

The Pragmatic Turn in Clinical Research: in Search for the Real World

Olga Zvonareva

Maastricht University, Netherlands, Siberian State Medical University and
National Research Tomsk State University, Russia
/ o.zvonareva@maastrichtuniversity.nl

Abstract

How does knowledge obtained in clinical trials apply to the actual treatment of patients? This question has recently acquired a new significance amidst complaints about the limited ability of trial results to improve clinical practice. Pragmatic clinical trials have been advocated to address this problem. In this article, I trace the emergence of the pragmatic turn in clinical research, starting from the first mention of 'pragmatic trial' in 1967, and analyse the changes to how such trials have been conceived. I argue that contemporary version of pragmatism in clinical trials risks missing the mark by focusing exclusively on establishing similarity between the trial and the clinic for the purpose of greater generalizability. This focus eclipses the move for carefully aligning medical experimentation with conditions, needs and concerns in the clinic aimed at greater usefulness.

Keywords: clinical trials, pragmatic trials, RCT, real-world evidence, statistics

"Although randomized clinical trials provide essential, high-quality evidence about the benefits and harms of medical interventions, many such trials have limited relevance to clinical practice" James H. Ware and Mary Beth Hamel (2011: 1658) wrote in one of the 2011 issues of *The New England Journal of Medicine*. With this opening line, Ware and Hamel, a biostatistician and a clinical scientist, joined the increasingly prominent conversation within the health research community about the relations between the tightly controlled experimental apparatus of clinical trials and the messy realities of clinical practice.

Concerns about how knowledge obtained in clinical trials applies to the treatment of

patients have accompanied the very rise of the randomised controlled trial (RCT), which has been the key method of evaluating medical interventions since the 1960s, particularly pharmaceuticals (Bothwell et al., 2016; Cartwright, 2007; Moreira and Will, 2016; Timmermans and Berg, 2003). Yet, in the last two decades, these concerns became ubiquitous. Simultaneously, an explosion of interest in pragmatic trials took place, stimulated by the promise of this approach to designing and running RCTs to improve relevance of clinical trials to clinical practice. Proponents of pragmatic trials nowadays envision such improvement as an outcome of undertaking trials under the so-called 'real-world conditions' (Dodd et al., 2016;



Zwarenstein and Treweek, 2009). An ostensibly paradoxical move takes place, whereby the more capable the experimental clinical trial machinery is at excluding interfering factors and suppressing seemingly irrelevant details to produce reliable universal knowledge, the less useful for clinical practice the results appear to be.

In this paper, I trace the emergence and implications of the pragmatic turn in assessing new pharmaceutical treatments. Such focus makes visible the changing ways in which the relevance of clinical research is conceived and established and points to some crucial differences between efforts to produce generalizable knowledge and efforts to produce useful knowledge.

Analytical perspectives: purification and contextualisation

RCTs rely on random assignment of study participants to groups. There can be one or more intervention groups where participants receive a new treatment and a control group where participants receive existing treatment or, sometimes, placebo. Outcomes are compared across groups and to ensure fair comparison, blinding is often employed so that participants and investigators are not aware who is assigned to which group. Randomization, the use of control groups, and blinding form the methodological backbone of the RCT, held in high esteem for its ability to keep bias at check and to make cause-effect relationships between treatment and outcome more palpable. But adhering to this triad is not enough for the RCTs to deliver on their promise of reliable results. To minimize interferences that may obscure the cause and effect relations, traditionally RCTs have been characterized by the narrow precision of tested interventions, the tight control of the conditions under which these interventions are administered, and the highly scripted experimental procedures (Calvert et al., 2011; Tunis et al., 2003). These characteristics are meant to perform what can be called purification to borrow from Latour (1993), that is to keep the noise of the daily world outside the confines of the medical experiment, thus clearly and reliably distilling the experimental intervention's true effects. Therefore, traditional RCTs rely on establishing purified

experimental environment capable of isolating a number of critical variables to produce knowledge of causal relations held to be universally valid (Rosengarten and Savransky, 2019). Ironically, these same characteristics have given reasons for concern about the relevance of RCTs to clinical practice, where as some worry contextual dynamics kept at bay in a trial laboratory may reshape the expected results (Bower, 2003; Brass, 2010).

Criticism of traditional RCTs by contemporary advocates of pragmatic clinical trials centres on the difference between the purified experimental environment of RCTs and the diversity and contingency of clinical practice. According to critics, the very specificity of RCTs, which distinguishes them from clinical practice, while supposedly optimal for producing robust evidence, leads to "limited applicability of many trial results beyond the artificial, 'laboratory' environment of the trial" (Treweek and Zwarenstein, 2009). Those advocating the wider use of pragmatic trials as a remedy for this applicability problem argue that the little resemblance between the trial laboratory and real-life clinical practice affects relevance of RCTs in two ways. First, questions to be answered during a trial often have little to do with questions faced daily by patients, physicians and policymakers (Zuidgeest et al., 2017). Second, answers to the questions eventually asked may still fail to hold when transported from the secluded experimental site to the clinic with patients, physicians, routines and technologies which are unlike those in an RCT. What pragmatic trials are expected to do is to 'show the real-world effectiveness of the intervention' (Ford and Norrie, 2016), i.e. what an intervention can do for actual patients under far-from-ideal circumstances. In short, the medical research community has recently become vocally concerned with the apparent weakness of the connection between the knowledge produced in clinical trials and the contexts where this knowledge is meant to be applied. We can understand this as a concern with weak contextualization of trial-produced knowledge and a call to complement the strive for rigour and reliability through purification with the strive for applicability through contextualization.

These concerns contrast with critical social science scholarship which has long highlighted the already contextually engaged nature of clinical

research. The picture of the RCT as insulated from the clinic and the world outside a trial is challenged by the Science and Technology Studies (STS) scholars. First, analyses of trial conduct in diverse locations highlight how the apparently standardised neatness and stable orderliness of an RCT are enabled by modifications, creative acts and negotiations performed by those doing the work that goes into successful accomplishment of a trial (Simpson and Sariola, 2012; Zvonareva et al., 2017). Second, STS works suggest that in designing their RCTs, especially in public health and health services research, trialists do modify the pure world of the experiment by selectively incorporating elements of the outside world, for instance, by coordinating delivery of an experimental intervention with existing organisational routines in the test sites (Will, 2007). This dual concern with purification and contextualisation at the same time serves to ensure interest and buy-in from those collectives whose cooperation is necessary for a trial to proceed and for its results to reach clinical practice. Third, scholars argue that clinical trials affect clinical practice not just through dissemination of findings after research completion, as is typically assumed. The very practice of research itself alters the operation of the healthcare organisations where trials are conducted, already during the preparation and running of medical experiments (Petty and Heimer, 2011). For instance, infrastructure gets built and renovated for trials, the relationships within a clinic change, and tests, drugs and other artefacts are shipped in.

