

CEB Working Paper

Centre Emile Bernheim
Research Institute in Management Sciences

 **Solvay Brussels School**
Economics & Management

Ethics of Randomized Controlled Trials: Should Economists Care about Equipoise?

Michel Abramowicz and Ariane Szafarz

Equipoise is defined by Freedman (1987, p.141) as a "state of genuine uncertainty on the part of the clinical investigator regarding the comparative therapeutic merits of each arm in a trial." This principle is grounded in the ethical motivation that any ex-ante preference for a given option would undermine the interests of those who are offered another. Randomized controlled trials (RCTs) in development economics disregard the equipoise requirement by typically disadvantaging the control group. This paper investigates how the equipoise principle is formalized in the medical literature and discusses whether and how it should be taken into consideration by economists. It argues that equipoise is especially relevant when double (or even single) blindness is excluded and when the control group includes already vulnerable individuals. More generally, this paper advocates for developing a vibrant ethics conversation on the design and fairness of RCTs in social sciences.

Keywords: Equipoise, fairness, ethics of RCTs, control group, placebo

JEL Classifications: C93, I14, B41, O16, D63

CEB Working Paper N°19/017
August 2019

Ethics of Randomized Controlled Trials: Should Economists Care about Equipoise?*

Michel Abramowicz and Ariane Szafarz**

July 2019

**This article is forthcoming in Bédécarrats, F., Guérin, I., & Roubaud, F. (Eds.)
Randomized Control Trials in the Field of Development: A Critical Perspective, Oxford
University Press.**

Keywords: Equipoise, fairness, ethics of RCTs, control group, placebo
JEL Classifications: C93, I14, B41, O16, D63

* The authors thank Cécile Abramowicz, Britta Augsburg, Marie Brière, Andres Garchitorena, Marek Hudon, Marc Labie, Jonathan Morduch, Tim Ogden, Martin Ravallion, the participants to the AFD Workshop *Randomized Control Trials in the Field of Development: The Gold Standard Revisited* (Paris, March 2019) and the three editors of the book, Florent Bédécarrats, Isabelle Guérin, and François Roubaud, for valuable comments and discussions.

** Corresponding author: Université Libre de Bruxelles (ULB), Solvay Brussels School of Economics and Management, CEB & CERMi, 50, av. F.D. Roosevelt, CP 114/03, 1050 Bruxelles, BELGIUM, Email: aszafarz@ulb.ac.be, Phone +32 2 650 48 65.

Ethics of Randomized Controlled Trials: Should Economists Care about Equipoise?

Abstract

Equipoise is defined by Freedman (1987, p.141) as a "state of genuine uncertainty on the part of the clinical investigator regarding the comparative therapeutic merits of each arm in a trial." This principle is grounded in the ethical motivation that any ex-ante preference for a given option would undermine the interests of those who are offered another. Randomized controlled trials (RCTs) in development economics disregard the equipoise requirement by typically disadvantaging the control group. This paper investigates how the equipoise principle is formalized in the medical literature and discusses whether and how it should be taken into consideration by economists. It argues that equipoise is especially relevant when double (or even single) blindness is excluded and when the control group includes already vulnerable individuals. More generally, this paper advocates for developing a vibrant ethics conversation on the design and fairness of RCTs in social sciences.

Randomized controlled trials appear to annoy human nature—if properly conducted, indeed they should.

Schultz (1995: 1456)

1. Introduction

Is lack of resources a good reason for providing the treated and control groups in randomized controlled trials (RCTs) with unequally promising options? Apparently, physicians answer “no” to this question but economists tend to say “yes.” Equipoise is an inescapable building block of medical RCTs. Strangely, many economists performing RCTs never heard about it. This paper fills the gap and investigates how the equipoise principle is formalized in the medical literature, and subsequently whether and how it should be taken seriously into consideration by economists.

Equipoise is defined by Freedman (1987, p.141) as a "state of genuine uncertainty on the part of the clinical investigator regarding the comparative therapeutic merits of each arm in a trial." The author considers this principle as "an ethical necessary condition in all cases of clinical research." Equipoise requires that before the trial starts, there is equal ignorance about the benefits and drawbacks of the treatment options. This requirement is grounded in the ethical motivation that any ex-ante preference for a given treatment option would undermine the interests of those who are offered another. And since the typical medical procedure in RCTs relies on double-blind treatment allocation, failing to fulfill equipoise would potentially hurt all the trial participants. In that sense, the equipoise requirement reinforces the 1964 World Medical Association's Declaration of Helsinki¹ stating, among other things, that control groups

¹ Carlson et al. (2004) discusses the later revisions of this Declaration. See also the 1982 International Ethical Guidelines of the Council of International Organizations of Medical Sciences for Biomedical Research Involving Human Subjects (CIOMS, 2002).

must receive the best existing treatment. This requirement is absolute, i.e., it applies regardless of the study's specific conditions, including its location.

Yet, the topic of equipoise is still controversial as its practical implementation raises key issues such as the balance between the opinions and preferences of the clinical community, the individual investigator, and the treated patient (Lilford & Jackson, 1995). Evidently, the appreciation of therapeutic merits may vary according to the sensitivity of the people involved in the implementation of the trial, and so lead to ethical dilemmas (Schafer, 1982), such as the thorny question of balancing a doctor's duties to her patient and to the advancement of science (Botros, 1990). Miller and Joffe (2011) however contest that the stakes are confined within the doctor-patient relationship. The authors place the debate in a wider context that relates to health policy. They weigh the interest of individual patients against the knowledge needed for drug approval. By doing so, they link equipoise to the typical public-health trade-off that opposes individual liberties to social justice (Kass, 2001; Childress et al., 2002). In at least this respect, there is a clear connection between the ethics of RCTs in the fields of medicine and economics.

