



OPEN ACCESS

Do we really know how many clinical trials are conducted ethically? Why research ethics committee review practices need to be strengthened and initial steps we could take to strengthen them

Mark Yarborough

Correspondence to

Dr Mark Yarborough, Bioethics Program, University of California Davis, Sacramento, CA 95817, USA; mayarborough@ucdavis.edu

Received 11 December 2019

Revised 19 May 2020

Accepted 22 May 2020

ABSTRACT

Research Ethics Committees (RECs) play a critical gatekeeping role in clinical trials. This role is meant to ensure that only those trials that meet certain ethical thresholds proceed through their gate. Two of these thresholds are that the potential benefits of trials are reasonable in relation to risks and that trials are capable of producing a requisite amount of social value. While one ought not expect perfect execution by RECs of their gatekeeping role, one should expect routine success in it. This article reviews a range of evidence showing that substantial numbers of ethically tainted trials are receiving REC approvals. Many of the trials are early phase trials that evidence shows have benefits that may not be reasonable compared with their risks and many others are later trials that evidence shows may lack sufficient social value. The evidence pertains to such matters as methodologically inadequate preclinical studies incapable of supporting the inferences that REC members must make about the prospects for potential benefit needed to offset the risks in early phase trials and sponsorship bias that can cause improperly designed, conducted, analysed and reported later phase trials. The analysis of the evidence makes clear that REC practices need to be strengthened if they are to adequately fulfil their gatekeeping role. The article also explores options that RECs could use in order to improve their gatekeeping function.

INTRODUCTION

Ongoing COVID-19 research currently underway reminds us why we prize biomedical research. We want improved health and healthcare for people, regardless their station in life. This noble aspiration is reflected in the mission statements of major research organisations. The American Association for the Advancement of Science mission is to ‘advance science ... and innovation throughout the world for the benefit of all people’¹ while that of the US National Institutes of Health (NIH) ‘is to seek fundamental knowledge about the nature and behaviour of living systems and the application of that knowledge to enhance health, lengthen life, and reduce illness and disability.’²

More than its humanitarian impulse provides biomedical research with its special status. It is the result as well of the dispassionate science meant to drive the design, conduct and reporting of the clinical investigations that embody it. Anthony Fauci recently reminded us of this at a recent White House briefing when he said, amid speculation

about the drug chloroquine, ‘... my job as a scientist ... is to ultimately prove without a doubt that a drug is not only safe but that it actually works...’³

Finally, the sacrifices of the people who serve as its lifeblood also contribute to biomedical research’s high status. Most often patients in clinical investigations, these individuals place themselves at risk and inconvenience to be objects of study in order to make pursuit of better healthcare possible in the first place. The bulk of these investigations, and the ones focused on here, are clinical trials of drugs and devices¹ undertaken for a range of reasons. Trials can investigate new drugs, determine whether use of already approved ones can be expanded to new indications, study new forms or doses of existing ones, and compare drugs against one another.⁴ Each trial imposes risks and inconveniences on volunteers. Among other things, all volunteers may be subject to the risks and discomforts of monitoring procedures and non-trivial opportunity costs.⁵ Participants in early trials can be subject to unknown toxicities and other side effects, while in later phase trials they may be randomised away from a drug known to be both safe and effective for their disease. Such sacrifices remind us anew why getting the science in clinical trials right is critical. Lack of fealty to the scientific principles and methods that make it possible for investigations to generate information useful to the societal aims of research means that the sacrifices of clinical volunteers may all be for naught.

The combined essential role of proper scientific standards and use of human volunteers who have critical welfare and moral interests at stake explains the safeguards, such as regulatory oversight, that are built into the clinical trials endeavour. Arguably no safeguard is more critical than Research Ethics Committees (REC). These serve as the final gatekeeper for clinical trials. They are tasked, at times uniquely so, with ensuring that trials meet ethical benchmarks articulated in foundational documents such as the World Medical Association’s Declaration of Helsinki,⁶ the Council for International Organizations of Medical Sciences International Ethical Guidelines for Health-Related Research Involving

¹Hereafter for reasons of brevity I will confine myself for the most part to the term ‘drug’ rather than ‘device,’ yet readers should note that the following discussion is as applicable to clinical trials studying devices as it is clinical trials studying drugs.



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Yarborough M. *J Med Ethics* Epub ahead of print: [please include Day Month Year]. doi:10.1136/medethics-2019-106014

Humans,⁷ and the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research's Belmont Report⁸ that collectively have created an international consensus around how clinical trials are ethically conducted.

These benchmarks include such matters as ensuring that research represents sufficient social value such that there is enough 'anticipated benefit to society in the form of knowledge to be gained from the research,'⁸ and determining that risks are 'reasonable in relation to anticipated benefits, if any, to subjects and the importance of the knowledge that may reasonably be expected to result'⁹ in the clinical trials that RECs approve. RECs are entrusted both with assessing when individual clinical trials are capable of meeting these benchmarks and when they are not and with sanctioning the former while preventing the latter.

This gatekeeping role establishes RECs as part of the bedrock of the pursuit of better health and healthcare. Thus, it behoves us all to assess how well they perform this role. While we should not expect perfect execution of this role from RECs, we should be able to expect routine success. In what follows, I review evidenceⁱⁱ that may shake our confidence that RECs are adequately preventing clinical trials that do not meet one or more of the just referenced ethics benchmarks.ⁱⁱⁱ While RECs by no means are responsible for the problems the evidence shows taints an alarming number of clinical trials, they nevertheless are responsible for striving to not let through their gate those trials that do no merit passage.

Given space constraints, the evidence discussed relating to the reasonableness of benefits in relation to risks pertains to early phase trials while the evidence relating to adequate social value pertains to later phase trials. This does not mean that risk/benefit ratios are of lesser concern in later phase trials or that social value is not an important criterion for early phase ones. Both are vital in all clinical trials. I also explore options for strengthening REC review practices so that we can be more confident in REC gatekeeping going forward.

