

EQUIPOISE AND INTERNATIONAL HUMAN-SUBJECTS RESEARCH

ALEX JOHN LONDON

ABSTRACT

This paper examines the role of equipoise in evaluating international research. It distinguishes two possible formulations of the equipoise requirement that license very different evaluations of international research proposals. The interpretation that adopts a narrow criterion of similarity between clinical contexts has played an important role in one recent controversy, but it suffers from a number of problems. An alternative interpretation that adopts a broader criterion of similarity does a better job of avoiding both exploitation of the brute fact of social deprivation and the exploitation of needy populations for the benefit of more well-off populations. It also holds out the promise of reconciling the need to find interventions that can be employed in developing world contexts with the cluster of moral values that must constrain the way such research is carried out.

The question of how to formulate the equipoise requirement in the context of international human-subjects research touches on some of the most fundamental issues in research ethics, yet it has received surprisingly little explicit and systematic discussion. For instance, in the recent controversy over international trials of a short-course of zidovudine (AZT) for the prevention of maternal-infant HIV transmission, the earliest and most vocal critics argued that the use of a placebo controlled design was unethical because it violated the equipoise requirement. They then argued that the short-course regimen should be tested against the current standard of care in the developed world, known as the Aids Clinical Trial Group 076 protocol.¹ In the relatively

¹ M. Angell. The Ethics of Clinical Research in the Third World. *NEJM* 1997; 337: 487-9 and P Lurie, SM Wolf. Unethical trials of interventions to reduce perinatal transmission of the human immunodeficiency virus in developing countries. *NEJM* 1997; 337: 853-6.

acrimonious debate that ensued, defenders of the placebo controlled design vigorously challenged the second of these claims, namely, that research conducted in a developing country should be governed by the same standard of care that prevails in the developed world.² Surprisingly, however, relatively little explicit attention was paid to the more fundamental point from which it was derived: that as proposed the short course trials would violate the equipoise requirement. Without a careful elucidation of the role of equipoise in evaluating international research it has been difficult to locate the crux of several important disputes and I have argued elsewhere that the disagreement over the interpretation of the standard of care for international research is a case in point.³

On a more fundamental level, however, the absence of a careful and sustained analysis of equipoise in this context has given rise to the uncritical acceptance of a particular way of applying the requirement to international research. In what follows, I will argue that the received position suffers from a number of problems. I will also argue that there is an alternative way of formulating the equipoise requirement that licenses very different evaluations of some research proposals. In particular, the received interpretation of the equipoise requirement results in restrictions on international research that are either much more stringent or much more permissive than either side of these recent debates may recognize. Furthermore, it frames the question of equipoise in a way that exaggerates the appearance of intransigent conflict and the need for making 'tragic choices' between important moral values.

The alternative conception of the equipoise argument that I sketch below avoids these problems. I will show that the difficulties that remain are either not unique to this interpretation alone, or that they are less problematic than they first appear. In the end, I hope that articulating these competing conceptions of equipoise and discussing their respective strengths and weaknesses will provide a framework in which the relationships between a cluster of important values can be more clearly charted.

² R.J. Levine. The 'Best Proven Therapeutic Method' Standard in Clinical Trials in Technologically Developing Countries. *IRB* 1998; 20: 5-9.

³ A.J. London. The Ambiguity and the Exigency: Clarifying 'Standard of Care' Arguments in International Research. *Journal of Medicine and Philosophy* 2000; 25: 379-397.

THE BASICS OF EQUIPOISE AND A CAVEAT

In its most basic formulation equipoise represents a state of genuine and credible doubt about the relative therapeutic merits of some set of interventions that target a specific medical condition. The requirement that equipoise exist as a necessary condition for the moral acceptability of a clinical trial comparing these interventions is motivated by two interlocking ideas. First, when equipoise obtains it is morally permissible to allow an individual's medical treatment to be assigned by a random process because there is no sufficiently credible evidence to warrant a judgment that one intervention is superior to the other(s). Second, clinical trials that are designed to break or disturb equipoise provide information that will enable the medical community to improve its existing clinical practices.⁴ The requirement is thus seen as a way to reconcile the need to improve the state of medical knowledge and clinical practice with the duty to ensure that the welfare of individual subjects is not knowingly sacrificed for the welfare of future patients or greater scientific understanding.

The equipoise requirement is thus a normative standard that articulates important scientific objectives, relating to the value of the data such trials should produce, as well as ethical boundaries that constrain the way this data may be obtained. In this latter respect, equipoise links together several important moral concepts. First, it stipulates that medical research must not violate what is called the 'duty of personal care,'⁵ or the 'therapeutic obligation.'⁶ Second, the concept of equipoise underscores an important epistemological aspect of this duty, namely, that the content of the obligation – what is required of a physician or researcher in some instance – depends in part on our ability to predict or foresee possible outcomes with an appropriate degree of certainty. For there to be a positive duty to provide a subject with a specific intervention there must be evidence of sufficient weight to license the judgment that it is likely to advance that person's interests. Finally, these considerations help to ensure that a clinical trial is just or fair by mandating that the interests of individual subjects are valued equally. A trial that begins in

⁴ B. Freedman. 'Placebo-Controlled Trials and the Logic of Clinical Purpose' *IRB* 1990; 12 (Nov/Dec): 1–6.

