

# Learning health systems, clinical equipoise and the ethics of response adaptive randomisation

Alex John London

## Correspondence to

Dr Alex John London, Center for Ethics and Policy, Carnegie Mellon University, Pittsburgh, PA 15213-3890, USA; [ajlondon@andrew.cmu.edu](mailto:ajlondon@andrew.cmu.edu)

Received 30 August 2017  
Revised 6 November 2017  
Accepted 14 November 2017  
Published Online First  
24 November 2017

## ABSTRACT

To give substance to the rhetoric of 'learning health systems', a variety of novel trial designs are being explored to more seamlessly integrate research with medical practice, reduce study duration and reduce the number of participants allocated to ineffective interventions. Many of these designs rely on response adaptive randomisation (RAR). However, critics charge that RAR is unethical on the grounds that it violates the principle of equipoise. In this paper, I reconstruct critiques of RAR as holding that it is inconsistent with five important ethical principles. I then argue that these criticisms rest on a faulty view of equipoise encouraged by the idea that a RAR study models the beliefs of a single rational agent about the relative merits of the interventions being studied. I outline a view in which RAR models an idealised health system in which diverse communities of fully informed experts shrink or grow as their constituent members update their expert opinions in light of reliable medical evidence. I show how a proper understanding of clinical equipoise can reconcile this conception of RAR with these five ethical principles. This analysis removes an in-principle objection to RAR and sheds important light on the relationship between clinical equipoise and transient diversity in the scientific community.

One ambition of learning health systems is to integrate the practice of medicine and the generation of reliable medical evidence in a way that will promote both continuous learning and evidence-based medical practice.<sup>1</sup> This ambition is driving an interest in novel trial designs such as basket trials, umbrella trials, platform trials and trial pipelines that seek to reduce inefficiencies from early-phase research through postmarketing monitoring and clinical evaluation.<sup>2–4</sup> Proponents of these study designs tout their ability to reduce study duration, eliminate delays between research stages and reduce the number of participants allocated to ineffective interventions.<sup>5–9</sup>

Response adaptive randomisation (RAR) is a core component of many of these approaches. Critics of RAR have challenged empirical claims about the merits of this design feature.<sup>10–12</sup> And there appears to be consensus that in the two arm cases, RAR is less efficient than an equal, fixed randomisation allocation (FRA).<sup>13</sup> However, in other settings, such as trials with more than two arms, a shared control group and clinical endpoints that manifest relatively soon after treatment delivery, studies that use RAR appear to have very favourable operating characteristics.<sup>14</sup> As a result, the coming years are likely to see a volley of simulation studies exploring the contexts in which RAR does, or does not, have

a place as a component of efficient clinical trial design.

Apart from this debate about empirical properties of designs incorporating RAR, a more foundational, in-principle criticism holds that RAR is objectionable because it violates the principle of equipoise.<sup>15</sup> This objection, formulated at length by Saxman, is a deeper problem since it would impugn the ethics of the panoply of novel study designs that incorporate RAR, even if those designs have the various efficiencies that their proponents claim.<sup>16</sup> The concern that RAR violates equipoise also illustrates some ethical challenges that arise from more closely integrating research and treatment activities. Understanding whether and how RAR is consistent with clinical equipoise thus has important implications for how we think about learning health systems.

I begin below with a brief explanation of clinical equipoise and five ethical principles with which it is connected. I then reconstruct critiques of RAR as holding that it violates these principles. I argue that these criticisms rest on a faulty view of equipoise encouraged by the idea that randomisation weights in an RAR study reflect the beliefs of a single decision maker about the relative merits of the interventions being studied. In contrast, I show how RAR can be reconciled with these five ethical principles once we understand clinical equipoise as a principle for dealing with conflicting or diverging opinions among diverse groups experts and we see RAR as reflecting how groups of experts in an idealised learning health system would change in size in light of reliable evidence about the relative merits of specific treatment approaches.

Understanding RAR as a model for how idealised groups of experts would change in light of clinical evidence and seeing clinical equipoise as a principle for dealing with transient diversity of opinion among medical experts removes an in-principle objection to RAR. At a deeper level, however, it also illustrates how learning health systems can pursue the promise of eliminating inefficiencies in research and practice while respecting a set of important values.

## THE ROLE OF CLINICAL EQUIPOISE

Charles Fried first introduced the concept of equipoise as a way to resolve a conflict between the scientist's commitment to advancing social welfare by generating reliable medical evidence and the clinician's fiduciary duty to advance the health interests of her individual patient.<sup>17 18</sup> For our purposes, we can frame this as a conflict between two moral principles: promoting research that has social value without compromising a fundamental concern for individual welfare.



