

Bad *JAMA*?



Stephen Senn



Declaration of Interest

- I consult regularly for the pharmaceutical industry
- I used to work for the pharmaceutical industry
- My career is furthered by publishing
- I take malicious pleasure in exposing woolly arguments
- In short, I am a very bad person
- See http://www.senns.demon.co.uk/Declaration_Interest.htm to find out exactly how bad

Bad PHARMA



- Written by medical journalist Ben Goldacre
 - Published autumn 2012
- Hard hitting exposé of the pharmaceutical industry
- Also very critical of drug regulators
- Looks at the problem of missing studies in detail

What I Shall Assume

- Journal editors base their decision as to what to accept (at least in part) on quality
 - Other things being equal, higher quality papers are more likely to be accepted
 - Editors may or may not take study *outcome* into account
- Authors base their decision as to what to submit (at least in part) on quality
 - Other things being equal, higher quality papers are more likely to be submitted
 - Authors may or may not take study *outcome* into account

Goldacre's Thesis

‘But to be kind, for the sake of completeness, and because industry and researchers are so keen to pass the blame on to academic journals, we can see if the claim is true....Here again the journals seem blameless: 74 manuscripts submitted to the *Journal of the American Association (JAMA)* were followed up, and there was no difference in acceptance for significant and non-significant findings.’ Bad PHARMA

- Negative studies are less likely to be published
- This is not because editors are less likely to accept them
 - The evidence shows otherwise
- Authors are less likely to submit them
- The pharmaceutical industry is particularly bad at submitting negative studies to journals

That *JAMA* study

- Olson et al, 2002
 - Nine authors
 - Prospective study of manuscripts submitted to *JAMA*
 - Covered February 1996 to August 1999 inclusive
 - Concluded there is no bias against negative studies
- What Goldacre doesn't tell you
 - Four of the authors were *JAMA* editors
 - The finding of lack of bias depends on the definition of positive study

JAMA 2x2

	Accepted	Rejected	Total
Positive	78	305	383
Negative	55	307	362
Overall	133	612	745

Statistic based on the observed 2 by 2 table :

Binomial proportion for column <Negative > : $\pi_1 = 0.1519$

Binomial proportion for column <Positive > : $\pi_2 = 0.2037$

$$\text{Odds Ratio} = \frac{(\pi_2)/(1-\pi_2)}{(\pi_1)/(1-\pi_1)} = 1.427$$

My analysis using StatXact. The authors quote relative risk.

Results:

Method	P-value(2-sided)	95.00% Confidence Interval
Asymp (Mantel-Haenszel)	0.0662	(0.9765 , 2.087)
Exact	0.08020	(0.9605 , 2.130)

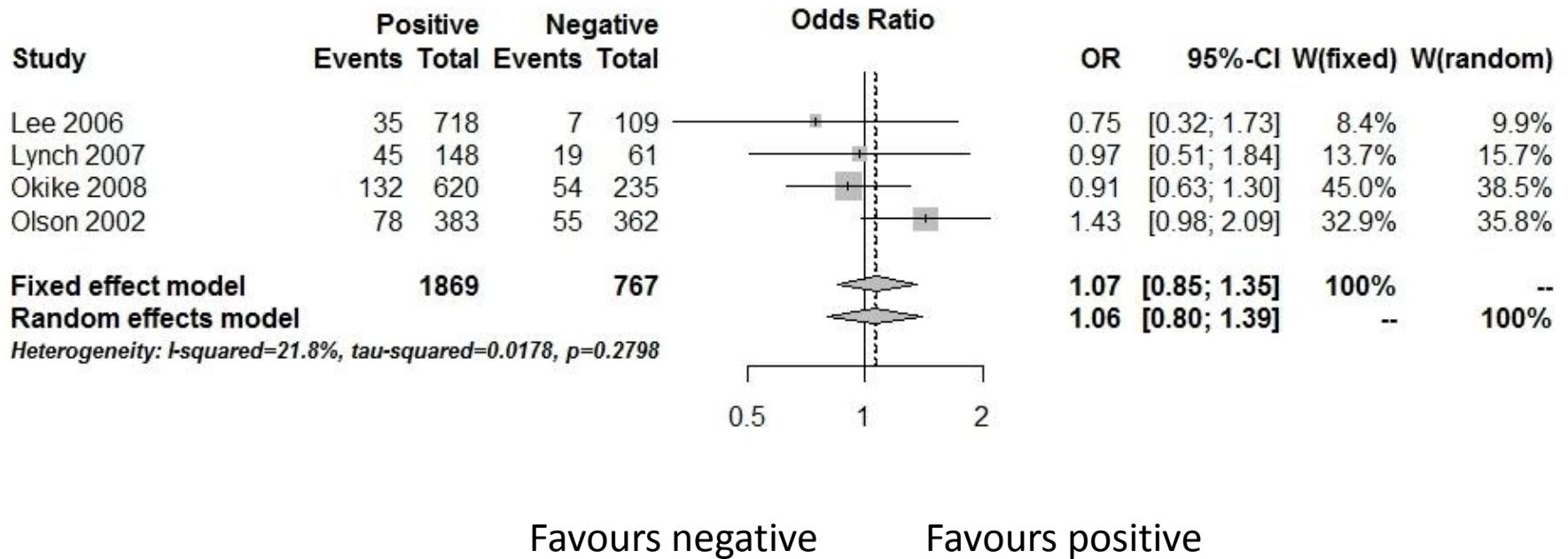
JAMA Logistic

Variables	Odds ratio(confidence limits)	P-value
Positive results vs negative	1.30 (0.87 - 1.96)	.21
Multi-centre vs single	1.60 (1.02 - 2.52)	.04
United States enrollment versus not	2.06 (1.20 - 3.52)	.008
Any funding versus no reported	1.42 (0.69 - 2.90)	.34
No. of participants >100 vs <100	1.38 (0.83 - 2.30)	.22
Sample size calculation vs none	1.90 (1.23 - 2.95)	.004
Randomisation described vs not	0.99 (0.65 - 1.52)	.98
Blinding vs not	1.51(0.98 - 2.34)	.06
Withdrawal reported vs not	1.35 (0.66 - 2.78)	.41
Analysis by treatment assignment vs not	1.22 (0.80 - 1.84)	.36