⁵ C. Fried. 1974. *Medical Experimentation: Personal Integrity and Social Policy*. Amsterdam. North-Holland Publishing.

⁶ D. Marquis. Leaving Therapy to Chance. *Hastings Center Report* 1983; 13: 40–47.

equipoise gives equal consideration to the interests of all subjects, even though that trial may ultimately show that one treatment option is superior to another.

Within the US and other technologically and economically developed nations the equipoise requirement continues to play an important role in the evaluation of human-subjects research, even though it remains the subject of searching philosophical criticism.⁷ In particular, there is considerable debate over different interpretations or specifications of the concept of equipoise itself, as well as a debate over whether or not any of these interpretations succeed in reconciling the goal of advancing clinical knowledge with the duty to protect the interests of individual trial participants. Different interpretations of equipoise are individuated by the way they specify a range of interlocking variables that determine how the concept will be employed in practical decision making. So, for instance, we need to specify *who* is to weigh the relevant evidence in order to decide whether or not equipoise exists. Is this a judgment that is up to individual physicians, individual patients, the medical community as a whole, or some larger community that includes patients and possibly others? Likewise, we need to specify the kind of evidence that will warrant a judgment that equipoise has been achieved or disturbed. Here, possible answers can range from evidence as thin as an individual's personal hunch to as strict as data from double-blind randomized clinical trials.

These are but two in a range of important disagreements. For the purposes of the present discussion, however, I am going to focus on a facet of equipoise that has special relevance to the international context. To the extent that different stances on the issues mentioned above add an additional degree of complexity to this issue, it strengthens, rather than detracts from, the point that the equipoise requirement in international research deserves more careful consideration than it has currently received. Although I do in fact believe that there are good, independent reasons to think that these broader issues are more tractable than they are sometimes made out to be, this is a subject that will have to be dealt with on another occasion.

⁷ For very clear and current review of several of these controversies, see F. Gifford. Freedman's 'Clinical Equipoise' and 'Sliding-Scale All-Dimensions-Considered Equipoise.' *Journal of Medicine and Philosophy* 2000; 25: 399–426.

THE RECEIVED CONCEPTION OF EQUIPOISE

Perhaps the most prominent reference to the concept of equipoise in the recent debate over international research was made by Marcia Angell in her provocative critique of the short-course AZT trials. Angell's editorial in the *New England Journal of Medicine* begins by recounting the importance of equipoise for both the scientific and ethical acceptability of a clinical trial. She notes that there should not be solid evidence that one proposed intervention will be superior to the other and that:

[1] If there is, not only would the trial be scientifically redundant, but the investigators would be guilty of knowingly giving inferior treatment to some participants in the trial. The necessity for investigators to be in the state of equipoise applies to placebo-controlled trials as well. [2] Only when there is no known effective treatment is it ethical to compare a potential new treatment with a placebo. [3] When effective treatment exists, a placebo may not be used. Instead, subjects in the control group of the study must receive the best known treatment.

[4] All except one of the trials employ placebo-control groups, despite the fact that zidovudine has already been clearly shown to cut the rate of vertical transmission greatly and is now recommended in the United States for all HIV-infected pregnant women.⁸

Here Angell spells out in some detail a position that is frequently espoused more tersely by subsequent commentators. For example, one writes: 'The only way that these placebo-controlled trials should be allowed [in the developing world] is if there is a genuine doubt about the benefits of AZT. No such doubt exists in the United States.'⁹

In a moment I will question the validity of Angell's argument on the grounds that her understanding of its conclusion [3] is inconsistent with her understanding of the equipoise requirement itself, as expressed in [1] and [2]. In order to demonstrate this inconsistency, however, it will be necessary to investigate one facet of Angell's conception of the equipoise requirement in a bit more detail.

⁸ M. Angell. The Ethics of Clinical Research in the Third World. *NEJM* 1997; 337: 487–9 at 487. The numbers in brackets do not appear in the original.

⁹ P.A. Clark. The Ethics of PlaceboControlled Trials for Perinatal Transmission of HIV in Developing Countries. *The Journal of Clinical Ethics* 1998; 9: 156–166 at 162.

Consider the claim that it is unethical to compare a short-course of AZT to a placebo in the context of the developing world because researchers in developed countries have shown that the 076 protocol can substantially reduce the rate of perinatal HIV transmission in the context of the developed world. If this claim could be established then let us grant, for the sake of the argument, that researchers conducting such a trial would indeed be guilty of knowingly allowing one group of subjects (members of the control group) to suffer foreseeable and (practical limitations aside) preventable harms. What needs to be considered, however, is what it takes to show that the results of research and clinical practice conducted in the context of the developed world are relevant to the context of the developing world.