To cite: London AJ.  
*J Med Ethics*  
2018;44:409–415.

**Social value:** Scientific research is only justified if it is reasonably expected to generate information that supports the development of better interventions or improvements in clinical practice.

**Concern for welfare:** It is impermissible to knowingly provide a person with an intervention that is known or credibly believed to be worse than another available option.

On Fried's view, equipoise is a state of uncertainty in the mind of the individual investigator regarding the relative merits of interventions A and B for some population of patients. When investigators are in such a state of uncertainty they do not knowingly disadvantage patients if they allow treatments to be allocated by a process that supports reliable medical inference, such as randomisation.

Benjamin Freedman largely accepted Fried's view of the importance of reconciling these values, but argued that Fried's equipoise (which Freedman called 'theoretical equipoise') was too narrow. Freedman argued that it was also permissible to randomise patients when there is 'honest, professional disagreement among expert clinicians' about the relative clinical merits of interventions A and B for a particular patient population.<sup>19</sup> Freedman called this condition 'clinical equipoise', and his contrast between theoretical and clinical equipoise is often recast as individual versus community equipoise.

Clinical equipoise can be operationalised as holding that if there is a set of interventions, each of which would be recommended for a patient P by some group of medical experts, then each of the interventions in this set is a permissible treatment for P.<sup>20–22</sup> So randomising P to receive an intervention in this set does not violate concern for welfare. But if some treatment C is not recommended for P by even a reasonable minority of expert clinicians, then it cannot be included as an option to which P might be randomised. We can codify this prohibition as follows.

**No impermissible gambles:** It is impermissible to directly give intervention C to a person (give it to them with probability 1) when some other intervention A is available, then it is impermissible to include C as an option in a design that would randomise that person to B with any positive probability when A is available.

Finally, Freedman followed Fried in seeing equipoise as a means of resolving a conflict between the demands of specific role-related obligations: the researcher's duty to generate social value and the clinician's fiduciary duty to individual patients. Others have argued that this view of the moral foundation of equipoise is too narrow—not all research is carried out by physicians and the prerogatives of various social roles are rightly limited by larger moral and social values. As a result, London has argued for a view in which clinical equipoise is not grounded in a role-related obligation, but in a general requirement that social roles and institutions must respect the status of people as free and equal through a commitment to equal regard.<sup>21–22</sup> For our present purposes, we can explicate this requirement as follows.

**Equal regard:** All relevantly similar persons should receive the same care and concern—it is not permissible to show less care and concern for the interests of some study participants in order to advance scientific progress or promote the welfare of others.

Some proponents of learning health systems object to designs that maintain an FRA throughout the study on the grounds that they are insufficiently responsive to the welfare interests of study participants. They see RAR as ethically attractive because they claim that updating the randomisation allocation in light of participant outcomes allows the study to apportion more

participants to better performing interventions than designs that maintain an FRA throughout the course of the study.<sup>4–8i</sup> If this empirical claim is true, it reduces in expectation the number of participants exposed to interventions that are ultimately found to be inferior to some alternative. RAR is thus presented as superior to FRA on the following ground.

**Rational expectation:** If in expectation a participant has a greater probability of being allocated to what turns out to be a superior intervention in study design F than in design G, it is rational for that participant to prefer F to G.

A design that satisfies rational expectation is attractive because it is supposed to preserve our commitment to advancing science while giving study participants a better chance of advancing their individual welfare.

### CRITICISMS OF RAR

Critics argue that even if RAR studies begin in clinical equipoise, 'equipoise is disturbed as soon as data are available from the first group of patients enrolled into the study and the randomisation is adapted to favor the "better" treatment arm'.<sup>16</sup> The idea is that an initial 1/x probability of being allocated to one of x interventions reflects equipoise about the relative merits of those interventions.<sup>11 23 24ii</sup> Altering the randomisation allocation in light of participant responses disturbs equipoise because, as proponents of RAR themselves suggest, the updated allocation weights reflect the relative performance of the interventions in question. Once the randomisation weights become unbalanced, the study has a preferred treatment and allocating participants to treatments regarded as inferior violates concern for welfare.

