Bad JAMA?

Stephen Senn, Strassen, Luxembourg

‘The tricks and distortions documented in these pages are beautiful and intricate and fascinating in their detail.’ Bad Pharma, Pxi

Warning
The views expressed here are mine alone and should not be ascribed to any organisation with whom I am associated.

Introduction
‘This has all been perpetrated by ordinary people but many of them may not even know what they’ve done.’ Ben Goldacre, Bad Pharma, Pxi.

Bad Pharma is a book by Ben Goldacre that is going to make the stir that its author and his publishers always hoped it would[1]. It is a damning indictment of the pharmaceutical industry and much of what it claims is true. I also agree with much of what Goldacre suggests should be done to improve matters. However, it is not, despite its author’s claims and hopes, an objective account. It is not a judgement or a summing up; it is not even an investigation; it is a case for the prosecution. Such cases have a legitimate place in any debate but they should not be presented for what they are not.

When reviewing evidence, distortions can occur as a result of sins of commission or sins of omission. You will find both in Bad Pharma. As regards the former I am going to take one early story of Goldacre’s and show you how the author employs a number of distortions, in some cases similar to those he criticises, to make you believe what he wants you to. As regards the latter I will cover one very important story that might give a rather different impression had it been included in Bad Pharma.

Although I think that Ben Goldacre is far from being an ordinary person (see section quote) I am going to concede that it is quite possible that he does not know he has misrepresented the evidence. The style of Bad Pharma is ‘lively’ and I too am going to be lively. I am not going to imitate Goldacre’s style, he is a good writer and this would be difficult to do, but I am going to try my best to be lively. Above all I am going to be unfair. In this I am copying Goldacre.

You should read what follows as an illustration of how easy it is to produce a case for the prosecution, not as a judgement on Goldacre.
What you need to know about me

‘If you like – if it’s in your nature, and if that is how you’d like others to see you – you can point out mistakes with self-righteous fury.’ *Bad Pharma* P372

I am a very bad person. I used to work in the pharmaceutical industry, I consult regularly for them. As a result of having worked in the industry I have shares in the successor company (merged) of one I used to work for. That’s bad enough and if you want to know exactly how bad all that is I maintain a declaration of interest page here http://www.senns.demon.co.uk/Declaration_Interest.htm and you can check. However, there is a far worse aspect of my character. I have a naturally disputative nature. I am a nit-picker par excellence with a long list on my CV of letters to the editor pointing out the mistakes of others. (Many medical journals don’t admit to mistakes easily and the list would be even bigger if that were not the case but we will come to that in due course.) I don’t quite fulfil the description at the end of *Bad Pharma* quoted at the beginning of this section because it is not self-righteous fury that motivates me (that is something Goldacre does very well) but something far worse, something I can only describe as *malicious pleasure*.

It gets worse. I used to teach medical students and I occasionally read the newspapers. These two activities have left me with the firm prejudice that doctors and journalists are two classes of person who cannot be trusted with evidence. (As regards the former class, I am very happy to share with anyone who writes to me, my technique for dealing with collaborating cheating cliques of medical students, from getting them to confess, to handing out the punishment.) In my cynical view someone who is doubly cursed, a *medical journalist*, hasn’t a hope of being objective about evidence. Thus, with all that baggage, the chances that I could write an objective account of Ben Goldacre’s book are no greater than that he could write one of the pharmaceutical industry. Anything that either of us writes will be full of *confirmation bias*: we will easily find what we are looking for and ignore what doesn’t fit the picture. When ‘investigating’ something you dislike it is just so much more interesting to make a note of all the errors, foibles and little tricks of the trade that you are criticising (There’s an example of a trick, by the way. Did you notice how I did that?) than to attempt a dispassionate, fair and balanced overview. You should bear that in mind and I shall come back to what you should do about it at the end. In the meantime let’s have a look at a problem but just before getting on to it I’d like you to take a little minute to think about something.

A year is a long time in medicine

‘Ask for a written guarantee that the main outcome of the trial will be published within a maximum of one year after completion.’ *Bad Pharma*, p359

Do you agree with Goldacre that the main outcome of a clinical trial should be published within a maximum of one year after completion? I might quibble at the length of time allowed and I might point out that it takes two to tango and you can take a horse to water but you can’t make it drink, especially if tangoing at the time. You can submit a paper to several journals but this doesn’t guarantee, as I know, alas, that it will be published. However, I do applaud the sentiment even if I have a quite different solution.