Put in its most general terms, in order for the fact that no doubt exists in the US concerning the benefits of AZT to be relevant to the design of a trial in the developing world, we must assume that the context of treatment in the developing world is *relevantly similar* to the context of treatment in the developed nations in which the 076 protocol demonstrated its effectiveness. This is because differences in the context of treatment bear on our ability to reliably project the knowledge of causal relationships from the one context into the other. Whether or not this is a controversial assumption, however, will depend on how we understand the criteria by which relevant contexts of treatment should be compared.

There are, however, narrow and broad interpretations of the criterion for similarity between the relevant contexts of treatment in the above argument. According to the narrow interpretation, the context of treatment in the US and Uganda, for example, is relevantly similar just in case we have no credible reason to doubt that AZT, as successfully administered in the 076 protocol, would have the same biological effects in both populations. On this interpretation, the relevant criterion of similarity is the physiological equivalency of the two populations. So understood, the only way that equipoise would exist between a placebo and the short-course of AZT in Uganda, when equipoise does not exist between them in the US, is if there were physiological differences between these populations of sufficient significance to generate a credible uncertainty as to whether AZT would behave the same way in the bodies of Ugandans as it does in Americans. In the absence of such physiological differences, proponents of this position argue that the short-course regimen should be compared against the 076 protocol rather than a placebo.

Angell appears to embrace this narrow interpretation of the context of treatment when she dismisses as subterfuge the idea that information gained in the industrialized world may not be relevant to developing countries because 'diseases and their treatments are very different in the Third World'. Instead, she argues, 'unless there are specific indications to the contrary, the safest and most reasonable position is that people everywhere are likely to respond similarly to the same treatment'.¹⁰ These remarks are aimed at countering claims that there were in fact important physiological differences between treatment populations in industrialized nations and test populations in the developing world. In particular, it was argued that the 076 protocol was tested in a well-nourished population with a low incidence of anemia whereas populations of the developing world are frequently malnourished and anemic. Given that AZT can exacerbate anemia, it was argued that a placebo group was necessary to ascertain the relative safety of AZT in such populations.¹¹ This is probably the clearest instance of a straightforward clash over the question of equipoise in the debate over the short-course AZT trials and it is carried out against the backdrop of the narrow interpretation of the context of treatment. Both Angell and her critics accept the narrow interpretation of the clinical context; they simply disagree over the empirical issue of whether or not we have sufficient warrant for doubts about the effects of AZT in the bodies of malnourished populations of the developing world.

In the sections that follow I challenge the narrow interpretation of the clinical context on several grounds.

One implication of the narrow interpretation is that it undercuts the alternative approach to research in the developing world that Angell herself supports. Angell argued that equipoise did not exist between the short course and a placebo because there were strong indications that even a short course of AZT would be better than nothing at all. From this premise she then argued that unless members of the control group received the 076 protocol, researchers would be guilty of sacrificing the welfare of one

¹⁰ Angell, *op.cit.* p. 848.

¹¹ H. Varmus, D. Satcher. Ethical Complexities of Conducting Research in Developing Countries. *NEJM* 1997; 337: 1003–5 at 1004; M.H. Merson. Ethics of Placebo-Controlled Trials of Zidovudine to Prevent the Perinatal Transmission of HIV in the Third World (editorial). *NEJM* 1998; 338: 836; R.J. Simonds, M.F. Rogers, T.J. Dondero. Ethics of Placebo-Controlled Trials of Zidovudine to Prevent the Perinatal Transmission of HIV in the Third World (editorial). *NEJM* 1998; 338: 836–7.

group of subjects for the sake of medical knowledge. Yet, if we consider the design that Angell recommends it is unclear how this second point is supposed to follow from the first. Based on the available data, it was reasonable to believe that the short course of AZT would be better than nothing. But it was just as reasonable to believe that the short course would not be as effective as the full 076 protocol.¹² So it looks as though Angell's own argument applies to the trial design that she herself recommends, only here it is the welfare of members of the short-course arm whose welfare is being sacrificed for the sake of knowledge.

On what grounds would it be permissible to conduct the sort of trial Angell recommends? Both the placebo controlled design and Angell's alternative seem to violate the equipoise requirement as Angell articulates it and if the sheer need to find an effective intervention for developing world populations justifies her preferred trial design then it also justifies the placebo controlled design that she criticizes. Alternatively, if we are to take the requirements of equipoise seriously, and if we embrace her interpretation of them, then it becomes unclear how we could *ever* justify searching for less expensive, less cumbersome, more portable alternative interventions that might provide some significant but less than optimal degree of relief to populations of the developing world.¹³ Consider the following dilemma. Either we think that a proposed intervention may be as good as or better than the existing alternatives, or we do not. If we do, there would be no reason to place additional burdens on developing world populations by locating the research there since the trial could be ethically conducted in the developed world. If we do not have reason to think that the proposed intervention will be equivalent to or better than the current alternatives then equipoise would not exist and it would be unethical to conduct the trial anywhere in the world.