STS research makes it clear that the picture of the traditional RCT as fully disconnected from clinical practice does not do justice to the complexity of the interactions involved. Trialists' actual practices do involve purification to ensure methodological adequacy and contextualisation for the sake of making trials relevant for those who conduct them and may use their results. However, contextualisation has not been an explicit consideration within the trials field itself until recently. The difficulty of direct transfer of evidence from RCTs to the clinic, while acknowledged, has been cast predominantly as the issue of practice being inferior to the RCT because of difference in resources and skill or as the issue

of the implementation gap where practice is lagging behind the results of trials (Dopson et al., 2003; Sanders and Haines, 2006). Efforts to address this difficulty, therefore, have focused on improving clinical practice by informing and technologically equipping it. But when proponents of pragmatic trials now state that "real-world evidence is needed" (Zuidgeest et al., 2017: 7), they appear to approach the problem of applicability from a different angle by criticising precisely the secondary importance given to contextualisation in the process of knowledge production. What is new and significant here is an emerging turn within the health research field itself towards explicitly reconsidering the connection between clinical experimentation and clinical practice in order to reform the RCT itself accordingly.

This development appears to be in line with the wider shifts towards greater contextualisation in contemporary knowledge production, which Nowotny, Scott and Gibbons (2001) described. Contextualisation, for these authors, involves the growing role that society and its diverse concerns play in science, but also a "shift within science from the search for 'truth' to the more pragmatic aim of providing a provisional understanding of the empirical world that 'works'" (Gibbons, 1999: 82). In their analysis, the authors mention medical research as one of the fields marked by strong contextualisation. However, according to pragmatic trial enthusiasts, such shift towards greater sensitivity to the needs of clinical practice and greater focus on the usefulness of results is not yet an accomplishment, but rather a task at hand.

In this paper, I investigate the new kind of balance pragmatic trials attempt to strike between what is considered a health intervention and what is considered its context and how exactly the relationship between an experiment and the real world is being reshuffled. Aiming to produce real-world evidence, pragmatic trials seek to reshape the classic RCT methodology, but which elements of this methodology are open for change and which are non-negotiable? If pragmatic trial departs from conceiving the RCT as a sterile and controlled laboratory, what then accords its epistemic robustness? And, most importantly, how promise of greater contextualization is

being fulfilled? To answer these questions, I first delve into the origins of the notion of pragmatism in relation to clinical trials. I analyse how the pragmatic attitude in clinical trials was conceived in the very first article on this topic published in 1967 and why this publication attracted significant attention from the medical research community only some 30 years later. Further, I follow the explosion of interest in pragmatic trials at the end of 1990s and focus on the ways in which pragmatism has been reinterpreted and on the strategies used to stabilise its contemporary version. In the concluding section, I discuss the implications of this pragmatic turn and argue that contemporary version of pragmatic trials risks missing the mark by allowing the focus on establishing similarity between the trial and the clinic environments for the purpose of greater generalizability to eclipse the move for carefully aligning medical experimentation with conditions, needs and concerns in the clinic for greater usefulness.

What problem are we solving?

Pragmatic trials were first distinguished by two French statisticians, Daniel Schwartz and Joseph Lellouch. They articulated their views in the 1967 article 'Explanatory and pragmatic attitudes in clinical trials'. Schwartz and Lellouch stated, may be aimed at solving two radically different types of problems. Trials conceived and implemented without clear recognition of what type of problem they aim to solve end up yielding inadequate and even ethically indefensible results.

Let us take a look at one of the examples Schwartz and Lellouch (1967) provided to explain their point. Imagine that trialists would like to compare two anti-cancer treatments, one being radiotherapy alone and another being the same radiotherapy but preceded by a novel drug. This drug may sensitise patients to the effects of radiotherapy and is to be administered over a 30-day period. Stating simply that the trial aims to compare the two treatments, as is often done, is misleading. Instead of this single general formulation, Schwartz and Lellouch offer to select one of two different approaches to designing the trial.

One approach would centre on the question *Does the drug have a sensitising effect?* To answer

this question, investigators would split trial participants into two groups: 'drug + radiotherapy' group and 'radiotherapy alone' group. The drug + radiotherapy group begins receiving their intervention right away, that is, they undergo 30 days of taking the drug and then radiotherapy. Simultaneously, the radiotherapy alone group undergoes a 30-day blank period, so that at the end of this period radiotherapy is administered at the same time to both groups. This approach allows to entirely equalize the conditions of administering treatments, so that the two groups differ only in the presence or absence of the drug. This is what Schwartz and Lellouch called *explanatory* trial. In this case, the treatment studied is the drug; investigators are able to distil the effects of this key component, and aim at *understanding*. But what would the presence or absence of a drug's sensitising effect mean for treating actual patients? The explanatory version of this trial would produce practical implications only if the drug + radiotherapy intervention turned out to be no better than radiotherapy alone after a delay. In this case, there is no reason to use the drug prior to radiotherapy in clinical practice, since the combined treatment would be no better than immediate radiotherapy without delay. However, if drug + radiotherapy turned out to be better than radiotherapy alone, investigators would end up in a situation where the drug, despite being proven efficacious, "may be of no practical interest since it has only been compared with radiotherapy inefficiently administered" (Schwartz and Lellouch, 1967: 639).

Schwartz and Lellouch then described another approach to designing this same trial, which they termed *pragmatic*. The pragmatic trial is aimed at *decision* and would seek to answer the question *Which of the two treatments should we prefer?* In this case, the radiotherapy alone group would receive radiotherapy at once, without the 30-day blank period, at what is likely to be the optimal time for the radiation treatment to benefit patients. Instead of comparing the presence of a drug with its absence under equalised conditions, this approach allows for comparison between two modes of therapy provided under conditions optimised for each therapy to work in terms of timing, dosage, mode of administration, auxiliary care, etc. Where an explanatory trial provides

information on the effects of the key component, a pragmatic trial compares two complex treatments as wholes under the conditions in which these treatments are likely to be applied in practice. The former entails stripping the tested treatments of the context of their administration and equalising conditions of their provision, while the latter entails separately defining each of the tested treatments in a contextualised way to include their presumed optimal usage conditions in practice.

Schwartz and Lellouch went on to stress that while treatments compared in a pragmatic trial are much more broadly and flexibly defined than in an explanatory trial, this does not constitute a violation of the essential experimental procedures:

The basic principle that two treatments must be compared in two groups which are in every other respect comparable is in no way contradicted by optimisation of the contextual factors. Instead, these factors become themselves part of the therapies to be compared and are thus distinguished from non-contextual factors for which comparability must be assumed. It is characteristic of the pragmatic approach that the treatments are flexibly defined and “absorb” into themselves the contexts in which they are administered. (Schwartz and Lellouch, 1967: 638)

Thus, the distinction is drawn differently between an experimental intervention and its context in explanatory and pragmatic trials with the latter being much more contextualised. In Schwartz and Lellouch’s terms, contextualisation refers to considering tested treatments in a broad sense, together with the particularities and conditions of their administration in clinical practice. Yet, contextualisation necessarily proceeds within the experimental framework. To produce reliable answers, a trial has to be controlled, meaning it must involve comparison between reasonably similar experimental and control groups.