As a result, critics charge that the means that a study uses to satisfy rational expectation are inconsistent with concern for welfare. To satisfy the former, the study uses updated randomisation weights to ensure that more people receive what the study regards as the best performing intervention. But as Saxman puts it 'although in the end more patients may be allocated to the "superior therapy," the trial continues to assign patients to a treatment for which there is an increasing statistical probability that it will prove to be inferior'.<sup>16</sup>

Critics also charge that the dynamic nature of RAR entails that it also violates the principle of equal regard. According to Saxman, 'patients who enter the study early bear more of the risk and burdens of the study than patients who enter later in the trial. This permits inequalities, since there is not an even distribution of risk and benefit across the otherwise equal participants'.<sup>16</sup> In order for RAR to satisfy the principle of rational expectation, more participants must be allocated to better performing arms. This is an empirical claim about how such designs will perform in practice. But allocating more participants to interventions that generate beneficial outcomes favours the interests of those participants over the interests of patients allocated to underperforming arms. The point of this objection is that RAR essentially requires researchers to knowingly show less concern for the interests of participants randomised to underperforming arms than for those allocated to better performing arms.

<sup>i</sup>Advocacy of adaptive designs is predicated on the belief that such novel designs will result in fewer numbers of subjects having to participate and receive an "inferior" treatment during the research process.<sup>15</sup> p. 192.

<sup>ii</sup>Under the equipoise principle, which states that all treatments are likely to be equally effective, subjects are randomized equally across treatments.<sup>11</sup>

Finally, assume that RAR offers a better bet in expectation than a trial with FRA because a >50% chance of receiving a better performing intervention is better than a 50% chance of receiving it. Critics of RAR hold that this involves a violation of no impermissible gambles. The reason is that responsible moral decision making requires performing the act that is judged to be best for a person, all things considered. If it is the considered judgment of an informed, expert clinician that intervention A is 70% likely to be better for you than intervention B, then it is wrong to give that person treatment B. Reducing the probability that they will be allocated to B does not somehow make the possibility of being randomised to B ethically acceptable.

On the reasoning I have laid out here, there appears to be an inherent tension between rational expectation and no impermissible gambles: once the allocation weights are imbalanced, a smaller chance of receiving B is worse than just being given A when that is what an informed, conscientious decision maker believes is the better option.

### FLAWED PRESUPPOSITIONS

The argument of the previous section holds that RAR necessarily violates clinical equipoise and that, as a result, it cannot reconcile the principles of social value, concern for welfare, equal regard, rational expectation and no impermissible gambles. This argument is predicated on two flawed assumptions, both of which seem natural when the randomisation weights in RAR designs are treated as reflecting a single agent's beliefs about the relative merits of the interventions being tested in a study. First is that equipoise is a state of uncertainty in the mind of a single decision maker. Second is that equipoise is a state of belief in which the relevant probabilities are equally balanced.

The idea that a clinical trial can be thought of as an idealised agent is itself a fairly natural view to take when allocation weights are adjusted according to Bayesian updating. In other words, modelling a trial as a Bayesian agent whose preferences over a set of interventions are updated in light of emerging data makes it natural to assume that equipoise must be a property of this agent's beliefs about the relative merits of the interventions in question. It is somewhat ironic, therefore, that when proponents of RAR present randomisation weights in this way, they inadvertently encourage a view of equipoise on which RAR is ethically untenable.

These remarks suggest two possible interpretations of the critics of RAR. On a less generous view, these critics simply adopt a faulty view of equipoise and their criticisms can be addressed by pointing out the flaws in individual equipoise and showing how they are alleviated by adopting clinical equipoise.<sup>iii</sup> However, a more charitable view holds that even if critics regard clinical equipoise as the appropriate conception of equipoise—as Saxman claims—they nevertheless cannot see how to apply it to studies using RAR because these studies effectively model a meta-agent whose beliefs about the relative merits of study interventions guides treatment allocation. A proper response to this objection requires demonstrating how RAR is consistent with clinical equipoise. Before showing how this is the case, therefore, it is worth being explicit about three problems with the view of equipoise as a state of equally balanced beliefs of a single agent.

<sup>iii</sup>The idea that equipoise requires individual uncertainty and an equal balance of probabilities persists in the literature. For example, Eyal and Lipsitch define equipoise in these terms and then argue that it should be rejected as a requirement for research.<sup>33</sup> This illustrates the importance of clearly demonstrating the operational utility of clinical equipoise, as I am doing here.