But I would like you to reflect on another question. How long should a journal be allowed before it retracts a paper that has been shown to be wrong? Let’s take as an example a journal that appears weekly and a paper based, say on incorrect data? For example, the paper might have been one
whose incorrect ‘results’ were being used to decide on treatment for patients in cancer. Note that here it is not a tango that’s involved but a break dance. It is something that the journal can decide on its own. And here’s another one for you. What would you think of a paper in which results from some patients had simply been counted twice so that what were really results from 197 patients appeared to be 394? Or suppose that there was an organisation devoted to summarising evidence and making it available to doctors online to make decisions and supposing it was informed that data had been counted twice. How long should it take to post a warning online that something is wrong: a month, six months, two years, three years? Write down your answers and we can look at them later.

It’s not the journals’ fault

‘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, p34

The story I am going to look at is this one. Goldacre has a passionately written first chapter about missing evidence. He shows quite rightly that it is a problem and that a particular source of bias is that negative studies, ones that don’t show a significant difference between treatments, are far less likely to be published than positive ones in which the experimental treatment for acne, say, knocks the spots off the competition. He wants you to understand that the main culprits are the pharmaceutical industry so he needs to deal with a side issue. Could it really be journal editors who are to blame? Are they refusing to publish negative studies?

‘No’, says Goldacre and gives you some references. Let’s look at the first one, the subject of the quotation above, a paper in the respected medical journal, JAMA and entitled ‘Publication bias in editorial decision making’[2]. Because I am going to take a hard look at this paper, I want to make it clear that this is not a bad paper. On the contrary it is a model of clarity and circumspection and also, and I mean this quite sincerely, a rare example of a study on the process of publication that takes ethics seriously. In fact it has only one fatal flaw, which in my experience is very few for a paper in a medical journal, and this flaw is one that is very difficult to spot even for a statistician (except, of course, one of exceptionally penetrating insight).

I will leave the fatal flaw for a bit and concentrate on something else. A problem with this paper is not the paper itself but how Goldacre uses it. He cites it as evidence that in general, as regards the work of editors deciding on what gets published, all is well. (Later in the book he changes his tune regarding editors but that is also a story I will have to postpone for the moment.)

However the JAMA authors were very careful never to pretend this. They never regarded it as being an examination of all medical publishing and they specifically warn in the comment section against assuming the results are generalisable. However, you would have had to have read the paper to see this. The abstract states: ‘Prospective cohort study of manuscripts submitted to JAMA from February 1996 through August 1999.’ ([2]p 2825 my emphasis ). Thus, there is one amusing thing about this example. Enrollment finished at the end of August 1999 but the study was not published until the beginning of June 2002. This means that it took two years and 9 months for the article to appear and it thus exceeded the ‘Goldacre limit’ of one year by 175% which seems like quite a lot,
even allowing for the fact that the authors will have had to wait some time to see if a paper submitted to *JAMA* in August 1999 was eventually finally rejected or accepted. We know of course that Goldacre must be allowing a lot less than one year for the business of deciding on rejection or acceptance of a submitted paper because part of the Goldacre limit has to be taken up with entering the data, checking the values, analysing the results, writing the paper and finally submission, a process that Goldacre describes elsewhere as being ‘a pig’ (p290), and of course if you submit to *JAMA*, for which the average acceptance rate was only 17.9% you had better, as a precaution, leave plenty of time to resubmit to other journals. So being *very*, *very*, generous and allowing *JAMA* six months for a decision (which only leaves six months of the Goldacre limit for everything else) everything should have been done and dusted first quarter 2001 rather than mid-2002 at the very latest. (Of course, it is always theoretically possible that the study was not truly prospective but that the data were collected some time later but we can dismiss this because the authors would have mentioned it and they didn’t.)

So it took a long time and this is despite the fact that the omens for an early acceptance were particularly propitious but you would not know why. Although you could find out by reading *Bad Pharma* that this was a paper about *JAMA* you wouldn’t find out that it was by *JAMA*. Four *JAMA* editors carried out a study with the help of five co-authors and discovered that *JAMA* editorial decisions were excellent. That, by the way, is not the fatal flaw. I have absolutely no problem with this. If you read the paper, and I am assuming that Goldacre did, it is absolutely clear there is no claim about journals in general. I have no problem with *JAMA* editors examining their own work and boasting about it in their journal. I *would* have had a problem if the editors of *JAMA* had done a comparative study of the quality of *JAMA* and *The Lancet*, say, and had not bothered to invite *The Lancet* editors to join them. Do you agree with me? If so, is this a general principle? If, for example, members of the Cochrane Collaboration carry out a review comparing studies carried out by the Cochrane Collaboration and the pharmaceutical industry, where do you stand?