While the medical literature fiercely debates the relevance of various specifications of the equipoise principle, research in economics is still silent on the topic. Of course, RCTs in economics are usually reviewed by ethical committees. Yet, these committees are typically local. Large-scale ethical requirements, including reference to (any sort of) equipoise, are still missing. We aim to break the apparent indifference of economists to an ethical concern that is key for medical experimentation. In line with Baele (2013) and Petticrew et al. (2013) who advocate the development of "social equipoise," this paper intends to initiate an equipoise conversation within the economic RCT community.

2. What is Equipoise?

The use of human beings as experimental subjects has created difficult ethical problems. The practice of assessing medical treatments with controlled experiments dates back from ancient times, but the equipoise principle is far more recent. It was formalized in the 20th century following the designing of randomization with placebo control groups and concealed assignments (Di Tillio et al., 2017).

Most modern codes of medical ethics are guided by the principles of the classical Hippocratic Oath devoted to the obligations of physicians to their patients. (Orr et al., 1997; Miles, 2005). A central tenet of the Oath concerns the duty of providing the best available treatment. Specifically, if the doctor has good reasons² to believe that treatment A is better than treatment B, then she cannot prescribe B instead of A to any of her patients (Shaw & Chalmers, 1970). Likewise, she should refrain in participating in any scientific study that would lead to giving treatment B rather than treatment A. This strong restriction imposed by medical ethics can hinder the development of large-scale medical studies based on the comparison of treatments A and B. To address this issue, Freedman (1987) introduced the concept of *clinical equipoise* that embeds the need of sufficient statistical evidence to conclude that treatment A does not dominate treatment B.³ The underlying idea is to place the so-called genuine uncertainty about the comparative therapeutic merits of the two treatments in the hands of the expert medical community rather than in the hands of an individual investigator (Freedman, 1987).

² The doctor's beliefs mix scientific knowledge with subjective experience and personal thinking. The subjective component inevitably adds complexity to formalizing the equipoise principle.

³ Development economists could argue that some poor populations have access neither to treatment A nor to treatment B, so that even treatment B would improve their condition. Section 4 addresses this argument often used to justify the inferior treatments provided to the control groups. We content that comparing life under a controlled experiment to regular life conditions ignores the role of the investigators, who can affect people's behaviors and feelings significantly. In that regard, RCTs in social sciences are on an equal footing with medical RCTs, which should prevent them from adhering to lower ethical standards.

Without affecting the fundamentals of RCTs, clinical equipoise helped relax practical constraints that had sometimes dictated stopping studies prematurely when early results indicated that one treatment is better than the other, at least in the short term. Freedman's idea was to leave enough time to the scientific community to build strong evidence based on large studies. Meanwhile, Freedman argues, clinical equipoise addresses, at least partially, the low take-up associated with the reluctance of physicians to enroll their patients in studies they do not feel comfortable about (Taylor et al., 1984).

Overall, the operational principle of clinical equipoise proved itself fruitful to the development of large-scale medical studies by contributing to the design and applicability of RCTs. By so doing, equipoise helps reconcile the rights of study participants and the quest for scientific breakthroughs (London, 2017).

The initial steps of the implementation of the equipoise concern into medical RCTs can inform the economists about the challenges and stakes involved. The cardiological clinical and investigational community was the first to implement large-scale medical RCTs aiming to inform the practical dispensing of patient care. One such pioneering study was the Beta-Blocker Heart Attack Trial (BHAT, 1982). Beta-blockers are heart drugs developed in the 1960s, whose discovery was rewarded in 1988 by the Nobel Prize for Medicine granted to Sir James Black. These drugs were initially popular for the treatment of hypertension. The BHAT trial randomized survivors of a recent myocardial infarction to either Propranolol, the first widely available beta-blocker, or placebo (Yusuf et al., 1985). The results showed a very significant reduction (7.2% vs. 9.8%) in medium-term total mortality (the average follow-up was 24 months). To this day, 36 years after the publication of this seminal study, beta-blockers are still the cornerstone of secondary prevention after myocardial infarction. Beta-blockers are also the most active drugs against angina pectoris—the chronic painful condition caused by partially

blocked cardiac arteries, a total blockage typically resulting in an infarction—and they significantly reduce mortality in heart failure (McMurray, 2010).

Before embarking on the BHAT, there was genuine equipoise among the clinical and scientific medical community about the potential protection offered by beta-blockade after a myocardial infarction (Nies et al., 1973; Shand, 1975). Animal experiments had shown enhanced survival, and patients with angina pectoris fared well with the drug, but there was concern that the associated blood pressure lowering would result in a net harm (Theroux et al., 1974).