First, treating equipoise as a state of equally balanced beliefs of a single agent creates a standard so strict that it prevents initiating a wide range of valuable research. In cases such as the controversy over high-dose chemotherapy with autologous bone marrow transplant for end-stage breast cancer, many clinicians were convinced that this aggressive therapy was superior to standard chemotherapy while other clinicians regarded it as more likely to be harmful than beneficial. If no clinician is uncertain in the sense required by individual equipoise (ie, believing that the probabilities of success for each approach are 50:50), it would not be permissible to conduct a randomised controlled trial of any kind to settle the conflict in expert opinion within the clinical community. In fact, individual equipoise prohibits all research in which different clinicians have definite but conflicting preferences for particular interventions—including established interventions and novel interventions.<sup>iv</sup> Clinical equipoise was proposed as a standard precisely because there is social value in resolving such conflicts among expert clinicians and randomisation in the face of such conflicting opinions is consistent with respect for welfare.

Second, as Freedman<sup>19</sup> noted 30 years ago, even if such research could be initiated, individual equipoise is so fragile that it would require the study—regardless of the strategy for randomisation—be terminated long before sufficient evidence has been generated to change practice in the clinical community. The reason is that early evidence will move the individual decision maker's beliefs off of 1/n. But this view of equipoise is incapable of reconciling concern for welfare with social value regardless of the randomisation strategy used because studies will have to stop well before sufficient evidence has been generated to alter clinical practice.

Third, an unrecognised problem with individual equipoise is that it prohibits socially valuable research and permits research that has little social value. The reason is that when clinicians have solid evidence that *n* interventions are of equivalent clinical value, a 1/*n* probability of receiving any such intervention represents a state of equipoise and satisfies the conditions of concern for welfare and equal regard.<sup>v</sup> If clinicians' beliefs about the equivalency of these interventions are true, such studies could be run to completion. But in the face of solid prior evidence of equivalent clinical utility, clinical research that merely confirms what the clinical community already knows is of little value and a waste of resources. This is a problem, in part, because the condition of equipoise should help to establish that running a trial has social value such that disturbing equipoise is likely to lead to an improvement in clinical practice.

### COMMUNITIES OF EXPERTS, CLINICAL EQUIPOISE AND RAR

If individual equipoise is morally flawed, then it remains to demonstrate how RAR is consistent with clinical equipoise and how this interpretation avoids the above pitfalls. On the view I

<sup>iv</sup>It might be argued that even if clinicians have strong preferences in favour of particular treatments they may nevertheless be humble enough to recognise that they might be wrong. But in cases where clinicians lack this humility, the original objection still applies. The deeper point is simply that individual uncertainty is not necessary for randomisation to be ethically permissible.

<sup>v</sup>Lilford and Jackson state, 'Equipoise is the point where there is no preference between treatments, i.e. it is thought equally likely that treatment A or B will turn out to be superior... At this point we... would take odds of 1:1 in a bet'.<sup>24</sup> This presentation of equipoise is consistent with a state of indifference between A and B that is grounded in solid evidence of their clinical equivalence. Hence, this view is open to the objection stated here.

am proposing here, a clinical trial with RAR should be seen as modelling an idealised health system in which diverse communities of fully informed experts who disagree about the relative merits of a set of interventions shrink or grow as their constituent members update their expert opinions in light of reliable medical evidence. In this view, randomisation weights are not the beliefs of *any* agent; they are an idealised representation of the probability that a patient in such an idealised learning health system would encounter a practitioner from these communities if they were to be allocated to a clinician at random.

To explain this account, consider a case in which there is a medical condition with three alternative interventions: A, B and C. Among medical experts, there is a practice community that favours intervention A over B and C. There is also a community of experts that favours intervention B over A and C and a third community that prefers C over A or B. For present purposes, the relevant point of the example is that the members of these practice communities are each experts in the mainstream medical community, familiar with the relevant scientific evidence, recognised as providing diligent, expert medical care.

This diversity within the larger medical community means that when people fall ill with this medical condition, some patients encounter practitioners from community A, while others encounter practitioners from community B or C. If it is permissible for a patient to be treated in clinical practice by practitioners from each of these clinical communities, then it follows that A, B and C are all admissible treatment options for that patient.<sup>20–22</sup> That is, when experts have definitive, but opposing judgments about which intervention is best for a patient, it does not violate concern for welfare to follow the recommendation of one set of experts, even though different experts would disagree with that recommendation.

If it is consistent with concern for welfare for a patient to be directly treated with A or B or C (to receive that intervention with certainty), then it cannot violate concern for welfare if that patient is assigned to those interventions with any distribution of probabilities that sums to 1. Even though every clinician in these treatment communities has a strict preference over the available treatment options (nobody thinks the probability of success for each is 1/3), clinical equipoise exists between these treatment options, and no set of randomisation weights that sums to unity is impermissible.