Taken from Table 2 of Olson et al,2002

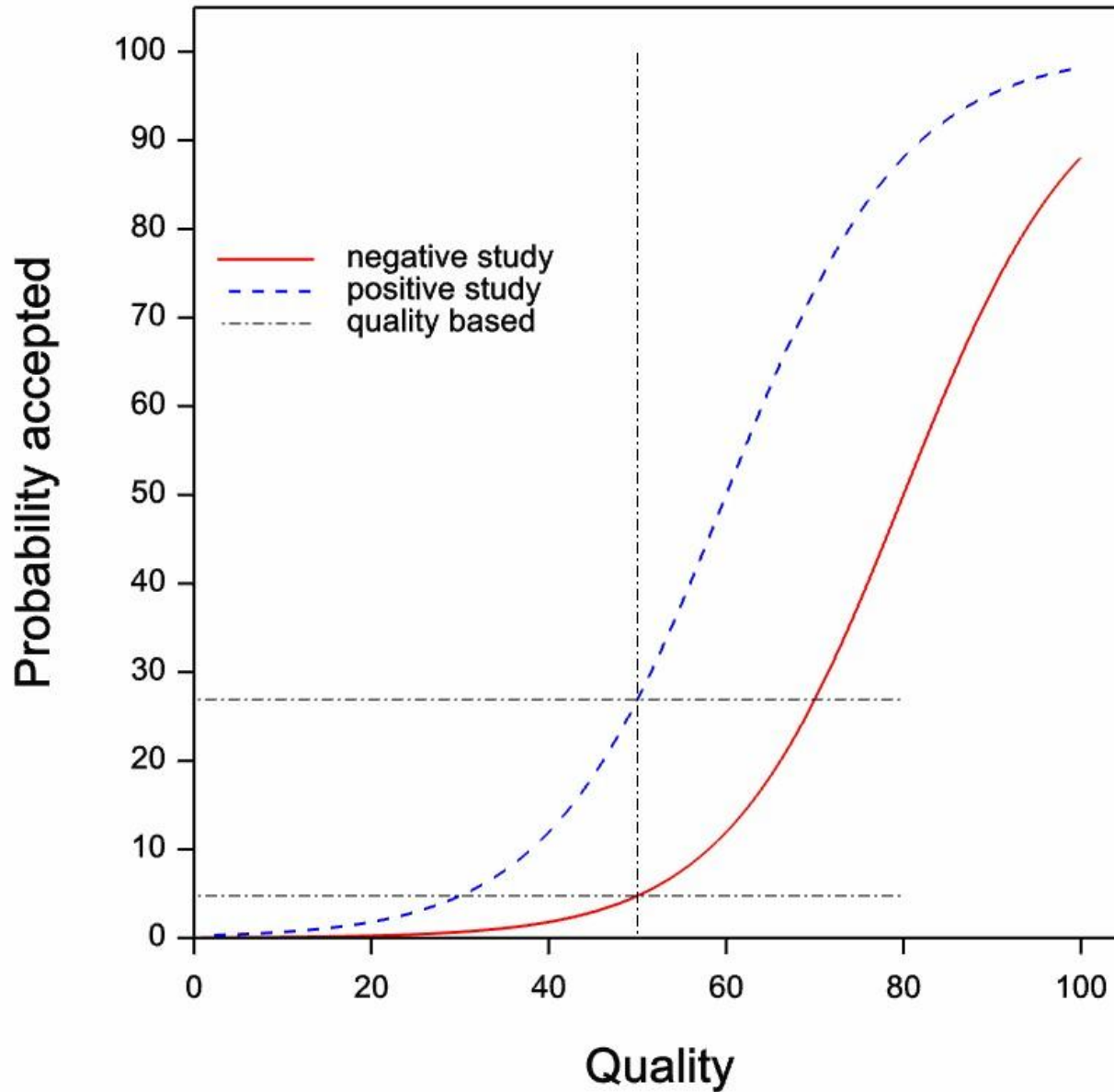
Summary of Studies

(Based on Song et al, 2009)