Apart from contextualising treatments to compare them under optimal rather than equalised conditions, Schwartz and Lellouch suggested that pragmatic and explanatory trials differ in several other respects. First, the difference

lies in how patients are included. For any given trial, suitable patients are selected from the class of all comers by means of inclusion and exclusion criteria. Within the explanatory approach, patients deemed suitable for a trial are strictly selected and made as homogenous as possible. Furthermore, some patients may discontinue participation during the trial because of side effects, changes in their schedules, unpleasant trial procedures, quarrels with personnel or other reasons. In an explanatory trial, the class of suitable patients is redefined *a posteriori* to exclude withdrawals. Under the pragmatic approach, trial participants are more heterogenous and selection is not taken too far so as to stay close to the class of all comers. Withdrawals are not excluded from the analysis, as the treatments under comparison are flexibly defined to absorb discontinued participation as well. Comparing the two approaches, Schwartz and Lellouch summarised:

[w]ith the explanatory approach, we compare strictly defined treatments on a relatively arbitrary class of patients; with the pragmatic approach, loosely defined treatments are compared on patients drawn from a predetermined class. viz. those to which the conclusions of the trial are to be extrapolated. We may say that in the first case the class of patient is defined to fit the predetermined treatments, while in the second the treatments are defined to fit the predetermined class of patients. (Schwartz and Lellouch, 1967: 643)

Second, the difference between explanatory and pragmatic trials lies in whether laboratory or normal conditions are adhered to. The first way implies more rigorous and intense procedures which could be performed only in the course of a trial (laboratory conditions). The second way adheres to conventions of the current clinical practice (normal conditions). Here, Schwartz and Lellouch view the clinic as an imperfect version of the laboratory, with the distinction between normal and laboratory conditions depending on the level of clinical practice and being able to vanish if this level were to rise. The distinction between normal and laboratory conditions is of the spectrum type in contrast with the optimal and equalised conditions of testing interventions, which

Schwartz and Lellouch viewed as “totally opposed concepts” (Schwartz and Lellouch, 1967: 639).

The third difference between explanatory and pragmatic trials lies in how the results of testing of the two treatments are compared. Since sample sizes are always finite, conclusions of any comparison are subject to a certain risk of errors. When the explanatory approach is adopted to discover whether a difference exists between two treatments, analysts are concerned with errors of the first kind where it is wrongly concluded that two treatments differ when in fact they don't and errors of the second kind where it is wrongly concluded that two treatments are equivalent whereas in actuality they differ. When the pragmatic approach is adopted to answer the question “Which of the two treatments should we prefer?”, the comparison proceeds differently. Errors of the first kind are irrelevant because when two treatments are equivalent, it does not matter which one is chosen. Furthermore, some difference is always assumed to exist between the two treatments, so probability of errors of the second kind is null. All attention instead is given to what Schwartz and Lellouch called errors of a third kind, which occur when it is concluded that one treatment is superior to another, whereas the opposite is the case. So, analysis within pragmatic trials focuses on errors of the third kind, since it is most undesirable to choose an inferior treatment, whereas analysis within explanatory trials ignores these kinds of errors.

The article “Explanatory and pragmatic attitudes in clinical trials” ended with a warning. Schwartz and Lellouch cautioned that trials could not be conducted adequately without specifying exactly what type of problem a trial was aimed at, i.e. a problem of understanding or a problem of decision-making, and consciously matching trial design to the type of problem. The two statisticians also called for a change in the dominant approach to designing clinical trials: “Most trials done hitherto have adopted the explanatory approach without question; the pragmatic approach would often have been more justifiable” (Schwartz and Lellouch, 1967: 648) and invited further discussion.

Indifference

Discussion, however, barely started at the time. The pattern of citations of Schwartz and Lellouch's seminal paper can serve as one indication of how interest in pragmatism in clinical trials and the ability of clinical trials to inform decision-making in clinical practice developed (see Figure 1). Data from Google Scholar suggest that within ten years from publication, the paper was cited only seven times, followed by a modest increase in the next decade. In 2019 however, “Explanatory and pragmatic attitudes in clinical trials” is cited more than 1,200 times. A sharp increase in cumulative citations is visible from the end of the 1990s, perhaps signalling a change in attitude towards traditional RCT and its ability to be a means to decision-making in health care rather than as formal hypothesis testing.

How should we understand the period of apparent indifference prior to the explosion of interest? Answering this question requires turning to the topic of statistics and its convergence with the ascent of RCT methodology to the dominant position it has enjoyed for the most part of the last fifty years.

In contemporary medical science statistics is ubiquitous. Yet, despite a number of examples of statistical analysis use in medicine throughout the past centuries, the involvement of statistics in clinical research started being visible only by the end of the 1940s (Higgs, 2000; Mainland, 1960). It is after the landmark British Medical Research Council's trial of streptomycin for tuberculosis in 1947-48 and similar trials of the US Public Health Service at the end of the 1940s and beginning of the 1950s that promoted RCT methods (Bothwell and Podolsky 2016; Marks 2000b) that statistical apparatus, propelled by the rise of RCT, solidified its place in medicine. This is not to suggest that the growing importance of statistics in general spilled into medical research and resulted in the rise of RCTs. It would rather be more accurate to say that proponents of RCTs recruited statistical expertise to support their efforts. Gain in prominence by RCT in consort with statistics was greatly aided by the movement for therapeutic reform most active in the US and the UK (Marks, 2000b; Podolsky, 2010). Medical scientists, academic physicians, journal editors and governmental officials who comprised

this movement were united in their conviction that more scientifically robust knowledge about drug effects would lead to better clinical practice. Scientifically robust knowledge was to be guarded from various kinds of biases introduced into medical research by participating patients but also by investigators themselves, from manipulations with patient assignment to favour particular therapies, to expectations influencing the reporting and analysis of experimental outcomes. For reformers, the RCT with randomised treatment assignments, use of control groups, and blinded assessment of outcomes presented an impersonal standard for keeping these biases in check and, thus, producing more reliable knowledge to guide clinical practice (Chalmers, 2001). And here an opening was presented for statisticians who forged an alliance with therapeutic reformers and aided the effort with procedures and ideas about experimental design developed in the field of statistics. Statisticians came to be in charge of weeding out weaknesses in trial design, eliciting risks of bias and policing quantitative aspects of study conclusions, contributing to the cause of the reform: to provide physicians with as decisive an answer as possible regarding the therapeutic merits of new treatments. Slowly but persistently, statistics became such a distinguishing mark of a well-designed trial that, as medical historian Harry M. Marks (2000a: 351) highlights, “[b]y the late 1960s, investigators would complain of ‘the benevolent tyranny’ statisticians held over therapeutic research”.

