Critical Assessment

Exploring the Asymmetrical Relationship Between the Power of Finance Bias and Evidence

Jeremy Howick

ABSTRACT Financial conflicts of interest can influence the design, conduct, and dissemination of medical trials. In attempting to resist "finance bias," critics and proponents have largely focused on trying to improve evidence. These efforts have led to successes ranging from the 1962 Kefauver-Harris amendments to the US Federal Drug and Cosmetic Act of 1938 to recent recommendations that all trials be published. However, there are two problems with the strategy of trying to improve evidence as a buffer against finance bias. First, without political teeth, rules of evidence can be ignored with relative impunity. This is because, as sociologist Bent Flyvbjerg has pointed out, there is an asymmetry between power (of finance bias) and rationality (evidence), tending towards victory of power in an open confrontation. Second, by improving the way evidence is produced, the process has become more expensive, and thus more susceptible to influence by finance bias. Unless they address the powers behind finance bias directly, critics and proponents may be doomed to lose the war against finance bias, even if they win some battles. For EBM to be effective, the ~~powerful forces~~ influencing the production and dissemination of evidence need to be addressed as a priority. This is starting to happen, with initiatives such as the AllTrials campaign, which identifies and exposes unpublished trials. On the other hand, there are reasons to be less optimistic, as Cochrane, the most trusted source of evidence, has become more susceptible to stronger influences from industry.

Faculty of Philosophy, and the Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford OX2 6GG, United Kingdom.
Email: *jeremy.howick@phc.ox.ac.uk*.
The author wishes to thank Dan Moerman and Bennett Holman.

Perspectives in Biology and Medicine, volume 62, number 1 (winter 2019): 159–187.

> *Power determines what counts as knowledge, what kind of interpretation attains authority as the dominant interpretation. Power procures the knowledge which supports its purposes, while it ignores or suppresses that knowledge which does not serve it.*
>
> **— Bent Flyvbjerg (1998)**

IN MANY IMPORTANT CASES, sociological forces, mostly financial conflicts of interest (I will call this "finance bias," defined below) are powerful enough to upstage evidence when it comes to deciding whether interventions are safe and effective. Using three widely cited historical cases that are hailed as successes of evidence vis-à-vis finance bias, I show that in fact the powers of finance bias eventually prevailed. These cases show that, by comparison, context-free philosophical debates about what counts as "best" evidence are becoming relatively unimportant, as far as the actual practice of medical research. If I am correct, then the power of finance bias needs to be redressed to ensure that good evidence is empowered. Most proponents and critics of EBM are either unaware of the need to do this, or resist doing so. ~~On the one hand,~~ EBM proponents have mostly focused on trying to improve methodologies. This approach is useful for ~~improving~~ how evidence is produced and reported. But without the power to enforce the methodologies they propose, the methods can be—and often are—ignored. Improving evidence standards also may inadvertently exacerbate the problem with finance bias, by making research more expensive. ~~On the other hand,~~ philosophical (and other) critics of EBM have generally favored an expansion of the EBM view of evidence to include evidence related to mechanisms and clinical expertise. This approach has not had any ~~evidence of~~ success in terms of improving *actual* medical practice. Also, as I will show, there are historical reasons to believe that emphasizing the evidential role of mechanisms ~~will~~ exacerbate the problem with finance bias, because this view plays into the hands of industry.

At least in part because they do not address finance bias directly, both critics and proponents ignore an important aspect of what it might take to save what has been called a crisis within EBM, and to prevent EBM from being "hijacked" (Greenhalgh et al. 2014; Ioannidis 2016).

This is not the first time I have discussed the influence of finance bias, and I'm far from the first one to do so. Among many others, Peter Gøtzsche (2014) has noted what he calls the deadly influence of finance bias and likens it to organized crime; Tom Jefferson has fought tooth and nail (with some successes that are outlined below) to expose how hiding data skews evidence (Jefferson et al. 2014); James Robert Brown (2008) has advocated mixing science and method by eliminating patents; Bennett Holman (2017) has dubbed the relationship between finance bias and EBM an "arms race," with improved evidence standards reacting to finance bias; and Justin Biddle (2007) has lamented the lack of recognition of social epistemology in the search for good evidence. Summarizing his view, which I will take to be broadly representative of other philosophical

commentators on finance bias, Holman states: "because the effects of industry funding are both entrenched and pervasive, for an epistemology to be applicable these effects must be incorporated, not as an afterthought, but as a normal part of medical research" (Holman 2017, 2). I believe that Holman and others are correct, and my arguments in this paper support their positions. But Holman's view is compatible with the possibility that finance bias and EBM are on a par, with one sometimes winning out over the other. I suggest that in order to save evidence, we need to address the underlying structures that allow the power of finance bias to upstage it.

The difference between my arguments and most previous arguments about finance bias can also be interpreted in the light of work in the sociology of scientific knowledge (SSK). By and large, arguments about finance bias might be categorized as arguments that support versions of the weak program within SSK: they all show (correctly) that sociological forces influence what ends up counting as evidence. My arguments move beyond this position. The weak program implies that if we recognize the sociological forces, we can adjust our points of view and save knowledge. For example, Holman's (correct, in my view) account that there is an arms race between finance bias and EBM encourages EBM proponents to think about the next way to adjust the way evidence per se is reported or produced. In fact, this is what happens, with the increased standardization of evidence designed as a defense against the influence of finance bias. However, I contend that these attempts to counteract financial bias with purely epistemological tools are destined to be too late, too small, or simply ineffective. Finance bias extends beyond influencing methodology to guiding the choices of hypotheses to be tested (using any methodology), shaping the way results are presented, influencing whether evidence is implemented into guidelines, and influencing whether the guidelines are followed. The weak program, because it focuses on methodological tweaks, does not seem to be able to redress these effects of finance bias. As a result, academic attempts to improve evidence, which are not backed by nearly the same power as finance bias, relatively impotent. In order to recalibrate the ~~dynamic and~~ asymmetric relationship between finance bias and evidence in favor of evidence, the forces that cause the imbalance need to be redressed.

The position supported in this paper is one propounded by sociologist Bent Flyvbjerg. In his 1998 book *Rationality and Power*, Flyvbjerg argues that we need to face the challenges of power head on. His theory has 10 propositions, three of which I shall focus on here:

1. Power defines rationality (industry interests define what counts as evidence);
2. Power wins in an open confrontation ~~(industry wins in an open confrontation)~~; and
3. If we are to preserve rationality (evidence) we need to empower evidence within political structures.

In the medical context of this paper, EBM is "rationality," and finance bias is "power." If we apply Flyvbjerg's arguments to the relationship between the power of bias and evidence—and I shall provide reasons why we should—then we must address the asymmetry between rationality and power if we want evidence to prevail.

While they did not cite him, Patashnik, Gerber, and Dowling, in their 2017 book *Unhealthy Politics: The Battle Over Evidence-Based Medicine*, seem to support Flyvbjerg's theory. They use empirical evidence from surveys to show that financial bias influences patients (who often fail to recognize the influence of finance bias), health-care professionals (whose interests are not always the same as patients' interests), and governments (which are often influenced by very large campaign contributions). Patashnik, Gerber, and Dowling also note what Flyvbjerg's theory predicts: attempts to curtail the problems facing health care have failed, in large part because promoting EBM does not have clear returns for politicians. This means that the power asymmetry between industry and EBM is likely to continue, with EBM the (much) weaker party. They take a step towards acknowledging Flyvbjerg's point by advocating a political solution involving coalition building. However, they may not go far enough because they do not address the underlying structures that enforce the power asymmetry.

If the power of finance bias and evidence are on an unequal playing field, it implies that solutions must address the asymmetry. Such solutions must involve independently funded research and finding ways to align the interests of patients with the interests of profit. This has started to happen with social impact bonds and in some insurance companies.

I shall start with an outline of the influence of finance bias and a definition.

The Power of Finance Bias Defines What Counts as Good Evidence

Industry-sponsored trials are more likely to show a beneficial effect of the intervention being studied than non-industry funded trials (Bero et al. 2007; Jorgensen, Hilden, and Gøtzsche 2006; Leopold et al. 2003; Lexchin et al. 2003; Schulz, Altman, and Moher 2010; Yaphe et al. 2001). This bias can have paradoxical consequences. For example, Heres and colleagues (2006) examined randomized trials that compared different antipsychotic medications. They found that olanzapine beat risperidone, risperidone beat quetiapine, and quetiapine beat olanzapine! The relative success of the drugs was directly related to who sponsored the trial. If the manufacturers of risperidone sponsored the trial, then risperidone was more likely to appear more effective than the others.

Numerous methodological explanations account for the difference between trials funded by industry and independently. First, industry-sponsored trials that do not reveal a result that is favorable to the company's product may simply remain unpublished. This leads to what is known as "publication bias": studies with positive results are more likely to be published than those with negative re-

sults (Hopewell et al. 2009). To cite one example, Turner and colleagues (2008) examined all the antidepressant trials registered with the US Food and Drug Administration (FDA). They identified 74 studies, 38 (51%) with positive results and 36 with negative or questionable results. Of the trials with positive results, all but one was published, whereas of the 36 trials with negative results, 22 were unpublished. To overcome this problem, many advocate mandatory registration of clinical studies (De Angelis et al. 2004). If trials must be registered, it is more difficult (but not impossible) to suppress trials with negative results.