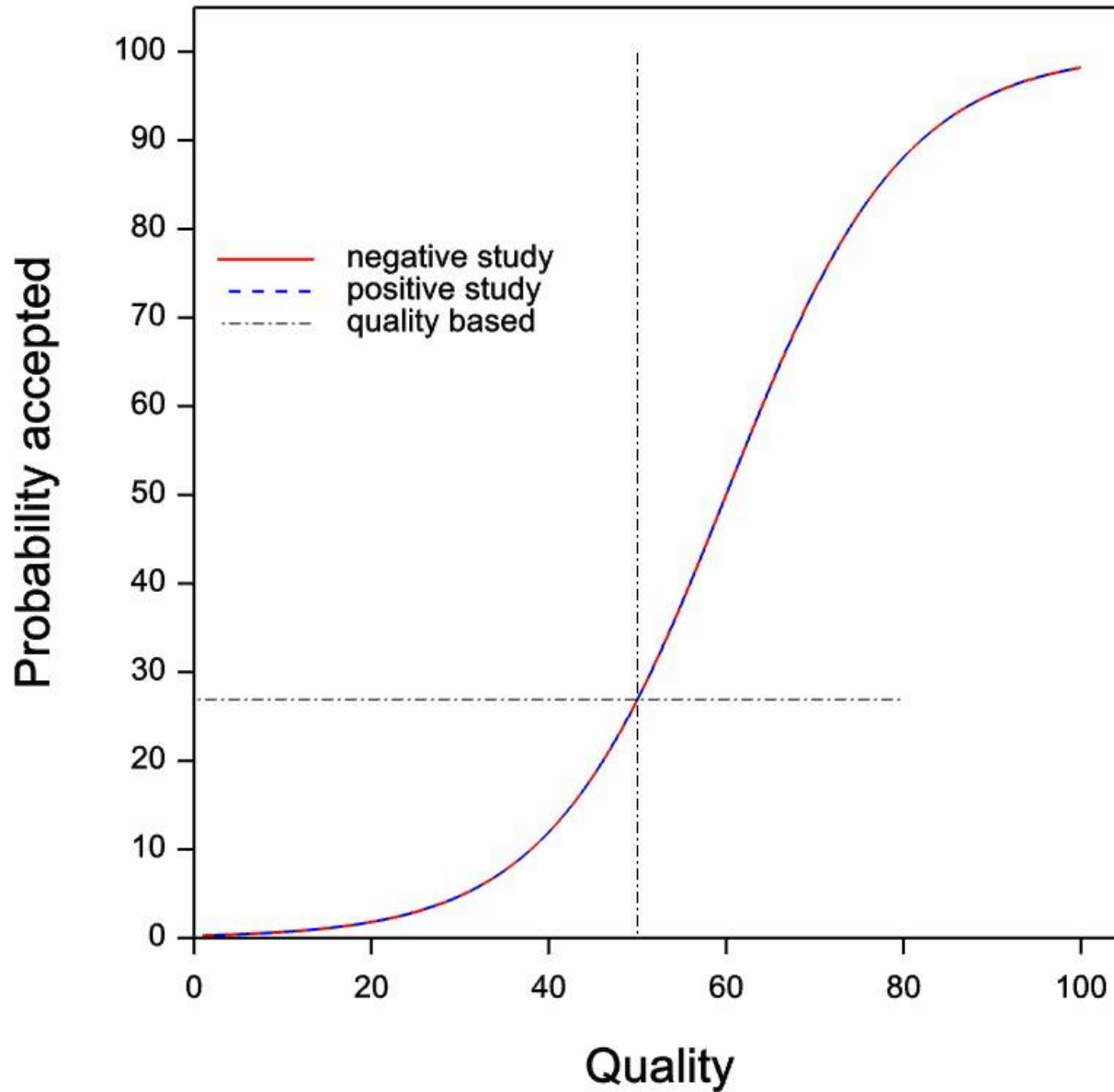


Analysis produced using Guido Schwarzer's *meta* package in R

Probability of paper being accepted v quality by result



Probability of paper being accepted v quality by result



Data Filter?

- However there may be the problem of data filtering
- Can the way that the data arrive be ignored?
- Is it safe to condition on what is seen and argue from that point onwards as if like were being compared with like?
- Or is there an inherent problem?
- Consider the following famous statistical story...

Abraham Wald (1902-1950)

- Rumanian/Hungarian/American, mathematical statistician
- Inventor of decision theory
 - brilliant and seminal paper of 1939
- Also innovator for sequential analysis
- Died in a plane crash in India
- Ironically, was employed by US military to advise on plane safety in World War II



Wald's Problem

- Returning planes were examined to see where they had been hit
- Engines were rarely hit
- Fuel tanks very often
- Extra armour could be placed but not everywhere
- Where should it be placed?

Wald

The Military and The Aircraft

- The US Military decided to reinforce the fuel tanks
- That was where the most shots were
- They argued that therefore the fuel tanks needed protection

Wald and the Aircraft I

- Wald argued that the pattern of shots received ought to be random
- The fact that it was not, indicated that this sample was not random
- If the shots hit the fuel tank, the plane returned safely
- If it hit the engine, it did not

Wald and the Aircraft II

- Hence, the fact that there were very few planes with shot engines indicated that engines were vulnerable
- Solution: reinforce the engines not the fuel tanks!

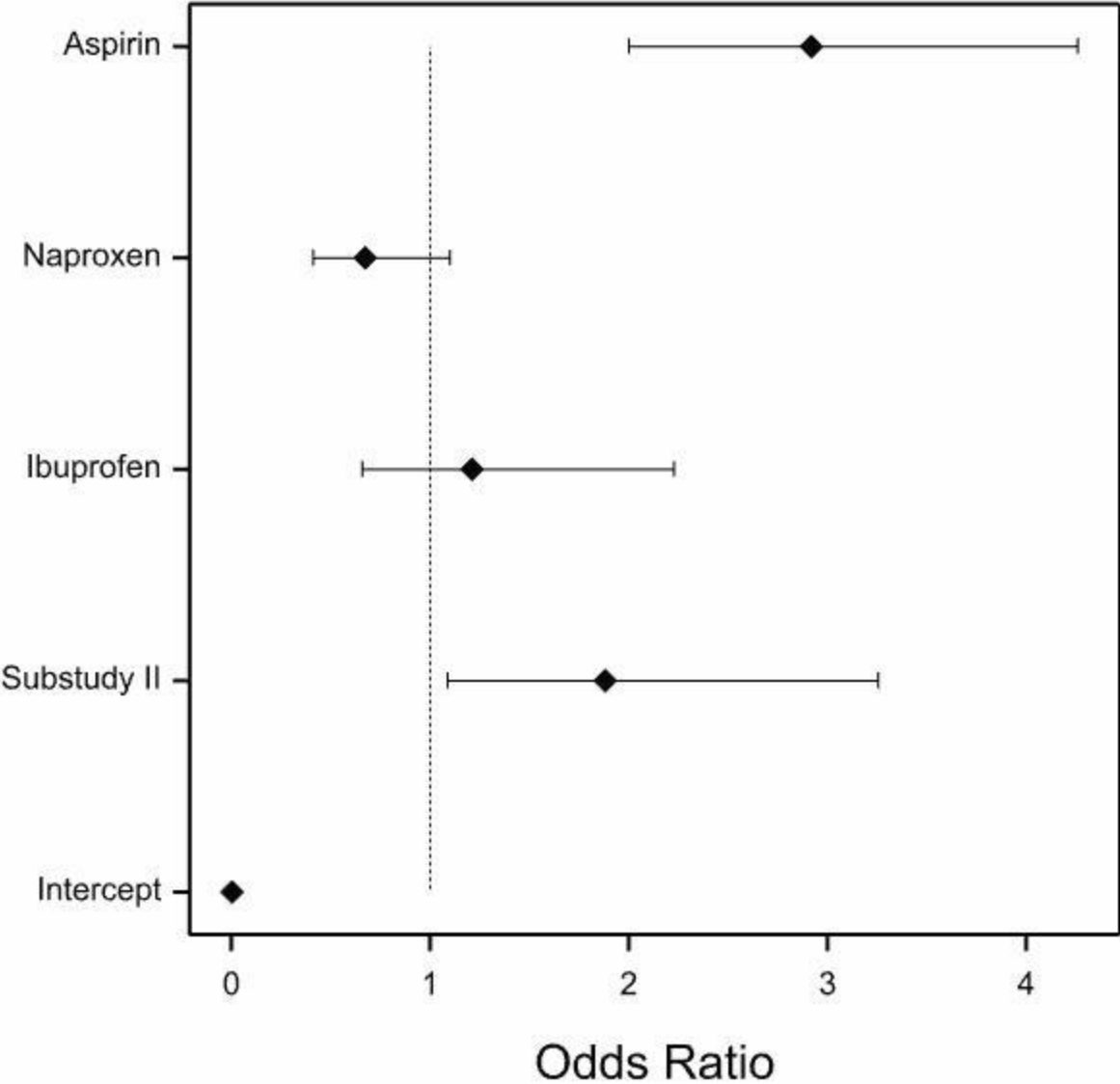
Data Filtering Some Examples

- Oscar winners lived longer than actors who didn't win an Oscar
- A 20 year follow-up study of women in an English village found higher survival amongst smokers than non-smokers
- Transplant receivers on highest doses of cyclosporine had higher probability of graft rejection than on lower doses
- Left-handers observed to die younger on average than right-handers
- Obese infarct survivors have better prognosis than non-obese