In a short time, statisticians became indispensable for planning and analysing medical experiments. Again, statisticians were not the primary driving force behind the ascent of RCT; but still they played an important role because they provided their expertise and tools to the movement of therapeutic reformers and, later on, to medical researchers who gradually came to rely on RCT to conduct their studies. Yet, while they were the owners of reliable tools for judging strengths and weight of evidence, they were also aware and not infrequently reminded that medicine was not their domain, it belonged to medical researchers (Marks, 2000b). Statisticians were eager to mark the territory of their expertise and to avoid venturing into areas where their knowledge could be challenged. In 1976, prominent American statistician Jerome Cornfield, one of the first sympathetic commenters on Schwartz and Lellouch’s work, reflected on how statisticians, by then firmly entrenched in the clinical trials field, distanced themselves from problems of decision-making in clinical practice in an attempt to adhere to the erected boundaries. Their first move, according to Cornfield, was to delineate how statistics as a field related to questions of decision-making in general: “It is not universally accepted that the theoretical analysis of decision making is a useful part of statistics. The Fisherian view is that it may be fit for business and tyranny, but surely not for the high, free purposes of science ...” (Cornfield, 1976: 409). Engaging with problems of making decisions in practice

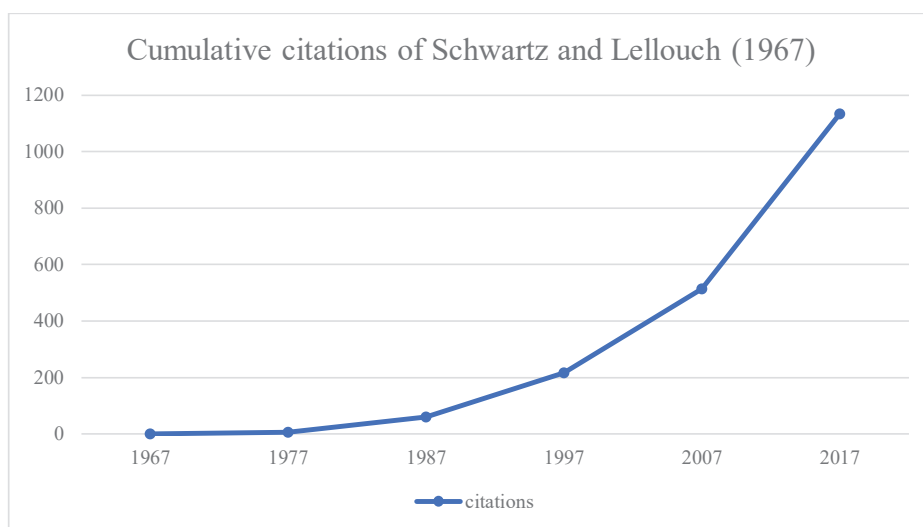


Figure 1. Cumulative citations of Schwartz and Lellouch (1967).

was not a generally accepted component of statistical expertise back then. The second move was to distinguish issues that pertained to the domain of statistics specifically in clinical trials and decouple them from the issues of decision-making in clinical practice:

A common attitude towards these problems [of decision-making] may be paraphrased as follows: "Decisions, although important, involve non-statistical issues and should be distinguished from the purely statistical issues, which consist of asking what the data show and how certain are the conclusions they will support. Once these are known, decisions and their costs can be considered, but preferably by someone else." (Cornfield, 1976: 410-411)

Therefore, statisticians, being the primary audience Schwartz and Lellouch appealed to, were reluctant to answer the call. Considering pragmatic questions such as "Which treatment should we prefer?" as Schwartz and Lellouch (1967) proposed, would require a major revision of the field's self-conception. It is not surprising that "[t]he existence of a decision-making, or as Schwartz and Lellouch ... put it, pragmatic function in clinical trials was almost entirely neglected in the original formulations [of RCTs by statisticians]" (Cornfield, 1976: 408).

Another group that could have answered Schwartz and Lellouch's call were therapeutic reformers themselves. However, those aspiring to elevate the scientific standards for judging the effects of medical treatments were busy with their own quest (Matthews, 1995). They led a campaign to persuade medical researchers to use methods of modern statistical experimentation and to convince medical practitioners to rely on RCTs as yardsticks for measuring claims of pharmaceutical companies. This campaign relied on straightforward messages meant to impress the medical audience with the opportunities opened up by statistical methods for achieving greater certainty and objectivity. In such endeavour, there was little space for delving into the subtleties behind statistical procedures.

It is illustrative here how reasons for randomisation were discussed among statisticians and how they were originally conveyed to the

medical audience. Ronald Aylmer Fisher (1926), whose work became a cornerstone of the statistical theory of experimental design, proposed to use randomisation for assigning treatments to be able to estimate random error variance and obtain a measure of uncertainty that characterised the experimental results and not at all to ensure homogeneity across and, hence, comparability of the groups in an experiment. For him, randomisation allowed establishing the validity of inference² (Armitage, 2003). Fisher conducted most of his experiments in agriculture, not in medicine, though. The entry of randomisation into clinical trials was aided by another statistician and epidemiologist, Austin Bradford Hill, who strived to make it attractive to medical audiences. Hill relied on a set of completely different arguments (Chalmers, 2011). Randomisation, he wrote, "ensures that neither our personal idiosyncrasies (our likes or dislikes consciously or unwittingly applied) nor our lack of balanced judgement has entered into the construction of the different treatment groups —the allocation has been outside our control and the groups are therefore unbiased" (Hill, 1952: 115). That is, the reformers offered randomisation to medical community as a technique to avoid prejudice and free researchers from the pains of ensuring comparability of the groups in an experiment. Randomisation, when used in an RCT in conjunction with other recommended techniques such as blinding, was basically presented as an assurance that results are safeguarded of bias and, therefore, trustworthy.

That such promises steered clear from statistical theory and were presented as a matter of common sense certainly added to their appeal. Yet, in the pursuit of an impact, the campaign for placing clinical practice on a scientific basis by means of the RCT swept under the carpet the complexity and limitations of statistical methods. Admission that statisticians disagree, let alone an engagement in discussion of conflicting approaches to the RCT, could temper the emerging enthusiasm for the RCT and potentially undermine the movement. Moreover, making the controversy public about just how much relevance clinical trial results have for making decisions in clinical practice, would damage the very central claim

of therapeutic reformers that RCTs are useful precisely for physicians. Therefore, advocates of the RCT were not keen to acknowledge the call for pragmatism and all the challenges involved in it.