Neither side of this debate seems to recognize the importance of this dilemma. In particular, it seems to undermine the

¹² This suspicion appears to have been confirmed by an equivalency study recently completed in Thailand: M. Lallemand, G. Jourdain, S. Le Coeur, et al. A trial of shortened zidovudine regimens to prevent mother-to-child transmission of human immunodeficiency virus type 1. *NEJM* 2000; 343: 982-91.

¹³ One implication of the present paper is that ethical international research is *not* premised on finding sub-optimal but affordable interventions for the developing world. Rather, ethical international research is premised on locating *optimal* interventions where these are the most effective interventions that can be implemented and maintained over time within a treatment context that is practically attainable in a population.

compatibility of two claims that each side seems to embrace, namely, that (1) equipoise should play an important role in evaluating international clinical trials and (2) that there are a range of cases in which conducting clinical trials aimed at finding interventions that can be widely utilized in the developing world can be a legitimate and morally permissible means of addressing certain distinctive and pressing healthcare needs of those populations. It is reasonable to believe that a more sophisticated conception of equipoise should help to identify some of these cases.

It is difficult to see how those claims can be reconciled if we adopt biological equivalency as the relevant criterion of similarity between research populations. To amplify this point, consider how the narrow interpretation construes the relationship between the question of equipoise and important, practical considerations that pertain to our ability to implement a protocol in a particular place.

THE ROLE OF BROADER, PRACTICAL CONSIDERATIONS

A narrow view of equipoise focuses on physiological equivalency as the criterion for similarity between the research context in countries such as the US and Uganda. This means that in the absence of credible doubts about the physiological similarity of these two research populations, it cannot be the case that equipoise exists for a research protocol in one country and fails to exist for the same protocol in the other. If we return to the case of the short-course AZT trials, proponents of the placebo design frequently pointed to a variety of important, practical obstacles to implementing the 076 protocol in the host nations, not only on a large-scale basis, but perhaps also within the context of a clinical trial. The most obvious and important obstacle was the sheer poverty of the developing nations in which the short-course trials were proposed. At approximately \$800 per mother, the 076 protocol was far beyond the reach of the \$8 average per-capita health care expenditures of these developing countries. Even if the AZT for clinical trials was donated by pharmaceutical companies, it was generally recognized that the developing nations in question could not afford the staggering cost of fully implementing the 076 protocol on a widespread basis. In fact, it was the recognition of this fact in conjunction with the high incidence of HIV in the developing world that motivated researchers to look for an alternative intervention in the first place.

The cost of the 076 protocol was not its only drawback, however. It was also pointed out that the 076 protocol requires mothers to stop breast feeding, since this is a known pathway of HIV transmission. In many of the developing nations in question, however, this requirement would be practically unachievable in part because of the scarcity of clean water. Additionally, the burden of having to purchase infant formula would add to the cost of an already expensive intervention.¹⁴ Similar objections were raised with respect to the availability of the kind and quality of staff and facilities that the 076 protocol requires. Most of the nations in question lack the sort of well established healthcare infrastructure that has become the norm in the developed world. This generated doubts about the ability of researchers to implement effective screening programs at a sufficiently early stage of pregnancy for the 076 protocol to be implemented, especially in light of the fact that most pregnant women in these countries do not appear in a clinical setting until fairly late in pregnancy.¹⁵

These sorts of practical concerns do not bear directly on the permissibility of clinical trials if we adopt the narrow interpretation of equipoise.¹⁶ If they are relevant at all, it is because they generate a set of additional, practical problems, that have to be weighed alongside of, and possibly against, the independent question of equipoise. On this view, Angell's interpretation of the equipoise requirement is accepted and, if she is right about the biological similarity of the two populations, then we are left with two basic options. First, if these various practical concerns are in fact insurmountable and it would thus be practically impossible to implement the 076 protocol in these countries, then it might be that the question of equipoise simply becomes irrelevant. That is, one might admit that Angell's narrow interpretation of equipoise is correct but argue that since we cannot be obligated to do what is practically impossible (or perhaps so difficult as to become supererogatory), her objections lose their normative

¹⁴ R.A. Crouch, J.D. Arras. AZT Trials and Tribulations. *Hastings Center Report* 1988; 28: 26-34 at 26.

¹⁵ Varmus and Satcher, p. 1004.

¹⁶ It has been pointed out, for example, that in the disputes over the short-course trials, 'the vehement emphasis on the 'best proven drugs' eclipsed considerations of whether the drug regimen could be safely applied in different settings.' S.R. Benatar, P.A. Singer. A new look at international research ethics. *BMJ* 2000; 321: 824-826 at 824. One purpose of the present paper is to give an account of equipoise that directly links issues of implementation and effectiveness.

force. Alternatively, however, given the same facts, one might take the opposite view and argue that the equipoise requirement cannot be set aside because of practical obstacles that bar people in the developed world from access to top-flight medical care. In this case, one might argue that it is not permissible to conduct clinical trials in the developed world until the conditions for equipoise can be achieved. The time and resources of western researchers, agencies, and governments should therefore be spent trying to ameliorate the very basic conditions that make effective therapies unattainable in those countries rather than conducting clinical trials in which they knowingly allow some members of the trial to suffer foreseeable and preventable harms.