**Representing research**

‘One study from 2007 took 179 representative asthma patients from the general population and looked at how many would have been eligible to participate in a group of asthma treatment trials. The answer was 6 per cent on average.’ *Bad Pharma*, p177

But perhaps, I am using unnecessarily strict standards of generaliseability. Well actually, I am much more relaxed on this issue than Goldacre himself. I don’t think much of the idea of representativeness in clinical research and have said so in many places (here is a particularly fine paper of mine on the subject[3]) and not because the pharma industry pays me to say so. The reason is far worse. It’s because I know that I am one of the few people that really understands this and I enjoy telling others how stupid they are. I now know that Goldacre is one of those who doesn’t understand. I will spare you the references to all my other papers but am happy to provide them if you wish but the main reason is that trials can use randomisation to eliminate biases of comparison even if this of itself does nothing to address biases of representativeness. (In fact right across all the sciences and right throughout scientific history, scientists, use and have used highly artificial and unrepresentative but controlled experiments to untangle cause and effect and medicine is no different.) One of Goldacre’s many complaints about pharma industry trials is that they are ungeneralisable. But the study in *JAMA* is not a trial and that means that any problem of
unrepresentativeness would apply in spades, in particular because it is being used by Goldacre to show what cannot be the explanation. So here is a question for Dr Goldacre. What percentage of the annual total of papers published in the medical literature appear in JAMA, 6 in a hundred, 6 in a thousand, 6 in ten thousand, and how ‘typical’ are they?

Goldacre also fails to distinguish between evidence of absence and absence of evidence[4]. In the days when I taught medical students I always explained the difference like this: someone who has been acquitted on a charge of child-molesting is not necessarily a good choice of baby-sitter. To prove that there is a bias in favour of positive studies requires a fairly high standard of proof. Failure to condemn by this moderately high standard does not prove that editors are innocent.

However, I digress. As regards what is claimed in the JAMA paper, again the authors are blameless. They never claimed that they proved there was no bias in favour of publishing positive results. In fact they conceded that the results if taken at face value would show a 30% increase in the odds of being published if the study was positive. However, as they explained, this increase was perfectly compatible with a chance explanation. The much-maligned (in Bad PHARMA) FDA will not allow you to claim no effect in this way and neither did the JAMA authors, only Goldacre did that. The authors are even disarmingly honest about one particular limitation of their study findings. Had they defined a positive study differently there would have been a ‘significant difference’.

Now you see it, now you don’t

Overall, the pharmaceutical industry spend around half a billion dollars a year on advertising in academic journals. The biggest –NEJM, JAMA – take $10 or $20 million each…. Bad Pharma, p305

However, these are minor problems. I will give Goldacre the benefit of the doubt here. It is a common mistake to regard absence of evidence as evidence of absence[4] and although Goldacre waxes lyrical about the problems posed by definition of endpoints (pp200-201), I am so used to the FDA being much stricter about this than journal editors that I can’t get excited about what might potentially be an issue in a published paper regarding a choice of two endpoints. According to the authors, depending upon which endpoint you use, there is a significant difference between positive and negative papers and the one Goldacre (implicitly) uses is more favourable to his case but by the standards of what gets published in the medical literature then this representation of the JAMA study is unremarkable.

However, as regards the use of this paper in Bad Pharma there is worse and this is harder to explain away. The very same study that is cited in Bad Pharma as evidence that the editors do not favour positive papers could be used (by the standards of evidence Goldacre employs) to show that there is no bias in favour, despite all the mouth-watering financial inducements, for publishing papers from the pharmaceutical industry compared to those from government (United States or foreign).

This is much more serious because much later in Bad Pharma, in chapter 6, Goldacre’s purpose is no longer to show that pharmaceutical industry is biasing the record by hiding negative papers, at which point he needs to show that editors are angels¹, but that it is biasing the record by bribing the

¹I really should resist making a joke here but if Pope Gregory the first couldn’t, why should I? All editors may not be angels but for some years an editor of the NEJM was an Angell whose first name was Marcia. She gets a mention in Bad Pharma.
editors to publish bad stuff, at which point he needs to show that the editors are sinners. (See the section quotation above.)

Table I of the JAMA paper (p2826) gives you some crude figures on papers submitted and published and you can have a look at various possible factors that might influence publication. For example 19 of 189 or 10.1% of papers submitted made it to publication if no subjects in the study came from the USA whereas 114 out of 556 or 20.5% made it if at least some of the subjects came from the USA. A difference of this size would not easily arise if chance were the only possible explanation. Or how about the main objective of the study itself? Here you have that 78 out of 383 positive studies, or 20.4% were accepted as opposed to 51 out of 341 negative studies or 15%. This difference is, however, the sort that could arise easily by chance. On the other hand, if we compare pharma industry trials to government trials how do things stack up? Then we have 63 out of 298 or 21.1% versus 73 out of 368 or 19.8%. Of course this difference, which is much smaller, whatever scale you measure it on, could also easily arise by chance. (These two categories of paper, pharma industry and government are not the only categories so the total is less than that of all papers.)