The second methodological reason industry-sponsored trials are more likely to demonstrate a benefit is bias that enters during data analysis. I know many prominent researchers have acknowledged in confidential conversations that they hire two or three statisticians to analyze data from their trials, then choose the one who produces the results they like the best. An obvious way a statistician can influence the results is to purposely—or, more charitably, erroneously—make a calculation error and make a treatment appear more effective than a control when in fact it is not. ~~There is strong evidence that data analysts can influence reports of studies.~~ Peter Gøtzsche (1989, 1990) analyzed reports of 196 trials comparing new nonsteroidal, anti-inflammatory drugs (NSAIDs) with established NSAIDs and found that the new drugs were five times more likely to appear more effective than the established drugs. This was surprising, based on historical evidence that NSAIDs are, on average, equally effective. Where possible, Gøtzsche re-analyzed the data and found that the apparent benefits of the new drugs were errors, and in all cases the erroneous reporting of results favored the new drugs. That is, the errors were not random, and it seems likely that the funding source influenced what errors ~~were~~ allowed. Gøtzsche (1989) also found that calculations of adverse effects were biased by selectively including or excluding patients who had withdrawn because of adverse effects. Gøtzsche is an evidence purist, who believes that the solution to the problem of biased data analysis is to mask the data analysts (Gøtzsche 1996). If the analysts do not know which is the experimental therapy, they will not be able to predictably make systematic errors or omissions that favor the apparent benefit of the experimental intervention.

However, blinding the data analysts wouldn't suffice. In some cases, trials suggest no benefit of a new drug, but the trial report is written as if the drug were beneficial. In Turner's (2008) review of antidepressant trials cited above, 11 of the 36 negative trials were presented in a way that conveyed a positive outcome. This means that a naïve review of the published antidepressant trials would suggest that 94% indicated a positive result, whereas in fact only about half (51%) actually suggested that the drugs had beneficial effects. The difference here is between virtual unanimity (94%) and a coin toss (51%). Other studies have replicated this finding (Bero et al. 2007). Stelfox and colleagues (1998) examined the relationship between the likelihood of supporting beneficial effects of calcium channel blockers and their financial relationships with the pharmaceutical industry. They found

that 96% of supportive authors had a financial relationship with the manufacturer, compared with 60% of neutral authors and 37% of critical authors.

Parallel problems arise with systematic reviews and (so-called) evidence-based guidelines. Industry-supported systematic reviews end up with industry-favorable conclusions (Jorgensen, Hilden, and Gøtzsche 2006). Lamenting the influence of finance bias, evidence guru John Ioannidis (2016) recently claimed that clinical medicine has become "finance-based medicine."

Gøtzsche (1996) has proposed a solution to spinning trials results that involves blinding manuscript authors, so that the people who write the papers do not know which treatment was the placebo and which was the control. He also believes that content experts should not be allowed to produce research in their domains. For example, if I were an expert on placebo treatments, Gøtzsche would say that I should not be allowed to produce evidence on placebo treatments, because to him it would be impossible for me to remain unbiased. His suggestions have not been adopted.

Defining Finance Bias

Finance bias can be defined as follows: a systematic distortion of evidence design, conduct, interpretation, or implementation (Aronson 2018), caused by judgments and choices that favor the interests of industry and profit in a way that is not consistent with what independent assessors would deem to be good evidential practice. It can run contrary to the interests of patients and the public. Finance bias can arise at any stage of the research process, including but not limited to: choosing the research question and methods for discovery, choosing the research design, analyzing and interpreting the results, and publishing and disseminating the findings. Also, it is important to note that neither the existence of finance bias, nor my emphasizing the need to tame it, implies that I am against industry. On the contrary. Entrepreneurship and industry are required for developing, producing, and expanding the treatment options available to us (Ioannidis 2016).

Readers might wonder why the definition doesn't simply state that finance bias is any influence of research by industry. This definition is too broad because, as I shall point out towards the end of this paper, it is possible to align the interests of industry with the interests of patients, in which case industry bias need not arise.

The qualification that finance bias influences evidence in a way that runs contrary to what an independent assessor might claim is difficult to satisfy. How, after all, do we identify independent assessors? This question is a good one, and more research is required to investigate answers. At the same time, at least in the cases I list below, independent assessors are easily identified. ~~In many important cases,~~ the lines between people who are being paid by industry and those who are not are clearly drawn. I therefore adopt a pragmatic approach and take an independent assessor to be an expert in evidence who has no financial interest.

A Note About Scope

The data I present here do not suffice to show that power wins out against rationality in all cases. In order to demonstrate a *general* power asymmetry between finance bias and evidence, I would have to do a systematic review of instances where the two conflict, and show that finance bias won out, at least in a majority of cases. I hope that such a systematic review is eventually conducted, perhaps starting with a certain domain such as publication bias. However, before such research is conducted, the question being asked must be clarified, and some motivation must be provided to make such a project worthwhile. This article provides an insight into a new way of looking at the relationship between the power of finance bias and the rationality of evidence that can guide future research. It also provides preliminary data that support this new way of looking at finance bias and its relative power compared with the interests of evidence. These preliminary data also point the way towards what must be done in order to address the problem with finance bias.

Power versus Rationality, in Three Acts

In this section I shall outline three historical episodes that have been viewed by many commentators as victories for evidence.

Act I. Thalidomide: Power Defines Rationality

Thalidomide [α-(*N*-phthalimido)-glutarimide] was responsible for maiming over 10,000 babies, and led to an unspecified number of miscarriages (Vargesson 2015). Researchers at Chemie Grünenthal discovered the drug in the early 1950s and found it to be an antiemetic effective in managing morning sickness. In 1957 the company launched thalidomide under the name Contergan (Distaval in the United Kingdom), and it was marketed aggressively. For example, one advertisement in the *British Medical Journal* from 24 June 1961 stated that the drug was "both highly effective . . . and outstandingly safe. 'Distaval' (*thalidomide*) has been prescribed for over three years in this country . . . but there is no case on record in which even gross overdosage with 'Distaval' has had harmful results. Put your mind at rest. Depend on the safety of 'Distaval'" (Distillers Company Limited 1961). The drug was a huge commercial success for Chemie Grünenthal, and by the late 1950s it was being marketed in 46 countries under 37 different trade names. Thousands of women took it.

At the time, it was not necessary to test new drugs in humans for safety and efficacy. It sufficed to test the drugs—often in an uncontrolled way—in some animal models. However, in some animals, including rats and rabbits, fetuses that develop abnormalities are reabsorbed (Flores et al. 2014). Hence the effects on the fetuses of animal models did not predict what would subsequently happen to human fetuses.

There were, however, observational reports that the drugs produced peripheral neuropathy in some patients (Fullerton and Kremer 1961). Pointing this out, physician Frances Kelsey blocked approval of the drug by the FDA in the US, where the disaster was largely avoided. Kelsey was subsequently awarded the President's Award for Distinguished Federal Civilian Service by President John F. Kennedy.

It turns out that Chemie Grünenthal were aware of the harmful effects of the drug for some time before they withdrew it from the market. The first to suffer from the drug was the wife of a Chemie Grünenthal employee, whose child was born without ears in 1956; soon after, a German journal reported a case of phocomelia in 1959 (Weidenbach 1959). Chemie Grünenthal denied that the disease was linked to the drug (Sjöström and Nilsson 1972). By 1959, Grünenthal had received over 100 complaints from German doctors and thalidomide distributors. In December 1960, the *British Medical Journal* published a description of four cases of peripheral neuritis attributed by the author to thalidomide (Florence 1960). As their profits from the drugs soared in parallel with the number of victims, the company continued to deny the birth defect connection.

Eventually, a case report from Australia started to link thalidomide to birth defects more explicitly (McBride 1961). Initially, Chemie Grünenthal ignored the reports, claiming that the drugs could not cross the placenta (which was known, at least by some, to be mistaken at the time; see Greek, Shanks, and Rice 2011). A few months later, however, they took thalidomide off the market.

In a long court case in Germany, Chemie Grünenthal was never found guilty. In 2012 the company offered an apology, but stopped short of admitting liability. They agreed to pay about $22,000 to each living thalidomide victim in Germany, and very little to victims outside Germany (Evans 2014). The reason the company escaped prosecution was because its testing met the standards of the time: in a response to an article written by a *Guardian* investigative journalist, Grünenthal stated that its conduct "was consistent with . . . the prevailing standards for the development and testing of the pharmaceutical industry at that time" (Evans 2014). If we put aside the fact that Chemie Grünenthal ignored the evidence, the company's defense was *partly* justifiable. At the time it was not required for new drugs to be tested in rigorously controlled clinical trials before being given marketing approval. In fact, the thalidomide disaster was a key motivation for revolutionizing the way drugs are tested.