TARGET study

- Trial of more than 18,000 patients in osteoarthritis over one year or more
- Two sub-studies
 - Lumiracoxib v ibuprofen
 - Lumiracoxib v naproxen
- Stratified by aspirin use or not
- Analysis of cardiovascular(CV)events at one year

TARGET odds ratios CV event



Moral

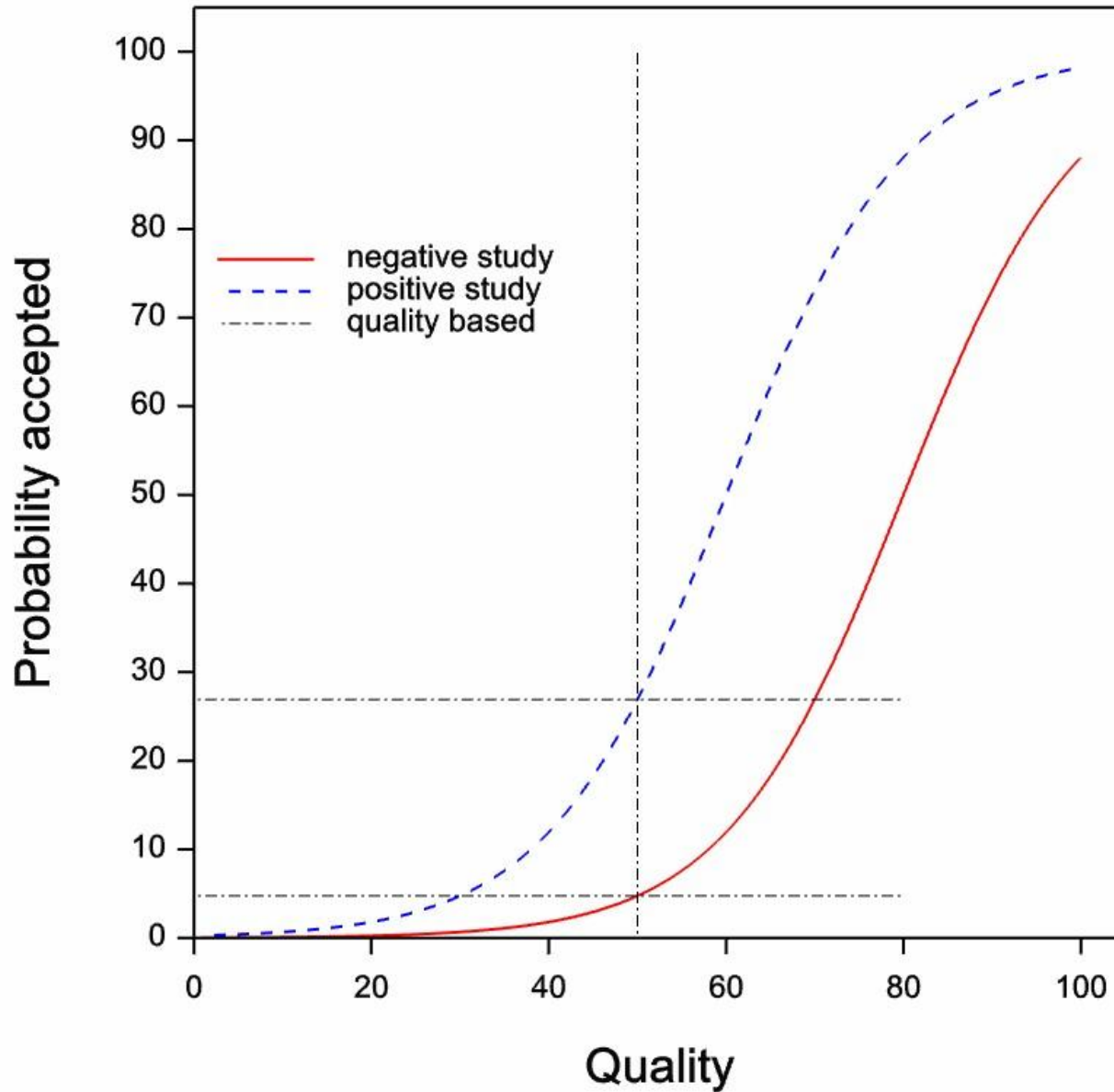
- What you don't see can be important
 - Actually this is a point Goldacre makes over and over but forgets
- The data may have arisen in a way which would bias naïve analyses
- We need to think about the publication bias carefully