EVIDENCE ABOUT BENEFITS AND RISKS THAT RAISES CONCERNS ABOUT CURRENT REC REVIEW AND APPROVAL PRACTICES OF EARLY PHASE TRIALS

Introductory considerations

Determining when there is a reasonable ratio between the risks and benefits of an early trial is a task that typically falls only to RECs.^{iv} This may catch some readers by surprise, given the

ⁱⁱSome of the evidence reviewed below I have referenced in prior publications, so readers may note redundancies in the descriptions of the evidence.

ⁱⁱⁱIn the comments that follow that highlight apparent inadequacies in REC review practices, it is important to note that there is no intent to criticise the individuals who serve on those committees. With more than 3 decades of service to and work with RECs, I am personally familiar with the selfless dedication REC members can bring to the REC mission of protecting the welfare, rights, and interests of research participants. Instead, the intent is to focus on the effectiveness of REC review practices themselves.

^{iv}It is important to note that one cannot make a blanket statement here about whether efficacy data are included in FDA review of IND applications. For example, in its review of IND applications for cellular and gene therapy products, the FDA recommends a preclinical studies programme that includes proof-of-concept animal studies capable of producing data about 'observable functional/behavioural' changes, as well as in vitro and/or in vivo studies assessing biodistribution of candidate cellular and gene

vetting done by regulatory agencies such as the US Food and Drug Administration (FDA) and the European Union (EU) European Medicines Agency. Phase 1 trials in either jurisdiction cannot launch unless the relevant regulatory agency issues permissions. These agencies vet the safety data produced by preclinical studies pertaining to new treatment modalities and sanction early trials only when they are satisfied that the preclinical data indicate that something is safe enough to try in humans. However, these agencies do not condition their permissions on review of efficacy data.^{10 11} They typically assess efficacy when they review licensure applications that contain supporting phase 3 data. RECs, however, in order to satisfy the ethical requirement of proportionality between risks and benefits, must consider the evidence on both sides of the risk/benefit ledger. Hence, preclinical efficacy data are just as critical as preclinical safety data in any and all REC attestations that benefits and risks ratios in early trials are reasonable.

These attestations stand or fall on the reliability of the scientific evidence that informs inferences REC members make about the potential for risks and benefits. The evidence comes from preclinical studies, the vast majority of which consists in animal studies. An advisory committee for the development of certain new therapeutics has recently reminded us how central these studies are to clinical translation efforts. They are used to both develop and test 'novel therapeutics, including small molecules, biologics, gene modifiers and cell therapies', as well as inform 'optimal dosing regime[s] and route[s].'¹² They cross a broad spectrum of investigations, ranging from 'identify[ing] new druggable targets and [helping with the] design or test[ing of] specific therapeutic modalities, as well as detect[ing] diagnostic or therapeutic biomarkers.' [Finally, they] allow the comparison of pathological phenotypes with the human disease, and [provide] insight into the underlying reasons for differences at varying levels of complexity, [thereby] allowing the identification of protective pathways in animals that could be enhanced in humans.¹²

It should be noted that many are deeply sceptical about the value of animal studies to clinical translation, given how vast the differences can be between species.¹³ Whether one shares this scepticism or not, these are the studies RECs must rely on to support inferences committee members make regarding the nature and potential of both benefits and risks in early trials. The strength of those inferences is determined by the scientific quality of the studies that generate them, since 'a well-designed experiment is a fundamental criterion for reliable information and for generating any benefit at all.'¹⁴ Let us now turn to the evidence about their quality.

Relevant evidence about preclinical research

The strength of inferences about safety and efficacy can turn on several factors of preclinical studies. To begin with, problems with the construct, internal and/or external validity of individual studies can undermine their reproducibility or generalisability. Construct validity refers to whether an experiment is capable of studying the outcome it is meant to measure.¹⁵ For example, studying a chronic human disease in an acutely injured animal can cause investigators to 'mischaracterize the relationship

therapies.⁶⁸ Biodistribution would speak at least indirectly to potential efficacy as well as safety. However, even for these novel therapies, these are non-binding recommendations and the issuance of a Phase 1 trial permission by the FDA does not entail a weighing by the FDA of risks and benefits to see if the ratio is reasonable.

between a preclinical study and an ensuing [clinical] trial.¹⁶ Internal validity refers to multiple study factors that influence what can be gleaned from the experimental population being studied.¹⁵ Factors include matters like a study's design and analysis. Problems such as inappropriate selection of control groups, lack of blinding and randomisation procedures and incorrect statistical analysis plans¹⁵ can produce 'spurious causal inferences'¹⁶ that result in erroneous research reports and irreproducible findings. External validity refers to 'experimental design features that [support] reproducibility and generalisability of the expected results' outside of the experimental setting that produced the results.¹⁵ For example, studies conducted on animals may discover things about a phenotype that get wrongly construed to a genotype instead.¹⁷ It is concerns about external validity that fuel much of the aforementioned scepticism about the value of animal studies.

Should RECs worry that there is more than the occasional preclinical study that lacks construct, internal and/or external validity? Ample evidence shows they should. It reveals substantial as opposed to occasional numbers of preclinical studies reporting unreliable findings. For instance, consider how mislabelled cell lines can pollute preclinical research findings.¹⁸ This results in preclinical studies meant, for example, to investigate one type of cancer getting conducted with cell lines from a different type of cancer instead, which can lead to an early phase trial that lacks a sound scientific basis and thus ethical risk/benefit ratio.¹⁸ Though this problem has been known about since 1950,¹⁹ it is estimated that fully a quarter of NIH sponsored research projects on cell lines may still be using misidentified or contaminated cell lines.²⁰ NIH at long last initiated efforts to prevent researchers from using mislabelled cell lines,²¹ but its efforts have limited reach, meaning that the problem still persists.