Neither of these options is easy to accept because each appears unable to account for important moral intuitions about the case at issue. On the one hand, the view according to which the question of equipoise becomes irrelevant to international research fails to do justice to the considerations that underwrite the equipoise requirement in the first place. It allows research to proceed without articulating moral boundaries that require, among other things, that those who design and carry out such trials prevent foreseeable harms from befalling an identifiable group of people. On the other hand, the alternative view – which prevents research from going forward on the grounds that equipoise does not exist – seems content to sacrifice the welfare of the literally thousands of people who might benefit from the results of such research to the glacial pace of international justice and social change.

These alternatives, neither of which is satisfactory, characterize the present state of the debate about the standards that should govern international research. It is important to stress, however, that they are predicated on the narrow interpretation of the criterion of similarity between clinical contexts. That is, they presuppose that the question of whether or not equipoise exists between two proposed interventions in a specific population can be settled independently of the practical considerations that bear on the degree to which those interventions can be implemented within a population. The possibility that such practical concerns might be relevant to the question of equipoise itself has been left largely unexplored. To a certain degree this may be because they look like two different, and possibly incommensurate, sets of concerns: equipoise deals with what we know (e.g., that the 076 protocol has proven highly effective in the developed world) whereas practical considerations deal with what we can, or cannot, do (e.g., whether we can effectively implement the 076

protocol in the context of a developing nation). To see the way in which these questions are intimately connected with one another, we must explore some of the virtues of the broad interpretation of the criterion for similarity between contexts of treatment.

A BROADER CONCEPTION OF EQUIPOISE

From a clinical standpoint, the practical considerations that bear on our ability to successfully implement a treatment protocol in a particular population are of fundamental importance when evaluating the impact that such a protocol might have on the health of individuals in that population. This point was recognized by Freedman when he argued that equipoise should be a '*portmanteau* measure including all the elements that contribute to the acceptance of a drug within clinical practice.'¹⁷ Among other things, Freedman argued that the question of equipoise should be framed around an intervention's 'net therapeutic advantage' where this is 'a compendious measure of a treatment's attractiveness.' This measure includes physiological considerations such as an intervention's direct impact on disease reduction, symptomatology, and ability to function, discounted by its particular side effect profile. However, it also includes broader, more practical considerations relating to differences in mode of delivery and ease of administration.

Factors that are ancillary to the brute biological characteristics of an intervention are important for several reasons. For Freedman, to weigh the attractiveness of competing interventions requires a comparison of the 'dynamic balance' created by a host of factors that are relevant to their clinical profile. For instance, he suggests that an antibiotic that is attractive because of its specific microbial action may have a lower net therapeutic advantage than alternatives that have fewer toxic side effects *or* that do not require intravenous delivery and constant medical monitoring. This example is particularly interesting in the present context because it reveals one way in which an intervention's net therapeutic advantage is influenced by the nature of the context in which the intervention is to be implemented.

In order to answer the question of how effective a particular intervention is likely to be relative to some alternative in a particular setting, we must first answer a host of practical

¹⁷ B. Freedman. Placebo-Controlled Trials and the Logic of Clinical Purpose. *IRB* 1990; 12: 1-6 at p. 5. The following quote is from p. 2.

questions about the nature of the context in which those treatments can be effectively administered and our ability to create and sustain such a context in a particular place or community. For this reason, a more robust conception of equipoise adopts a broader and more flexible criterion for similarity between treatment contexts, what I will call the criterion of 'clinical comparability.' This conception of equipoise can be stated formally as follows. Let an intervention I represent a treatment for a problem P and a protocol for its implementation. Let a treatment setting S represent an identifiable population and the set of background conditions within which that population lives and receives medical care. Such background conditions may include unique physiological characteristics of that population, certain social and cultural norms operative within the population, material resources available in that population including infrastructure and other social resources, and other conditions that may be relevant as well.

Principle of Equipoise: Equipoise exists between interventions I_1 and I_2 relative to problem P in a treatment setting S , just in case credible doubts exist about the relative net therapeutic advantage of I_1 and I_2 for treating P in S and there is no intervention I_3 that is preferable to either or both I_1 and I_2 for treating P in S .

Credible doubt about the relative net therapeutic advantage of I_1 and I_2 for P in S exist just in case there is no treatment setting S^* such that both (1) S and S^* are clinically comparable and (2) good evidence exists for the superiority of I_1 or I_2 or some I_3 for P in S^* .

Clinical comparability: S and S^* are clinically comparable with respect to an intervention I for a problem P just in case both (1) for the set of identifiable conditions C^* that are judged to be necessary for realizing the effectiveness of I as a treatment for P in S^* , a functionally equivalent counterpart C can be practically attained in S and (2) the functional equivalency of C in S can be practically sustained over time.