I want to make it quite clear here that I am not claiming this as evidence of lack of bias. This is because I think that there is a fatal flaw with this kind of research. I am simply claiming that if Goldacre thinks that this sort of study has any relevance to showing in chapter 1 of Bad Pharma why the editors are not to blame if negative studies are less likely to get published he needs to explain why it has no relevance to the material of chapter 6 in which he claims the pharma industry are bribing their way into publication.

Let’s take a side-step to examine Goldacre’s ability to report research accurately.

**Sauce for the goose**

‘Research reviewing a long series of FDA votes found that experts are slightly more likely to vote in a company’s interest if they have a financial interest to that company.’ Bad Pharma, p126

Let’s take a gander at this particular story. It describes some research[5] by Lurie et al published in a journal that makes quite an appearance in the pages of Bad Pharma. Yes, I am referring to JAMA. Think about it and tell me which of the following statements about the research described you think you might expect from Goldacre’s description.

1. Six different general approaches to analysing results were performed and in some cases applied to a number of different aspects of conflict. (For example an individual might have a conflict of interest because having worked for a sponsor of the drug being considered or for having worked for a rival manufacturer.) As a result rather a large number of different statistical analyses were performed. None was declared as primary and there were no adjustments for multiple testing.
2. The researchers found that given the actual votes cast, in none of any of the cases they reviewed would it have made any difference to the majority decision if the votes of the conflicted (by interest) person had been discounted.
3. The third of the six approaches would have shown a less favourable result for the drug in question if the conflicted persons’ votes had been eliminated.
4. Using the sixth approach to analysis there was no significant association between outcome and conflict of interest.

5. Using the fifth analysis there was a significant effect but it was paradoxical. Conflicted parties whose conflict should have led them to vote for the drug were less likely to vote for the drug in question (but not significantly so). On the other hand where the conflict should have led them to vote against the drug there was a slight but significant apparent bias in favour of the drug.

Well, perhaps you might have expected something like point 3. But in view of the other points and in particular point 5, don’t you think that Goldacre’s description is just a teeny bit misleading?

If you don’t believe me, look the paper[5] up.

Multiple Errors

“So now, let’s imagine you’re running a trial where you measure ten different independent outcomes. If we set the cut-off for statistical significance as ‘one in twenty’, then even if your drug does nothing useful at all, in your single trial you’ve still got a 50/50 chance of finding a positive benefit in at least one of your outcomes, simply from random variation in your data.” Bad Pharma, pp201-205

Actually, Goldacre is wrong here. Using this logic, if you had 40 independent different outcomes you would expect a 200%, ‘chance of finding a positive benefit in at least one of your outcomes,’ which is clearly absurd. However, the general message is correct: other things being equal, the more you measure the greater the chance that something will turn up.

However, these wise words of Goldacre are certainly relevant to his reporting of the Lurie et al paper ‘…if you haven’t got a positive result you can just spin harder’ Bad Pharma p216

Again, I want to make it clear, that the fault with this story is in Goldacre’s reporting and not the original research.

Hitting the flaw

‘If you are not statistically minded, then you’re not missing out; at least no more here than elsewhere in life’ Bad Pharma (p154)

This is the most offensive and also the most wrong-headed remark in the whole of Bad Pharma. Without being statistically minded not only are you liable to make the sort of error reported in the previous section you also have no hope of spotting the fatal flaw and when you see it, it is really quite interesting and you will have the added satisfaction of knowing more about the subject than a quartet of JAMA editors, five co-authors and a medical journalist. Admittedly, this does not set the bar very high.

So, now, at last, I am going to tell you what the fatal flaw is. However, I must warn you that this section is not going to be easy and you are going to have to put your thinking caps on. However, I can give you this by way of encouragement: Statistics is a subject that many medics find easy but
most statisticians find difficult. If you find what follows difficult you are in good company, or at least in company I think is good.

I am going to suppose that the probability of having a paper published depends on its ‘quality’. If this is not the case and you disagree, then this is a great finding that will streamline the whole process of peer review and make it much easier for authors to reach the Goldacre limit of publishing in one year. We can then dispense with peer review altogether. Actually, I am perfectly serious in saying that I think that this might be a very good thing indeed but since I am frequently asked to review for medical journals and even sometimes agree to do so, I know that their editors don’t agree and I appear to be in a minority on this.