In response to the public uproar surrounding the thalidomide disaster, Congress enacted the Kefauver-Harris amendments to the Federal Food, Drug and Cosmetic Act in 1962. The salient feature of the amendments for present purposes is that in order to be approved by the FDA, evidence of effectiveness had to be based on:

> evidence consisting of adequate and well-controlled investigations, including clinical investigations, by experts qualified by scientific training and experience

> to evaluate the effectiveness of the drug involved, on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports. (US Congress 2007)

The Kefauver-Harris Drug Amendments constituted a major milestone in the development of what eventually became the evidence-based medicine (EBM) "philosophy" of evidence (Howick 2011b). Importantly for the purposes of this ~~article,~~ the amendments were not merely designed as a methodological innovation, they were explicitly intended to protect consumers against the power of corporations to abuse patients. Estes Kefauver, who drafted the amendments, was a consumer protection, antitrust, organized labor supporter, and he was accused of being a communist. (The Federal Bureau of Investigation held several files on Kefauver.) His aim was clearly for this methodological innovation to be a powerful protection to humans. The Kefauver-Harris amendments are rightly hailed as a win for rationality and evidence. ~~But~~ I shall provide reasons to question the extent of the victory in the next section.

Act II. Antiarrhythmic Drugs: In an Open Confrontation, Power Wins

Antiarrhythmic drugs were widely used in the 1980s to treat people who had suffered from myocardial infarction (heart attack). Myocardial infarction often damages the muscle and electrical system in the heart, leaving it susceptible to arrhythmias. A common type of arrhythmia, ventricular extra beat (VEBs), occurs when the left ventricle contracts before it has time to fill completely. The heart then fails to pump sufficient blood. Without treatment, lung, brain, and kidney damage ensue~~s~~. Worse, VEBs can also degenerate into ventricular fibrillation and death. Large-scale epidemiological studies suggested that between 25 and 50% of sudden cardiac deaths are associated with arrhythmias. Based on this understanding of the underlying mechanisms, several drugs were developed and found to be successful for regulating VEBs. The drugs became widely prescribed in the belief that they would reduce cardiac deaths. However, when a randomized trial that used *mortality* as an endpoint was conducted, it was discovered that the drugs increased mortality relative to placebo (CAST 1989). Given the widespread use of the drugs, it has been estimated that tens of thousands of people were killed by the drugs each year.[4] Had the initial trials used mortality as an outcome, they

[4] ~~It's not just improved EBM standards that have done this. Improved evidence standards (connected to EBM of course) have also contributed to the problem, as have diminishing returns. Treatments have become more and more effective, with the result that larger and larger trials are needed to determine smaller and smaller beneficial effects. If a treatment reduces mortality in a condition from, say, 10% to 2%, you need an enormous study to improve upon that. Another problem is that there has been increased emphasis on rarer (orphan) disease studies, which are more expensive to conduct because they tend to be multicenter. Then there are diseases that do not respond in some people to otherwise effective treatments. For example, the treatment of migraine has been revolutionized by the invention of triptans, a huge pharmacological success story. But in some people triptans are contraindicated and some do not respond. So an expensive monoclonal antibody has been invented for that small group of individuals.~~

would have quickly discovered that the drugs increased mortality, and the antiarrhythmic disaster would have been avoided. Why didn't they? An important part of the answer is: they weren't required to.

The pharmaceutical companies that were developing antiarrhythmic drugs in the 1980s knew they had to conduct controlled studies (because of the 1962 Kefauver-Harris amendments), but they were unsure what kind of controlled trials the FDA would require. Demonstrating that the drugs reduce VEBs was much easier, cheaper, and less risky than demonstrating that the drugs reduced mortality. And that is just what the industry pushed for.

On 8 and 9 October 1980, the companies developing the drugs convened a consensus conference and invited industry representatives, academic researchers, and members of the FDA (Morganroth et al. 1981). Detailed accounts of how the industry representatives manufactured consent at the conference has been published elsewhere (Holman 2017), so I shall highlight only a few points during the meeting that suffice to demonstrate how finance bias won the day.

First, the cards were stacked in favor of industry interests before the meeting began~~: they set the agenda and had an industry-friendly chairperson~~. The stated agenda was to develop "guidelines to determine how to evaluate such new antiarrhythmic drugs for both efficacy and safety in the most expeditious manner. This symposium will not address important issues of whether or not [VEB] suppression is definitely necessary to prevent sudden death" (Morganroth et al. 1981, 2). This was an odd assertion, given that the very purpose of the meeting was to determine *whether* a surrogate endpoint (VEB suppression) was appropriate. Second, the industry had Joel Morganroth chair the meeting. Morganroth had had industry ties since at least 1976, and after the meeting was paid millions of dollars by 3M to travel around the world promoting antiarrhythmic drugs. When questions were raised about whether reduction in VEBs sufficed, Morganroth consistently suppressed them.

At numerous points during the discussions that followed the presentations, the question about using mortality as an endpoint was raised, then dismissed. For example, Dr. Irving Herling, co-director of the Likoff Cardiovascular Institute in Pittsburgh noted: "we've seen that you can suppress ectopy and still have the individual sustain sudden death" (123). Morganroth replied that "we cannot address the question in this symposium . . . this means using the definition for drug efficacy as the statistical elimination [of VEBs]" (123).

In the same discussion session, Dr. Robert Temple of the FDA worried that it was "troubling to think of that [reduction in arrhythmias] as the definition of effectiveness." Using very similar language, Morganroth replied: "the real crux of your question I think is, how do we know what the best definition of efficacy should be unless we know whether or not it prevents sudden death. Unfortunately, we can't answer that part of the question, so we have dropped back to try to answer the question how do we know whether the drug is doing something at all [i.e., suppressing VEBs]" (128).

Temple's questions were followed by a presentation from the then-Director of the Bureau of Drugs, Dr. J. Richard Crout. Crout's presentation was more of an editorial about FDA policy and history. Unlike most of the other presentations, he did not provide any references to support his claims. Crucially, he stated: "Sometimes, however, with other classes of drugs, there is a conscious decision to accept pharmacological effectiveness as sufficient. Examples here are . . . antiarrhythmic drugs" (176–77). Needless to say, once the surrogate endpoint was deemed acceptable by the then-Director of the Bureau of Drugs, the (relatively simple) trials with surrogate endpoints were done and the drugs were aggressively marketed. Eventually, however two randomized trials were conducted to address growing concern among cardiologists, and these trials discovered that the drugs were killing people (CAST 1989; CAST II 1992). Soon after the trials were published, the drugs were removed from the market.

The antiarrhythmic case is a cause célèbre for the EBM movement, who cite it in textbooks (Sackett 1997, 2000; Straus et al. 2005), and a book detailing the damage it caused was published shortly after the CAST trials (Moore 1995). EBM proponents use the example to emphasize the need for properly reported and conducted randomized trials that include clinically relevant endpoints. And they have had some success in these efforts, particularly in three important areas: (1) standardized reporting guidelines; (2) insisting on clinically relevant endpoints; and (3) disclosure of financial conflicts of interest.

Until recently, there was no standardized way of reporting the methods and results of randomized trials. Trial reports often failed to report *how* they randomized; whether or how they concealed allocation; which groups were blinded; and which statistical methods were used. This incomplete reporting left anyone wishing to critically appraise these documents in the dark. Better reporting of trials, if the reporting accurately reflects the actual conduct (more about this below), improves trial quality. Hence, it is not surprising that a year after the announcement of EBM as a new movement in 1992, 30 researchers gathered in Ottawa, Canada, to discuss how to assess the quality of randomized trials. They developed what became the Consolidated Standards of Reporting Trials (CONSORT) Statement, which was first published in 2001 and updated in 2010 (Altman et al. 2001; Schulz, Altman, and Moher 2010).

The initial aim of the CONSORT group was to improve the quality of the trials, not the quality of the reports. But they quickly changed their focus. In their words:

> during preliminary discussions, participants felt that many of the suggested scale items were irrelevant because authors did not regularly report them. In fact, there was accumulating evidence that the quality of reports of RCTs was less than optimal. Therefore, unanimous agreement steered the remainder of the workshop to focus on ways to improve the reporting of RCTs. (CONSORT 2018)

The CONSORT Statement and related efforts have had some success. Most reputable journals now require that randomized trials report trials using the CONSORT method. Using similar approaches, reporting guidelines now exist for most types of study, including observational studies, systematic reviews, economic evaluations, animal studies, and more. Additionally, many of the better-established guidelines have extensions. The CONSORT statement, for example, has 23 extensions, which are all collated and promoted in an umbrella organization called Enhancing the QUAlity and Transparency Of health Research (EQUATOR) (Equator Network 2018).

As a consequence of CONSORT and related efforts, the quality of trial reporting has improved (Moher, Jones, and Lepage 2001). And since there seems to be a positive correlation between reporting and trial quality (Savovic et al. 2012), it is likely that improved reporting has improved the quality of trials. All this is good, but it remains unclear whether these efforts have done much to stem the effects of finance bias.

Another way in which EBM proponents have improved things in a way that might avert future antiarrhythmic drug scandals is to promote clinical endpoints. I have argued elsewhere that the very definition of "clinical effectiveness" includes evidence that an intervention affects (or does not affect) a clinically relevant endpoint (Howick 2011b). This suggestion is reflected in the most widely used system for ranking the quality of controlled studies (randomized or observational) in systematic reviews. The Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) system ranks evidence in two steps. First, an a priori ranking of high or low quality is assigned to a study, with observational studies ranked low and randomized trials high. Then, there are many considerations that lead the study to being rated up or down. Finally, after being adjusted up or down, the a posteriori quality of the study is assigned. This can be high, moderate, low, or very low.