This formulation of equipoise is much more explicit about possible sources from which credible doubts about the effectiveness of an intervention may arise. For example, they may exist because its net therapeutic advantage over a placebo has yet to be demonstrated or effectively measured. Additionally, however, even if an intervention I is known to be an effective treatment for

P in one treatment setting (S^*), the reliability of our judgments about its likely net therapeutic advantage over alternatives in another treatment setting (S) depends crucially on whether the conditions under which I is known to be effective in the first setting (S^*) can be replicated in the second (S). This is true regardless of whether S^* is the context of treatment in the US and S is the context of treatment in a developing world population or S^* is the context of treatment established in clinical trials of I (in the US, say) and S is the context of treatment in which I will be used in clinical practice (in the US or elsewhere). As the degree of clinical comparability that can be achieved between these contexts of treatment increases or decreases, so does the reliability of our judgments about the likely net therapeutic advantage of I in the new treatment setting.

Even within developed nations, clinicians may have difficulty achieving in clinical practice results that can be attained within the context of well run clinical trials. In fact, this is a general problem that experimental research of any kind must grapple with: ensuring that the conditions under which something is known to be effective can be replicated in its use outside of that context. This is an especially important point for international research, however, where a wide variety of differences between treatment contexts can affect the calculation of an intervention's net therapeutic advantage.

Differences between treatment settings are only counted as relevant to the extent that they may foreseeably influence the measure of an intervention's net therapeutic advantage. Nevertheless, the likely net therapeutic advantage of an intervention may vary significantly across treatment settings. Some of the reasons for these variations are rooted to different degrees in economic differences between developed and developing nations. It would be a mistake, however, to construe all economic issues in this context as questions of a people's ability to purchase expensive interventions.¹⁸ In addition to the costs of procurement, the economic profile of health care interventions includes the level of infrastructure required for their implementation including requirements relating to technology, availability of facilities and staff of a certain quality in appropriate quantities. The fact that an

¹⁸ For instance in, U. Schüklenk, R. Ashcroft. International research ethics. *Bioethics* 2000; 14: 158–172 at 167. It is pointed out that in many cases, what we can achieve in a particular population may hinge on the kind of reductions we can carve out of the pricing structures of extremely profitable international pharmaceutical companies. The equation of availability and affordability is then put into the mouth of a possible consequentialist position at 168.

intervention requires intensive staff supervision, or a particularly high degree of skill to implement, may be a substantial problem for populations in which relatively few medical personnel must strive to meet the needs of comparatively large numbers of people. Furthermore, the economic profile of an intervention also includes the burdens that patients incur from possible opportunity costs related to the treatment. The fact that one population of people is nomadic and highly migratory while another is highly immobile and agrarian may not bear on the question of equipoise if the intervention in question has a benign side effect profile and can be made sufficiently portable. But it may be relevant if the intervention in question requires prolonged hospitalization, frequent and lengthy treatment visits, or has a side effect profile that would prevent travel for long periods of time.

For these reasons, even treatments with relatively low procurement costs may have significantly different economic profiles and the net therapeutic advantage of such interventions may vary widely relative to different contexts of treatment. This means that in some cases, to make an intervention 'available' in a community will not (only) require lowering its procurement cost, but increasing or otherwise modifying the surrounding infrastructure so that it can support its effective implementation.

Although some of the relevant differences between contexts of treatment will have an important economic aspect to them, this need not always be the case. Differences in the convenience of long term treatments may affect their relative net therapeutic advantage, especially if they generate significant differences in compliance. Such factors can also be magnified by differences in social or cultural norms. In agrarian populations where people must work away from their homes and the careful measurement of time is not as important as in urban or developed world settings, potent but complicated drug regimens that must be taken on a strict schedule throughout the day may be less attractive than less potent, single dose options. Or again, even within the developed world, certain identifiable groups may maintain deeply held religious convictions that prevent them from accepting certain forms of medical intervention. The net therapeutic advantage of artificial blood products relative to transfusion may differ significantly between the general American population, say, and the population of Jehovah's Witnesses. This is due not to some unique physiological constitution of the Witnesses but to their beliefs about the religious significance of blood products. These beliefs are clinically relevant to the extent that they dispose Witnesses to

reject blood transfusions and to accept artificial alternatives. This may be an exceptional case, but it illustrates another possible way in which a population's beliefs or cultural norms may themselves be relevant to the question of equipoise.

With these points in mind, it is particularly important to emphasize that the central issue for the broad interpretation is not the degree of clinical comparability that *already exists* between two treatment settings, but the degree of clinical comparability that is *practically attainable and sustainable*. This is because the latter question links up directly with the epistemological aspect of the duty of personal care. Regardless of whether I is currently available in a population, if we know that the clinical context in which I can be effectively implemented in a population is *practically attainable* then it would violate the duty of personal care to conduct a clinical trial in which an alternative to I is tested against a placebo there.¹⁹ To do so would be to knowingly give unequal consideration to the equal interests of the participants in such a trial. It would violate the same duty of care to conduct a clinical trial of any alternative to I in a population unless there are credible doubts about the net therapeutic superiority of I to such an alternative in a treatment setting that is practically attainable in the population in question.