Perhaps more concerning is a finding based on meta-analyses and simulation of published preclinical studies as a whole. It looked at various features of the studies, such as their design and analyses features like sample sizes and data analysis plans, and reported that 'most research findings are false for most research designs and for most research fields.'²² Meta-analyses of preclinical studies for individual disease areas report similarly discouraging results. For example, a 2013 assessment of 4445 animal studies for 160 candidate treatments for a range of neurological disorders that were subsequently tested in humans found that, while over 1700 of them reported positive findings, 'only 919 studies would a priori be expected to have such a result.' This finding led the authors to conclude that 'only eight of the 160 evaluated treatments should have been ... tested in humans.'²³ Yet all 160 studies received REC approval. Finally, a 2020 review of preclinical studies related to neuromuscular disorders reports that almost 30% had deficient use of control groups, blinding or randomisation.¹² Such studies help to explain the findings of a recent survey of biomedical researchers showing that the vast majority of respondents believe there is a 'reproducibility crisis' in biomedical research.²⁴

One might question the direct relevance of these aggregate findings to RECs since they have to review individual trials, not aggregate challenges of preclinical research generally. While true, these aggregate findings suggest there are genuine reasons to worry that any given preclinical study, including those used to support REC inferences about potential risks and benefits of early trials, may well report a false-positive finding. A recent analysis of failure rates in clinical trials of acute stroke bears this out. Its authors estimate that a majority of the individual reports of positive findings from animal studies meant to inform clinical studies of acute stroke are actually false positive results.²⁵

A likely explanation for this prevalence of false-positive findings is because, as Kimmelman *et al* have explained, the vast majority of preclinical studies are best viewed as hypothesis generating, as opposed to hypothesis confirming, studies.²⁶ This means that, while those studies may have produced potentially promising findings, it is not known whether those findings are false, which would predict likely failure in human studies. This helps explain why researchers are increasingly realising that 'most preclinical experiments do not represent a true preclinical efficacy study'¹² capable of supporting reliable REC inferences about efficacy that need to be factored into REC assessments of risk/benefit ratios in early trials. However, rather than requesting additional data from follow-up preclinical studies that would be needed to rule out the possibility that the initial positive findings were false-positives, RECs too often permit researchers to proceed straight to clinical trials.²⁶

To appreciate the ethical import of this, consider remarks made at an annual meeting of the American Society of Clinical Oncology by Glenn Begley, coauthor of a landmark study about the irreproducibility of preclinical studies.²⁷ He stated that preclinical studies should be treated like single case reports of clinical findings. Just as one ought not alter the standard of patient care on the basis of a single case report, RECs should never approve a clinical trial on a similarly slim basis. '[These hypothesis generating trials] should not, under any circumstances, trigger a clinical trial. I think there are thousands of patients [who] have been treated inappropriately based on publications [about such studies] in the top journals.'²⁸

One might grant the ubiquitous nature of such problems in individual studies as well but counter that what is really important to REC risk/benefit assessments are the specific preclinical efficacy studies reported in the investigator brochures (IB) that RECs receive when reviewing individual trials. These are the actual studies that an early trial will be based on and thus the only studies that an REC should concern itself with. Unfortunately, a recent study of IBs supporting the early phase trials approved at three research sites between 2010 and 2016 revealed severe problems that suggest the need for concern regarding the reliability of REC approvals based on the specific preclinical studies referenced in the IBs.²⁹

The study examined 109 IBs that reported on 708 preclinical efficacy studies. Less than 5% of them contained any information about such basic study characteristics as randomisation, sample size calculation and blinded outcome assessment that would help RECs make reliable inferences from study findings. Most alarming is the fact that no publications were referenced for almost 90% of the studies. Thus, even if an REC was truly curious about the reliability of a study, this examination of IBs suggests the REC would likely be unable to ascertain it because it would be unable to examine design details about the individual preclinical efficacy study. And when a REC does not ascertain the quality of any given preclinical efficacy study, then the reliability of assessments regarding whether risk/benefit ratios are reasonable and thus whether it is ethical to conduct a given early phase trial remain unknown. All this combined evidence about preclinical studies raises alarms about whether RECs are adequately preventing early trials with unacceptable risk/benefit ratios from passing through their gate.

EVIDENCE ABOUT THE SOCIAL VALUE OF CLINICAL TRIALS

Introductory considerations

Let us now move further along the clinical translation pathway to phase 3 and 4 trials to see how REC review fares at this end

of the spectrum. It is here where we can, for example, meaningfully assess whether a new drug actually works, see if a drug proven to work for one health condition might work in others, or see if one drug works better than other approved ones. All of these types of investigations can advance biomedical research's goal of improved health and healthcare. However, here too this outcome is contingent on the extent of fealty to the demands of good science, fealty on display in the design, conduct, analysis and reporting of the clinical trials.

Knowing the extent of fealty about any particular trial can prove far from straightforward. Returning to early phase trials, the objectives of them are more narrowly constrained than they are for many later phase trials: can we learn enough about safety and potential efficacy to see if a new modality should progress to later phase studies? Later phase trials pose different questions and can have different challenges. Ideally, we would like to know such things as whether a new drug can extend life or reduce suffering and life-limiting symptoms. However, this societal objective usually gets interwoven with trial sponsor interests, especially since private companies fund the majority of later phase trials. Their interests can include such matters as, can we produce a dataset that will sway a regulatory agency to approve a new drug application or, if a drug is already approved, can we produce a dataset that will result in peer-reviewed articles or impactful educational presentations that help persuade physicians to prescribe our drug instead of another?

It is worth noting this distinction between early and later phase trials because the clinical trials that get done are the clinical trials that someone is willing to pay for. Organisations like the NIH can rarely afford to do more than try to shape research agendas that later phase trials can investigate while regulators like the FDA review the applications that sponsors choose to bring them. This places sponsors, the majority of whom are profit-oriented companies, in the driver's seat of most of the clinical trials related to licensing and postlicensure use of drugs.

There is, however, a critical similarity between early and later phase trials to note: just as it is difficult to study human disease in animal models, it is often difficult to adequately study drugs in humans. When chronic diseases like diabetes can take years to shorten a life or may or may not cause grave damage to organs, conducting trials that interrogate a drug's effects on these endpoints may prove impossible, due to such matters as feasibility, expense or drug adherence. So, we frequently are left with trials that study secondary endpoints that, while they may or may not be able to tell us if a drug can truly improve healthcare, are well suited for creating datasets suitable for licensure and postlicensure marketing purposes.