"Indirectness" is one of the considerations that can lead to down-rating the quality of the evidence. For a controlled study to be deemed of high quality, it must, among other things, provide "direct" evidence. Direct evidence means (again, among other things) that the outcome should be a clinically relevant outcome rather than a surrogate outcome. In the GRADE Working Group's words: "the use of a surrogate outcome requires rating down the quality of evidence by one, or even two, levels" (Guyatt et al. 2011). The GRADE Working Group does, to be sure, offer examples of surrogates that might not result in lowering the quality of evidence (mainly cholesterol concentrations as a good surrogate for heart disease). Exceptions notwithstanding, the group downgrades evidence for not using a clinical endpoint.

Yet another way in which EBM efforts have attempted to solve antiarrhythmic type scandals is to insist on disclosure of financial conflicts of interest. Most major journals have editorial policies that require authors to declare any relevant

conflicts of interest. The forms can be tedious and need to be signed by all authors. At medical conferences, it is common for presenters to declare any conflicts of interest before starting their talk—this is also required by many major medical conferences. Declaring conflicts of interest has the benefit of increasing transparency. However, there are no rules that prevent people with declared conflicts of interest from publishing their papers, and indeed they often do so. To verify this, it suffices to check the financial disclosure section of most medical trials, where it is common to see a host of conflicts, including conflicts directly relevant to the paper written. A notable exception to this was the Cochrane Collaboration, who until recently forbade reviews funded by anybody with a conflict of interest associated with commercial sponsorship (Cochrane 2014). However, the Cochrane stance on finance bias has been relaxed, which has threatened to pull the organization apart (see below). Moreover, Cochrane does not have a mechanism for preventing individual trials that are produced by researchers with conflicts of interest from being included in the reviews (more on this below, as well).

Another important EBM move to improve evidence is the World Health Organization's call for all trials, including those with results that do not support industry interests to be published (Moorthy et al. 2015). This is being supported by other groups, including the AllTrials campaign (Goldacre 2015). EBM researchers recently demonstrated their ability to enforce trial publication in the updated Cochrane Review of neuraminidase inhibitors.

Act III. EBM Obtaining Hidden Data: The Oseltamivir Case

Neuraminidase inhibitors are recommended for use against influenza and related complications in interpandemic years and in a pandemic. Anticipating the H1N1 flu pandemic, governments stockpiled the drugs. The US spent more than $1.3 billion on reserves of antivirals such as oseltamivir (Tamiflu), while the British government spent almost £424 million ($703 million) stockpiling some 40 million doses of oseltamivir. But do neuraminidase inhibitors such as Tamiflu work? The EBM answer to that question should have been resolved in 2006, when a team of researchers led by Tom Jefferson published a systematic review of randomized trials of the drugs (Jefferson et al. 2006). In their review, they found that oseltamivir had a small statistically significant protective effect against influenza (relative risk 0.33, 95% confidence interval 0.18 to 0.59), and that it had a similarly small effect for treating influenza (hazard ratio 1.20, 95% confidence interval 1.06 to 1.35). In the discussion, however, Jefferson and colleagues noted missing data in some of the trial reports, and failed attempts to obtain complete data from the manufacturers.

After a long battle, Jefferson eventually obtained the missing data, in the form of clinical study reports (CSRs) from GlaxoSmithKline, Roche, and the European Medicines Agency, which reported trials of their neuraminidase inhibitors, including trials that had not previously been published. Clinical study reports are

clinical documents that record all aspects of the trial. They are very detailed and very cryptic to the uninitiated. There were over 150,000 pages of data in what might be considered a "data dump." Together with Carl Heneghan (director of the Oxford Centre for Evidence-Based Medicine), Jefferson's group expanded their team of authors to examine the study reports and produced an updated version of the review in 2014 (Jefferson et al. 2014). The arduous troll through the reports showed what we would have expected because of finance bias: the positive benefits of oseltamivir and other antivirals were smaller, and the risks of harms greater, than previous estimates that did not include the hidden data. For adults, there was no significant reduction in hospitalization or serious complications (defined as those leading to study withdrawal). The antiviral drugs had a small benefit in prophylaxis (for every 51 people who took the drug for prevention, one person who would otherwise have developed influenza did not). The drugs also induced complications such as nausea, headaches, and renal events.

The neuraminidase study was an EBM success story, and it has been rightly hailed as a landmark in the development of methods for synthesizing evidence (Clarke 2015). ~~In the end, the suspicions of publication bias were confirmed: in the previously published trials, the benefits of the drugs had been exaggerated while the harms had been understated.~~ A hope is that this trial will make governments weigh proper evidence more carefully before spending billions of dollars (or pounds) on interventions.

Summary of the Play: An Illusion of Rationality Winning

On the face of it, it seems that the three anecdotes described above provide evidence for the progress of EBM and rationality. Thalidomide led to the need for better trials, the antiarrhythmic drugs fueled the entire EBM movement, and the unpublished trials of neuraminidase inhibitors were eventually examined. This is true, and the accounts I have provided here show it. There were, however, flies in the ointments of all the apparent EBM successes.

THE HIDDEN FAILURES OF EVIDENCE TO CURTAIL THE POWER OF FINANCE BIAS

> *The normative emphasis on rationality leaves the modern project ignorant of how power works, and therefore open to being dominated by power. Relying on rationality therefore risks exacerbating the very problems modernity attempts to solve.*
>
> **—Bent Flyvbjerg (1998)**

I shall now argue that the focus on improved methods of reporting threatens to be toothless unless parallel efforts are made to redress the imbalance between the power of finance bias and the interests of evidence. Even worse, methodological improvements in evidence production may have had the paradoxical effect of exacerbating finance bias.

The Hidden Failure of the Kefauver-Harris Amendments

The antiarrhythmic case shows that the Kefauver-Harris Amendments have not been implemented properly, because clinical endpoints were not used. This practice continues: a recent study of cancer trials showed that about half of new cancer drugs are approved without demonstrating benefits on any clinical outcome (Davis et al. 2017). In some other areas, the Amendments are ignored altogether. For example, medical devices are approved without any trials at all (Cohen and Billingsley 2011). The Kefauver-Harris Amendments were certainly an improvement that we should all be grateful for, but the extent to which they improved the intervention evaluation process cannot be affirmed without additional systematic studies. Anecdotal evidence presented here suggests that the improvements may have been small.

The Hidden Failure of the Antiarrhythmic Drug Story

EBM efforts to avert antiarrhythmic-type disasters have mostly focused on trying to improve trial quality and trial reporting. Unfortunately, ~~things don't always happen this way. T~~he improved quality of reporting has not yet had any measurable benefit in terms of reducing finance bias. Similarly, the insistence on disclosing financial conflicts of interest does not prevent such interests from influencing evidence. Moreover, the requirement to disclose financial conflicts of interest is flouted. As I was drafting this paper, Dr. José Baselga, one of the world's top cancer doctors, was found to have failed to disclose millions of dollars of payments, including over $3 million from Roche (Ornstein and Thomas 2018). Baselga had been putting positive spins on trials whose sponsors had been paying him large sums of money. And Baselga's case does not appear to be unique: a third of oncologists do not report financial conflicts of interest (Wayant et al. 2018).

The focus on reporting quality, as opposed to trial quality, also allows the powers of finance bias to get off relatively lightly. There is no restriction placed on those with conflicts of interest when it comes to producing evidence: there is only a request that those conflicts be declared.

More worryingly, an unintended consequence of better reporting is that it may have exacerbated the influence of finance bias by making research much more expensive. Whereas before the era of improved trial conduct and reporting single authors could produce randomized trials with no external funding, phase III trials now cost millions of pounds or dollars and take teams of people years to produce. Systematic reviews have also increased in complexity, as have the resources required to produce them. The Cochrane handbook began as a document of less than 100 pages; it is now thousands of pages long. A systematic review also requires teams of professional reviewers an average of 23 months to complete.

The higher costs make it relatively easy for (rich) companies and relatively difficult for more independent university investigators to do proper trials. As a

result, universities are becoming increasingly dependent on corporate financing. This blurs the line between industry funded trials and (supposedly independent) trials conducted within academic institutions (See Biddle 2007 for a comprehensive account of the trends towards increased industry entrenchment within academia and elsewhere.) And we don't know how serious the problem of finance bias influencing academia is, because there is little transparency in this area. While financial transactions between medical practitioners and industry are somewhat regulated financial transactions between academia and industry are hardly regulated at all (ABPI 2018; CMS 2018; IOM 2009). Thus, although EBM efforts have undoubtedly helped improve the quality of evidence reporting and conduct, this has led to increased costs of research, making academia more susceptible to finance bias.[2]

Still worse, because the powers of finance bias are not addressed directly by purely methodological innovations, it is not clear whether insisting on clinical endpoints would have prevented the antiarrhythmic disaster, let alone whether it would avoid such a disaster in the future. Recall that the director of the FDA (Crout) asserted that although clinical endpoints were usually desirable, surrogate outcomes sufficed in some cases, and that antiarrhythmic drug cases represented such a case. Had GRADE been used at the time, one can imagine that Crout might have pointed to the fact that GRADE allows exceptions and made the same arguments. Establishing counterfactual claims about past events is problematic, but the fact that GRADE ~~promotes clinical endpoints in most~~ cases does not imply that their methodological rules would have prevented the antiarrhythmic drug disaster. And as was pointed out above, many drugs ~~are~~ approved without clinical endpoints.