Figure 1 shows the probability of acceptance for two types of paper as a function of ‘quality’ measured on an arbitrary scale from 0 to 100. Quality is shown on the horizontal axis. The two types of paper are ‘result positive’ and ‘result negative’. The probability of acceptance is shown on the vertical axis. The way that you use this diagram is as follows. Choose a given value of the quality of a paper and locate that quality on the horizontal axis. Draw a vertical line until it intersects a curve. If your paper is positive, use the upper curve and if it is negative use the lower curve. Now draw a horizontal line from the point of intersection to the vertical axis and read off the probability of acceptance on the vertical axis.

I have done this for studies with a quality of 50 and you will see that the probability of being accepted is about 50% for a negative study of such quality and just under 90% for a positive study. I have shown the two curves as being very different but they could always in theory be identical and in a sense, what the JAMA authors are trying to show is that they are. If that were the case it would make no difference to a submitted study’s chance of being published if it were positive or negative, although its quality would, of course, be highly relevant.
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

Now, ideally we would like to be able to construct these curves by using data but we can’t. That’s because we don’t see all the data, a point that Goldacre repeatedly makes. However if it were the case, and it’s a big if, that the distribution as regards quality of submitted papers to a journal were identical for negative and positive studies then if the acceptance rate were similar this would be a strong indication that the curves (which we don’t see) were identical. However, you have to believe that authors’ decision as to whether to submit or not is based on quality only and not on study outcome.
If the authors make a decision based on outcome then the situation is quite different. But who would believe that authors would do such a thing? Well Ben Goldacre for one. See the section quotation. In fact that’s a major plank of chapter 1. In other words, if Goldacre stopped to think and analyse, rather than just rushing on eagerly to triumphantly produce yet another nail for the coffin, he would realise that what he claims about authors makes it impossible to judge editors in the naive way that all those evidence-based medicine crusaders have done.

So what might happen to the data we see in journals if what Goldacre believes about authors and the process of submission of article and their publication? Now, I am afraid, you are going to have to think even harder.

**Flying pigs**  
After the paper is written, the horror begins. Several colleagues who have their names on it will have small comments, suggestions tweaks. ....Finally the submission process is a pig. Every academic journal has different pernickety requirements; every one wants the references to be formatted in a different way...and so on. Bad Pharma, p291

Here it is necessary for me to get personal. I occasionally get involved in writing grant proposals. This is a terrible activity, mind-numbingly boring, worse even in my opinion than reading grant proposals and that is really bad. In other words, there is a cost. If you were to tell me that the probability of my proposal being successful was 1% then, even if I never had to write another if successful because the reward was so great, I would not write this one. On the other hand, if you tell me that the probability is 99% I will probably write it, although it does, of course, depend on the money at stake.

Now suppose that you are a medical researcher faced with a submission proposal that is, to quote, ‘a pig’. If you act rationally you will not target a high prestige journal unless you think you have a reasonable chance of success. Rational behaviour is to decide on submission for a given journal if the probability of acceptance is greater than or equal to some threshold that depends on the prestige of the journal. In that case what you would see is what you would get in figure 2. In other words despite the fact that the two curves are very different you would get identical threshold probability of acceptance for the two types of study. What would differ is not the threshold probability of acceptance but the quality of the study at the threshold.

Remember also that you have a huge choice of journals. You might like to get a publication in JAMA but if you assess that the chances are 1 in a 1000 whereas if you submit to the Aleutian Islands Medical Gazette your chances are 99 in a 100, you might decide to sacrifice your chances of getting the world to recognise your genius for a near cert of an admittedly geographically limited fame in the Bering Straits.

So suppose, and it is at least as (and I would argue more) plausible than what the JAMA authors implicitly assumed, that probability of acceptance is the basis for submission. In that case it is Figure 2 that is relevant and not Figure 1. Now, we have to be a bit careful here because although this threshold probability would be the same for the two curves, the papers that would be submitted would be a mixture of papers above the threshold and these mixtures could be different. This would still lead to different overall probabilities of acceptance. However, the way this would pan out would be quite difficult to predict, a number of scenarios are possible, and what makes it even more
difficult is that it is quite plausible that the quality of the study would have an impact on the probability of it having a positive result in the first place. In fact, if this is not the case there is not much point in consulting a medical statistician when you plan your study and that you should do so is one of the things that the FDA, the MRC the pharmaceutical industry and even journal editors seem to agree on. Of course if Goldacre’s most offensive remark is anything to go by he may not agree. (I will come back to this point because it is relevant to another of Goldacre’s errors.)

Figure 2

This makes the whole thing extremely complex, so complex, in fact, that even I, who think of myself as not lacking in self-confidence (others, such is the petty-mindedness of humanity, label this as arrogance) don’t know exactly how the data we have should be analysed. What I do know, however, is this. Never mind quartets of medical editors, a whole orchestra of such persons with a medical journalist as a soloist couldn’t produce intelligible music out of this.
Of quality and quibbles

"...you cannot simply say, 'I saw it in a peer reviewed paper, therefore it is true’ “Bad Pharma, P304

I agree entirely with Goldacre here and I would say that this applies a fortiori when even the peer-reviewed paper didn’t say it was true. Certainly you can’t say ‘I saw it in Bad Pharma therefore it’s true’.