These considerations are mentioned in order to highlight both how the safety and efficacy endpoints in early phase trials can be less fungible than the endpoints for later phase trials and why, in light of the high financial stakes following a drug's approval, it is unwise for RECs to be incurious about the ways that sponsor pecuniary interests can get woven into the planning, design, conduct and reporting of later phase trials. While drug companies can earn immense profits from drugs approved on the basis of datasets about surrogate endpoints that may not really improve health or lessen disability, society has an interest in avoiding approving ineffective new drugs or having physicians and consumers preferring more expensive ones that may in reality be no better than less expensive ones. Either outcome provides little social value or may cause social disvalue. RECs are our last and, at times, only line of defence for preventing clinical trials that lead to these unwelcome outcomes.

Are such unwelcome outcomes frequent enough, though, that RECs need to concern themselves with them? There is certainly ample scholarship that casts many pharmaceutical companies' clinical trial practices in a negative light. There is space here only to note this scholarship. Some work looks more at the industry as a whole and finds it lacking^{30 31} while some looks more at troubling episodes when sponsors withheld critical safety data about drugs from licensure application datasets.^{32 33}

More germane to the task at hand is a body of work³⁴⁻³⁷ from Science and Technology Studies scholar Sergio Sismondo describing what he calls 'ghost managed medicine,' which refers to how drug companies socially construct the knowledge that informs physician prescribing behaviours. This social construction occurs in the planning, design and conduct of clinical trials, as well as the reporting of their findings through scientific journals and other dissemination strategies which influence physician prescribing practices and thus drive corporate profits. Sismondo artfully describes how drug companies use the leeway they have to decide what they will study, how they will study it, and what will get said about it by whom, all within the scientific strictures of clinical research and the legitimacy those strictures confer.

A brief illustration is in order. The gold standard of clinical trials is a randomised double-blinded trial where neither the physician nor the patient/research participant knows which of two or more treatments being compared is given to whom. Findings from these kinds of trials are the ones that physicians typically prize the most because they can possess the requisite rigour needed to confer high reliability on the results such trials produce. This no doubt was the kind of trial that Anthony Fauci had in mind during his remarks quoted above in the Introduction.

Even with the strictures of a randomised double-blinded trial in place, sponsors can nevertheless select, for example, not only what drugs but at what doses, get compared. They also decide who is eligible for inclusion in the studies. These and other choices about a trial's design can bias a trial's outcome in a particular direction. But the sponsor's influence does not end here. They also own the data, control data analysis and determine what data to include in the datasets that journal readers see. This extent of leeway also permits them to run multiple analyses on data subsets or entire datasets when preparing drug applications. These choices can mean the difference between approval or denial of licensure. In other words, ghost management can effectively sabotage the regulatory review process.^{37 38}

Post regulatory approvals, companies also control what journal articles contain and who the authors are. Many drug companies contract with publication planning companies to do the actual writing. These companies are quite adept at getting studies through the peer-review process of leading medical journals.³⁷ They can also allow sponsor approval of authorship, which typically comes in the form of lead authors who are well-regarded physicians who may have had no or little say over a trial's design or data review and analysis. Further, they can withhold authorship from those most responsible for the design, analysis and reporting of the trial and its outcome, while still complying with the authorship policies of most all major medical journals.³⁹ Such concerns are borne out by a 2018 study documenting the outsized, and often hidden, role of industry employees in the design, analysis and reporting of industry sponsored clinical trials. Among the more alarming findings of the study were those showing that only 40% of the academic authors were involved in data analysis of trials and that only a third of academic authors reported having final say on trial design.⁴⁰

Relevant evidence about the social value of later trials

If Sismondo's contention that corporate 'ghost management' of clinical trials can be at odds with good science is correct, there should be evidence substantiating it. Again, due to space constraints, we can only note some of it. First is work ongoing since 1980⁴¹ by numerous scholars about what has been dubbed 'sponsorship bias,' wherein sponsors 'distor[the] design and reporting of clinical experiments to favour the sponsor's aims.'⁴² More recent examples of the scholarship document sponsors both selectively reporting data to regulatory agencies and over-reporting positive as opposed to negative results in the published literature.⁴³⁻⁴⁶ While much of this body of work looks at aggregate findings that do not necessarily prove that there is a preponderance of industry-sponsored trials that produce reports used to elevate company interests over the social value that clinical trials are meant to serve, they certainly lend credence to Sismondo's claim that too many clinical trials are 'sophisticated marketing disguised as disinterested science.'³⁶

A look at additional evidence found in two recent studies looking at the impact of company marketing strategies on clinical trials strengthens these suspicions. The first study examined a group of head-to-head trials comparing the effectiveness of competing drugs.⁴⁷ In theory, such trials are among the most important trials to be conducted. They can tell us, for example, whether one of those drugs is truly superior to the other or whether a cheaper one is comparable to a more expensive one. When such knowledge is achieved, future patients and society as a whole benefit. Unfortunately, the study revealed that few of the trials reviewed in it rose to this level of social value.⁴⁷ Its authors conducted a random review of 319 randomised trials that were published in 2011 that compared drugs or biologics. The vast majority (82%) was sponsored by industry. Only 3 among them involved 'truly antagonistic comparisons' that would be capable of producing accurate comparisons of the drugs or biologics. Most alarming for our purposes is that the study also found that all but 2 of the industry-sponsored trials reported favourable results. These results, we will see shortly, can boost marketing efforts and future prescribing practices for products that, for all we know, may in fact be inferior, that is, neither non-inferior or superior, ones.⁴⁸

The other study looked at 194 drug trials, also published in 2011, in 5 leading medical journals.⁴⁹ The study's authors wanted to determine whether trials might have been conducted largely, if not exclusively, for marketing purposes in order to get physicians comfortable with prescribing a new drug. To test whether this was the case, the authors looked at studies whose results were published in the *New England Journal of Medicine*, *The Lancet*, *Annals of Internal Medicine*, *PLOS Medicine*, and *BMJ*. The authors concluded that 41 of the 194 trials were found to be 'suspected marketing' and that all the suspicious ones were industry funded. They reached the conclusion about suspected marketing in large part due to the fact that those studies tested drugs on patients from an average of 171 different geographical areas, compared with just 13 different geographical areas for the trials deemed not to be marketing trials. Such additional geographic dispersion assured that there were vastly higher numbers of prescribing physicians giving the study drug to their patients in the suspected marketing trials compared with the other trials. Physicians who recruited their patients into the trials would be more likely to prescribe the study medication to their future patients after completion of the trial and many other physicians who read about the trial results will be persuaded to prescribe these new medicines since the studies appear in such

prestigious medical journals and, as the authors of the study report, mask not only the marketing nature of the studies but the role of sponsors in designing and analysing them.⁴⁹