¹⁹ It is an innovation of the present paper that it links this emphasis on practical attainability with the epistemological requirements of equipoise. Recently, some researchers have begun to adopt the language of the 'highest practically attainable' standard of care in the context of international research. For instance, one recent consensus statement repeatedly emphasizes that in the developing world participants in clinical trials, 'should be assured the highest standard of care practically attainable in the country in which the trial is being carried out' (Science, ethics, and the future of research into maternal infant transmission of HIV-1. *The Lancet* 1999; 353: 832–835 at 833.) A few lines later, however, the same statement argues that:

Where there is no antiretroviral therapy currently available in the host country, and no reasonable expectation of its availability during the time frame of the planned trial, it is imperative to test and identify rapidly a regimen that is more effective than no anti-HIV-1 intervention and more affordable and implementable than the proven ZDV regimens. A no intervention controlled design may be ethically justified in host countries where there is no antiretroviral therapy currently available and no reasonable expectation of its availability during the time frame of the planned trial (p. 834).

This suggestion seems to violate the epistemological aspect of the duty of personal care. I would also suggest that this position can only be seen as a viable alternative if the notion of a 'standard of care' is separated from the requirements of equipoise. For problems with this general move, see London.

This point is of particular importance because it provides a safeguard against exploiting the mere fact that some population does not currently have access to needed healthcare interventions. Economically disadvantaged, socially isolated or oppressed groups may lack access to a host of inexpensive, easily administered interventions for a variety of health care conditions precisely because they are disadvantaged, isolated, or oppressed.²⁰ By focusing on the degree of clinical comparability that can be practically achieved between two clinical contexts, we provide a greater degree of protection than if we require only that the level of care provided to members of the control group not fall below the level of care that they would otherwise receive within their community.²¹ The latter standard licenses clinical research in any situation where social or economic deprivation results in lack of access to basic medical care. The former standard only licenses clinical research when the circumstances of a particular population are such that a reasonable effort would not be able to achieve the relevant degree of clinical comparability between their clinical situation with respect to a particular intervention, and the context in which that intervention has proven to be effective.²² Since it is likely that this degree of clinical comparability already exists, or can easily be achieved, in most developing world populations with respect to basic health care interventions, this standard will only license clinical research in cases that are 'unique' in some identifiable, clinically relevant way.

Whether it appears this way on the surface or not, both conceptions of equipoise outlined in this paper are committed to certain views of the role of economic considerations in formulating the equipoise requirement. By focusing on biological equivalence the narrow interpretation frames the question of equipoise in a way that most closely resembles clinical contexts in which sufficient economic resources are available to overcome the myriad obstacles that may hinder the effective implementation of an intervention. Wealthy nations with robust healthcare infrastructures and socio-political mechanisms that attempt to provide fair access to well established healthcare systems can *act as if* economic considerations are transparent to the question of equipoise because it is only in fairly exceptional circumstances

²⁰ See important examples in George J. Annas, Michael A. Grodin. Human Rights and Maternal-Fetal HIV Transmission Prevention Trials in Africa. *American Journal of Public Health* 1998; 88: 560-563.

²¹ For criticism of this standard, see London.

²² I will return to the question of what constitutes a 'reasonable effort' at the end of the paper.

that such nations would be unable to maintain the appropriate clinical context to support an intervention. This does not explain, however, why nations with significantly less developed healthcare infrastructures, whose healthcare systems must operate within much tighter resource constraints, should be required to act as if this is the case for them as well.

By requiring that clinical comparability be practically sustainable over time, this standard also provides a check against a second kind of exploitation, namely, the use of developing world populations solely for the purpose of gaining knowledge to be used for the benefit of citizens of the developed world. As a result, it provides a greater degree of protection than the narrow interpretation defended by commentators such as Angell. For example, as some critics have recently pointed out, by insisting that the short-course of AZT be compared against the 076 protocol, Angell is in effect insisting that the alternative intervention be compared against a baseline clinical context that is widely available in the developed world and that, if it is practically attainable at all, is nevertheless practically unsustainable in the developing world.²³ The results of such a trial would be more immediately relevant to the clinical context of populations in the US and Europe rather than Africa and South East Asia so that even if it is not conducted with the intent of using developing world populations to answer questions that are most relevant to developed world populations, it may nevertheless have this effect.

This last problem with the narrow interpretation simply reiterates in a more concrete fashion the dilemma that I posed earlier. While physiological equivalence may seem on its face like an elegantly simple and clear interpretation of equipoise, it provides no moral guidance on how to conduct the very research that would be most relevant to populations of the developing world.

By focusing on clinical comparability, rather than biological equivalence, the broader conception of equipoise requires trials to address questions that are specific to the needs of the target population. It does so, however, precisely because it recognizes that the health care needs of developing world populations are intimately bound up with a network of complex social and economic issues. Furthermore, by factoring those issues into the question of equipoise it ensures that they are not neatly swept

²³ R.J. Levine. The Need to Revise the Declaration of Helsinki. *NEJM* 1999; 341: 531–534 at 533.

under the moral rug. By having to articulate the limits to what can be done to improve the baseline situation of a particular people with respect to some proposed intervention, we are forced to confront difficult issues in a way that is public, and makes clear the decisions for which we are accountable. These decisions may not be popular, but public controversy and earnest deliberations are surely preferable to a procedure that also confronts these issues, but in a way that leaves their role unarticulated and hidden from public scrutiny.