Now, one could argue, that I have ignored a crucial point about the JAMA study: the editors who were judging how well they had done, also included in the analysis some other factors. As they put it: ‘We extracted information on study characteristics examined by others for association with publication and on objective indicators of study quality and reporting transparency.’ (p2826) This meant that they were also able to do an analysis of factors that might be predictive of acceptance apart from outcome (positive or negative). For example, one factor they included, as we have seen, was source of funding. Some of these factors are quality factors and so one could argue that the authors would be able to control for the quality aspect and thus disentangle the data sufficiently to see whether there was a bias in favour of positive or negative studies.

However, quite apart from the fact that the authors were apparently completely unaware of the possible problem represented by Figure 2, they must have believed that their attempts to measure quality completely objectively were a failure. If they were a success then all they needed to do was make reviewing a pure check-box activity. I suspect in fact the score that they developed has poor discriminating power. In other words, it does not predict very well what gets published and what does not get published. But if what gets published in JAMA does not reflect quality, what are we talking about?

One could also argue, that I have ignored other studies that Goldacre cites. However, these were summarised in an energetically prosecuted meta-analysis published in PLOS medicine[6]. That summary does indeed find no reduced probability of publication for negative studies but it also inherits the fatal flaw. And in any case one of the studies summarised, and which Goldacre cites, has a different major finding he (and the meta-analysts) appear to have overlooked. I quote from the abstract[7]. ‘Studies with a negative outcome were of higher quality (p = 0.003) and included larger sample sizes (p = 0.05).’ This is exactly what you would expect from figure 2 and also ‘Studies with a negative outcome, although seemingly superior in quality, fared no better than studies with a positive outcome in the peer-review process.’ (p1010) This is also what you would expect if authors were correctly predicting that there was a bias against negative studies. They judge it is not worth submitting a negative study unless the study is of above average quality because journals are biased against negative studies, and, it seems, they might well be right.

Finally, what experimental evidence, there is is against Goldacre and against Figure 1 and in favour of Figure 2. This evidence comes from studies submitting manuscripts of fake studies in ‘positive’ and ‘negative’ versions for reviewers to judge. Here he cites a number of studies most of which found a bias in favour of positive papers. However, in one of them, by Epstein[8] the difference of 35% acceptance for positive studies versus 26% for negative, the results was not significant, so Goldacre pulls his ‘absence of evidence is evidence of absence’ trick to inform the reader ‘that this difference was not large enough to be statistically significant’(p35) and then finally delivers this magisterial 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’. (p36)
*Size matters*

Before we get going, we need to establish one thing beyond any doubt: industry funded trials are more likely to produce a positive, flattering result than independently funded trials. *Bad Pharma*, P1

Before I leave this story, I draw attention to one interesting finding of the *JAMA* study. If you regard this sort of study at being at all useful about saying something about this issue, then there was one characteristic in the papers submitted to *JAMA* that leaps out (apart from whether Americans were studied or not) as being predictive of success and that is whether a sample size calculation was performed. Here it is the case that 73/287 (25.4%) papers with such a calculation were accepted as opposed to 60/458 (13.1%). A difference of this size would not easily arise by chance.

Now, as it happens, I am not very keen on judging the results of completed trials using sample size calculations. I personally think they are a planning issue not a reporting issue and have said so in print[9]. This is not, by the way, because the pharmaceutical industry has bribed me to say so, and in fact such calculations are nearly always done for regulatory studies. No, it reflects that other part of my bad nature. It’s schadenfreude not veniality. I was telling others that they did not understand the subject.

However, I do think that performing such a calculation is useful as a planning exercise. Why is what is useful for planning a trial not useful for judging it? I used to put it like this, ‘I would not advise you to attempt to row across the Atlantic in a 20 ft open dory. However, the fact that I would give you this advice would not cause me to deny that John Ridgway and Chay Blyth did exactly that in 1966.’

So sample size calculations are important when planning a study. I am sure I used to stress that when teaching at a certain London medical school even if, when heading up the statistics clinic at the same medical school, I was frequently faced with clients who thought that it was something you did once the study was over. I can’t pretend, therefore, that I thought it was impossible to graduate from that medical school without understanding it, so I can hardly claim the attitude in *Bad Pharma* is a surprise.