The Hidden Failure of the Neuraminidase Story

The neuraminidase story is a story of EBM winning a battle in the war against finance bias to overcome publication bias. However, the relevance of the story to the war against finance bias is unclear, for at least three reasons. First, the victory was hollow. The antiviral drugs had already made the companies millions, and in 2016 the patents began to expire, making the issue relatively irrelevant as far as profits were concerned.

Second, the labor-intensity of the review makes it difficult to replicate. Even setting aside the long battle to obtain unpublished trials (which has no guarantee of success ~~in the future~~), the review itself took a large team of people a long time to produce. I was part of the team, and without strong leadership motivating us to do long hours of work, including on weekends, it would not have been completed. Because of this labor-intensity, such methods are not feasible as a general

[2] Gøtzsche's expulsion was reminiscent of the University of Toronto rescinding David Healy's offer of employment after he gave a talk about unsuspected harms of psychotropic drugs (Spurgeon 2002).

method for completing systematic reviews. To wit, the method used for this "super-review" has not been replicated since it was published in 2014, and future replications seem destined to be rare at best. Instead of doing more super-reviews, efforts are now focused on trying to enforce mandatory publication of trials so that hunting for unpublished trials is not required in the future. The AllTrials campaign is enjoying some success in achieving this, and their early successes provide reasons for optimism about the future. But as of the present, the problem of publication bias ~~has not been~~ solved.

Third, even the eventual super-review did not provide decision-makers with a complete picture of what happened in the trials. This is because even the clinical study reports were incomplete. Many of them had large swathes of text redacted, and there was evidence of selective reporting. Beyond that, none of the studies properly defined pneumonia, which led to a paradox. Oseltamivir significantly reduced self-reported, investigator-mediated, unverified pneumonia, but no oseltamivir treatment studies reported beneficial effects on radiologically confirmed pneumonia. Then, there was evidence that the placebos contained active ingredients (intended to mimic the adverse effects of the drugs). This is useful to make the placebo and treatment more similar and thus preserve the blind. However, it also rendered the measurement of drug-induced adverse events (defined as the difference between adverse event rates in the treatment group and the adverse event rates in the placebo group) biased to minimize the suspected adverse effects. So even this super-review was limited by known (but hard to rule out) biases.

Moreover, even if the super-review ~~had been complete~~, it would still not be enough to counteract the forces that caused the publication bias in the first place. Moving from evidence to action requires that we move from an *is* to an *ought* (Hume 1738–40). EBM, at best, can inform what *is*, but not what *ought* to be, done. Even if EBM efforts were completely successful and 'perfect' evidence were produced (whatever that is), finance bias can influence what ought to be done about it. Deciding what ought to be done requires that we engage in deliberations that include extra-evidential factors, such as personal, collective, and economic values. In a democracy, these value judgments are deliberated in conferences, much like the consensus conferences in the pre-EBM era. As such, they fall prey to all the influences of financial bias loathed by EBM proponents and illustrated in the antiarrhythmic drug case described above. ~~Failure of EBM to engage more clearly with the forces that shape these discussions greatly diminishes the force of evidence. Indeed this is just what happened in the neuraminidase case.~~ Following the publication of the updated review that contained previously unpublished trials, industry-funded authors rushed to publish responses in high-impact journals such as *The Lancet,* defending the effects of the drugs (Muthuri et al. 2014), and members of parliament in the UK agreed that they would stockpile the drugs again if they had a chance (O'Dowd 2014). Of course, politicians might vote to stockpile the drugs even if they believed that the drugs

were not very effective, due to the precautionary principle. However, we see here that addressing the problems with evidence alone does not suffice.

I shall illustrate this point further with a personal anecdote (names withheld and some details changed to preserve anonymity). I was at a medical conference a few months ago where I met three researchers who work for one of the UK institutions charged with checking the evidence of effectiveness and cost-effectiveness. They presented the results of their detailed investigations to the National Institute for Health and Care Excellence (NICE). If my colleagues deem that the new medical technology is cost-effective, NICE gives a "thumbs up." On the other hand, if they do not, they give a "thumbs down," and the new technology is not approved to be provided by the National Health Service. This seems relatively straightforward, and, in the context of scarce resources, fair.

The problem is that the drug companies don't take thumbs down lightly, and often appeal. The appeals are painful, and often accompanied by press campaigns where (often: industry-funded) patient groups lash out at NICE for not paying for treatments that would help them. To their credit, NICE often upholds their decisions in spite of this pressure. At the same time, there are many examples of what appear to be NICE U-turns in the face of pressure. To name just a few: in 2018, NICE reversed their initial view on the cost-effectiveness of tocilizumab (RoActemra), a drug to treat giant cell arthritis (McKee 2018); in 2017, they reversed their initial assessment that brentuximab vedotin (Adcetris) was not a cost-effective treatment for Hodgkin's lymphoma (McKee 2017b); and in the same year, they reversed their initial assessment that trastuzumab emtansine (Kadcyla) was cost-effective for treating advanced breast cancer (McKee 2017a). In some cases these apparent U-turns might be for evidential reasons (the drug company might provide more evidence that establishes its cost-effectiveness more clearly). It is also possible that they are caving in to pressure from industry. Stories of industry twisting regulators' arms are, after all, not uncommon. In the US, it happened famously with the approval of flibanserin ("female Viagra) and rofecoxib (Vioxx), with the latter leading to an estimated tens of thousands of deaths (Biddle 2007; Cronin and Stone 2005; Holman and Geislar 2018). The problems with financial bias influencing guideline development have been described in detail elsewhere (George, Vesely, and Woolf 2014; Graham, Alderson, and Stokes 2015).

Another Act: Peter Gøtzsche Excommunicated from Cochrane

A few weeks before I submitted ~~this~~ paper, something major happened in Cochrane that supports Flyvbjerg's observation that in an open confrontation, power beats rationality. I mentioned above that Peter Gøtzsche is one of the few people in the EBM community who confronts the powers of finance bias directly. I also cited his empirical studies of the influence of industry bias. He was a founding member of Cochrane, and in 2017 was elected to its governing board.

He chose to stand for election because of what he reported to be Cochrane's increasing coziness with industry. He was also upset that Cochrane leadership was shying away from taking stances that might upset industry. In a letter initially published on the Nordic Cochrane website but subsequently removed (it was preserved elsewhere), he stated:

> Cochrane executive leadership has even refused to comment publicly on new health technology policies, open access policies and other key advocacy opportunities despite the fact that an auditing of Cochrane fulfilment of objectives has shown a total failure to comply with Cochrane advocacy objectives. (Gøtzsche 2018)

In short, Gøtzsche was recommending that Cochrane should take steps to redress the asymmetry between the power of finance bias and evidence. Gøtzsche's attack on Cochrane heated up shortly after Cochrane published a systematic review of randomized trials of human papillomavirus (HPV) vaccines for preventing cervical cancer in women (Arbyn et al. 2018). Most of the Cochrane authors on the published protocol for the review had conflicts of interest related to the HPV vaccine manufacturers, ~~and most~~ of the authors of the eventual review ~~also had conflicts of interest~~.

Together with Lars Jørgensen and Tom Jefferson, Gøtzsche published a critique of the review, claiming that the benefits of the vaccine had been exaggerated, and the harms understated (Jorgensen, Gøtzsche, and Jefferson 2018). The main evidence for their arguments was that data from some 20 studies and CSRs had not been included in the review. They had informed Cochrane editors about their concern about missing studies before the review was published, and Gøtzsche also sent a letter to the European Medicine's Agency (EMA), outlining his views about the review.

The Cochrane editors published a reply to Jefferson and Gøtzsche's critique (Tovey and Soares-Weiser 2018), which Jefferson and Gøtzsche subsequently claimed, ignored their main point, namely that harms had been inadequately reported. The editors admitted that some data had been omitted but claimed, "addition of these data makes little or no difference to the results of the review for the main outcome." Even if it were true that the omission of data did not affect the results (which could only be confirmed if it were actually analyzed within the review), this response has questionable relevance. The fact that the editors allowed a ~~faulty review, and where the faults appear to arise, at least in part, due to author financial conflicts of interest, to be published in the first place is the important point~~.

Shortly after this war of words, in September 2018, at Cochrane's 25th anniversary annual general meeting, Gøtzsche was expelled from the board. Much remains to be learned of the details, but views converge on the fact that Gøtzsche's vociferous objection to finance bias was a key factor in the dispute.[3]

As with most dramas that have human characters, there is more to this one than meets the eye. Gøtzsche can be abrasive, and his critique of the HPV view emboldened some anti-vaxx groups, which may have an unintended harm. However, it is important to avoid the red herring accusation that Gøtzsche and Jefferson are "anti-vaxxers." Neither Jefferson nor Gøtzsche was against the vaccine. Instead, they were critical of the evidence. In fact, their willingness to contribute to the vaccine debate hornets' nest is further evidence that they were unafraid to take on power directly. And in their defense, they state that the public needs accurate information about intervention efficacy and safety in order to make informed choices. One might equally blame the Cochrane editors for further inciting the vaccine debate: by allowing a review that didn't contain data that it should have, the editors fueled the conspiracy that the "system" is hiding things.

It remains to be seen whether Gøtzsche's expulsion will energize those within the EBM community to empower evidence, or whether Cochrane's actions will make those who wish to confront finance bias scared to stand up.