Readers may think that the fact that the articles passed the high peer-review bar set by such prestigious medical journals with very low acceptance rates allays concerns about the quality of the science behind the trials in question. However, it is at the very least equally plausible that the 'ghost' or company-affiliated authors with little or no actual involvement in the studies who take the lead in writing the text and coaching the first and/or senior authors are quite talented in navigating the peer-review process, as their considerably higher than average acceptance rates attest.³⁶ This interpretation is supported by a recent analysis of the latest version of the 'Good Publication Practices 3' (GPP3), a set of voluntary guidelines established by the International Society for Medical Publication Professionals to strengthen medical publications.⁵⁰ The analysis contends that GPP3, despite its lofty aims, actually 'provides a de facto manual for how marketing through academic journal content can be conducted in compliance with contemporary editorial standards.'⁵¹

All of this evidence suggests that sponsors of ghost-managed clinical trials derive financial benefit by cloaking themselves in both the scientific and ethical trappings of clinical trials. Their reports of the trials they sponsor enjoy both the scientific imprimatur regulatory approvals and peer-reviewed journals confer and the ethical imprimatur REC approvals confer on the trials themselves.

This evidence, like the evidence about preclinical research, raises concerns about REC practices. At a minimum, we see how REC approvals facilitated recruitment of hundreds of thousands of people⁴⁷ into ethically suspect trials because the REC reviews failed to flag the potentially flawed findings to be produced by the sponsorship bias that can result in deficient designs, analyses and reports. REC approvals also facilitated the subsequent disproportionate influence of those findings on future clinical guidelines and prescribing practices and the social disvalue as opposed to value that influence can cause.

COUNTERARGUMENTS AND COUNTERMEASURES

The evidence just reviewed paints a sobering portrait of REC performance in their gatekeeping role meant to assure, for example, that the potential benefits of conducting a trial are reasonable in relation to the risks it may cause and that trials are capable of producing a requisite amount of social value.^{6 52} What, then, might we do at this juncture? At least three options come to mind.

One is for RECs to just sit by and hope for structural changes in biomedical research so that the problems highlighted here will wane. This would mean that, going forward, fewer potentially ethically deficient clinical trial applications would find their way to RECs in the first place. Many will find such hope hard to come by. It would require corporate CEOs to so completely share the same scientific values as Fauci that their companies would fully support and never try to subvert the efforts of regulatory agencies to bring safe and effective drugs to market, but no others.

That hope might be equally hard to come by with respect to the ecology of preclinical research, although here there are at least some encouraging developments to note. For example, guidelines meant to improve the reliability of animal studies⁵³ have now been adopted by more than 1000 journals, though that endorsement to date has had a modest impact. A 2017 study reported that the vast majority of researchers were either

unaware of or ignored the guidelines.⁵⁴ More encouraging are the results of an individual journal that now requires authors to use reporting practices tailored to the journal that are in fact producing substantial improvements.⁵⁵⁻⁵⁶ Much time and many additional changes in preclinical research practices will be needed, though, to see if the problems in preclinical research reviewed above are to ever be sufficiently ameliorated. Given such modest cause for hope, it is not clear how RECs could have the luxury of just waiting. Some action on their part to strengthen their review practices thus seems in order.

This brings us to the second option, which is to adopt new practices that could combat problems that the considerations and evidence reviewed above substantiate. Many thoughtful commentaries and recommendations have already been proffered in order to strengthen REC practices. For example, many commentaries explain the importance of REC review of the scientific quality and value of both early⁵⁷⁻⁵⁹ and later phase⁶⁰⁻⁶¹ trials. Nor is there a shortage of recommendations for how RECs could do this. For example, in order to help prevent launching a phase 1 trial on the basis of a false positive finding, Kimmelman *et al* suggest that RECs not approve phase 1 trials in the absence of confirmatory studies supporting efficacy claims,²⁶ while Binik and Hey have recently justified and described a method RECs could use to assure the social merit of clinical trials.⁶¹

New REC practices meant to combat sponsorship bias that can be of legitimate concern in some industry-sponsored trials may also be worth considering. While there are federal regulations and journal practices that have been implemented out of concerns about financial conflicts of interest in clinical trials, they are limited in scope. Federal regulations target investigators rather than sponsors. This leaves sponsor financial conflicts of interests and the sponsorship bias they may cause unaddressed. So too journal editorial policies target authors, not sponsors.

It is not clear what protections against sponsorship bias RECs should think such policies afford. One recent study suggests that disclosure has little if any effect on how manuscript reviewers rate the quality of manuscripts where the conflicts are present.⁶² Further, there is worry that both federal and journal disclosure policies aggravate rather than ameliorate the worries raised by financial relationships, since disclosure requirements may encourage disclosers to think that transparency is all that is required of them to guard against the problems that financial relationships can cause.⁶³

Such considerations about the limitations of current disclosure practices suggest that strengthened REC practices to combat sponsorship bias are needed. RECs could do this by, for example, treating sponsors like they treat individual investigators. Since RECs often require investigators of investigator-initiated trials to sequester themselves from analysing data in early phase studies sponsored by start-up companies they have substantial ownership interests in, they could similarly try to assure independent review of data of corporate sponsored later trials. While they could not sequester sponsors from their data, they could nevertheless insist on additional independent review of the data. This would afford greater assurance than we have today that sponsorship bias is not undermining the scientific quality and social value of ghost managed trials.