Yet, the level of uncertainty about treatment superiority can change during the course of an RCT for several reasons, including the partial results of the trial itself and the publication of meaningful results by other teams. The data and safety monitoring boards (DSMBs) have the responsibility of determining whether “equipoise has been sufficiently disturbed during the course of a trial to warrant stoppage” (Dickert & Emanuel (2015; 31). This is a difficult decision to make since, on the one hand, early stoppage can harm the overall validity of the study, and on the other hand, the continuation of the study with a compromised equipoise can harm its subjects. For example, there were rumors that the Second International Study of Infarction Survival (ISIS-2, 1988) ignored the ongoing findings published in 1987 by an Italian competing research team, the Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico (GISSI). Precisely, the ISIS-2 trial investigated 17,187 patients admitted in the Coronary Care Unit (CCU) with a working diagnosis of acute myocardial infarction (AMI). The ISIS-2 patients were randomized between Streptokinase, a clot-dissolving drug that was hoped to better the prognosis by reducing the infarct size, and placebo. In the course of the ISIS-2 study, GISSI partial results strongly favoring Streptokinase began to emerge suggesting that ISIS-2 was unfair toward its placebo patients, i.e. the control group, with respect to their treated counterparts. Yet, this episode unfolded between 1985 and 1988, a period during which the

equipoise concept was still little known in the medical research community. Ultimately, the ISIS-2 study went on as planned and its results confirmed the substantial mortality benefits of Streptokinase. Nowadays, 30 years after its youth waywardness, equipoise belongs to the core ethical standards of medical RCTs.

3. Equipoise vs. Blindness

Even though there is a large consensus in the medical community that equipoise belongs to the ethical standards of the profession, the practical implementation of this principle raises several practical issues. The most basic problem, and perhaps the most challenging one in the field, is how to prove that a given trial satisfies clinical equipoise. Several methods can be used to bring convincing evidence. They include references to previous studies and testimonies of divergences about treatments within the clinical community. Specific conditions can however compromise the implementation of clinical equipoise. This section comments on two major issues that may sound familiar to social scientists: the lack of experimental blindness and the decision to enroll a patient with a pre-existing condition.

Double blindness has become the norm of medical RCTs. In economic RCTs, by contrast, it is hardly implemented, for alleged reasons of practical unfeasibility. This section shows that, like in economics, implementing blindness is not always feasible in medical subfields, such as surgery.⁴ But in line with the medical literature, we content that even with partial or no blindness, the principle of equipoise should be followed. For instance, the GISSI (1987) study was an unblinded RCT “of which protocol specified three interim analyses, at 3000, 6000, and 9000 recruited patients. Results from these were presented to the ethics committee only; a difference in mortality exceeding three standard deviations or an unacceptably high incidence

⁴ Young et al. (2004) suggest assessing the feasibility of equipoise-based RCTs in surgery by organizing opinion polls among surgeons in the field. Such polls would gauge the possibility of testing a surgical treatment against another surgical treatment.

of adverse reactions to SK would have led the committee to call an early halt to the trial” (GISSI 1987, p. 398). In fact, the absence of experimental blindness makes equipoise an even more imperative requirement.

Nowadays, medical RCTs are typically conducted in a blinded fashion. In a single-blinded study, the patient is unaware of which treatment she receives. Double-blindness means the subject and the field investigator (who treats the patients, follows them and transmits their end-points results to the trial steering committee) are both unaware (Days & Altman, 2000). The purpose of blindness is to limit bias, since knowing which treatment is applied could result in unconscious misinterpretations of the results. Also, the subjective experience of trial patients can be colored not only by their knowledge of which drug they are receiving, but also by their feelings about their doctor’s expectations. Patients often think that the experimental medication will perform better than the reference drug or the placebo—even though with equipoise they are mistaken—and will tend to report less symptoms when they know that they are receiving the perceived better option. Podsakoff et al. (2003) summarize the potential sources of common biases observed in behavioral research. For example, the bias of social desirability stems from the patients’ desire to respond according to perceived social acceptability rather than to their true feelings. To some extent, the same biases apply to medical providers. Importantly, single blindness can favor—often subconscious—investigators’ self-serving biases (Camfield et al., 2014), hence the advantage of double over single blindness. This argument is particularly relevant for soft or subjective tested outcomes, which are common in social sciences. More generally, the harder the end point, the lesser the interest of blindness. In the extreme case, it is difficult to misinterpret death. But even so, when feasible, double blindness has become the norm in medical RCTs.

While most medical RCTs deal with drugs, some evaluate lifestyle interventions or surgical procedures. For these studies, finding a placebo that is consistent with using double-blind (or even simple-blind) trials is far from obvious. Surgery-based studies raise the tantalizing question of whether to perform a sham operation. Here is the dilemma. First, if the patients in the control group receive no surgery, they obviously know that they are untreated, and the placebo effect is lost. Alternatively, all the patients can be brought to the operating room, resulting in the unoperated patients entering the recovery ward still half-anesthetized and with a fresh surgical scar. The second scenario would undeniably add validity to the trial in that the only difference between the control and treated groups will be the actual surgical treatment, but at the same time the trial will have done real harm to the controls, in having subjected them to totally unnecessary surgical (albeit not therapeutic) and anesthetic risks. One can see the dilemma as a trade-off between individual interest and the greater good, or as equipoise vs. single blind. In social science, where experimental blindness is practically hard to reach, if not impossible, the trade-off should, somewhat paradoxically, play in favor of equipoise.

In addition to the pure ethical rationale, the so-called Hawthorne effect provides an argument showing that the absence of blindness in experimental design reinforces the need for using some equipoise principle. The Hawthorne effect, uncovered by sociologists (Roethlisberger & Dickson, 1939), refers to the situation observed in Chicago's Hawthorne Works of the Western Electrical Company from 1924 to 1927. The researchers observed that regardless of whether they tested an increase or a reduction of the plant's artificial illumination intensity, they obtained a positive effect on workers' productivity. The psychological experience of benefitting from a change in working conditions appeared to be more important than the nature of the change itself. Such psychological benefit can only take place in a treated group where people are aware of being granted a change in conditions, and not in a control group where no change is operated. Hence, providing no placebo-type of change to the control group can harm not only

the individuals in the control group but also the reliability of the study by leading to falsely positive discoveries driven by the Hawthorne effect. More and more economists address this issue by providing some type of change to the control group.