Physicians, in turn, while having their interests most directly tied to pragmatism in clinical trials, tended to be too unfamiliar with statistical foundations and reasoning to consider the difference between the explanatory and pragmatic approaches relevant. The limitations of explanatory approaches to trials appeared to be embedded in the arcana of statistical theory which was rarely a significant part of physicians' education or subject they would regularly encounter in their daily work. Consequently, all audiences who could potentially take part in the discussion proposed by Schwartz and Lellouch either lacked interest or would have their own agendas directly threatened by such discussion.

The pragmatic turn

While the ideas of Schwartz and Lellouch initially failed to give rise to discussion, the notion of pragmatism in clinical trials did get traction some thirty years later. In 2003, when the rise of attention to pragmatic trials became visible, a group of primary care researchers wrote:

To a great extent the conduct of pragmatic trials is a recent phenomenon. While one of the earlier descriptions of pragmatic versus explanatory trials was by Schwartz in 1967 ... most of the published editorials considering pragmatic trials as a methodology have been since 1998 ... A Medline search ... yielded 34 articles reporting on pragmatic clinical trials. All 34 were published since 1995 and 26 of them were published since 2000. (Godwin et al., 2003)

A number of shifts enabled this turn to pragmatic trials. For one, during the years following Schwartz and Lellouch's publication, the prevalent thinking among statisticians about the mission of statistics in general and its role in medical research in particular changed. Peter Armitage, a past president of the Royal Statistical Society, expressed the newer attitude in the following way: "We can accept ... the implied limitations of statistical investigation, without in any way depreciating

the contributions of statistical investigations, and clinical trials in particular, to the technology of therapeutic medicine — as helping to show what is *useful*, rather than what is *true*" (Armitage, 1998: 2677, italics in original). This emphasis on usefulness signalled a departure from adherence to the narrowly conceived 'territory of statistics' and the willingness to engage with clinical practice and its concerns. While statisticians did not constitute the major driving force propelling the rise of interest in pragmatic trials, the reversal of the field's self-conception created an opening for engaging with pragmatic questions and contributing to a long overdue discussion.

What appears to have been decisive for making the time ripe for the pragmatic turn is the unlikely convergence of patients' actions for recognition of their needs, the slower than expected uptake of medical research findings by physicians and the consolidation of efficiency-focused healthcare management approaches. Since the 1980s, groups such as HIV/AIDS activists entered the relatively insular world of clinical research and demanded a place in designing and carrying out clinical trials along with medical researchers and statisticians. Their actions triggered changes in drug approval standards to increase access to experimental drugs and facilitated modifications in trial procedures to increase flexibility and responsiveness of trial protocols and to use outcome measures meaningful for patients (Epstein, 1996). Addition of these new participants in research planning made pragmatic questions such as "Which of the two treatments should we prefer?" not only legitimate but urgent for clinical research.

Pressure to make trials more 'useful' for making decisions in actual practice also came from those concerned with the fate of evidence-based medicine (EBM), a powerful movement that therapeutic reformers of the past intellectually flowed to. Physicians' enthusiasm for timely incorporation of the results of well-designed experiments into their practice appeared to lag behind what proponents of scientific medicine had hoped for (Pope, 2003). The medical research community's reflections on the reasons for this disappointment tend to come back to the crucial obstacle: ordinary physicians rarely see these results as relevant (Cranney et al., 2001; Haynes et al., 1997).

The diversity of patients, conditions, and circumstances physicians face leads them to doubt the applicability of research results in their daily work. Consequently, for those concerned with sustaining EBM's momentum, one central course of action has focused on reshaping clinical research to allow physicians to recognise the realities of their work in medical experiments. It is important to highlight here that the original proponents of RCTs also saw their efforts as directed at providing clinical practice with useful knowledge. However, usefulness of this knowledge was meant to stem mostly from avoidance of bias in its production through the use of traditional RCT methodology. Such unbiased knowledge was meant to substantiate decision-making in clinical practice. But the growth of interest to pragmatic trials signalled the emergence of thinking that keeping bias at check was not sufficient to ensure relevance and usefulness of trial-produced knowledge for making decisions in practice.

Last but not least, changes in how healthcare is organised have also made the pragmatist challenge more pertinent. Recent decades have seen the evolution of managerial approaches to governing clinical practice, with a growing group of decision makers taking upon themselves the task of ensuring uniform quality of services provided to patients, while keeping expenditures at bay (Calvert et al., 2011; Muir Gray, 2004). Ever-rising healthcare costs placed matters of choice on top of these decision makers' agendas. Which drugs should be reimbursed given that reimbursing every drug that a physician may want to prescribe is not feasible? Which treatments should necessarily be offered by health providers for specific conditions? Which procedures need to be excluded from treatment plans as not providing additional advantages commensurate with their higher costs? Consequently, more requests began to arrive for research to evaluate medical treatments taking into account parameters important for making such choices. Pragmatic trials, with their aspiration to improve the link between clinical research and decision-making in clinical practice, appear to have affinity with the concerns of this group of healthcare managers as well.

Reinterpretation

With the alignment of these different actors' interests around making clinical trials more useful, the Schwartz and Lellouch's notion of pragmatism was not only dusted off and put to service, but also reinterpreted. The reinterpreted version, while perhaps more palatable to the diverse members of the pragmatic trials bandwagon, bears little resemblance to the two statisticians' thinking in 1967.

We can trace the change in thinking about designing trials for informing clinical practice through close reading of the three recent publications that have been central to shaping the contemporary views on pragmatic trials. The first of these publications presented an extension of the influential Consolidated Standards of Reporting Trials (CONSORT) endorsed by multiple medical journals and editorial organisations (CONSORT group, n.d.). This extension guides investigators in preparing reports of findings from specifically pragmatic trials (Zwarenstein et al., 2008). The second and third of these publications offered readers a tool titled PRECIS (PRagmatic Explanatory Continuum Indicator Summary) for distinguishing parameters suitable for pragmatic and explanatory trials. The tool was presented in two versions: PRECIS-1, very similar to the CONSORT extension, and PRECIS-2 that developed the tool further (Loudon et al., 2015; Thorpe et al., 2009). The guideline and the tool were created by health services researchers, such as Merrick Zwarenstein and Shaun Treweek and a few statisticians, such as Kevin E. Thorpe and Douglas G. Altman together with an international group of trialists. From the beginning, the authors packaged their views in such formats (a standard and a tool with clear-cut design options) that invited practical application and accorded to additional influence and reach to their work. Zwarenstein wrote on his personal web page: "This guideline, which forms part of the internationally recognized ... CONSORT statement has influenced the way they [pragmatic trials] are described and published" (Zwarenstein, n.d.). In such ways, Zwarenstein and others have popularised particular characteristics as hallmarks of pragmatic trials and made certain considerations almost obligatory for those who like to conduct a pragmatic trial. Widely cited, their contribu-

Table 1. Questions for explanatory and pragmatic trials.

