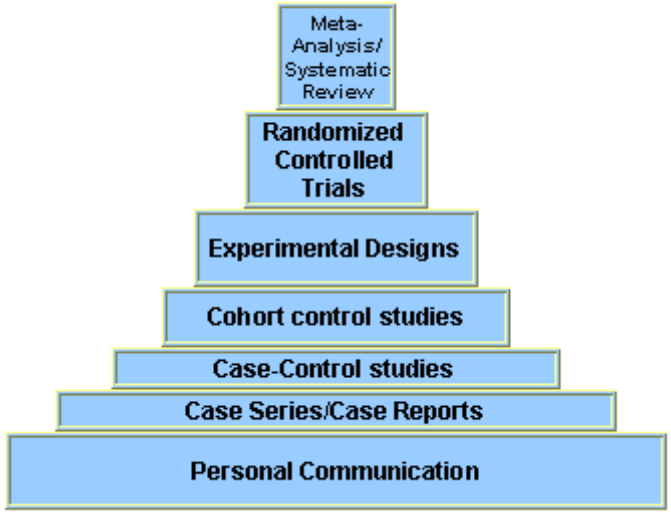
“Evidence-Based Medicine”

- thyroxine for myxoedema
- insulin for diabetic ketoacidosis
- vitamin B12 for pernicious anaemia, etc,etc
- appendicectomy for acute appendicitis …
- along with a large range of accepted surgical procedures: aortic aneurism repair, stenting after cardiac arrest,..
- Not to mention more humble therapeutic procedures like the Heimlich Manouevre…

“Evidence-Based Medicine”

- Retreat/clarification
- Clinical expertise/ ‘patho-physiologic rationale’ to be incorporated not ignored..
- (How?)
- Other kinds of trial can provide evidence (though largely only if no RCTs are available on the therapy concerned)
- RCT retains a very special role
- Evidence Hierarchy

“Evidence-Based Medicine”



“Evidence-Based Medicine”

- But only one of many
- 2002 study found 40 different hierarchies
- 2006 study added 20 more
- All agree that the RCT remains king (trump, no overall evaluation)
- But also differences
- a. some put meta-analyses top, others omit them
- b. cohort/case control

“Evidence-Based Medicine”

- Also
- 1. Explicit concession that RCTs are not needed for ‘dramatic’ effects (Glaziou et al)
- 2. Fleeting recognition that it seems odd to hold that the most clearly efficacious treatments do not have ‘best’ evidence
- 3. And odd that various unfortunate a priori judgments seem to be endorsed

“Evidence-Based Medicine”

- Also also
- Swing in the frequentist/Bayesian balance
- Finally
- One very influential voice:

Sir Michael Rawlins

“Evidence-Based Medicine”

- Rawlins certainly pro evidence in general sense
- But scathing about ‘anorak EBM’

“Evidence-Based Medicine”

- 1. Evidence hierarchies internally unjustified since they overrate RCTs:
- “The notion that evidence can be reliably placed in hierarchies is illusory. Hierarchies place RCTs on an undeserved pedestal for .. although the technique has advantages it also has significant disadvantages. Observational studies too have defects but they also have merit.”

“Evidence-Based Medicine”

- 2. Whole idea a mistake:
- “Hierarchies attempt to replace judgement with an oversimplistic, pseudo-quantitative, assessment of the quality of the available evidence. [In fact d]ecision makers have to incorporate judgements, as part of their appraisal of the evidence, in reaching their conclusions.”

“Evidence-Based Medicine”

- Groundhog day?
- Whole rationale for EBM distrust of judgment…
- Certainly need to try to codify the judgment
- And in particular give some account of how to arrive at overall judgments of the weight of evidence
- Rigid hierarchies surely not the way
- Especially if accompanied by the clear-cut but naïve ‘trumping’ rule.

“Evidence-Based Medicine”

- Overall:
- Of course medicine must be based on the evidence
- But the issue of what the evidence is saying is altogether more complex than EBMers (and others) made it seem
- Certainly easy ‘black and white’ rules (cp significance testing) are no good
- But too easy to say ‘we have to bring in judgment’
- And leave it at that
- Need to characterise what good judgment is
- Maybe the position we are going on to can help?