What Philosophers Have to Say—and What They Don't

In the introduction I mentioned several philosophical colleagues who have addressed the issue of finance bias and how it corrupts evidence. I hope to have added to their work by suggesting that the power asymmetry between finance bias and evidence may require a solution that addresses power more directly. At the very least, methodological or epistemic solutions need to be empowered so that they are implemented rigorously.

Many ~~other~~ philosophers of science have criticized EBM by proposing that evidence about mechanisms should be incorporated into the decision-making process (Clark et al. 2014; Dragulinescu 2012; Glennan 2017; Russo and Williamson 2007). Some proponents of this view have started a group called "EBM+" (2018), whose aim is to "improve the ways in which evidence-based medicine handles evidence of mechanisms."[4] Their argument infers *from* alleged epistemic problems with EBM *to* a claim that evidence of mechanisms will help solve these problems (Parkkinen et al. 2018). I have commented on their arguments elsewhere (Howick 2011a; Howick, Glasziou, and Aronson 2013). For the purposes of this paper, it is only important to note that the EBM+ proposal is entirely silent on the issue of financial conflicts of interest. To wit, their 131-page textbook does not mention the terms *conflict* or *financial* at all, and does not discuss finance bias. By ignoring the problem, they cannot possibly solve it. Worse, if their proposals are taken up, they might make the situation worse by further

[3] I believe that any attempt to include relevant evidence is worthwhile. At the same time, it is fair to note that their methods for adding evidence of mechanisms to EBM have not been tested. Their textbook Evaluating Evidence of Mechanisms in Medicine Principles and Procedures devotes less than two pages to the only "worked example" of "probiotics and dental caries."

[4]

empowering finance bias. In fact, they explicitly note that their proposal might help get drugs to market more quickly. Their handbook states: "By considering evidence of mechanisms in conjunction with clinical study evidence, decisions can be made earlier: one can reduce the time taken for a drug to reach market" (Parkkinen et al. 2018, 15). Importantly, by "clinical study evidence," they do not mean randomized trials or systematic reviews of randomized trials: in their view, observational studies alongside evidence of mechanisms suffice. Such a position would certainly have supported the view that antiarrhythmic drugs could be approved based on a surrogate endpoint, and as such, their proposals are certainly music to the ears of industry, in whose interest it is to circumvent the EBM standards that slow down the process of getting new treatments approved.

In sum, it is fair to say that many prominent philosophical proponents and critics of EBM are united in their focus on methodological solutions to problems within EBM. If the thesis in this paper is correct, then attention to methodological points, while somewhat useful, is not likely to prevent finance bias from influencing evidence. More attention needs to be placed on addressing the structures that will empower evidence.

Objections

One possible objection to my argument is that if EBM methods were policed to perfection, then finance bias would not be a problem. ~~There are two answers to this problem, both of which support my thesis that the asymmetry between evidence and finance bias must be redressed by moving beyond purely methodological solutions. The first answer is that~~ even if EBM methods sufficed to rule out the influence of finance bias (which I have shown that in at least some important cases, they do not), evidence cannot be implemented without being empowered. Hume's is/ought distinction, mentioned above, that arises when we move from evidence to guidelines, shows ~~how this happens in the guideline production process~~.

~~Another example of the lack of teeth in purely methodological attempts to confront finance bias is the EBM movement's failure to demonstrate that it has improved human health, on a population level.~~ EBM proponents were initially confident that evidence of the average benefits of EBM would emerge over time. The first edition of the EBM textbook (1997) did not mention the need for rigorous evaluation of EBM at all, presumably assuming that such a benefit was obvious. The second edition of the same textbook (2000) acknowledged the need for rigorous evaluation of EBM and claimed that the lack of randomized trials testing its patient benefits were limited by pragmatic and ethical concerns. They also acknowledged that "evidence that EBM "works" has been late and slow to come" (7). The most recent edition (2010) gave up the idea of evaluating EBM altogether, claiming that we should move our focus to the most efficient

ways to teach EBM. To support this position, EBM proponents cite *individual* trials (such as the antiarrhythmic drug trial). This response, however, seems unsatisfactory. If EBM were helpful, one might reasonably expect to see some improvement in human health resulting from its adoption. However, during EBM's reign between 1995 and 2015, life expectancy has carried on rising more or less at the same rate it was before the advent of EBM. In fact, starting in 2014, life expectancy in the United Kingdom has tapered off, while in the US it is falling (WHO 2018). The change in life expectancy is caused by many factors, so it would be mistaken to infer anything about EBM directly based on this crude statistic. Meanwhile, other medical revolutions such as clean water and washing hands produced measurable population benefits.

Another objection is the argument that EBM proponents are best placed to stick to what they are good at: improving methods. According to this view, revolutionary efforts to confront the powers behind finance bias are best left to revolutionaries. This is true: academics in general are not good politicians, and they have less money behind them than industry (unless industry funds them). However, it is also true, if my thesis here holds, that *unless* academic proponents of EBM become more like Peter Gøtzsche, in the sense that they are willing to address the problem with finance bias more systematically, efforts to improve methods may yield little fruit or even have unintended harms. It is also worth noting that some of the initial proponents of EBM *were* revolutionary. Iain Chalmers has never been able to find a permanent home in a university (in part, perhaps, because he is not comfortable within the confines of academia); Dave Sackett was routinely attacked by senior professors as he attempted to embed EBM within the Oxford curriculum (Heneghan 2015); and Ben Goldacre calls out anyone who doesn't publish trials (Goldacre 2015). The future of EBM might depend on its proponents being more revolutionary, not less.

The True Remedy for EBM Must Address the Real Causes of Its Problems

I've argued here that a main cause of the current problems with evidence seems to be, at least in the cases I have presented, an asymmetry of the relationship between rationality (evidence) and power (financial bias), with financial bias being by far the stronger, and strong enough to beat evidence. It follows that redressing this imbalance must be part of the solution to the problems with evidence. According to Flyvbjerg, this must involve changing the structures in which finance bias and the interests of evidence operate. It is beyond the scope of this essay to provide a detailed map for such a solution; however, my diagnosis of the problem does permit me to provide a blueprint for what kinds of solutions are required.

One easy way to engage directly would be to downgrade evidence if researchers with conflicts of interest produced it. I once suggested this at a GRADE

meeting, where my suggestion was not taken seriously, for the alleged reason that focusing on more purely methodological markers of trial quality would suffice. The anecdotal evidence presented here suggests that GRADE's reasons for rejecting my proposal were mistaken. Purely methodological solutions to finance bias cannot be effective because they don't address the powers behind finance bias.

Until recently, Cochrane insisted that independent researchers produce reviews, and the "EBM Manifesto" called for independent evaluation of evidence (Heneghan et al. 2017). These efforts are part of the solution, but they need to be followed up with more direct efforts to redress the asymmetry between the power of finance bias and the interest of industry. Moreover, as I described in relation to the Gøtzsche drama, the separation between Cochrane and the powers of finance bias has become blurred. In order to tackle the problem with financial bias we must promote a system where the production and promotion of evidence is explicitly empowered within stable structures.

An obvious way to redress the imbalance between the power of finance bias and the interests of evidence is independent evaluation. Ioannidis (2016) puts the reason for this succinctly: "corporations should not be asked to practically perform the assessments of their own products. If they are forced to do this, I cannot blame them, if they buy the best advertisement (i.e., 'evidence') for whatever they sell" (84). As Gøtzsche suggested, the need for independent evaluation applies across the board to anyone with a conflict of interest, not just those with industry conflicts of interest.

Another way to stabilize relationships between evidence and financial bias could be to take finances out of the equation altogether, for example by rejecting capitalism. However, even in the unlikely case that such a solution were deemed desirable and could be implemented, one needn't go that far: aligning the interests of industry with the interests of patients would make a difference. Some insurance companies and patients are beginning to facilitate this. It is in the financial interest of insurance companies to collect insurance premiums and not pay money to treat patients~~, and at~~ least in principle, this means it is in the interest of insurance companies for people to be healthy. Again just in principle, they should therefore support evidence that benefits patients in the most cost-effective way possible. ~~There are other forces at play as well. Since insurance companies make most of their money investing premiums, it is not obviously in their interest to waste time on the expensive business of making people eat less and exercise more. Still, it is *an* interest of insurance companies to have healthy payers, so making investments that promote health and profits makes sense.~~

Taking this insight into account, an insurance company called Vitality Health is starting to incentivize patients to have healthier lifestyles. They offer clients an Apple Watch for £29 (about $35US, or about a tenth of the full retail price), and then they offer "vitality points" if customers reach certain activity goals. For

example, walking 10,000 steps per day earns customers five points. These points can then be redeemed at Starbucks or Cineworld. To be sure, these rewards are problematic: customers might buy a processed sugary snack at Starbucks, and watching a movie hardly promotes physical activity. Still, the idea is right: the insurance company will save money if customers have healthier lifestyles, so they are thinking of interventions that are likely to improve patient health while making them more money.

A third way to align the interests of profit with health is social impact bonds. Social impact bonds (also called "pay for success bond") are becoming increasingly popular. They work when financiers, including industry, pay money up front for a new intervention that they believe will lead to better health. When and if savings as a result of better health emerge, the financiers get their money back plus a pre-agreed profit. For example, Social Finance US, a nonprofit organization focused on financing social impact bonds, launched a pilot project aimed at better management of asthma in Fresno, California. They will measure the outcomes carefully, and if the intended benefits and cost-savings materialize, they will get a return on their investment from the service providers.