So I think that having a decently sized trial, having conducted pilot studies, having planned the trial with the help of all sorts of highly qualified persons including statisticians, data-base managers, clinical research associates and of course physicians, should increase your chances of success. If it doesn’t, I have a very good way to make drug development cheaper. So I don’t find it at all surprising that the success rate for pharmaceutical studies is higher. If not, something is going wrong.

Finally, I want to make it quite clear that I do not believe that everything is fine as regards publishing studies, I do not believe all is well with the pharma industry’s approach to publication but I also do not believe all is well with academic research. In fact, despite the fact that I think that the Cochrane Collaboration is a good thing I also do not think it is beyond reproach. I do think the medical literature is very poor. I think it is only the drug regulatory process (a quite separate matter from publication) that is half-decent. I also believe, however, that as this story shows *Bad Pharma* cannot be trusted in its summary of the evidence.
Duke credit

‘Researchers at Duke University, North Carolina, then surveyed the contracts between medical schools and industry sponsors, and found that this edict was flouted as a matter of routine.’ Bad Pharma P43

For many of us the biggest story on the subject of data quality in medical research in the last few years was the ‘Duke University scandal’. Keith Baggerly and Kevin Coombes, statisticians at the MD Anderson spent a frustrating amount of time (well in excess of 1000 man hours) trying to repeat some results on biomarkers in cancer that had been published by scientists at Duke University. This is far more time, I can assure you, than Ben Goldacre ever spent trying to get data out of the FDA.

It had been claimed that there was a way of deciding which cancer-patients would benefit and which would not using these biomarkers. In fact patients were even being treated according to these recommendations. Baggerly and Coombes eventually decided that the results were plain wrong. They had to give up trying to get the medical journals concerned to rectify the situation and published their findings in a statistical journal[10]. I have some sympathy with them for doing this. Several years ago I tried to get the Cochrane Collaboration (CC) and a certain medical journal to correct the results of a meta-analysis in which patients had been counted twice. There was no question of fraud, it was just plain methodological error and 197 patients under a control treatment were entered twice as if they came from two separate trials. (This sort of mistake was made several times.) This error was traceable to an unfortunate way that meta-analyses were carried out generally in CC studies and this is not the only such study in which I have found it.

Call me old-fashioned but I think that counting data twice is a very bad thing. I often put it likes this. The CC is worried about missing data whereas I am worried about data that are present that don’t even exist. I think where this happens it requires instant ‘action to retraction’. In the case of the CC this action had to wait more than three years. In fact, as far as I am aware, the study was never retracted it was just eventually replaced. As regards the medical journal concerned I am not going to say which it was but leave this as a tantalising missing data problem for you to solve. However, I will say that the correspondence between me and the editor in charge did not peter out until I had been accused of unprofessional behaviour by not having communicated the mistake earlier! As regards any retraction it is now four years and I am still counting. You can compare these values with the figures you wrote down earlier.

Now, let’s return to the Duke University story[10]. Some while after the statisticians published their results, the medical journals started slowly, very slowly to retract the papers. You might, like for example to get hold first of one of the original papers[11], which was published in JAMA, as it happens, and then the retraction notice[12]. When I tried this on 10 October 2012 I found the former much easier than the latter. How long between the appearance of the Baggerly and Coombes paper and the JAMA retraction? More than twice the Goldacre limit to get a paper published. Please compare this to your answers to my question at the beginning.

And here’s a quote from Keith Baggerly himself that runs counter to everything Goldacre claims in Chapter 3 Bad Regulators, ‘It seems likely that several of the problems identified with the Duke trials would have been caught by an FDA review....’[13] Interestingly, JAMA has a policy of not publishing pharma industry trials unless the analysis is checked by an independent academic
statistician, a point to which Goldacre refers (p326). But just think how much embarrassment they could have saved themselves if they had had all academic analyses checked by the industry!

Well actually, Amgen did something like this. They had a look at 53 landmark studies in cancer. They could only reproduce results for six[14]. You won’t find this story in Bad Pharma. Or how about the Bayer scientists who had a similar experience? They reviewed 67 studies and could only reproduce about a quarter[15]. Again no mention in Bad Pharma. But don’t forget how bad I am. You will find both of these companies mentioned in my declaration of interest. Note, I am not claiming here that any incorrect analyses were involved in the original papers, still less that any fraud was involved. What I can say, however, is this. When you submit a study to the FDA you provide the data-files and may be asked for your programs. I have reviewed for medical journals for over 15 years now and have never been given either.