This shows that the key consideration at the initiation of a study is not that it offers participants a 1/n chance at receiving one of n interventions but that each of the interventions in the study is an admissible intervention for the particular patients randomised into the study. In the present case, the notion of ‘admissible intervention’ is operationalised with the condition that there exists at least one informed, expert practice community that would recommend that intervention for that patient.<sup>20–22</sup>

Because this last point is somewhat subtle, it warrants a brief further comment. The upshot of the previous remarks is that if n interventions are admissible for a patient, then any randomisation scheme that includes only those n interventions (where the probabilities sum to unity) is consistent with clinical equipoise. This is a claim about *prima facie* ethical permissibility. From this claim, it follows only that there is no in-principle ethical objection grounded in equipoise to any randomisation scheme that meets the above conditions. This includes adaptive randomisation, fixed 1/n randomisation or fixed but imbalanced randomisation schemes such as 1:2:2. From this claim, it does *not* follow that there are no other grounds for ethical or methodological concern with some of these allocation schemes. The operating characteristics of every trial should be carefully evaluated to

ensure that they are as efficient and rigorous as possible. The upshot of the previous point is that equipoise does not provide a shortcut that identifies some trial designs as impermissible on the basis of the randomisation weights they use, thereby forestalling the need for such a careful, case-by-case, evaluation of the operating characteristics of different designs. As I explain below, the same reasoning applies for the principles of social value, concern for welfare, equal regard, rational expectation and no impermissible gambles.

Once an initial block of participants has been randomised, the allocation probabilities in RAR are adjusted on the basis of the observed outcomes. Critics hold that equipoise no longer exists in this situation, but I have argued that this is because they view allocation probabilities as the beliefs of a single agent, they treat equipoise as referring to the state of those beliefs and they take equipoise to require a state of belief that assigns equal probability to the merit of each intervention. On the view I am defending here, clinical equipoise remains, even when the allocation weights are adjusted to favour one intervention over others, so long as every participant still receives an intervention that is regarded as best for them by at least one well informed, medically expert treatment community.

To see how this is possible, consider that treatment communities do not change their practices all at once. Experts within these communities have different prior beliefs about disease pathology, mechanism of action and treatment properties. As a result, their views about the relative merits of the interventions in question change at different rates when confronted with the same evidence. While some clinicians may alter their practice on relatively little evidence, others will not, and it may take significant evidence from multiple sources to persuade the most committed members of that community. So, after seeing outcomes of the initial block of patients that favour one intervention (eg, A) over the others, some members of the communities that favour B or C may change their treatment recommendation for a particular patient, but others in that community will not. As a result, there are still members of practice communities that favour B or C who would recommend those interventions over A for their patients. As long as this is the case, interventions A, B and C remain admissible options for patients with this condition and it is permissible to randomise that patient to interventions in the set of admissible interventions.

As evidence emerges, treatment communities whose recommendations are supported by observed outcomes accrue adherents while those whose recommendations are not supported by evidence lose them. The relative size of these different communities of practitioners is reflected in the trial’s randomisation weights. This is the sense in which they do not reflect the beliefs of any agent, but instead represent the probability that a patient in such an idealised health system would encounter a treating physician from one of these diverse communities if their treating clinician were allocated at random.

In this model, judgments about how to treat patients are left to medical experts whose recommendations reflect the totality of the available medical evidence. Because there is diversity among such experts, there is no single agent whose views of all of these experts is treated as sacrosanct and elevated to special prominence. Random allocation reflects a reasonable response to such a situation of diversity.

A study that begins in clinical equipoise, in the sense defined here, satisfies concern for welfare because every participant in the trial receives an intervention that is regarded as the best available option by experts from at least one informed, expert practice community. It satisfies equal regard because no participant is

shown a lesser degree of concern than would be shown for their interests if they were treated by a representative of the community that prefers the intervention to which they are ultimately randomised. In other words, no participant in the study is allocated to a treatment that everyone recognises as inferior to some available option in order to advance science.

Altering the randomisation weights does not violate concern for welfare, equal regard or no impermissible gambles as long as every participant receives an intervention that is regarded as best for them by at least one treatment community. This is consistent with updating randomisation weights in light of the accumulation of medical evidence because even if rational inquirers recognise that initial evidence from a clinical trial supports the clinical merits of one intervention (A) over the others (B or C), that evidence may not be strong enough to lead responsible experts to alter their recommendation, or to alter the recommendation of every expert in that community. As a result, there will still be experts within those communities that continue to regard B or C as preferable to A. As long as this is the case, B and C remain admissible interventions and the study remains in clinical equipoise.