Bayesianism

- Bayesianism is a completely general account of scientific rationality
- At least in orthodox approaches, it is very simple, based on just two requirements for an 'agent' to be rational:
- 1. At any given time, her degrees of belief over the claims that she considers satisfy the probability calculus
- 2. If between time t_1 ("old") and time t_2 ("new"), all that the agent learns of epistemic relevance is that previously uncertain e in fact holds, then her new degrees of belief are related to the old via the **principle of conditionalisation**:
- For any hypothesis h, $p_{new}(h) = p_{old}(h/e)$
- Bayesianism often described as unacceptably 'subjectivist' but in fact 1. is highly normative: real agents can only approximate an ideal Bayesian-rational agent

Bayesianism

- As for 2, notice that this ***updating rule*** is not Bayes' theorem
- Though of course that is how we will standardly calculate $p_{old}(h/e)$ [as remember $p_{old}(e/h).p_{old}(h)/p_{old}(e)$]

Bayesianism

- Notice that this ***updating rule*** is not Bayes' theorem
- Though of course that is how we will standardly calculate $p_{old}(h/e)$ [as remember p(e/h).p(h)/p(e)]
- This yields a simple and intuitively appealing account of the confirmation of theory by evidence:
- e confirms h if (and to the extent that) e raises h's probability
- That is,
- e confirms h if p(h/e) > p(h) (i.e. $p_{new}(h) > p_{old}(h)$)
- And (on at least one natural measure) the extent to which e confirms h is given by p(h/e) - p(h)
- e is neutral wrt to h if p(h/e) = p(h)
- And e disconfirms if p(h/e) < p(h)

Bayesianism

- Particularly straightforward for deterministic hypotheses
- Suppose we think of h as the whole accepted theoretical framework based on 'central' theories like classical physics or GTR
- ***A. Falsification***
- Suppose the actually observed e is inconsistent with h (h entails e' which e entails is false) then $p_{new}(h) = p_{old}(h/e) = [p_{old}(e/h).\ p_{old}(h)]/p_{old}(e)]$
- But $p_{old}(e/h) = 0$ [$p_{old}(\neg e/h) = 1$ and $p_{old}(e/h) = 1 - p_{old}(\neg e/h)$]
- So $p_{new}(h) = 0$
- The rational Bayesian agent ascribes zero credibility now to h, since she knows it to be false.

Bayesianism

- ***B. Confirmation***
- Suppose h has predicted e (i.e. e is a logical consequence of h)
- Initially e is of course not known to hold ($p_{old}(e) \neq 1$) but then it turns out that e is in fact correct (so that $p_{new}(e) = 1$))
- Then probability $p_{new}(h) = p_{old}(h/e) = [p_{old}(e/h).\ p_{old}(h)]/p_{old}(e)$
- Since e is entailed by h, $p_{old}(e/h) = 1$ and so this reduces to
- $p_{new}(h) = p_{old}(h)/p_{old}(e)$
- Since $p_{old}(e) < 1$ there is bound to be some confirmation of h
- How much depends on
- 1. How likely we think the hypothesis was ahead of its success with e
- 2. How 'severe a test' e represented – how unlikely it was to occur according to other (plausible) theories.

Bayesianism

- 2. as we have often seen is a welcome and important result
- 1. looks initially a bit iffy from an ‘objectivist’ point of view
- But in fact ...
- (notice that these do not feel like ‘merely subjective’ judgements)
- (Current ‘prior’ will standardly be an even older posterior …)

The Bayesian approach

- How about Bayesianism applied to statistical hypotheses?
- Bayesian statistics is the only developed alternative to the classical frequentist approach.
- It assumes that we start out (always) with some 'prior probability' distribution over the various hypotheses that we are considering
- Where these are best interpreted as plausibility ratings in the light of background knowledge
- And we simply modify those probabilities – from the 'prior' distribution to the 'posterior' in the light of
- (i) the 'likelihoods' ($p(e/h_i)$) and
- (ii) prior of the evidence ($p(e)$)

The Bayesian approach

- So let's take an artificial case just for simplicity's sake:
- Suppose we have good evidence that some machine turns out coins that are either fair (p(head) = 0.5) or heavily biased in favour of heads (p(head) =0.9) with the same frequency
- So now we have before us one coin produced by that machine
- In that case we are interested in two hypotheses about the coin h_1 and h_2
- Since the machine turns out both with equal frequency, it seems reasonable to set the 'priors' as $p(h_1) = p(h_2) = 0.5$
- Now we toss the coin 4 times and observe e = 3 heads out of 4