	Explanatory	Pragmatic
Schwartz and Lellouch	Does the drug have a specific effect? (Schwartz and Lellouch used sensitising effect to radiation in their example)	Which of the two treatments should we prefer?
Authors of CONSORT extension and PRECIS tools	Can this intervention work under ideal conditions?	Does this intervention work under usual conditions?

tions do not simply reflect a general consensus regarding properties of pragmatic trials, but also disseminate specific ideas about what pragmatic approach entails. These ideas involve implicit assumptions regarding the nature of experiment, the ‘real world’ and the relations between the two.

Let us first take a look at how the questions to be answered by pragmatic and explanatory trials are formulated by the authors of the CONSORT extension and PRECIS tools. They write: “Pragmatic trials seek to answer the question, “Does this intervention work under usual conditions?,” whereas explanatory trials are focused on the question, “Can this intervention work under ideal conditions?”” (Thorpe et al., 2009: 464). Compare these questions with the questions envisioned by Schwartz and Lellouch (see Table 1).

For Schwartz and Lellouch the difference between explanatory and pragmatic trial is the difference between distinguishing a causal connection in a laboratory and making a decision in clinical practice, all things considered. Whereas for contemporary pragmatists, the difference between pragmatic and explanatory trial collapses into a difference between the conditions under which an intervention is tested.

What is understood here as ‘ideal’ and ‘usual’ conditions? The contemporary authors clarify that ideal conditions that characterise explanatory trials conducted in laboratory settings are such that “give the initiative under evaluation its best chance to demonstrate a beneficial effect” (Loudon et al., 2015: 1). Ideal circumstances that maximise the chances of success, according to the authors, include trial participants most likely to adhere and respond to an intervention, highly trained and experienced practitioners delivering an intervention, well-resourced setting and strict

standardisation of an intervention and its delivery. Illustrative here is a statement with which the authors convey that irrespective of the amount of efforts invested, trialists can never carry out an entirely explanatory trial: “no patients are perpetually compliant, and the hand of the most skilled surgeon occasionally slips, so there can never be a “pure” explanatory trial” (Thorpe et al., 2009: 465). So, this is what characterises the ideal conditions that the explanatory trial ostensibly aspires to maintain: full adherence, comprehensive knowledge, no mistakes and complete availability of all necessary resources.

The usual conditions which exist outside of laboratory, in real-world settings, in contrast, are marred by all possible imperfections and variation, which interfere with the performance of the intervention being tested. To achieve its purpose of determining “the effects of an intervention under the usual conditions in which it will be applied” (Thorpe et al., 2009: 464), the pragmatic trial is to preserve these imperfections and variation. Instead of aiming to cancel out the noise of doctor-patient relationships, patients’ life circumstances, physicians’ attitudes and organisational routines to achieve a clean picture of causes and effects, pragmatic trials need to preserve the messiness of the usual conditions to see how an intervention would behave in the wild. Will it be able to withstand the adverse conditions? According to contemporary pragmatists, this task exceeds by far in difficulty the challenges met by those conducting an explanatory trial. On one hand, the difficulty here appears to be in engaging the clinical practice conditions into the experiment and running a trial in such a way that it changes these conditions as little as possible:

... the act of conducting an otherwise pragmatic trial may impose some control resulting in the setting being not quite usual. For example, the very act of collecting data required for a trial that would not otherwise be collected in usual practice could be a sufficient trigger to modify participant behavior in unanticipated ways. (Thorpe et al., 2009: 465)

On the other hand, this advocated absence of control to preserve the usual conditions still appears to be in need of strict control. Messiness that ends up being engaged in a trial may diverge from the messiness in "the settings for which a trial is intended to provide an answer" (Thorpe et al., 2009: 467):

For some interventions what is usual for each domain may vary across different settings. For example, the responsiveness and compliance of patients, adherence of practitioners to guidelines, and the training and experience of practitioners may be different in different settings. (Thorpe et al., 2009: 467)

Contemporary pragmatists offer to read pragmatic trials as requiring release of the strict control that is characteristic of explanatory trials to open a door to messiness characteristic of practice, but doing it in a controlled fashion to ensure that the imperfections and variation cherished now within the trial correspond to those that are usual for a particular target setting. Control here is directed at ensuring that messiness within a trial is the correct kind of messiness.

The PRECIS tools in essence are meant to help trialists with exactly these tasks: to establish and maintain similarity between trial conditions and the real world (i.e. conditions of actual clinical practice, according to pragmatists) or, one can say, what is deemed usual for a particular segment of the real world. Let us take a closer look at the PRECIS-2, the latest version of the tool, and see how it offers to ensure that the experiment is conducted under the correct kind of usual conditions. PRECIS-2 presents nine domains, each corresponding to a range of choices that can move a trial closer to or farther from what is considered the real world, thus making a trial more or less pragmatic. These domains include

eligibility, recruitment, setting, organization, flexibility (delivery), flexibility (adherence), follow-up, primary outcome, and primary analysis and were visualised by Loudon et al (2015) as a wheel:

Each domain encourages trialists to think about their trial and the recipients in the usual care situation in which their results might be applied if the intervention proves beneficial. If trialists are aiming for high applicability (that is, a pragmatic approach to design decisions), then we would expect the match between trial and usual care to be very good. (Loudon et al., 2015: 3-4)

The tool offers trialists to consider how pragmatic or explanatory their choice in each domain should be for the purposes of their trial, from 1 (very explanatory) to 5 (very pragmatic).

Applicability of trial results, Loudon et al. (2015: 2) wrote, "is the outcome of these choices, which affect the ease with which the trial results can be applied to and by the usual community of users of the intervention in the settings in which the trial designers envisioned it being used". In the contemporary reinterpretation of pragmatism in clinical trials, shaped to a large extent by the authors of the CONSORT extension and PRECIS tools, pragmatic trial aids clinical practice through maximising applicability of its results. We can understand the nine domains of the PRECIS-2 tool, then, as control points investigators are encouraged to use to juxtapose a trial with usual clinical practice. Through establishing similarity between the two, applicability of trial results to a particular segment of the real world is to be established.