As a result, it can be the case that a trial that satisfies no impermissible gambles also satisfies rational expectation. To see this, consider two alternatives. In the first, as critics of RAR suggest, when evidence from the initial block of participants favours A over B or C, the trial is terminated. Some clinicians who once recommended B or C will change their practice and favour A, but others will not. As a result, the clinicians that favour B or C will continue to provide those interventions to patients in the health system without the opportunity to gather reliable evidence to clearly distinguish their clinical merits. In the second scenario, the weights in an RAR study are updated to favour A, but because some clinicians in B and C continue to recommend those treatments for patients over intervention A, the study continues.

In both scenarios, patients in the health system continue to receive A, B or C, but in the second scenario, patients are allocated to interventions in a way that generates reliable medical evidence. Generating this evidence allows more members of the relevant medical communities to change their practice. However, because in an ideal health system the probability of encountering members of a community shrinks as evidence against their recommendations grows, patients in a RAR study have a better chance of being treated with what is ultimately recognised as the best treatment for their condition, satisfying rational expectation. As long as shrinking communities are regarded as medically expert and are permitted to treat patients, including the interventions they recommend in a trial satisfies no impermissible gambles.

Finally, when fully informed medical experts have conflicting judgments about which interventions are likely best for a particular indication, a study that clarifies the relative clinical merits of those interventions has a strong *prima facie* claim to social value. The reason is that the trial is a necessary step in reducing unwarranted variation in clinical practice and potentially improving the care of a substantial proportion of the patients with the disease in question.

## OBJECTIONS

It might be objected that the model I have outlined here is unstable for reasons that Leonard Savage attributed to Woodbury.<sup>25</sup> That is, as evidence accumulates favouring one option over the others, and more clinicians join the favoured practice community, we can regard the randomisation weights as

expressing the preference of the medical community for the best performing intervention. In effect, taking a weighted average of the opinions of the different treatment communities generates a meta-agent whose views should guide treatment recommendations. In this case, the objection holds that the view I have proposed collapses back into a single agent whose treatment recommendations violate equipoise as soon as randomisation weights move off of  $1/n$  for  $n$  interventions.

This objection is alluring because it appeals to the idea that rational inquiry requires an agent and it treats randomisation weights as though they are the beliefs of a social agent. This social agent takes diverse experts in the community and aggregates their judgments into a higher-order decision model. In essence, it assigns a weight to the likelihood that each expert is correct and then chooses in a way that maximises expected value. This is a concrete example, in microcosm, of a larger view of scientific consensus that many find intuitively appealing, namely, that the goal of scientific consensus is to take in the diversity of beliefs in the scientific community, assign them weights and form a single all-things-considered model out of this diversity.

This proposal faces several problems. To begin with, it is a version of the linear opinion pooling rule for combining individual judgments into a group or social judgment. But the social agent constructed by assigning weights to the views of individual experts can make recommendations that conflict with the recommendations of each of the experts from which it is created. For example, each expert may regard certain events (the temperature in Beijing today and whether to use treatment A or B for a certain patient in New York) as probabilistically independent and, as a result, would not base treatment decisions on what he or she recognises as an irrelevant event (no medical expert will decide the merits among rival treatments for a patient who resides in New York by asking what the weather is that day in Beijing). But these relationships of probabilistic independence are not generally preserved in the linear opinion pool. As a result, the 'social agent' can change its treatment recommendations on learning the weather in Beijing, even though no particular expert would do so.<sup>vi</sup>

In addition to the defects of the judgments of this social agent, there is the additional problem that the social model is not normative for practice communities or their individual members. From the fact that this group agent is no longer indifferent between interventions it does not follow that any particular expert should or would alter their treatment practice, let alone that all experts should or would. After all, every expert is assumed to know *ex hypothesi* that many other, equally well-credentialed and informed experts hold treatment views that conflict with their own. As a result, if studies are stopped when this agent comes to have a slight preference for one intervention, studies would be terminated long before members of the relevant practice communities are willing to alter their practice. This would undercut the social value of research without

<sup>vi</sup>Genest and colleagues establish that in order to be 'Externally Bayesian' the pooling rule has to be of the 'logarithmic' form.<sup>34</sup> For the purposes of the present argument, it is sufficient to see that one aspect of the Externally Bayesian condition is that when experts regard two events (A & B) as independent, this condition requires preserving the expert judgments that each event is irrelevant to the probability for the other ( $P(A | B) = P(A)$ ), after pooling. The proposal discussed above involves creating a social consensus model by taking weighted averages of different treatment communities. But weighted averages are convex combination and as Seidenfeld and colleagues show, a convex combination of expert judgments is not a logarithmic rule.<sup>35</sup> As a result, the social agent discussed above cannot preserve the judgments of experts about which events are relevant to treatment decisions.

advancing anyone's welfare since it would still be the case that members of practice communities that prefer what appear to be disfavoured interventions will continue to regard them as best for their patients.