Negative Thinking

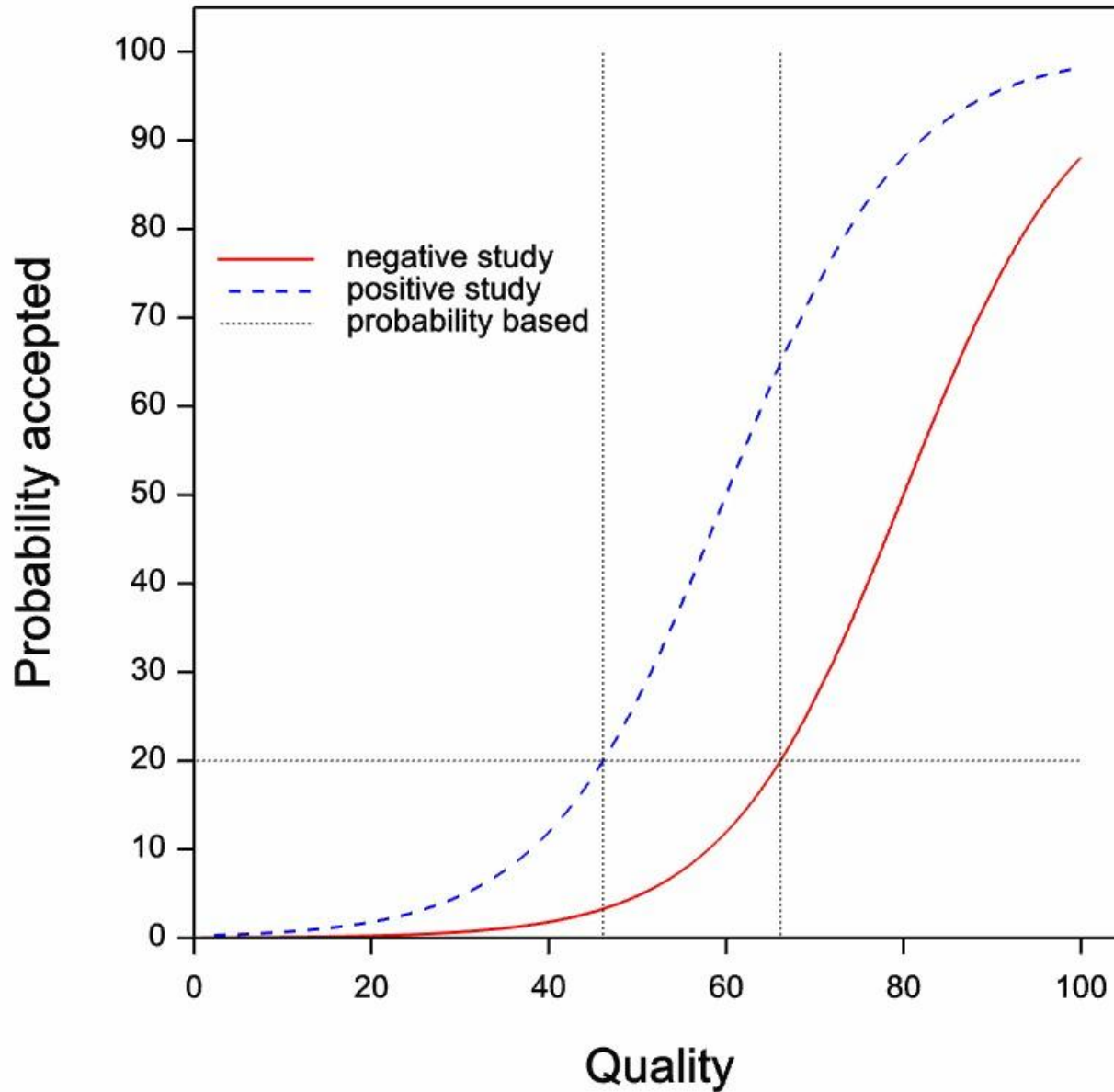
‘In the light of all this, the data on what researchers say about their own behaviour is very revealing. In various surveys they have said that they thought there was no point in submitting negative results, because they would just be rejected by journals.’ *Bad Pharma*, p36

- If authors act rationally they will make a decision based on the estimated probability of acceptance whether to submit or not
- In that case the relevant threshold for submission *is not a quality threshold but a probability threshold*
- Return to our graphical model of paper acceptance...

Probability of paper being accepted v quality by result



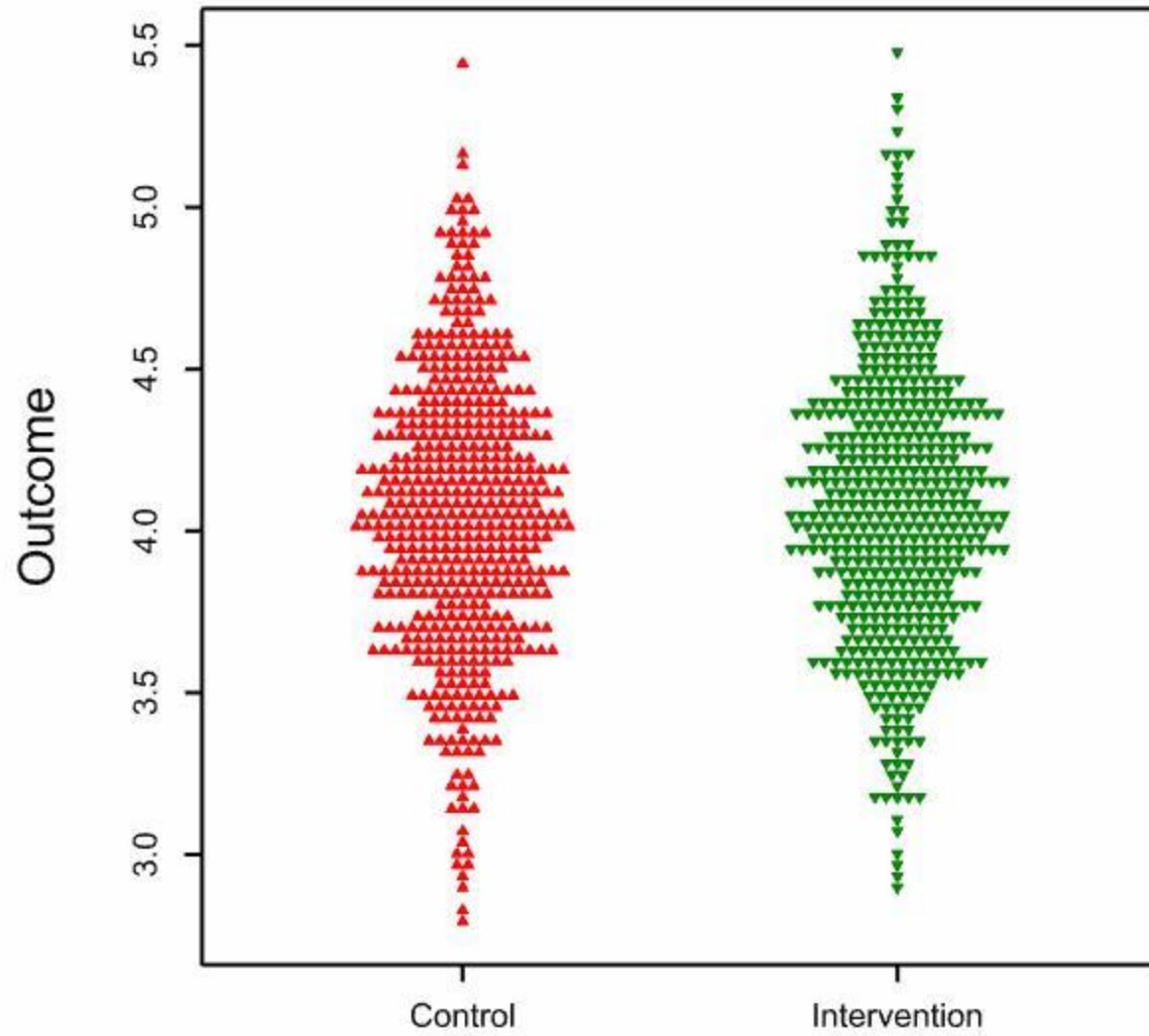
Probability of paper being accepted v quality by result



A Clinical Trial Example

- Cholesterol lowering agent being prepared to placebo
- LDL measured at outcome
- Data are as shown on the next slide

LDL at outcome



No Effect?