Another moral tension stems from the decision to enroll a patient with a given condition in a study. The tension opposes again the physician's commitments both to her patient and to scientific progress. According to Weijer et al. (2000, p. 756), "the answer seems to depend greatly on which side of the Atlantic you reside. In the United Kingdom, the individual uncertainty principle is widely endorsed. However, in North America, clinical equipoise—reflecting collective uncertainty—is the dominant ethical basis". The so-called uncertainty principle works like equipoise applied specifically to each patient considering her individual pre-existing condition. Under that principle, the interests of each patient must be examined individually by the doctor before any enrollment in the study can be contemplated. A potential consequence is poor recruitment in studies, where ultimately clinical equipoise is put at risk. By contrast, the U.S. National Bioethics Advisory Commission (NBAS, 2001, p.77) prescribes that Institutional Review Boards (IRBs) "determine whether the relation between risks and potential benefits is reasonable. To do so, IRBs should determine whether the procedures meet the criteria of research equipoise (...) in addition to being justified in terms of the potential knowledge gain for society. Investigators and IRBs should understand that the term *research equipoise* applies to any type of research involving interventions or procedures that offer the prospect of direct benefit to participants (...)." Typically, subjects whose conditions have proven therapies benefit from studies designed as 'standard treatment plus placebo' versus 'standard treatment plus experimental drug.'

4. Should Economists Care about Equipoise?

While economists feel comfortable with RCTs in which the control group receives nothing, medical scholars consider that a placebo, which is better than nothing, is no ethically admissible option when the most recent state of science has uncovered a better drug. Following the equipoise requirement, this stance holds even if the better drug is expensive and the tested population cannot afford it.

How come that the ethical debate surrounding the application of equipoise has remained so confidential in the economic RCT community? Due to the relatively recent emergence of RCTs in economics, it could well be that economic research is still in its pre-Freedman stage. This possibility however has little credibility given that economists have largely borrowed from medical experimental designs in drafting their own studies. Researchers from both the medical and economic communities are therefore likely aware of the ongoing debate about unethical medical studies (Halpern et al., 2002), including those organized in developing countries (Gulhati, 2004; Jintarkanon et al., 2005; Milford et al., 2006).⁵ And yet, economists and other social scientists tend to disregard the equipoise requirement by typically disadvantaging the control group.

Take, for instance, the recent debate that opposed Professor Megan Stevenson from George Mason University, who investigates the impact of bail money through RCTs, to the Massachusetts Bail Fund (MBF), which pays up to \$500 bail for low-income people. Each side developed a number of insightful arguments, and the following extracts from their Twitter conversation epitomize the essence of the equipoise controversy.⁶

⁵ For instance, the Indian Council of Medical Research showed concern that trials ensure compliance with ethical guidelines (Chatterjee, 2008). Mudur (2005, p. 1044) quotes Prof. Falguni Sen from Fordham University, New York, saying that “Given the vulnerability of uneducated and poor patients, India has a long way to go in ensuring adequate protection to human subjects.”

⁶ We are grateful to Tim Ogden for having called our attention to this Twitter conversation.

MBF (March 8, 2019): “RCTs randomly assign a “new treatment” (in this case assistance from the bail fund) to one group while another group receives the “standard treatment” (in this case no help with bail, resulting in pre-trial incarceration). Terrified yet? WE KNOW THE IMPACT OF MONEY BAIL. The research EXISTS: people have better case & life outcomes when they can fight their case from freedom. We believe people when they tell us the difference it made to have their bail posted. It is WRONG to randomly pick some people to receive a lifesaving treatment while sending some people to jail. It is RACIST to engage in this kind of “research” when you know racial disparities in our courts and jails are longstanding and stark”.

Prof. Stevenson (March 14, 2019): “Until you have already served every single client (for the bail fund, (this would mean having the capacity to bail out every single person that needs it) there are some people who are “randomly” not receiving services. An RCT just makes this random process more explicit. For instance, say you only have resources to staff the courts 5 days a week. Randomize which days are staffed. Or say you don't have time to meet with every defendant during a shift. Start by meeting defendants with odd docket #'s, then move on to even #'s if there is time. Both of these methods are RCTs and would create very valuable research! Because although you think you know that your program is extremely effective, you really don't. Are you reaching the clients that need you the most? Those who would suffer were it not for your help?”

Redrafted along the equipoise line of thought, the argument of MBF is that assigning randomly a bail to the treatment group and no-bail to the control group violates the equipoise principle because the control group is disadvantaged. Prof. Stevenson dismisses this argument on the grounds that previous on-field knowledge is insufficient to assess confidently the superiority of the treatment (“although you think you know (...) you really don't”). She puts the search for

rigorous scientific validation above the concern that equipoise is not fulfilled.⁷ She also refers to the typical rationale of development economists that real life is already unfair/uncertain for RCT subjects (“*Until you have already served every single client*”). Basically, the idea is that the RCT is important for science and for future generations.

Thus, a realistic possibility is that economists consider their own studies as benign, meaning that organizers don’t need to bother about the potentially negative consequences for participants of the control group who fail to receive any benefits of the tested treatment. For instance, the argument goes, failing to supply a farmer with the loan she could need is no big harm. In economic terms, the situation where 50% of the farmers receive the loan and 50% do not is Pareto-improving with respect to the status quo where loans are fully absent from the picture.