**Meat and poison**

‘...as industry is keen to point out, where people have compared the methods of independently-sponsored trials against industry-sponsored ones, industry trials often come out better. This may well be true, but it is almost irrelevant...’ Bad Pharma, p171

Now let us return to the investigation by Baggerly and Coombes[10]. This is a serious story but Goldacre chooses not to treat it. It can’t be because of some particular prejudice against Duke University; see the opening quotation for the previous section. Perhaps the clue is here, ‘...outright fraud is almost the only source of distortion that receives regular media coverage, simply because it’s easy to understand. That’s reason enough for me to leave it alone and move on to the meat.’ (p175-176) I am not suggesting that the Duke data are fraudulent, although others have. However, I am suggesting that if fraud does not disturb you, then data that, for whatever reason, are completely meaningless won’t bother you either. Perhaps this attitude is shared by others. Perhaps this is why it is so difficult to get retractions of double-counting.

Well, one man’s poison is another man’s meat and I am not unconcerned about fraud. In fact having encountered, during the eight years I worked in the industry, examples of non-pharma medics making up data in the trials I was involved in, I am very concerned about it indeed. Goldacre has his personal history of interaction with pharma reps and I don’t disagree with this but I have my anecdotes too. For some reason, the case of the patient who was treated days after her death in a car accident, of which her doctor was unaware, sticks in the mind.

Of course, you may say to me, ‘what has any of this to do with Goldacre’s book?. OK, so Goldacre cited JAMA and Duke University in support of claims he made. But just because some JAMA authors made an error does not mean that JAMA authors cannot be trusted. Just because some Duke University authors published results that turned out to be wrong doesn’t mean that other Duke scientists cannot be trusted.’ I agree. But is this a universal principle? If you substitute ‘GSK’ for ‘Duke’ or ‘JAMA’ are you still happy?

And Goldacre also attacked the regulators and he did this without recognizing, which he should have done, that they do a much better job than the BMJ, The New England Journal of Medicine (NEJM), The Lancet, JAMA and the Aleutian Islands Medical Gazette.
As regards the story of the Duke data, you might like to GOOGLE this example. You will learn quite a lot about the medical press that is not in *Bad Pharma*.

**Speaking seriously**

In the unlikely event that any example I have used is simply wrong, there will be another to slot in its place. *Bad Pharma* P372

Have I been unfair? Yes. Has Goldacre been unfair? Yes. Have I carried out an unbiased view of *Bad Pharma*. No. Do I disagree with all of what Goldacre recommends? No. In an article I wrote in 2000 I said ‘No sponsor who refuses to provide end-users with trial data deserves to sell drugs’ [16](p26) and I have made similar statements since elsewhere[17]. However, unlike Goldacre I believe the medical press is beyond redemption. The pharma industry problem is only the tip of the iceberg. Peer review as a condition of publication has to go. Pharmaceutical sponsors should not be allowed to hide behind regulators[16]. It’s the data we need, not the papers. Insisting on publication within a year is impractical if this means publishing in journals. These should not be used for publishing primary research but for commentaries. One should make data available on the web although we need to get together and think how we are going to manage this. However, we can’t exempt academics from this and I predict a very long and bitter battle where they are concerned.

Am I being postmodern and do I think that there is no such thing as truth? No. I think it’s out there but it’s complex and it’s hard to find. The literature on medicine is undoubtedly skewed by the pharma industry but it’s also skewed by academics desperate for a publication and the literature on methodology is skewed by researchers who are just out of their depths and have more to offer in the way of self-righteous and prejudiced indignation than they do in expertise. (I could write a whole book, for example, on meaningless, irrelevant, and in some cases downright wrong investigations into methodology for dealing with subgroups by people who are deeply confused and have succeeded in infecting chief executive officers of the pharmaceutical industry with their confusion.)

Do I think that *Bad Pharma* should never have been published? No. Do I think it’s no good? No. It is good if you understand what it is. It is a case for the prosecution. Science is a debate. We need people to shake up our prejudices. Ben Goldacre as an Ivan Ilich is fine by me. I think that chapter 5 on *Bigger Simpler Trials* is very good and that chapter 6, *Marketing*, although hard to take for someone who has worked in the industry is also good. I think that chapter 4 *Bad Trials*, is downright misleading and chapter 3 on *Bad Regulators* is a disgrace.

Do I think you should not read *Bad Pharma*? No. Read it, but read it critically. Give the case for the prosecution a fair hearing but don’t mistake the prosecutor for the judge and hear what the defence has to say. Above all think!

Is there anything that is beyond the pale in *Bad Pharma*? Two things are not acceptable. You can’t lecture others about the problems of missing data and then say that that doesn’t apply to me: I’ll just replace the stories that failed by ones that work. You also can’t dismiss arguments by doing a Mandy Rice-Davies. I predict that that is a ploy that Goldacre will use in answering criticism but then he wouldn’t he?
References
9. Senn SJ. Power is indeed irrelevant in interpreting completed studies. *BMJ* 2002; **325**: 1304-.