- Distributions look very similar
- No apparent difference between groups
- t-test of difference at outcome has following results

Difference of means: -0.0278

Standard error of difference: 0.0273

95% confidence interval for difference in means: (-0.08128, 0.02567)

Test statistic $t = -1.02$ on 998 d.f.

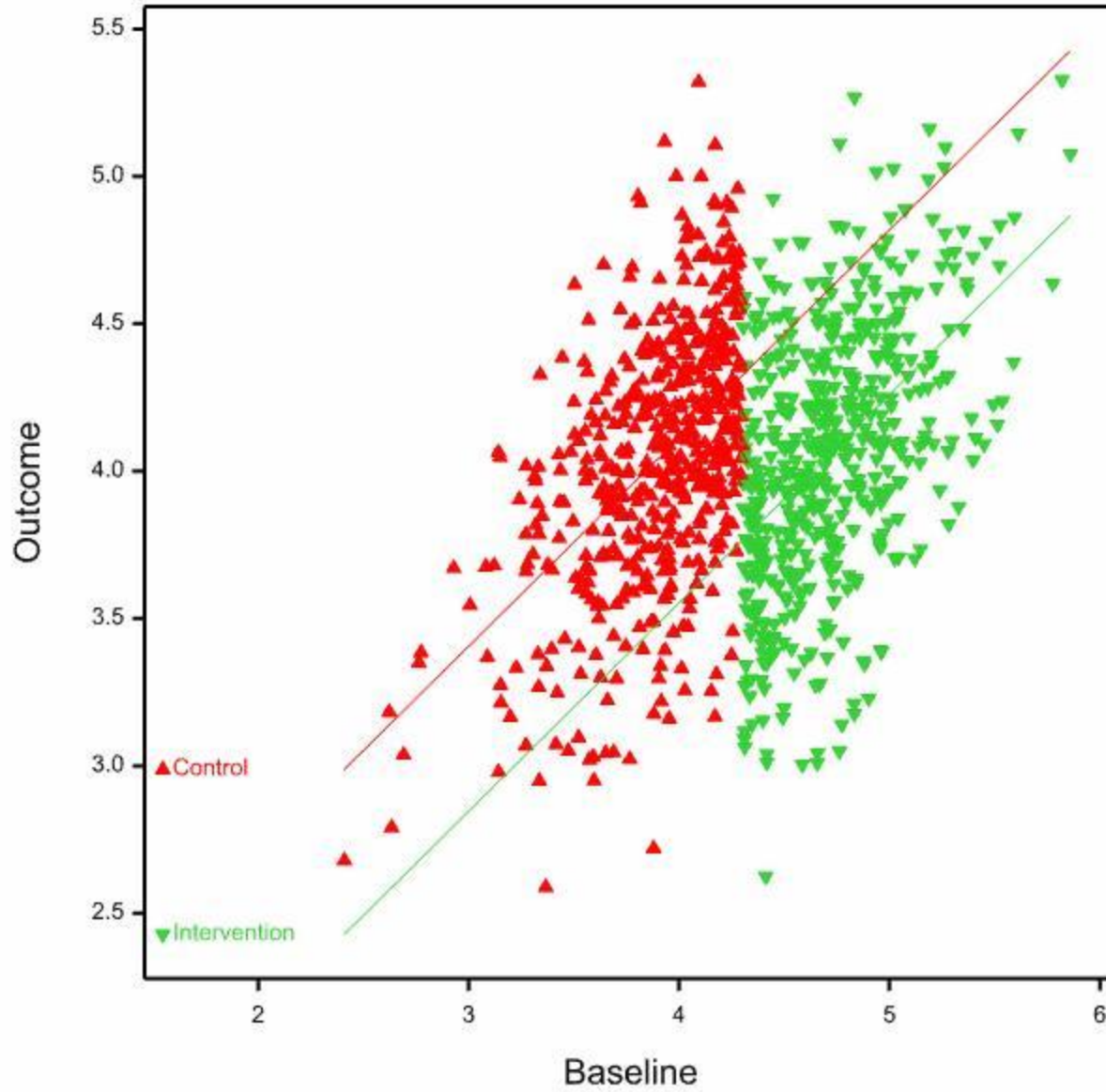
Probability = 0.308

- Conclusion of no (or at least no proven) effect seems reasonable

What I did not tell you

- This is *not* a randomised trial
- It is a regression discontinuity trial
- Patients were given the intervention *if* they had high cholesterol at baseline
- Were not given an intervention if they had low cholesterol
- Plotting outcome against baseline this is what you see...

LDL at baseline and outcome



Analysis of Covariance

Estimates of parameters

Parameter	estimate	s.e.	t(997)	t pr.
Constant	1.722	0.162	10.61	<.001
Baseline	0.5966	0.0415	14.39	<.001
Treat Intervention	-0.4486	0.0414	-10.84	<.001