However, people live in communities where social norms apply, and changes brought to a few individuals can have serious consequences for the whole society.⁸ A typical example stems from an AIDS prevention intervention that implies testing all the participants for HIV. In this case, and in many others, the study can have major destabilizing effects. This is particularly relevant in contexts where social norms are disrupted (Morvant-Roux et al., 2014). Backlashes are typically observed in the context of interventions intended to favor women’s empowerment in patriarchal societies (Schuler et al., 2018). As theorized by Rabin (1993), economic fairness has welfare implications. People can find it acceptable to get nothing when it is everybody’s fate but react negatively if they keep getting nothing while others are randomly, and thus “unjustly,” rewarded. Moreover, if the tested treatment is expected to make such a little difference, then maybe it does not deserve to be investigated with such a heavy (and expensive)

⁷ Garchitorena et al. (2019) mention that “many RCTs are carried out only to confirm results of observational studies.”

⁸ To mitigate this problem, randomization is sometimes clustered at the community level. Cluster-RCTs can still raise the concern of equipoise. For example, in some RCTs exploring the distribution of insecticide-impregnated bed nets, there is only one group that receives them for free (Müller et al., 2008; Tarozzi et al., 2014).

experimental design. Using RCTs implicitly supposes that the stakes are high enough to deploy a sophisticated and costly experimental design. Put boldly, failing to account for—at least some sort of—equipoise implies that the study either is a waste of money or causes “moral discomfort” (Baele, 1913; 4).

Unlike medical doctors, development economists are not used to see ethics interfere with their research methods. Generally, ethics is not mentioned as a concern by the “randomistas,” let alone as a goal in experimental design (Barrett & Carter, 2010). Their objectives are elsewhere, typically in ambitious policy implications to “solve poverty” (Karlan and Appel, 2011) and in testing economic theories. Banerjee and Duflo (2009: 156) view experiments as “a powerful tool for testing theories.” Both types of concerns deviate from that of medical investigators.

Arguably, policy recommendations are closer to the physician’s interest for the situation of her patient. There is, however, a notable nuance since economic policy is intended as a general treatment for a large group of people, some of which are not even seeking treatment. The take-up problem frequently met by randomistas occurs when participants are not interested in the treatment. The investigator’s typical response is triggering take-up with some sort of encouragement (White, 2013), which can in turn have unwanted consequences on the empirical results. By contrast, medical RCTs use local field investigators to assess the condition of the patients to be enrolled in the study. This procedure prevents researchers from including non-sick participants, both in the treated group and in the control group. For economists, proceeding in the same way without doing any harm would mean identifying the needs of all the participants and so building two types of options (possibly including some compensation) of *ex ante* equal interest to be provided randomly to the treated and control groups. Hence equipoise.

The contrast between the objectives of policy vs. theory is reminiscent of what Weijer et al. (2000) present as a UK vs. US controversy, or more broadly, as the science vs. patient trade-

off. Yet, in experimental economics there is no consensus on a middle-of-the-road option that would offer to the control group the equivalent of the “standard treatment plus placebo” in order not to harm the individuals who do not receive the experimental treatment. In fact, most economic RCTs go in the opposite direction by testing treatments about which the expectations of positive outcomes are the highest. This is the reason why the microfinance community was so disappointed to see that the results of the related RCTs were so modest (Duvendack et al., 2011). If the experiments were constrained by equipoise, this community would probably have been less disappointed following the mitigated outcomes of RCTs. But paradoxically, the modest impacts captured by these RCTs could also be presented as *ex post* rationalization that some kind of equipoise was met. A key issue in this field of research stems however from the fact that the outcomes checked vary across publications, leading to testing disparate impacts on subjects’ well-being. For instance, Attanasio et al. (2015) use a large set of outcomes including, e.a., increase in entrepreneurship, schooling, consumption, and repayment rates. This outcome heterogeneity adds another layer of complexity to the implementation of equipoise in social sciences.⁹

The short shrift given to ethics in economic RCT conversations echoes the past lack of reactions following unethical medical experimentation on disadvantaged populations. In the Tuskegee Study (1932-1972) sponsored by the U.S. Public Health Service, the group of interest was made up of poor African-American men with untreated syphilis (Caplan, 2001). The justification provided by the experimenters to give no treatment to the control group was that these poor men would not afford the treatment anyway (Angell, 1997). This argument may sound familiar to those who have questioned randomistas about the unfairness toward the members of their

⁹ The issue of multiple outcomes goes beyond the microfinance literature. A recent study by Schilbach (2019) on the impact of commitment devices on alcohol consumption among cycle-rickshaw drivers in India uses outcomes such as alcohol consumption, daytime sobriety, productivity, earnings, and savings. In any case, randomly assigning sobriety incentives is ethically questionable.

control groups. Interestingly, the Tuskegee Study stopped following an intervention of the media—namely the *Washington Star* and the *New York Times*—that embarrassed the Nixon administration (Angell, 1997).