The Bayesian approach

- The likelihoods are $p(e/h_1)$ and $p(e/h_2)$
- $p(e/h_1) = 0.25$
- $p(e/h_2) = 0.36$
- What is p(e)? well given that we take ourselves to know that one of the two hypotheses is true
- $p(e) = p(h_1)p(e/h_1) + p(h_2)p(e/h_2)$
- $p(e) = 0.5 \times 0.25 + 0.5 \times 0.36 = 0.305$
- So Bayes theorem tells us how to 'update'
- $p_{new}(h_1) = p_{old}(h_1).\ p(e/h_1) / p_{old}(e)$
- $p_{new}(h_2) = p_{old}(h_1).\ p(e/h_1) / p_{old}(e)$

The Bayesian approach

- $p_{new}(h_1) = p_{old}(h_1).\ p_{old}(e/h_1) / p_{old}(e) = 0.41$
- $p_{new}(h_2) = p_{old}(h_1).\ p_{old}(e/h_1) / p_{old}(e) = 0.59$
- So…
- Of course 4 tosses is a small sample
- If we had tossed it 10 times and got 8 heads then if we regarded h_1 as the null it would be rejected on Fisherian principles (p= 0.04)
- But on Bayesian principles it would ‘just’ have become less probable (and the bias hypothesis more probable)
- roughly 0.15 and 0.85 respectively

The Bayesian approach

- So we can bring in consideration of what we know already
- If there are good grounds from earlier evidence that a particular hypothesis is true then the fact that one particular piece of evidence is very improbable given h would not lead us to reject
- It would of course lead us to reduce our degree of belief in h (= $p_{new}(h)$) but, depending on the numbers involved, perhaps not by very much.
- Dogmatic?
- Not really: reflecting the fact that this particular piece of evidence does not exist in an ‘evidential bubble’.

The Bayesian approach

- For more representative cases, where some parameter of interest – e.g the average effect of some treatment – can take on any of a range of values – see Howson & Urbach reading
- But what happens in general terms is that your prior distribution of uncertainties over the possible values is turned, by the evidence you collect, into a posterior distribution – standardly with a shifted mean and a smaller variance.
- So suppose you think the likeliest value ahead of your clinical trial is zero (a sort of Bayesian null)
- And you observe a positive effect
- Then you won't 'reject' anything or 'accept' an alternative
- It is just that your new degrees of belief will make a positive effect more likely..

Advantages of the Bayesian approach

- No doubtful logic
- Stick with probabilities rather than introducing any rejection rule
- (Lottery paradox therefore avoided)
- Without loss – we don't need to know stuff in order to make reasonable decisions
- (Especially if we are only pretending to know)
- we can feed our Bayesian probabilities into a decision theory analysis
- Seems more honest
- Does not fall to the base rate fallacy ,as Fisher seems to, where there are well-defined priors (or pretty well defined intervals for priors)

Problems with the Bayesian approach

- Where do the priors come from *in general*?
- Aren't we opening ourselves up to bias?
- (This was Fisher's main concern)
- There are some 'convergence theorems' that show that subject to certain conditions, disagreements about priors get "washed out" in the limit
- But is agreement in the limit enough?

Problem with the Bayesian approach

- So, e.g., suppose someone said that they were a perfect Bayesian agent and always updated properly in accordance with the theorem
- However when they started their evidential analysis of the Creationism vs Darwinism debate they decided that their priors were
- p(Darwin) = .000000000000001; p(Creationism) = 0.999999999999
- Even though they accept, say, that e.g. p(fossils/Darwin) > p(fossils/ creationism) they still think (rationally!) that creationism is overwhelmingly likely to be true

Bayes or Bust??

- Overall: has many advantages
- Aside from anything else it clearly allows space for the sort of judgments argued to be necessary by Bradford Hill and by Rawlins
- That is, via assignment of priors
- But corresponding to the fact (?) that we don't want to give free rein to 'judgment'
- We need constraints on priors in order for it to provide a fully acceptable framework.
- Many have tried ..
- Time to try again ..?