These models, which align the interests of industry with those of patients, are in their infancy. However, they show that it is possible to resolve the power asymmetry between evidence and finance bias, by aligning their interests.

Conclusion

EBM practitioners have spearheaded some important improvements in the evidence base of medical treatments and devices. They have certainly led to standardized reporting of clinical studies. However, finance bias has too often trumped EBM attempts to improve the production, reporting, and use of evidence. Solutions proposed by critics and sympathizers of EBM alike have largely focused on more standards and different types of evidence. These solutions have some limited benefits, and also unintended harms. More importantly, they are limited *in principle,* because they leave a major sociological source of bias—finance bias—all but untouched. Effective solutions require addressing the asymmetry between the power of finance bias and the interests of good evidence. These solutions should include independent evaluation of medical interventions at a minimum. Other solutions, in the form of social impact bonds and creative insurance company strategies, are beginning to emerge.

References

Altman, D. G., et al. 2001. "The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration." *Ann Intern Med* 134 (8): 663–94.

Arbyn, M., et al. 2018. "Prophylactic Vaccination Against Human Papillomaviruses to Prevent Cervical Cancer and Its Precursors." *Cochrane Datab Syst Rev* 5: CD009069. DOI: 10.1002/14651858.CD009069.pub3.

Aronson, J. K. 2018. "A Word About Evidence: 5. Bias—previous definitions." In *Catalogue of Bias*. Oxford: University of Oxford.

Association of the British Pharmaceutical Industry (ABPI). 2018. *Disclosure UK. http://www.abpi.org.uk/ethics/ethical-responsibility/disclosure-uk/*.

Bero, L., et al. 2007. "Factors Associated with Findings of Published Trials of Drug-Drug Comparisons: Why Some Statins Appear More Efficacious Than Others." *PLoS Med* 4 (6): e184. DOI: 10.1371/journal.pmed.0040184.

Biddle, J. 2007. "Lessons from the Vioxx Debacle: What the Privatization of Science Can Teach Us About Social Epistemology." *Soc Epistemol* 21 (1): 21–39.

British Medical Journal (BMJ). 2014. "Medical Journals and Industry Ties." *BMJ.* DOI: 2014;349:g7197

Brown, J. R. 2008. "Politics, Method, and Medical Research." *Philos Sci* 75 (5): 756–66.

CAST Investigators. 1989. "The Cardiac Arrhythmia Suppression Trial (CAST)." *N Engl J Med* 321: 406–12.

CAST II Investigators. 1992. "Preliminary Report: Effect of Encainide and Flecainide Onmortality in a Randomized Trial of Arrhythmia Suppression After Myocardial Infarction." *N Engl J Med* 327: 227–33.

Centers for Medicare & Medicaid Services (CMS). 2018. "Open Payments." CMS.gov. *http://www.cms.gov/openpayments/*.

Clarke, B., et all.. 2014. "Mechanisms and the Evidence Hierarchy." *Topoi* 33 (2): 339–60.

Clarke, M. 2015. "History of Evidence Synthesis to Assess Treatment Effects: Personal Reflections on Something That Is Very Much Alive." *JLL Bull. http://www.jameslindlibrary.org/articles/history-of-evidence-synthesis-to-assess-treatment-effects-personal-reflections-on-something-that-is-very-much-alive/*.

Cochrane. 2014. "Conflicts of Interest and Cochrane Reviews." Cochrane Community, May.https://community.cochrane.org/editorial-and-publishing-policy-resource/ethical-considerations/conflicts-interest-and-cochrane-reviews.

Cohen, D., and M. Billingsley. 2011. "Europeans Are Left to Their Own Devices." *BMJ* 342: d2748. DOI: 10.1136/bmj.d2748.

CONSORT. 2018. "How CONSORT Began." *http://www.consort-statement.org/about-consort/history*.

Cronin, J., and A. Stone. 2005. "Conflicts of Interest on COX-2 Panel." Center for Science in the Public Interest. *http://www.cspinet.org/new/200502251.html.*

Davis, C., et al. 2017. "Availability of Evidence of Benefits on Overall Survival and Quality of Life of Cancer Drugs Approved by European Medicines Agency: Retrospective Cohort Study of drug Approvals 2009–13." *BMJ.* DOI: 10.1136/bmj.j4530.

De Angelis, C., et al. 2004. "Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors." *N Engl J Med* 351 (12): 1250–51.

Distillers Company Limited. 1961. "Distaval." *BMJ* (5242).

Dragulinescu, S. 2012. "On 'Stabilising' Medical Mechanisms, Truth-Makers and Epistemic Causality: A Critique to Williamson and Russo's Approach." *Synthese* 187 (2): 785–800. DOI: 10.1007/s11229-011-0011-9.

EBM+. 2018. EBMplus.org.*http://ebmplus.org/*.

Equator Network. 2018. *Developing a Reporting Guideline. http://www.equator-network.org/toolkits/developing-a-reporting-guideline/*.

Evans, H. 2014. "Thalidomide: How Men Who Blighted Lives of Thousands Evaded Justice." *Guardian*, 14 Nov.

Florence, A. L. 1960. "Is Thalidomide to Blame?" *BMJ* 2 (5217): 1954.

Flores, L. E., et al. 2014. "Early Detection and Staging of Spontaneous Embryo Resorption by Ultrasound Biomicroscopy in Murine Pregnancy." *Reprod Biol Endocrinol* 12:38. DOI: 10.1186/1477-7827-12-38.

Flyvbjerg, B. 1998. *Rationality and Power: Democracy in Practice*. Chicago: University of Chicago Press.

Fullerton, P. M., and M. Kremer. 1961. "Neuropathy After Intake of Thalidomide (Distaval)." *BMJ* 2 (5256): 855–58.

George, J. N., S. K. Vesely, and S. H. Woolf. 2014. "Conflicts of Interest and Clinical Recommendations: Comparison of Two Concurrent Clinical Practice Guidelines for Primary Immune Thrombocytopenia Developed by Different Methods." *Am J Med Qual* 29 (1): 53–60. DOI: 10.1177/1062860613481618.

Glennan, S. 2017. *The New Mechanical Philosophy*. Oxford: Oxford University Press.

Goldacre, B. 2015. "How to Get All Trials Reported: Audit, Better Data, and Individual Accountability." *PLoS Med* 12 (4): e1001821. DOI: 10.1371/journal.pmed.1001821.

Gøtzsche, P. 1989. "Methodology and Overt and Hidden Bias in Reports of 196 Double-Blind Trials of Nonsteroidal Antiinflammatory Drugs In Rheumatoid Arthritis." *Control Clin Trials* 10 (1): 31–56. DOI: 0197-2456(89)90017-2 [pii].

Gøtzsche, P. 1990. "Bias in Double-Blind Trials [Thesis]." *Dan Med Bull* 37: 329–36.

Gøtzsche, P. 1996. "Blinding During Data Analysis and Writing of Manuscripts." *Control Clin Trials* 17 (4): 285–90; discussion 290–93.

Gøtzsche, P. C. 2014. *Deadly Medicines and Organised Crime: How Big Pharma Has Corrupted Healthcare*. London: Radcliffe Publishing.

Gøtzsche, P. 2018. "A Letter from Peter C Gøtzsche Following His Expulsion from the Cochrane Collaboration & The Chaos Within the Group." *Age of Autism*. https://www.ageofautism.com/2018/09/a-letter-from-peter-c-g%C3%B8tzsche-following-his-expulsion-from-nordic-cochrane-centre-the-chaos-within-.html.

Graham, T., P. Alderson, and T. Stokes. 2015. "Managing Conflicts of Interest in the UK National Institute for Health and Care Excellence (NICE) Clinical Guidelines Programme: Qualitative Study." *PLoS One* 10 (3): e0122313. DOI: 10.1371/journal.pone.0122313.

Greek, R., N. Shanks, and M. J Rice. 2011. "The History and Implications of Testing Thalidomide on Animals." *J Philos Sci Law* 11 (3): 1–32.

Greenhalgh, T., et al. 2014. "Evidence Based Medicine: A Movement in Crisis?" *BMJ* 348: g3725. DOI: 10.1136/bmj.g3725.

Guyatt, G. H., et al. 2011. "GRADE Guidelines: 8. Rating the Quality of Evidence—Indirectness." *J Clin Epidemiol* 64 (12): 1303–10. DOI: 10.1016/j.jclinepi.2011.04.014.

Heneghan, C. 2015. "Reflections on David Sackett's time at the Centre for Evidence-Based Medicine." CEBM, June 1. *http://www.cebm.net/2015/06/david-sackett-and-the-centre-for-evidence-based-medicine/*.

Heneghan, C., et al. "Evidence Based Medicine Manifesto for Better Healthcare." *BMJ* 357: j2973. DOI: 10.1136/bmj.j2973.

Heres, S., et al. 2006. "Why Olanzapine Beats Risperidone, Risperidone Beats Quetiapine, and quetiapine Beats Olanzapine: An Exploratory Analysis of Head-To-Head

Comparison Studies of Second-Generation Antipsychotics." *Am J Psychiatry* 163 (2): 185–94. DOI: 10.1176/appi.ajp.163.2.185.