One might argue that such a step is unnecessary since, first, regulatory agencies will independently review the submitted datasets of industry-sponsored phase 3 trials when they review the drug applications that come before them. Second, there are new reporting requirements, such as the FDA's 'Final Rule for Clinical Trials Registration and Results Information Submission' issued in 2017, that requires the reporting of all clinical trials

in hopes that negative as well as positive results of all trials get reported.^v

Such arguments must be squared with some facts. Recall what we saw above showing that the datasets that regulatory agencies review may be limited to the data that sponsors choose to submit to them, which often are a far cry from the complete datasets needed to ultimately determine the findings of any given drug trial or set of trials. It is also important to note a recent study undertaken to document compliance with the new Final Rule referenced immediately above. The authors concluded that, while industry trial sponsors are more likely, compared with other sponsors, to post trial results, compliance with this rule is currently 'poor',⁶⁴ just as it is for a similar EU reporting requirement.⁶⁵ Thus, REC adoption of new practices of their own could certainly strengthen their gatekeeping role with respect to guarding against clinical trials that may be tainted by sponsorship bias.

One might agree that some new practices to combat sponsorship bias may be desirable but argue that, short of regulatory reform, independent review of data from corporate-sponsored later phase trials is a bridge too far for RECs. If so, then they could at least consider another reform that has been recommended, which is for RECs to require informed consent disclosures that would alert research candidates to the fact that there is no guarantee that the clinical trials they are considering joining will be appropriately analysed and reported.⁶⁶

Even if there were agreement around either or both of these recommended changes in REC practices in order to combat sponsorship bias, they likely would suffer the same fate as other published recommendations for reforming REC practices. Such recommendations rarely get put to the test to see if they can achieve their intended reforms. One might attribute this fate not to REC indifference to the issues but the inability of RECs to access critical information that would assist them in strengthening their gatekeeping role. Industry sponsors, as alluded to above, are unlikely to abandon their ghost management of later phase trials anytime soon by disclosing to RECs research planning, design and analysis information that would reveal which of their trials are essentially marketing masquerading as science. Nor are IBs likely to become more informative anytime soon about the design features, and thus potential quality shortcomings, of individual preclinical efficacy studies. Absent such vital information, one can be sceptical about the possibility of strengthening practices through REC adoption of published recommendations.

As reasonable as scepticism might be about obstacles that will need to be overcome, it does not remove RECs from their gatekeeping role, nor detract from the ethical importance of that role, which brings us to our third option. This option consists in gathering and examining available facts about REC practices in regard to their policies, instructions to both clinical trial applicants and reviewers and information found in their meeting minutes that can shed critical light on why so many trials tainted by the evidence reviewed above get approved by so many RECs year in and year out.

The information this proposed type of research would produce would show, for example, whether and to what extent

^vOne might also point to the role of Data and Safety Monitoring Boards (DSMB) as an additional layer of independent review of data. While DSMBs no doubt play a critical role in clinical trials, their roles are constrained by their charters that can focus on safety but not always efficacy and members often serve at the pleasure of sponsors.

RECs instruct applicants or reviewers of early trials about what the desired quality metrics of preclinical efficacy studies are that relate to any particular early phase trial. The facts would also show whether RECs provide instructions to their reviewers about how to conduct a risk/benefit assessment when reviewing phase 1 trials. So too they would show the frequency with which the social value of later phase trials is a topic of committee deliberations about applications for later phase trials and whether they have established policies or traditions to address those kinds of concerns when they get raised during deliberations.

The analysis of the evidence should be guided by the gatekeeping role of RECs that establishes their need to be accountable for how they discharge critical tasks, such as assuring proportionality between risks and benefits, that no other oversight bodies are tasked with. This unique gatekeeping status underscores the fact that prior regulatory approvals, such as the permissions needed to conduct phase 1 trials or the drug licensure needed for conducting later phase trials, permit investigators and sponsors to approach, but not pass through, the REC gate. Only RECs get to decide which clinical trials can actually pass through.

Ideally, large national and international studies of REC documentary evidence of the types just mentioned could describe and assess current REC practices and thereby highlight the strengths and weaknesses of them. Such studies will require sponsors but there are many national and private funders devoted to improved healthcare that are candidate sponsors. Either in lieu of or in conjunction with such formal research projects, individual RECs could do their own data gathering and analysis. This should also prove quite useful in highlighting the strengths and weaknesses of the practices of individual RECs.

Readers can now ponder which of the three options is the best way forward at this juncture. To choose the first would be to implicitly endorse a status quo that the evidence reviewed above shows to be clearly deficient. The third option is the preferred one at this juncture for this author. Although the second option clearly has merit, history shows that RECs need to be more motivated than they have been to date to instigate the reforms that the recommendations in the second option are meant to prompt. Option three is proffered as a means of supplying information about current REC performance in their gatekeeping role that might help to produce the needed motivation that is a precondition for RECs entertaining recommendations in the absence of new regulatory mandates.

RECs and REC reformers can take encouragement from other reform efforts in biomedical research that it is possible to proactively prompt important reforms, rather than waiting on regulatory agencies to first require them. The encouragement is found in the progress research reformers have made possible in the face of other daunting challenges. Approximately 20 years of work by preclinical research reformers to uncover problems are now paying dividends in such forms as greater insistence by sponsors on research rigour.⁶⁷ Reformer efforts to show why clinical trial results need to be more transparent were also of long duration but at least now there are trial reporting requirements and we likely will see improved compliance with them over time. Initial actions by RECs to try to strengthen their review practices may produce similar achievements that also could ultimately prompt comparable changes from regulatory bodies that would bolster RECs' abilities to strengthen their review practices. This, however, will require initial action from RECs instead of inaction, which brings us back to the third option. The only current barrier to RECs unilaterally instigating reforms is REC complacency with the current regulatory

status quo and option three is designed to help combat that complacency.

CONCLUSION

This essay has reviewed voluminous evidence showing, for example, that too many phase 1 trials may be launched on the basis of false-positive preclinical findings and that too many later phase trials may not only be incapable of improving healthcare but may actually end up making it worse. Clearly, RECs are not the cause of the false-positive findings or the socially wasteful or harmful later phase trials. RECs do continue to approve the trials though. The point of this essay has been to call attention to this fact in hopes both that RECs might be persuaded to make changes and to show a feasible first step toward change. Change is clearly needed if there are to be fewer ethically deficient trials going forward that enrol unwitting research participants into countless early trials who can be subjected to major risks that for all anyone knows may be disproportional to their benefits and into countless later trials that for all anyone knows may lack sufficient social value. Surely those people deserve better from the RECs with jurisdiction over those trials.