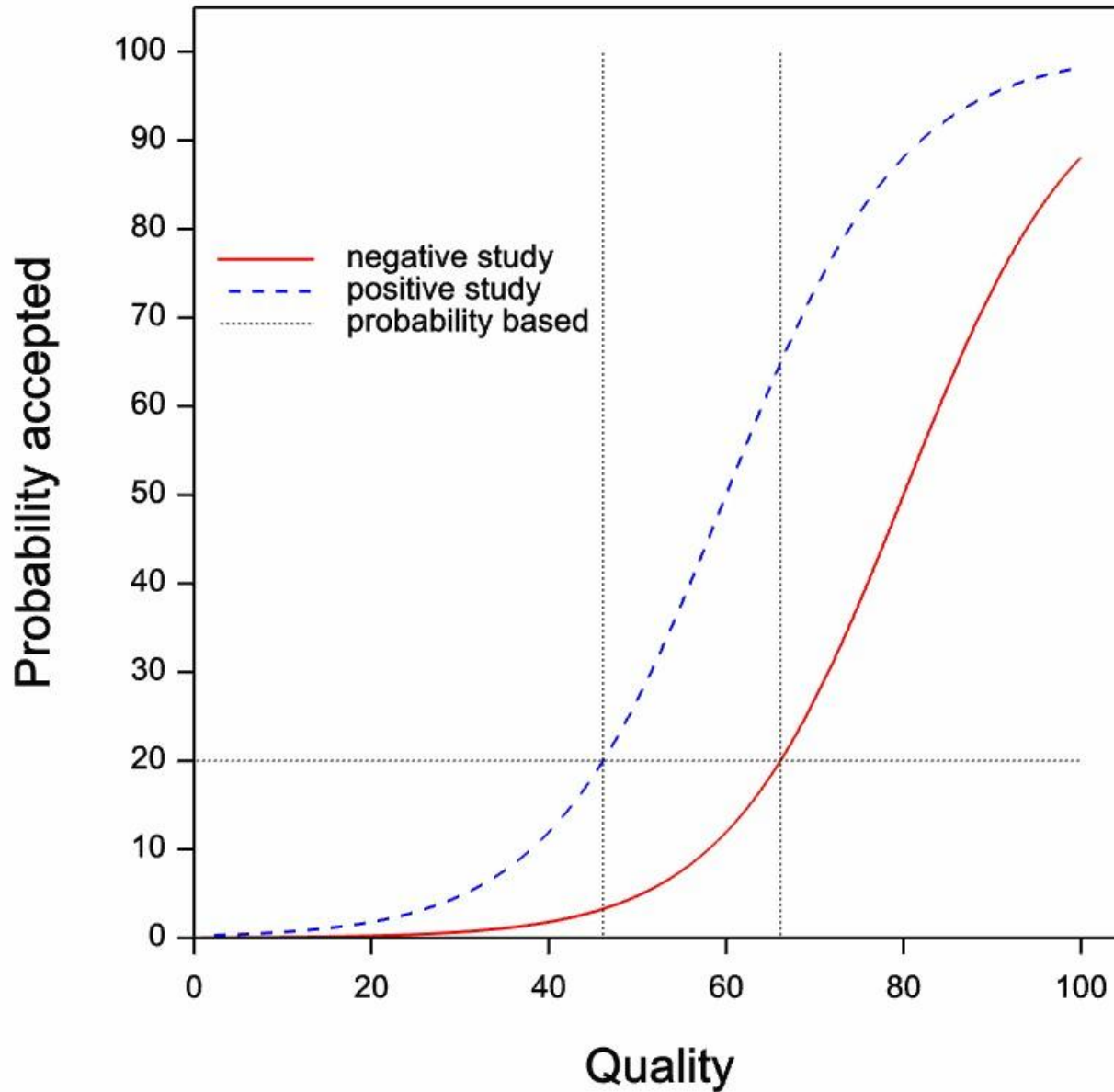
Parameters for factors are differences compared with the reference level:

Factor	Reference level
Treat	Control

The Evidence of Things not Seen

- It is dangerous to assume that things arise at random unless we have taken steps to make sure they do
- Here a non-random allocation process requires us to take the baseline into account
- When we do we get a quite different picture
- The cholesterol lowering treatment works after all

Probability of paper being accepted v quality by result



An Analogy

- Is there a sex bias in promotion to senior lecturer?
- Study a large number of applications
- Discover that women were just as likely to be promoted as men
- What we had overlooked was the following..
- The women were on average much more highly qualified

Possible Explanation?

- There is a sex bias against women
- The women know it
- They are not bothering to apply until they are much better qualified than their male colleagues
 - More publications
 - More grants
 - More teaching experience

Explanation of the meta-analysis

- Authors know there is a bias against negative studies
- Other things being equal they are much less likely to submit to a prestigious journal
- In consequence the *probability* of negative and positive studies being accepted is the same
- They differ in *quality*

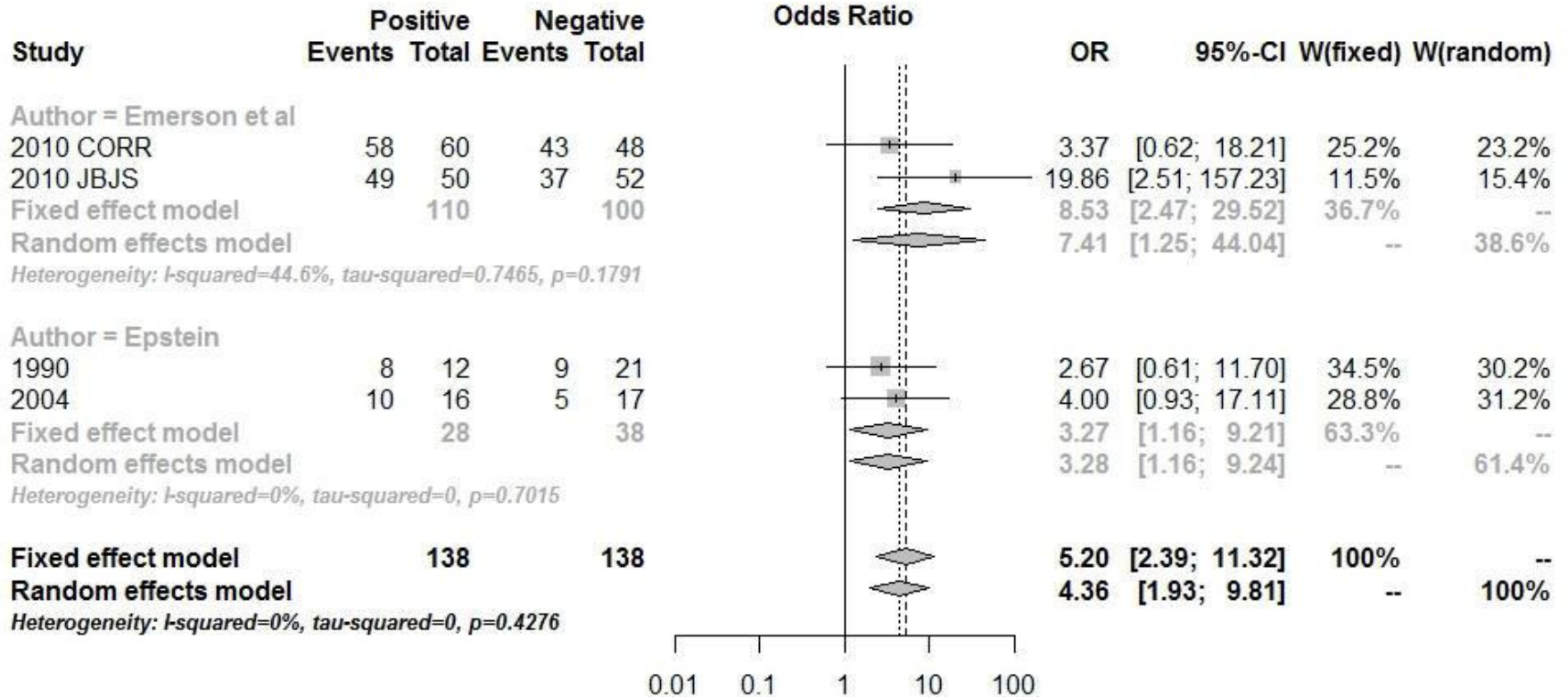
To Sum Up

- Comparing negative and positive studies the assumption has been made that (apart from the defining condition) like is being compared with like
- The data have been analysed as if they came from a randomised trial
- They don't
- There are data missing and these are missing 'not at random' to use the Rubin classification

How could we deal with the bias?