Finally, there is the problem that, given the impoverished nature of our understanding of the underlying causal structure of health problems, experts have a difficult time predicting which theories of disease or interventions are likely to be correct or best.<sup>26</sup> In this environment there are significant dangers of group-think, the situation in which every expert quickly converges to the same view of a problem. In part, this is because communities that behave this way are prone to accepting incorrect answers on the basis of spurious results that are bound to happen as a matter of chance. As a result, communities with more diverse opinions among experts are healthier and more productive in the sense that they are less prone to converging on false answers and more efficient at exploring alternatives and thereby locating better solutions to pressing problems.<sup>27–31</sup>

The approach that I have outlined here has the distinct advantage of recognising that reasonable, transient diversity among experts is not simply a descriptive feature of many actual scientific communities, but a normatively desirable feature that plays an important epistemological role in the health and fecundity of those communities. In this approach, the role of clinical research is to generate reliable medical evidence on the basis of which such experts are likely to alter their opinions. Recognising that the informed, scientific judgments of experts in diverse communities change at different rates allows the approach I have outlined to continue inquiry as long as those communities of experts are regarded as falling within the sphere of the responsible practice of medicine.

At this point, it might be objected that the approach I have described here requires an account of when we should stop regarding a minority of the medical community as reasonable and view their treatment preferences as no longer a part of the standard of care. This is indeed an important and pressing problem. But it is one that we face regardless of the conception of equipoise that we adopt. Moreover, this process is complicated by the fact that in some cases, minority opinions turn out to be correct and the received wisdom is wrong.

Nevertheless, it is an advantage of RAR that it can make such difficult social decisions more tractable by explicitly representing alternative perspectives within a single trial. In the model I am presenting here, RAR studies should reflect the way an idealised community of diverse, but fully informed scientific experts would alter their opinions in light of reliable medical evidence. It is an advantage of the model I present here that it highlights the importance of ensuring that the beliefs of the idealised communities reflected in the RAR study design capture the enthusiasm of some real-world clinician/researchers for novel interventions as well as the more conservative or sceptical views of other experts. Explicit decisions can then be incorporated into the trial about when a community's views should be regarded as no longer reflecting the practice of responsible medicine.

Finally, it might be objected that although equipoise may be sufficient for ensuring that the risks of a study are morally acceptable, it is not a necessary condition and that it therefore would not matter if RAR did systematically violate that requirement.<sup>32</sup> The main problem with this objection is that it is too quick to assume that conflicts between the principles described here are inevitable and that all clinical research must necessarily rely on the altruism of study participants. Even if it is true that research must ultimately rely on the altruism of participants, it is important that we not demand more altruism of people than

necessary. Furthermore, the view that I have proposed takes important steps to linking certain features of clinical research to the behaviour of well-functioning learning health systems. Understanding the symmetries between diversity in clinical practice in health systems and equipoise in clinical research is an important step in clarifying the ethical norms capable of reconciling the tensions inherent in learning health systems.

## CONCLUSION

The arguments presented here defuse the moral objection that RAR essentially involves a violation of clinical equipoise. They do not address difficult empirical questions about when, if ever, RAR represents a more efficient approach to clinical trials than traditional FRA designs.<sup>13</sup> Nevertheless, interpreting RAR as modelling the behaviour of an idealised learning health system and showing how clinical equipoise reconciles that design with five important ethical principles demonstrates the relevance of this requirement to efforts to improve the efficiency of both research and clinical practice.

**Acknowledgements** I thank Teddy Seidenfeld, Jonathan Kimmelman, Kert Viele, Susan Ellenberg, Kevin JS Zollman and Lindsay Renfro for their criticisms and advice.