Holman, B. 2017. "Philosophers on Drugs." *Synthese*: 1–28.

Holman, Bennett, and Sally Geislar. 2018. "Sex Drugs and Corporate Ventriloquism: How to Evaluate Science Policies Intended to Manage Industry Bias." *Philosophy of Science*Accepted February 17, 2018.

Hopewell, S., et al. 2009. "Publication Bias in Clinical Trials Due to Statistical Significance or Direction of Trial Results." *Cochrane Database Syst Rev* (1): MR000006. DOI: 10.1002/14651858.MR000006.pub3.

Howick, J. 2011a. "Exposing the Vanities—and a Qualified Defence—of Mechanistic Evidence in Clinical Decision-Making." *Philos Sci* (5): 926–40.

Howick, J. 2011b. *The Philosophy of Evidence-Based Medicine*. Oxford: Wiley-Blackwell.

Howick, J., P. Glasziou, and J. K. Aronson. 2013. "Can Understanding Mechanisms Solve the Problem of Extrapolating from Study to Target Populations (the Problem of 'External Validity')?" *J R Soc Med* 106 (3): 81–86. DOI: 10.1177/0141076813476498.

Hume, D.. 1738–40. *A Treatise of Human Nature*. Oxford: Oxford University Press, 2000.

Institute of Medicine (IOM). 2009. *Conflict of Interest in Medical Research, Education, and Practice*. Washington, DC: National Academies Press.

Ioannidis, J. P. 2016. "Evidence-Based Medicine Has Been Hijacked: A Report to David Sackett." *J Clin Epidemiol* 73: 82–86. DOI: 10.1016/j.jclinepi.2016.02.012.

Jefferson, T., et al. 2006. "Neuraminidase Inhibitors for Preventing and Treating Influenza in Healthy Adults." *Cochrane Database Syst Rev* 3: CD001265. DOI: 10.1002/14651858.CD001265.pub2.

Jefferson, T., et al. 2014. "Neuraminidase Inhibitors for Preventing and Treating Influenza in Healthy Adults and Children." *Cochrane Datab Syst Rev* 4: CD008965. DOI: 10.1002/14651858.CD008965.pub4.

Jefferson, T., et al. 2014. "Risk of Bias in Industry-Funded Oseltamivir Trials: Comparison of Core Reports versus Full Clinical Study Reports." *BMJ Open* 4 (9): e005253. DOI: 10.1136/bmjopen-2014-005253.

Jorgensen, A. W., J. Hilden, and P. C. Gøtzsche. 2006. "Cochrane Reviews Compared with Industry Supported Meta-Analyses and Other Meta-Analyses of the Same Drugs: Systematic Review." *BMJ* 333 (7572): 782. DOI: 10.1136/bmj.38973.444699.0B.

Jorgensen, L., P. C. Gøtzsche, and T. Jefferson. 2018. "The Cochrane HPV Vaccine Review Was Incomplete and Ignored Important Evidence of Bias." *BMJ Evid Based Med*. DOI: 10.1136/bmjebm-2018-111012.

Leopold, S. S., et al. 2003. "Association Between Funding Source and Study Outcome in Orthopaedic Research." *Clin Orthop Relat Res* (415): 293–301.

Lexchin, J., et al. 2003. "Pharmaceutical Industry Sponsorship and Research Outcome and Quality: Systematic Review." *BMJ* 326 (7400): 1167–70. DOI: 10.1136/bmj.326.7400.1167.

McBride, W. 1961. "Thalidomide and Congential Abnormalities." *Lancet* 278 (7216): 1358.

McKee, S. 2017a. "Major U-Turn Sees Roche's Kadcyla Backed for NHS Use." *PharmaTimes Online*. http://www.pharmatimes.com/news/major_u-turn_sees_roches_kadcyla_backed_for_nhs_use_1195881.

McKee, S. 2017b. "NICE U-Turn Recommends Some Access to Takeda's Adcetris." *PharmaTimes Online*. http://www.pharmatimes.com/news/nice_u-turn_recommends_some_access_to_takedas_adcetris_1192062.

McKee, S. 2018. "NICE U-Turn on Treatment for Giant Cell Arteritis." *PharmaTimes Online*. http://www.pharmatimes.com/news/nice_u-turn_on_treatment_for_giant_cell_arteritis_1226320.

Moher, D., A. Jones, and L. Lepage. 2001. "Use of the CONSORT Statement and Quality of Reports of Randomized Trials: A Comparative Before-and-After Evaluation." *JAMA* 285 (15): 1992–95.

Moore, T. J. 1995. Deadly Medicine: Why Tens of Thousands of Heart Patients Died in America's Worst Drug Disaster. New York: Simon & Schuster.

Moorthy, V. S., et al.. 2015. "Rationale for WHO's New Position Calling for Prompt Reporting and Public Disclosure of Interventional Clinical Trial Results." *PLoS Med* 12 (4): e1001819. DOI: 10.1371/journal.pmed.1001819.

Morganroth, J., et al. 1981. The Evaluation of New Antiarrhythmic Drugs: Proceedings of the Symposium on How to Evaluate a New Antiarrhythmic Drug: The Evaluation of New Antiarrhythmic Agents for the Treatment of Ventricular Arrhythmias, held at Philadelphia, Pennsylvania, October 8–9, 1980. The Hague: Martinus Nijhoff.

Muthuri, S. G., et al. 2014. "Effectiveness of Neuraminidase Inhibitors in Reducing Mortality in Patients Admitted to Hospital with Influenza A H1N1pdm09 Virus Infection: A Meta-Analysis of Individual Participant Data." *Lancet Respir Med* 2 (5): 395–404. DOI: 10.1016/S2213-2600(14)70041-4.

O'Dowd, A. 2014. "Government Says It Would Stockpile Tamiflu Again." *BMJ* 349: g6386. DOI: https://DOI.org/10.1136/bmj.g6386.

Ornstein, C., and K. Thomas. 2018. "Top Cancer Researcher Fails to Disclose Corporate Financial Ties in Major Research Journals." *NY Times*, Sept. 8.

Parkkinen, V.-P., et al. 2018. Evaluating Evidence of Mechanisms in Medicine Principles and Procedures. Cham, Switzerland: Springer Open.

Patashnik, E. M., A. S. Gerber, and C. M Dowling. 2017. *Unhealthy Politics The Battle over Evidence-Based Medicine*. Princeton: Princeton University Press.

Russo, F., and J. Williamson. 2007. "Interpreting Causality in the Health Sciences." *Int Stud Philos Sci* (2): 1157–70.

Sackett, D. L. 1997. *Evidence-Based Medicine: How to Practice and Teach EBM*. New York: Churchill Livingstone.

Sackett, David L. 2000. *Evidence-Based Medicine: How to Practice and Teach EBM*. 2nd ed. Edinburgh: Churchill Livingstone.

Savovic, J., et al. 2012. "Influence of Reported Study Design Characteristics on Intervention Effect Estimates From Randomized, Controlled Trials." *Ann Intern Med*. DOI: 10.7326/0003-4819-157-6-201209180-00537.

Schulz, K. F., D. G. Altman, and D. Moher. 2010. "CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials." *PLoS Med* 7 (3): e1000251. DOI: 10.1371/journal.pmed.1000251.

Sjöström, H., and R. Nilsson. 1972. *Thalidomide and the Power of the Drug Companies*. Harmondsworth: Penguin.

Spurgeon, D. 2002. "Psychiatrist Settles Dispute with Toronto University." *BMJ* 324 (7347): 1177.

Stelfox, H. T., et al. 1998. "Conflict of Interest in the Debate Over Calcium-Channel Antagonists." *N Engl J Med* 338 (2): 101-6.

Straus, S. E., et al. 2005. *Evidence-Based Medicine: How to Practice and Teach EBM*. 3rd ed. London: Elsevier: Churchill Livingstone.

Tovey, D., and K. Soares-Weiser. 2018. *Cochrane's Editor in Chief Responds to BMJ EBM Article Criticizing HPV Review*. Cochrane.org. Sept. 3. *http://www.cochrane.org/news/cochranes-editor-chief-responds-bmj-ebm-article-criticizing-hpv-review*.

Turner, E. H., et al. 2008. "Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy." *N Engl J Med* 358 (3): 252–60. DOI: 10.1056/NEJMsa065779.

United States Congress. 2007. Title 21 of the United States Code. Department of Health and Human Services. Washington, DC: US GPO.

Vargesson, N. 2015. "Thalidomide-Induced Teratogenesis: History and Mechanisms." *Birth Defects Res C Embryo Today* 105 (2):140–56. DOI: 10.1002/bdrc.21096.

Wayant, C., et al. 2018. "Financial Conflicts of Interest Among Oncologist Authors of Reports of Clinical Drug Trials." *JAMA Oncol*. DOI: 10.1001/jamaoncol.2018.3738.

Weidenbach, A. 1959. "Totale Phokomelie." *Zbl Gynak* 81: 2048–52.

World Health Organization (WHO). 2018. "Global Health Observatory Data: Life Expectancy." World Health Organization. *http://www.who.int/gho/mortality_burden_disease/life_tables/situation_trends/en/*.

Yaphe, J., et al. 2001. "The Association Between Funding by Commercial Interests and Study Outcome in Randomized Controlled Drug Trials." *Fam Pract* 18 (6): 565–68.