Lurie and Wolfe (1997) point out the different study designs used in the U.S. and in developing countries¹⁰ in medical RCTs testing new drugs intended to save the life of infants born to HIV-infected women. The authors oppose the situation of the two trials performed in the US, in which the study groups had access to antiretroviral drugs, to those that took place in developing countries, where most patients had no such access, likely for financial reasons. Lurie and Wolfe (1997) also report the anecdote of a Harvard researcher who applied to the US National Institutes of Health (NIH) to get funding for an ethically well-designed--with an actively treated control group--study in Thailand and received from the NIH the cost-reducing recommendation to run a placebo-controlled trial instead. It is only after the director of Harvard's human subjects committee replied that, for such a situation, placebo-controlled trials were unethical that the NIH accepted the argument. Beyond anecdotal evidence, Petryna (2009) observes that the share of medical RCTs organized in emerging markets¹¹ increased from 10% in 1991 to 40% in 2005. She questions the exploitative character of offshoring clinical trials, which can be used by pharmaceutical companies to encourage doctors in emerging countries to prescribe high-cost medicines, and so undermine the delivery of affordable treatments. Evidently, the bulk of experimental economists have no such profit-oriented motivations. Yet RCTs are costly to implement and so divert money away from other, often less-consuming, experimental designs. Moreover, the history of medical trials in developing countries pinpoints the reputational threats associated with experiments involving poor populations. These populations are easy to exploit

¹⁰ Ivory Coast, Uganda, Tanzania, South Africa, Malawi, Thailand, Ethiopia, Burkina Faso, Zimbabwe, Kenya, and the Dominican Republic.

¹¹ The author cites several examples involving Eastern Europe and Brazil.

because they are mostly unaware of their rights to full disclosure about the experimental design, and to subsequent informed consent or denial.

Another notable difference between medical and economic studies relates to the scope—rather than the nature—of the impact sought, since the economic perturbations brought to existing structures and traditions are believed to be minor compared to drug testing. In addition, most RCTs performed by economists in developing countries seek evidence that relates one way or another to the purpose of aiding these countries. These studies cannot be suspected of having a commercial agenda or using developing countries as a lab for treatments intended for developed economies. Yet, RCTs without equipoise allocate supposedly favorable situations to randomly chosen individuals. Following an argument often used when experimental drugs are rationed, one would expect that the treatment would be offered to those who are the most in need. Along this line of thought, Baele (2013; 19) claims that “randomization also violates the prioritarian moral principle of ensuring a certain level of well-being to the worse-off sub-population before thinking of either maximising in absolute terms the wealth of the population (consequentialist version) or ensuring individual freedom (liberal version).” Accordingly, even benign economic treatments are no reason for deviating from the ethical requirement of equipoise toward the control group.

If the economic profession finds the equipoise requirement suitable, its implementation would lead to constraining study design and imposing that the treated and control groups get qualitatively similar outcomes, where similarity must be backed by expert opinion. However, the funders’ preference for promising impact can put pressure in the direction opposite to equipoise (Ravaillon, 2018). Overall, the practical implementation of this procedure can be tedious, but at least it is expected to exclude blatantly unethical behaviors.

In addition, the medical examples described in this section reveal that even in medical sciences where the Declaration of Helsinki and the equipoise requirement are supposed to act as moral compasses, ethical relativism dies hard. Researchers performing RCTs in developing countries tend to lower the standards of care based on the financial argument that most people—and/or their governments—cannot pay for the best treatments. Dismissing such arguments, the editorial article by Marcia Angell (1997; 848) announced that the prestigious *New England Journal of Medicine* has decided not to publish studies reporting “unethical research, regardless of their scientific merit.” Will economic journals follow the same path? At the moment, there is no indication of such a move.

5. Conclusion

The time has come to ask RCT development economists, and other randomistas, to confront the ethical consequences of their doings with the ultimate goals of the trials they launch. Like in the medical community in the 1980s, more and more scholars are questioning the “gold standard” paradigm (Cartwright, 2007; Deaton, 2010; Bédécarrats et al., 2017) along several dimensions. Even though RCTs can affect lives significantly, criticisms based on their ethical dimension are so far among the least vocal ones. Arguably, this is due to the underlying good intentions, such as helping poor people address malaria prevention (Cohen & Dupas 2010) and sanitation threads (Duflo et al., 2015). At the same time, as Baldassarri and Abascal (2017; 62) put it, “field experimenters ‘play God,’ intervening in people’s lives in consequential ways.” Often, previous field knowledge makes negative outcomes predictable with some degree of confidence. In such cases, the lack of equipoise associated with the specific risks to which disadvantaged people are exposed result from either the indifference or the insufficient field experience of the experimenters. From an ethical standpoint, this situation is worse than that of studies randomly assigning presumably favorable treatments. Yet, both can have long-lasting

effects not only at the individual level, but also on interpersonal relationships within communities.

On the other hand, arguments focusing on the unfairness entailed by the experimental design neglect the potential long-term benefits brought by the novel knowledge that RCTs are built to deliver. Even in the medical sphere, some authors criticize the very notion of equipoise or propose to amend it (Fries & Krishnan, 2004; Ubel & Silbergleit, 2011). The main concern relates to failing to account for the future benefits that are disregarded when dealing with the therapeutic obligation of physicians toward their patients. Veatch (2007; 182) claims that “it is not anyone’s equipoise that is morally critical; it is whether the potential subjects consent to be randomized without being unduly coerced, manipulated, or exploited.” Miller and Brody (2007; 153) argue “that the ethical principles governing medical therapy are different from those governing clinical research”. The authors emphasize the importance of the returns expected from the study as a legitimate reason to deviate from the requirement of equipoise. This logic corresponds to accepting to sacrifice the welfare of some individuals (typically those in the control group) for the greater good of society and future generations. It is however mitigated by the double blindness that makes the sacrifice probabilistic rather than certain, and so spreads it over all the study participants. Economic RCTs mostly use no blindness, which makes it more difficult to refer to this convenient excuse. In any case, scientific studies treating unfairly or inflicting sacrifices to people, and especially to those already disadvantaged, will always be ethically questionable. And since morally debatable studies are preferentially organized in countries where legal protection is weak, the issue of imposing—at least some form of—equipoise to economic experimentation should go beyond the rhetoric.