The broad interpretation of equipoise thus has a number of important features to recommend it. Focusing on the specific needs of a particular population increases the likelihood of producing actual benefits for that population and decreases the likelihood that research will be conducted in developing world populations solely for the benefit of populations of the developed world. At the same time, however, because it focuses on the degree of clinical comparability that can be achieved between the relevant treatment contexts, it ensures that research targets only problems that cannot be addressed by means that could reasonably be implemented within that population. As a result, it licenses only research that targets those needs of a population that would remain after a reasonable effort has been made to improve clinically relevant aspects of their baseline situation. Finally, by requiring that clinical comparability can be practically sustained over time, this standard ensures that research is designed to address these problems in ways that the nations in question – perhaps with the continued support of third party funding, negotiated discounts on expensive equipment and supplies, or some combination of these – can reasonably be expected to maintain and continue to implement once the research in question has been completed. This will ensure that clinical research is designed in a way that it can be reasonably expected to make a lasting contribution to the medical needs of the populations in which the trial is carried out.

These considerations should make it clear that the broad interpretation of equipoise – unlike its alternative – is not a double standard for clinical research. It is, rather, the same standard applied the same way to substantively different clinical contexts in order to set the same ethical limits on a single end: finding healthcare interventions that can be implemented within conditions that are practically attainable in a treatment population. As a result, it offers an important middle way between two equally unacceptable alternatives. On the one hand, the fact that economically and technologically developed nations

can achieve a very high level of care for the overwhelming majority of their populations should not in itself preclude developing nations from undertaking the very programs of research that would facilitate their ability to better meet the healthcare needs of their own populations. On the other hand, the absence of effective health care interventions within developing world populations should not by itself justify using those populations as subjects of clinical research. The broad interpretation of equipoise alone articulates terms for navigating these extremes in a way that is both scientifically sound and ethically responsible.

DIRECTIONS FOR FURTHER INQUIRY

The arguments of the previous section have been directed at two ends. First, they have outlined significant deficiencies in the narrow conception of equipoise. Second, they have described a broader conception of equipoise that avoids these difficulties and holds out the promise of reconciling the need to find interventions that can be employed in developing world contexts with the requirements of substantive moral constraints that must set ethical boundaries on such research. If the debate over international research standards is to go forward in a way that is coherent and fruitful, much more careful and concerted attention will have to be paid to the details of the broader conception of equipoise outlined here.

In particular, a number of important issues require more detailed treatment than I have been able to provide here. Some are practical and involve our ability to implement this standard in a way that is coherent and principled. For instance, what are the standards that should be used to determine when some set of conditions is practically achievable in a treatment population and when they are not? At what point in the lobbying industry, international organizations, and governments can researchers or their sponsors claim to have made a sufficient effort to bring about possible improvements in the baseline situation of a population such that they may then initiate clinical research? Furthermore, what kind of assurances are necessary to establish the 'reasonable likelihood' that the benefits of international research will be implemented effectively over time in the host population, before that research may proceed? The debate over some of these issues has already begun and against the background of the present discussion their resolution takes on increased importance.

Prior to these more pragmatic issues, however, the broader conception of equipoise requires additional conceptual and technical clarification. In particular, as the number of variables relevant to the effective deployment of an intervention increases, so does the difficulty in reliably projecting our knowledge of the relevant causal relationships from a developed world context, into a developing world context. Overestimating these difficulties, or adopting an overly fastidious standard of proof, would inappropriately expand the boundaries of permissible international research. Underestimating them, however, or adopting an overly permissive epistemic stance, would restrict the boundaries of ethically acceptable research and perhaps result in the deployment of interventions that, although effective in developed world settings, do not actually improve the condition of those who receive them in the developing world. As a result, it is particularly important to find an epistemic standard that strikes an appropriate balance between these two possibilities.

Finally, although I cannot justify this assertion here, a possible implication of this conception of equipoise is that morally acceptable international research cannot take place in a vacuum. That is, to pass the equipoise requirement international research initiatives will have to be coordinated with, or at least responsive to, a nation's larger public health initiatives and political needs. This is an aspect of the broad interpretation that needs to be explored in much greater detail, in part, because it may provide a natural bridge between requirements for the ethical design of international clinical trials, and larger political issues relating to the value of international research as a non-paternalist means of assisting developing nations.²⁴

Alex John London

Department of Philosophy & Center for the Advancement of Applied Ethics

Carnegie Mellon University

135 Baker Hall

Pittsburgh, PA 15213

USA

ajlondon@andrew.cmu.edu

²⁴ Acknowledgment: I would like to thank Tracy E. London and Richard Scheines for their very helpful comments on this manuscript.