- One way would be to have randomised studies
- Provide reviewers with different randomised versions of the same manuscript
 - Conclusions changed to be positive or negative
- These are also mentioned briefly in *Bad Pharma*

Meta-analysis of Interventional Studies



Warning: I have not done an extensive search for these

Goldacre's Summing Up

'Overall, though, even if there are clearly rough edges in some domains, these results don't suggest that that journals are the main cause of the problem' *Bad Pharma*, p36

What about quality?

‘The same thing has been tried with papers submitted to the *BMJ*, *The Lancet*, *Annals of Internal Medicine* and the *Journal of Bone and Joint Surgery*. Again and again no effect was found.’ *Bad Pharma* p34

- Is Goldacre right?
- Let us have a look at the *Journal of Bone and Joint Surgery*

Lynch et al, 2007

- All manuscripts about hip or knee arthroplasty submitted to *Journal of Bone and Joint Surgery*, American volume, over 17 months were evaluated
- Study design, quality and outcome were noted
- 209 Manuscripts reviewed

Lynch et al 2

‘Commercially funded and United States-based research is more likely to be published; good-quality studies with negative outcomes are not’

- Title is very revealing!
- No difference found in probability of acceptance positive and negative studies
- However significant difference in quality, negative versus positive $p=0.003$

Aristotelian theories of causality?

- Why did the chicken cross the road?
 - Because in response to various bio-physical stimuli its legs propelled it forward orthogonally to the road direction? (Efficient cause.)
 - To get to the other side? (Final cause)
- Why did the author submit to this journal?
 - Because the quality was right for this journal?
 - Because the probability of acceptance was right for this journal?

Who's Guilty?

Authors or Reviewers?

- The Goldacre view is that journals are blameless and authors are guilty
- But there is something very strange about this
- By and large these are the same people
 - albeit in different roles
- Can we believe that researchers as authors are biased against submitting negative papers but researchers as reviewers are not?

In Summary

- Authors tells us they are reluctant to submit negative papers because they fear rejection
- Authors are also reviewers
- It is logical for them to submit by quality of acceptance
- The empirical evidence show that submitted negative papers are of higher quality
- The experimental evidence shows that reviewers are less likely to accept negative papers
- Equal probability of acceptance of positive and negative studies does not show lack of bias

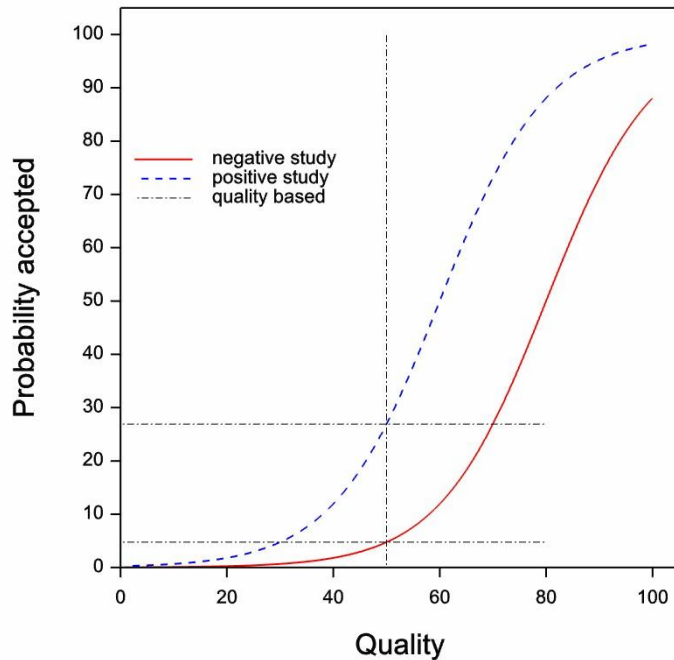
Minding your Ps and Qs

You Choose

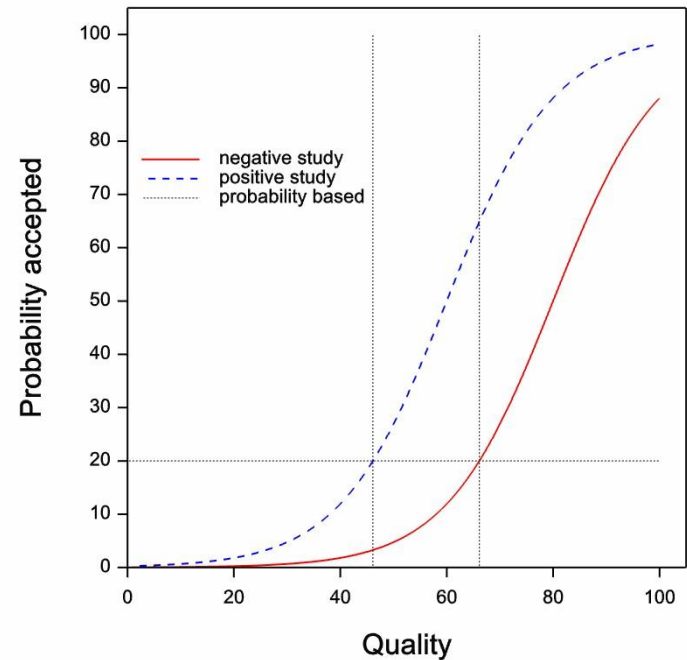
Quality based

Probability based

Probability of paper being accepted v quality by result



Probability of paper being accepted v quality by result



Finally

Statistics is a subject that many medics find easy but most statisticians find difficult.

Guernsey McPearson