Overall, pragmatic trial in its contemporary formulation broke in a number of significant ways with what Schwartz and Lellouch imagined. Contemporary authors chose to focus on what French statisticians called 'normal and laboratory conditions' as a primary demarcation criteria between pragmatic and explanatory trials and develop it further while putting aside other considerations offered in the 1967 article. This move is conscious. The PRECIS-1 publication indicated Thorpe and colleagues' awareness that, when introducing the idea of pragmatism, Schwartz and Lellouch were concerned with much more than 'normal and laboratory conditions': "Schwartz and Lellouch clearly linked the

ability of a trial to meet its purpose with decisions about how the trial is designed. ... [Yet,] how useful a trial is depends not only on design but on the similarity between the user's context and that of the trial" (Thorpe et al., 2009: 474). The contemporary authors warned to "not confound the structure of a trial with its usefulness to potential users" (Thorpe et al., 2009: 474) and in the rest of their writings about pragmatic trials proceeded to focus exclusively on the conditions within and outside the trial. In this interpretation, the pragmatic trial strives to achieve a similarity between the trial conditions and conditions in what is called the real world. In this way pragmatic turn seeks to change only conditions within the trial and not necessarily trial design principles and certainly not the methodological backbone of the RCT — randomization, the use of control groups, and, where possible, blinding.

Discussion: Pragmatic turn and contextualization

Tracing the changes in how pragmatic clinical trials are conceived, we can notice discontinuity in thinking about pragmatism in clinical trials. Schwartz and Lellouch's (1967) original notion had to do with dilemmas and choices that emerge in clinical practice. The specificity of these dilemmas and choices lies in that decision in practice is rarely about selecting one or another active pharmaceutical ingredient. The decision in practice tends to be between different modes or strategies of therapy, which include costs, ways of administration, additional care, particularities of the patient's condition and much more. Schwartz and Lellouch called all of these 'context' and proposed to make trials pragmatic by contextualizing them, which involves letting treatments being compared absorb the context. Trials that rely on such inclusive definitions of treatments together with other pragmatic strategies are useful because they can aid decision-making by helping choose a superior treatment, broadly conceived, as opposed to distinguishing whether a drug has a certain type of effect.

Three decades after Schwartz and Lellouch's article was published, health research community turned to the notion of pragmatism in clinical

trials and reinterpreted it. Contemporary pragmatists too argue in favour of contextualization as key to pragmatism but conceive it differently. Here contextualization involves making conditions in clinical trial similar to conditions in clinical practice, often in a particular location. Contemporary pragmatists' starting point is a fundamental difference between the laboratory (traditional RCT) and the real world (usual clinical practice) in terms of behaviour of a medical intervention. This difference threatens the usefulness of trial results because how a drug behaves in a laboratory-like explanatory trial may resemble little what it would do after being let loose in the clinic, which is just too imperfect to sustain the laboratory results. A pragmatic trial, then, is a trial conducted under what is considered the usual conditions in a setting where a tested intervention is to be used, as opposed to sanitised and orderly laboratory conditions. The greatest benefit of pragmatic trials thus conceived is that their results are deemed to be more applicable in clinical practice, since the imperfections of the world, such as suboptimal adherence, differential availability of resources and variability of physician treatment strategies, which plague the clinic, have already been factored in evaluating the effects of the experimental treatment.

It is not hard to notice that the two versions of pragmatic trials attempt to bring benefits to medical practice via very different routes. Contemporary enthusiasts strive for applicability understood as a synonym or, one can say, an outcome of generalizability. Increased similarity of conditions in a trial and in a clinic granted by a pragmatic design leads to greater generalizability since patients and routines appear to be more representative of the usual care. And the more generalizable trial results are, the more applicable they are taken to be as well. This is not to say that relations between generalizability and applicability in trials are always understood in this way, but this understanding is firmly embedded in the claims of contemporary pragmatists. The described line of thinking is reminiscent of a wider discussion in biomedical literature about efficacy (which is tested in traditional RCTs with their ideal conditions) and effectiveness (which contemporary pragmatic trials with their real-world condi-

tions attempt to test) (Flay, 1986; Gartlehner et al., 2006; Glasgow et al., 2003). But this line of thinking has nothing to do with Schwartz and Lellouch's proposal. In fact, Schwartz and Lellouch were not concerned with generalizability as such. Yes, they mention 'the usual conditions' that pragmatic trials need to involve, but this is only one and rather minor component of their proposal. They were concerned with making trials more *useful*, i.e. asking relevant questions, looking for outcomes that make a difference, and defining experimental treatments and comparators in a way that makes sense in clinical practice. In short, usefulness was to be achieved through defining and designing medical experiments in a way that engages with concerns of clinical practice. But when difference between traditional and pragmatic RCTs is casted simply as that of efficacy and effectiveness as is commonly done now, the question of usefulness is not on the table anymore because it is assumed that if trial results are more generalizable and, hence, applicable then they are also more useful. However, usefulness and applicability as it is currently conceived in pragmatic trials field are very different beasts and when the quest for usefulness is abandoned and only applicability is sought instead, the promise of more contextualized Mode-2 type of clinical research cannot be realised.

The comparison of the outlined versions of pragmatism in trials also makes visible just how much the contemporary version relies on separating controlled inside and uncontrolled outside in the medical experimentation. This is how STS scholars have theorised a laboratory: as a result of a process that distinguishes an inside, an environment where only those influences are allowed that are considered relevant for making a certain epistemic claim, and an outside, an environment full of noise and irrelevant disturbances (Guggenheim, 2012). Such separation implies analytical differentiation between nature (drug's true effects, for instance) and human culture (routines and relationships that constitute clinical practice). STS scholars also highlighted problematic character of this differentiation since it does not do justice to the inextricable connection between nature and culture and does not necessarily make either one more knowable (Callon et al., 2009; Jasanoff, 2011).

At least some of the challenges faced by clinical research now, such as results that would not hold and lack of trials designed to answer questions pertinent to practice, stem from taking for granted this long-established dichotomy. Contemporary pragmatists are aware of the consequences of keeping nature and culture strictly apart in clinical trials and seek to bridge the divide. In their daring attempt, however, they still do not seem to come far enough and abandon the divide. Instead, they extend the RCT's methodological backbone into clinical practice in order to involve messiness of the clinic as one more variable that cannot be ignored anymore and needs to be factored in. This move addresses the problem of external validity, making trial produced knowledge more generalizable to certain practice settings. But it does not necessarily make trial-produced knowledge as useful as it can be.

Schwartz and Lellouch's version of pragmatism, in contrast, starts with much less divisive notion, that of a decision that needs to be made in clinical practice. This decision is hybrid, necessarily combining elements of both nature and culture. In effort to decide, which treatment we should prefer, pragmatic trial seeks to align diverse elements such as data, interests of patients, experimental methodology, care strategies, side effects and ways to tackle them, and many more. Pragmatism here, instead of solving the problem of the great divide between the laboratory and the real world, avoids it altogether by locating itself in the space where elements of both intertwine. In doing so, early pragmatists opened the door to not only make trial produced knowledge more generalizable, but, first of all, to make it more useful by fully taking on board conditions, needs and concerns of clinical practice from the very beginning. In practice, in order to take their insights seriously and move towards more contextualized knowledge production, pragmatic trials could, apart from adhering to conditions usual for clinical setting, begin from research questions collaboratively defined by investigators together with those who are expected to use research results later on. In this way pragmatic trials would stem from choices physicians and patients have to make and, thus, provide answers more capable of making a difference in practice setting. Treat-

ments being compared in trials could be broadly and flexibly defined to include their optimal usage conditions in clinic to further enhance applicability of trial results. Also, conclusions about superiority and inferiority of investigated treatments could be made on a broad basis to include considerations relevant to different users, beyond narrowly understood efficacy. Discarding early pragmatists' insights now would mean losing

an opportunity to strengthen contextualization of clinical research in a sense of its societal embedding, responsiveness, and relevance to the diverse needs experienced in clinical setting.