**Contributors** AJL conceived and wrote this paper.

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## REFERENCES

- Institute of Medicine. *Rewarding provider performance: aligning incentives in medicare, pathways to quality health care*. Washington, DC: The National Academies Press, 2007.
- Renfro LA, Sargent DJ. Statistical controversies in clinical research: basket trials, umbrella trials, and other master protocols: a review and examples. *Ann Oncol* 2017;28:34–43.
- Trushheim MR, Shrier AA, Antonijevic Z, et al. PIPELINES: creating comparable clinical knowledge efficiently by linking trial platforms. *Clin Pharmacol Ther* 2016;100:713–29.
- Lewis RJ. The pragmatic clinical trial in a learning health care system. *Clin Trials* 2016;13:484–92.
- Berry DA. Adaptive clinical trials: the promise and the caution. *J Clin Oncol* 2011;29:606–9.
- Harrington D, Parmigiani G. I-SPY 2—A glimpse of the future of phase 2 drug development? *N Engl J Med* 2016;375:7–9.
- Lin J, Lin LA. A General Overview of Adaptive Randomization Design for Clinical Trials. *J Biom Biostat* 2016;07:294.
- Meurer WJ, Lewis RJ, Berry DA. Adaptive clinical trials: a partial remedy for the therapeutic misconception? *JAMA* 2012;307:2377–8.
- Saville BR, Berry SM. Efficiencies of platform clinical trials: a vision of the future. *Clin Trials* 2016;13:358–66.
- Freidlin B, Korn EL. Ethics of outcome adaptive randomization. *Wiley StatsRef: Statistics Reference Online* 2016 (Epub ahead of print: 15 Feb 2016).
- Lee JJ, Chen N, Yin G. Worth adapting? Revisiting the usefulness of outcome-adaptive randomization. *Clin Cancer Res* 2012;18:4498–507.
- Thall P, Fox P, Wathen J. Statistical controversies in clinical research: scientific and ethical problems with adaptive randomization in comparative clinical trials. *Ann Oncol* 2015;26:1621–8.
- Hey SP, Kimmelman J. Are outcome-adaptive allocation trials ethical? *Clin Trials* 2015;12:102–6.
- Cellamare M, Ventz S, Baudin E, et al. A Bayesian response-adaptive trial in tuberculosis: the endTB trial. *Clin Trials* 2017;14:17–28.
- Laage T, Loewy JW, Menon S, et al. Ethical considerations in adaptive design clinical trials. *Ther Innov Regul Sci* 2017;51:190–9.
- Saxman SB. Ethical considerations for outcome-adaptive trial designs: a clinical researcher's perspective. *Bioethics* 2015;29:59–65.
- Fried C. *Medical experimentation: personal integrity and social policy*. Amsterdam: North-Holland, 1974.
- London AJ. Clinical Equipoise: Foundational requirement or fundamental error? Steinbock B, ed. *The oxford handbook of bioethics*. Oxford, UK: Oxford University Press, 2007:571–95.
- Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med* 1987;317:141–5.

- 20 Kadane JB. *Bayesian methods and ethics in a clinical trial design*. New York: Wiley, 1996.
- 21 London AJ. Reasonable risks in clinical research: a critique and a proposal for the Integrative Approach. *Stat Med* 2006;25:2869–85.
- 22 London AJ. Two dogmas of research ethics and the integrative approach to human-subjects research. *J Med Philos* 2007;32:99–116.
- 23 Thall PF. Ethical issues in oncology biostatistics. *Stat Methods Med Res* 2002;11:429–48.
- 24 Lilford RJ, Jackson J. Equipoise and the ethics of randomization. *J R Soc Med* 1995;88:552–9.
- 25 Savage LJ. *The foundations of statistics*. New York: Dover Publications, 1972.
- 26 Kimmelman J, London AJ. Predicting harms and benefits in translational trials: ethics, evidence, and uncertainty. *PLoS Med* 2011;8:e1001010.
- 27 Hull DL. *Science as a process: an evolutionary account of the social and conceptual development of science*. Chicago Illinois: The University of Chicago Press, 2008.
- 28 Kitcher P. *The advancement of science: science without legend, objectivity without illusions*. New York: Oxford University Press, 1995.
- 29 Solomon M. Scientific rationality and human reasoning. *Philos Sci* 1992;59:439–55.
- 30 Muldoon R. Diversity and the division of cognitive labor. *Philos Compass* 2013;8:117–25.
- 31 Zollman KJS. The epistemic benefit of transient diversity. *Erkenntnis* 2010;72:17–35.
- 32 Miller FG, Brody H. A critique of clinical equipoise: therapeutic misconception in the ethics of clinical trials. *Hastings Cent Rep* 2003;33:19–28.
- 33 Eyal N, Lipsitch M. Vaccine testing for emerging infections: the case for individual randomization. *J Med Ethics* 2017:1–7.
- 34 Genest C, McConway KJ, Schervish MJ, *et al*. Characterization of externally Bayesian pooling operators. *The Annals of Statistics*:314–32.
- 35 Seidenfeld T, Schervish MJ, Kadane JB. Coherent choice functions under uncertainty. *Synthese* 2010;172:157–76.