References

- Angell, M. (1997). Editorial: The ethics of clinical research in the Third World. *New England Journal of Medicine* 337(12): 847-849.
- Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E., Harmgart, H. (2015). The impacts of microfinance: Evidence from joint-liability lending in Mongolia. *American Economic Journal: Applied Economics* 7(1): 90-122.
- Baele, S.J. (2013). The ethics of new development economics: Is the experimental approach to development economics morally wrong? *Journal of Philosophical Economics*, VII:1.
- Baldassarri, D., Abascal, M. (2017). Field experiments across the social sciences. *Annual Review of Sociology* 43: 41-73.
- Banerjee, A.V., Duflo E. (2009). The experimental approach to development economics. *Annual Review of Economics*. 1: 151–178.
- Barrett, C.B., Carter, M.R. (2010). The power and pitfalls of experiments in development economics: Some non-random reflections. *Applied Economic Perspectives and Policy* 32(4): 515-548.
- Bédécarrats, F., Guérin, I., Roubaud, F. (2017). All that glitters is not gold. The political economy of randomized evaluations in development. *Development and Change*, forthcoming.
- Beta-Blocker Heart Attack Trialists (BHAT) (1982). A randomized trial of propranolol in patients with acute myocardial infarction. *Journal of the American Medical Association* 247(12):1707-1714.
- Botros, S. (1990). Equipoise, consent and the ethics of randomised clinical trials, in P. Byrne (ed.) *Ethics and Law in Health Care and Research*, John Wiley & Sons, Chichester, UK, 9-24.
- Camfield, L., Duvendack, M., Palmer-Jones, R. (2014). Things you wanted to know about bias in evaluations but never dared to think. *IDS Bulletin* 45(6): 49-64.
- Caplan, A.L. (2001). Twenty years after: the legacy of the Tuskegee syphilis study. *Bioethics, Justice and Health Care*. Belmont, Ca.: Wadsworth-Thomson Learning, p. 231-5.

- Carlson, R.V., Boyd, K.M., Webb, D.J. (2004) The revision of the Declaration of Helsinki: past, present and future. *British Journal of Clinical Pharmacology* 57(6): 695-713.
- Cartwright, N. (2007). Are RCTs the gold standard? *BioSocieties* 2(1), 11-20.
- Chatterjee, P. (2008). Clinical trials in India: Ethical concerns. *Bulletin of the World Health Organization*, 86(8): 581-582.
- Childress, J.F., Faden, R.R., Gaare, R.D., Gostin, L.O., Kahn, J., Bonnie, R.J., Nieburg, P. (2002). Public health ethics: Mapping the terrain. *Journal of Law, Medicine & Ethics* 30(2): 170-178.
- Cohen, J., Dupas, P. (2010). Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *Quarterly Journal of Economics* 125(1): 1-45.
- Council for International Organizations of Medical Sciences (CIOMS) (2002). International ethical guidelines for biomedical research involving human subjects. *Bulletin of Medical Ethics* 182:17-23.
- Days, S.J., Altman, D.G. (2000). Blinding in clinical trials and other studies. *British Medical Journal*; 321:504.
- Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature*, 48(2), 424-55.
- De Dickert N.W., Emanuel E.J. (2015). Ethics in cardiovascular medicine. in Mann, D. L.; Zipes, D.P. Libby, P. Bonow, R.O.(Eds.) *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine*, tenth edition, Philadelphia, PA: Elsevier Saunders, pp. 29-34.
- Di Tillio, A., Ottaviani, M., Sørensen, P.N. (2017). Persuasion bias in science: Can economics help? *Economic Journal*, 127(605): F266-F304.
- Duflo, E., Greenstone, M., Guiteras, R., Clasen, T. (2015). Toilets can work: Short and medium run health impacts of addressing complementarities and externalities in water and sanitation National Bureau of Economic Research, WP 21521.

- Duvendack, M., Palmer-Jones, R., Copestake, J.G., Hooper, L., Loke, Y., Rao, N. (2011). What is the evidence of the impact of microfinance on the well-being of poor people? EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Fries, J.F., Krishnan, E. (2004). Equipoise, design bias, and randomized controlled trials: the elusive ethics of new drug development. *Arthritis Research & Therapy* 6(3): R250.
- Freedman, B. (1987). Equipoise and the ethics of clinical research. *New England Journal of Medicine* 317(3): 141-145.
- Garchitorena, A., Megan B. Murray, M.B., Hedt-Gauthier, B., Farmer, P.E. Bonds, M.H. (2019). Reducing the knowledge gap in global health delivery: contributions and limitations of randomized controlled trials. Forthcoming in Bédécarrats, F., Guérin, I., Roubaud, F. (Eds.) *Randomized Control Trials in the Field of Development: A Critical Perspective*.
- Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico (GISSI) (1987) . Long-term effects of intravenous thrombolysis in acute myocardial infarction: Final report of the GISSI study. *Lancet* 2: 871–874.
- Gulhati, C.M. (2004). Needed: Closer Scrutiny of Clinical Trials. *Indian Journal of Medical Ethics* 1: 4-5.
- Halpern, S.D., Karlawish, J.H., Berlin, J.A. (2002). The Continuing Unethical Conduct of Underpowered Clinical Trials. *Journal of the American Medical Association* 288(3): 358-362.
- Jintarkanon, S., Nakapiew, S., Tienudom, N., Suwannawong, P., Wilson, D. (2005). Unethical Clinical Trials in Thailand: A Community Response. *Lancet* 365(9471): 1617-1618.
- Kass, N.E. (2001). An ethics framework for public health. *American Journal of Public Health*, 91(11): 1776-1782.
- Karlan, D.S., Appel, J. (2011). *More Than Good Intentions: How a New Economics is Helping to Solve Global Poverty*. New York: Dutton
- Lilford, R.J., Jackson, J. (1995). Equipoise and the ethics of randomization. *Journal of the Royal Society of Medicine* 88(10): 552-559.