Acknowledgements I would like to thank anonymous reviewers, as well as Kenneth Goodman, Timothy Houk, Nanette C. Joyce, Meaghan O'Keefe, Robert Nadon, Sarah Perrault, Tina Rulli, and Hanno Würbel for comments and suggestions they made on previous versions of the manuscript. I would also like to acknowledge the support of Fondation Brocher.

Funding A portion of the author's time was supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Number UL1 TR001860.

Disclaimer The content is solely the responsibility of the author and does not necessarily represent the official views of the NIH.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Mark Yarborough <http://orcid.org/0000-0001-8188-4968>

REFERENCES

- 1 American association for the advancement of science. Available: <http://www.aaas.org/aboutaaas/>
- 2 National Institutes of Health. Mission and goals. Available: <https://www.nih.gov/about-nih/what-we-do/mission-goals>
- 3 Fauci A, 2020. Available: <https://www.youtube.com/watch?v=dqE3bKRgY6o> [Accessed May 20, 2020].
- 4 Prescrire international. 2019 Prescrire drug awards 2020.
- 5 Halpern J, Paolo D, Huang A. Informed consent for early-phase clinical trials: therapeutic misestimation, unrealistic optimism and appreciation. *J Med Ethics* 2019;45(6):384–7.
- 6 [The Helsinki Declaration of the World Medical Association (WMA). Ethical principles of medical research involving human subjects]. *Pol Merkur Lekarski* 2014;36(215):298–301.
- 7 Council for International Organizations of Medical Sciences. *International ethical guidelines for health-related research involving humans*. Geneva, Switzerland: Council for International Organizations of Medical Sciences, 2016.
- 8 National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. *The Belmont report*, 1979.
- 9 Revised common rule, 82 FR 7149 2017.
- 10 FDA. Investigational new drug (IND) application. Available: <https://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/InvestigationalNewDrugINDApplication/default.htm>