- London, A.J. (2017). Equipoise in research. Integrating ethics and science in human research, *JAMA Guide to Statistics and Methods*, 317(5):525-526.
- Lurie, P., Wolfe S.M. (1997). Unethical trials of interventions to reduce perinatal transmission of the human immunodeficiency virus in developing countries. *New England Journal of Medicine* 337(5): 853-856.
- McMurray, J.J.V. (2010). Systolic heart failure. *New England Journal of Medicine*, 362: 228-238.
- Milford, C., Wassenaar, D., Slack, C. (2006). Resources and Needs of Research Ethics Committees in Africa: Preparations for HIV Vaccine Trials. *IRB: Ethics & Human Research* 28(2): 1-9.
- Miller, F.G., Joffe, S. (2011). Equipoise and the dilemma of randomized clinical trials, *New England Journal of Medicine* 364(5): 476-480.
- Miller, F.G., Brody, H. (2007). Clinical equipoise and the incoherence of research ethics, *Journal of Medicine and Philosophy* 32(2): 151–165.
- Morvant-Roux, S., Guérin, I., Roesch, M., Moisseron, J.Y. (2014). Adding value to randomization with qualitative analysis: the case of microcredit in rural Morocco. *World Development* 56: 302-312.
- Mudur, G. (2005). India plans to audit clinical trials. *British Medical Journal* 331(7524), 1044.
- National Bioethics Advisory Commission (NBAC). (2001). *Ethical and Policy Issues in Research Involving Human Participants*. Volume I. Bethesda, MD.
- Müller O., De Allegri, M., Becher, H., Tiendreboogo, J., Beiersmann, C., Ye, M., Kouyate, B., Sie, A., Jahn, A. (2008) Distribution systems of insecticide-treated bed nets for malaria control in rural Burkina Faso: cluster-randomized controlled trial. *PLoS ONE* 3(9): e3182.
- Nies, A.S., Evans, G.H., Shand, D.G. (1973). Regional hemodynamic effects of beta-adrenergic blockade with Propranolol in the unanesthetized primate. *American Heart Journal* 85: 97-102.

- Petryna, A. (2009). *When Experiments Travel: Clinical Trials and the Global Search for Human Subjects*. Princeton, NJ: Princeton University Press.
- Petticrew, M., McKee, M., Lock, K., Green, J., Phillips, G. (2013). In search of social equipoise. *British Medical Journal* (Online), 347.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology* 88(5): 879.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review* 83(5): 1281-1302.
- Ravallion, M., 2018. Should the randomistas (continue to) rule? Forthcoming in Bédécarrats, F., Guérin, I., Roubaud, F. (Eds.) *Randomized Control Trials in the Field of Development: A Critical Perspective*.
- Roethlisberger, F.J., Dickson W.J. (1939). *Management and the Worker*. Volume 5 in Cambridge, Mass.: Harvard University Press.
- Schafer, A. (1982). The ethics of the randomized clinical trial, *New England Journal of Medicine* 307(12): 719-724.
- Schuler, S.R., Lenzi, R., Badal S.H., Nazneen S. (2018). Men's perspectives on women's empowerment and intimate partner violence in rural Bangladesh, *Culture, Health & Sexuality*, 20(1): 113-127.
- Schulz, K.F. (1995). Subverting Randomization. *Journal of the American Medical Association* 274(18): 1456-1458.
- Second International Study of Infarct Survival Collaborative Group (ISIS-2) (1988). Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction. *Lancet* 2: 349-360.
- Shand, D.G. (1975). Propranolol. *New England Journal of Medicine* 293: 280-284.

- Tarozzi, A., Mahajan, A., Blackburn, B., Kopf, D., Krishnan, L., Yoong, J. (2014). Micro-loans, insecticide-treated bednets, and malaria: Evidence from a randomized controlled trial in Orissa, India. *American Economic Review* 104(7): 1909-41.
- Theroux P., Franklin D., Ross J Jr. Kemper W.S. (1974). Regional myocardial function during acute coronary occlusion and its modification by pharmacologic agents in the dog. *Circulation Research* 35: 896-908.
- Ubel, P.A., Silbergleit, R. (2011). Behavioral equipoise: A way to resolve ethical stalemates in clinical research, *American Journal of Bioethics* 11(2): 1-8.
- Veatch, R.M. (2007). The irrelevance of equipoise. *Journal of Medicine and Philosophy* 32(2): 167–183.
- Weijer, C., Glass, K.C., Shapiro, S.H. (2000). Why clinical equipoise, and not the uncertainty principle, is the moral underpinning of the RCT. *British Medical Journal* 321: 756-758.
- White, H. (2013) an introduction to the use of randomised control trials to evaluate development interventions. *Journal of Development Effectiveness* 5(1): 30-49.
- Young, J., Harrison, J. White, G., May, J., Solomon, M. (2004), Developing measures of surgeons' equipoise to assess the feasibility of randomized controlled trials in vascular surgery. *Surgery* 136(5): 1070-1076.
- Yusuf, S., Peto, R., Lewis, J., Collins, R., & Sleight, P. (1985). Beta blockade during and after myocardial infarction: An overview of the randomized trials. *Progress in Cardiovascular Diseases* 27(5): 335-371.