- 11 Kimmelman J, Federico C. Consider drug efficacy before first-in-human trials. *Nature* 2017;542(7639):25–7.
- 12 Willmann R, Lee J, Turner C, et al. Improving translatability of preclinical studies for neuromuscular disorders: lessons from the TREAT-NMD Advisory Committee for therapeutics (TACT). *Dis Model Mech* 2020;13(2).
- 13 Pound P, Ritskes-Hoitinga M. Is it possible to overcome issues of external validity in preclinical animal research? why most animal models are bound to fail. *J Transl Med* 2018;16(1):304.
- 14 Brønstad A, Newcomer CE, Decelle T, et al. Current concepts of Harm-Benefit Analysis of Animal Experiments - Report from the AALAS-FELASA Working Group on Harm-Benefit Analysis - Part 1. *Lab Anim* 2016;50(1 Suppl):1–20.
- 15 Würbel H. More than 3Rs: the importance of scientific validity for harm-benefit analysis of animal research. *Lab Anim* 2017;46(4):164–6.
- 16 Kimmelman J, Henderson V. Assessing risk/benefit for trials using preclinical evidence: a proposal. *J Med Ethics* 2016;42(1):50–3.
- 17 Joyce NC. *Personal communication*, 2019.
- 18 Boonstra JJ, van Marion R, Beer DG, et al. Verification and unmasking of widely used human esophageal adenocarcinoma cell lines. *J Natl Cancer Inst* 2010;102(4):271–4.
- 19 Tadich P. Scientists Have Conducted Decades of Research on Mislabeled Cell Lines. motherboardvice.com [Internet]., 2017. Available: https://motherboard.vice.com/en_us/article/d7y4d7/scientists-have-conducted-decades-of-research-on-mislabeled-cell-lines [Accessed 25 Aug 2016].
- 20 Freedman LP, Cockburn IM, Simcoe TS. The economics of reproducibility in preclinical research. *PLoS Biol* 2015;13(6):e1002165.
- 21 Barrett T. BioSample. In: *The NCBI Handbook* [Internet]. 2nd edition. Bethesda (MD: National Center for Biotechnology Information (US), 2013. <https://www.ncbi.nlm.nih.gov/books/NBK169436/>
- 22 Ioannidis JPA. Why most published research findings are false. *PLoS Med* 2005;2(8):e124–701.
- 23 Tsilidis KK, Panagiotou OA, Sena ES, et al. Evaluation of excess significance bias in animal studies of neurological diseases. *PLoS Biol* 2013;11(7):e1001609.
- 24 Baker M. Is there a reproducibility crisis? *Nature* 2016;533(7604):452–4.
- 25 Schmidt-Pogoda A, Bonberg N, Koecke MHM, et al. Why most acute stroke studies are positive in animals but not in patients: a systematic comparison of preclinical, early phase, and phase 3 clinical trials of neuroprotective agents. *Ann Neurol* 2020;87(1):40–51.
- 26 Kimmelman J, Mogil JS, Dirnagl U. Distinguishing between exploratory and confirmatory preclinical research will improve translation. *PLoS Biol* 2014;12(5):e1001863.
- 27 Begley CG, Ellis LM. Drug development: raise standards for preclinical cancer research. *Nature* 2012;483(7391):531–3.
- 28 Begley CG. *An Unappreciated challenge to oncology drug discovery: pitfalls in preclinical research*. ASCO annual meeting, 2013.
- 29 Wieschowski S, Chin WWL, Federico C, et al. Preclinical efficacy studies in investigator brochures: do they enable risk-benefit assessment? *PLoS Biol* 2018;16(4):e2004879.
- 30 Angell M. *The truth about the drug companies: how they deceive US and what to do about it*. Random House, 2005.
- 31 Brody H. *Hooked: ethics, the medical profession, and the pharmaceutical industry*. Rowman & Littlefield Publishers, 2007.
- 32 Lemmens T. Leopards in the temple: restoring scientific integrity to the commercialized research scene. *J Law Med Ethics* 2004;32(4):641–57.
- 33 Light DW, Lexchin J, Darrow JJ. Institutional corruption of pharmaceuticals and the myth of safe and effective drugs. *J Law Med Ethics* 2013;41(3):590–600.
- 34 Sismondo S. Ghost management: how much of the medical literature is shaped behind the scenes by the pharmaceutical industry? *PLoS Med* 2007;4(9):e286.
- 35 Sismondo S. Ghosts in the machine: publication planning in the medical sciences. *Soc Stud Sci* 2009;39(2):171–98.
- 36 Sismondo S. Corporate disguises in medical science: Dodging the interest repertoire. *Bulletin of Science Technology & Society* 2011;31(6):482–92.
- 37 Sismondo S. *Ghost-manged Medicine: Big Pharma's Invisible Hands*. Mattering Press, 2018.
- 38 Jørgensen L, Götzsche PC, Jefferson T. Index of the human papillomavirus (HPV) vaccine industry clinical study programmes and non-industry funded studies: a necessary basis to address reporting bias in a systematic review. *Syst Rev* 2018;7(1):8.
- 39 Matheson A. The ICMJE Recommendations and pharmaceutical marketing—strengths, weaknesses and the unsolved problem of attribution in publication ethics. *BMC Med Ethics* 2016;17:20.
- 40 Rasmussen K, Bero L, Redberg R, et al. Collaboration between Academics and industry in clinical trials: cross sectional study of publications and survey of lead academic authors. *BMJ* 2018;363.
- 41 Hemminki E. Study of information submitted by drug companies to licensing authorities. *Br Med J* 1980;280(6217):833–6.
- 42 Jefferson T. Sponsorship bias in clinical trials: growing menace or dawning realisation? *J R Soc Med* 2020;113(4):148–57.
- 43 Bourgeois FT, Murthy S, Mandl KD. Outcome reporting among drug trials registered in ClinicalTrials.gov. *Ann Intern Med* 2010;153(3):158–66.
- 44 Lundh A, Lexchin J, Mintzes B, et al. Industry sponsorship and research outcome. *Cochrane Database Syst Rev* 2017;2.
- 45 Lexchin J, Bero LA, Djulbegovic B, et al. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;326(7400):1167–70.
- 46 Lexchin J. Those who have the gold make the evidence: how the pharmaceutical industry biases the outcomes of clinical trials of medications. *Sci Eng Ethics* 2012;18(2):247–61.
- 47 Flacco ME, Manzoli L, Boccia S, et al. Head-To-Head randomized trials are mostly industry sponsored and almost always favor the industry sponsor. *J Clin Epidemiol* 2015;68(7):811–20.
- 48 Matheson A. Marketing trials, marketing tricks - how to spot them and how to stop them. *Trials* 2017;18(1):105.
- 49 Barbour V, Burch D, Godlee F, et al. Characterisation of trials where marketing purposes have been influential in study design: a descriptive study. *Trials* 2016;17(1):31.
- 50 International Society of medical publication professionals. GPP3 guidelines, 2015. Available: <https://www.ismpp.org/gpp3>
- 51 Matheson A. Can self-regulation deliver an ethical commercial literature? A critical reading of the "Good Publication Practice" (GPP3) guidelines for industry-financed medical journal articles. *Account Res* 2019;26(2):85–107.
- 52 Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA* 2000;283(20):2701–11.
- 53 Kilkenny C, Browne WJ, Cuthill IC, et al. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol* 2010;8(6):e1000412.
- 54 Enserink M. Sloppy reporting on animal studies proves hard to change. *Science* 2017;357(6358):1337–8.
- 55 Ramirez FD, Jung RG, Motazedian P, et al. Journal Initiatives to Enhance Preclinical Research: Analyses of *Stroke*, *Nature Medicine*, *Science Translational Medicine*. *Stroke* 2020;51(1):291–9.
- 56 Minnerup J, Dirnagl U, Schäbitz W-R. Checklists for authors improve the reporting of basic science research. *Stroke* 2020;51(1):6–7.
- 57 Habets MGJL, van Delden JJM, Bredenoord AL. The social value of clinical research. *BMC Med Ethics* 2014;15:66.
- 58 Habets MGJL, van Delden JJM, Bredenoord AL. The unique status of first-in-human studies: strengthening the social value requirement. *Drug Discov Today* 2017;22(2):471–5.
- 59 Yarborough M, Bredenoord A, D'Abramo F, et al. The bench is closer to the bedside than we think: uncovering the ethical ties between preclinical researchers in translational neuroscience and patients in clinical trials. *PLoS Biol* 2018;16(6):e2006343.
- 60 Borgerson K. Redundant, secretive, and isolated: when are clinical trials scientifically valid? *Kennedy Inst Ethics J* 2014;24(4):385–411.
- 61 Binik A, Hey SP. A framework for assessing scientific merit in ethical review of clinical research. *Ethics Hum Res* 2019;41(2):2–13.
- 62 John LK, Loewenstein G, Marder A, et al. Effect of revealing authors' conflicts of interests in peer review: randomized controlled trial. *BMJ* 2019;367.
- 63 Cain DM, Loewenstein G, Moore DA. The dirt on coming clean: perverse effects of disclosing conflicts of interest. *J Legal Stud* 2005;34(1):1–25.
- 64 DeVito NJ, Bacon S, Goldacre B. Compliance with legal requirement to report clinical trial results on ClinicalTrials.gov: a cohort study. *Lancet* 2020;395(10221):361–9.
- 65 Goldacre B, DeVito NJ, Heneghan C, et al. Compliance with requirement to report results on the EU clinical trials register: cohort study and web resource. *BMJ* 2018;362.
- 66 Yarborough M. Rescuing Informed Consent: How the new "Key Information" and "Reasonable Person" Provisions in the Revised U.S. Common Rule open the door to long Overdue Informed Consent Disclosure Improvements and why we need to walk Through that door. *Sci Eng Ethics* 2019. doi:10.1007/s11948-019-00170-8. [Epub ahead of print: 23 Dec 2019].
- 67 Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. *Nature* 2014;505(7485):612–3.
- 68 U.S. Department of health and human services food and drug administration, center for biologics evaluation and research. preclinical assessment of investigational cellular and gene therapy products 2013.