Acknowledgements

This work was supported by Russian Science Foundation (project # 18-78-10016)

References

- Armitage P (1998) Attitudes in clinical trials. *Statistics in Medicine* 17(23): 2675–2683.
- Armitage P (2003) Fisher, Bradford Hill, and randomization. *International Journal of Epidemiology* 32(6): 925–928.
- Bothwell LE and Podolsky SH (2016) The emergence of the randomized, controlled trial. *The New England Journal of Medicine* 375(6): 501–504.
- Bothwell LE, Greene JA, Podolsky SH, et al. (2016) Assessing the gold standard — lessons from the history of RCTs. *The New England Journal of Medicine* 374(22): 2175–2181.
- Bower P (2003) Efficacy in evidence-based practice. *Clinical Psychology & Psychotherapy* 10(6): 328–336.
- Brass EP (2010) The gap between clinical trials and clinical practice: The use of pragmatic clinical trials to inform regulatory decision making. *Clinical Pharmacology & Therapeutics* 87(3): 351–355.
- Callon M, Lascoumes P and Barthe Y (2009) *Acting in an Uncertain World: An Essay on Technical Democracy*. Cambridge, Massachusetts: MIT Press.
- Calvert M, Wood J and Freemantle N (2011) Designing “real-world” trials to meet the needs of health policy makers at marketing authorization. *Journal of Clinical Epidemiology* 64(7): 711–717.
- Cartwright N (2007) Are RCTs the gold standard? *BioSocieties* 2(1): 11–20.
- Chalmers I (2001) Comparing like with like: some historical milestones in the evolution of methods to create unbiased comparison groups in therapeutic experiments. *International Journal of Epidemiology* 30(5): 1156–1164.
- Chalmers I (2011) Why the 1948 MRC trial of streptomycin used treatment allocation based on random numbers. *Journal of the Royal Society of Medicine* 104(9): 383–386.
- CONSORT group (n.d.) Welcome to the CONSORT Website. Available at: <http://www.consort-statement.org> (accessed 14 September 2019).
- Cornfield J (1976) Recent methodological contributions to clinical trials. *American Journal of Epidemiology* 104(4): 408–421.
- Cranney M, Warren E, Barton S, et al. (2001) Why do GPs not implement evidence-based guidelines? A descriptive study. *Family Practice* 18(4): 359–363.
- Dodd LE, Lantos JD and Pinheiro J (2016) University of Pennsylvania 8th annual conference on statistical issues in clinical trials: Pragmatic clinical trials (morning panel). *Clinical Trials: Journal of the Society for Clinical Trials* 13(5): 493–503.
- Dopson S, Locock L, Gabbay J, et al. (2003) Evidence-based medicine and the implementation gap. *health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine* 7(3): 311–330.
- Epstein S (1996) *Impure Science: AIDS, Activism, and the Politics of Knowledge*. Berkeley, CA: University of California Press.
- Fisher RA (1926) The arrangement of field experiments. *Journal of the Ministry of Agriculture of Great Britain* 33: 503–513.
- Flay BR (1986) Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine* 15(5): 451–474.
- Ford I and Norrie J (2016) Pragmatic Trials. *New England Journal of Medicine* 375(5): 454–463.
- Gartlehner G, Hansen RA, Nissman D, et al. (2006) A simple and valid tool distinguished efficacy from effectiveness studies. *Journal of Clinical Epidemiology* 59(10): 1040–1048.
- Gibbons M (1999) Science’s new social contract with society. *Nature* 402(S6761): C81–C84.

- Glasgow RE, Lichtenstein E and Marcus AC (2003) Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health* 93(8): 1261–1267.
- Godwin M, Ruhland L, Casson I, et al. (2003) Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. *BMC Medical Research Methodology* 3(28).
- Guggenheim M (2012) Laboratizing and de-laboratizing the world: changing sociological concepts for places of knowledge production. *History of the Human Sciences* 25(1): 99–118.
- Haynes RB, Sackett DL, Guyatt GH, et al. (1997) Transferring evidence from research into practice: Overcoming barriers to application. *ACP Journal Club* 126(3): A14-15.
- Higgs E (2000) Medical statistics, patronage and the state: The development of the MRC Statistical Unit, 1911–1948. *Medical History* 44(3): 323–340.
- Hill B (1952) The clinical trial. *The New England Journal of Medicine* 247(4): 113-119.
- Jasanoff S (2011) *Designs on Nature: Science and Democracy in Europe and the United States*. Princeton and Oxford: Princeton University Press.
- Latour B (1993) *We Have Never Been Modern* (tran. C Porter). Cambridge, Massachusetts: Harvard University Press.
- Loudon K, Treweek S, Sullivan F, et al. (2015) The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 350: h2147.
- Mainland D (1960) The use and misuse of statistics in medical publications. *Clinical Pharmacology & Therapeutics* July-August(1): 411–422.
- Marks HM (2000a) *The Progress of Experiment: Science and Therapeutic Reform in the United States, 1900-1990*. Cambridge: Cambridge University Press.
- Marks HM (2000b) Trust and mistrust in the marketplace: statistics and clinical research, 1945–1960. *History of Science* 38(3): 343–355.
- Matthews JR (1995) *Quantification and the Quest for Medical Certainty*. Princeton, N.J: Princeton University Press.
- Moreira T and Will C (2016) *Medical Proofs/Social Experiments: Clinical Trials in Shifting Contexts*. London and New York: Routledge.
- Muir Gray JA (2004) Evidence based policy making. *BMJ* 329(7473): 988-989.
- Nowotny H, Scott P and Gibbons, M (2001) *Re-thinking science: Knowledge and the public in an age of uncertainty*. Cambridge, UK: Polity.
- Petty J and Heimer CA (2011) Extending the rails: How research reshapes clinics. *Social Studies of Science* 41(3): 337–360.
- Podolsky SH (2010) Antibiotics and the social history of the controlled clinical trial, 1950–1970. *Journal of the History of Medicine and Allied Sciences* 65(3): 327–367.
- Pope C (2003) Resisting evidence: the study of evidence-based medicine as a contemporary social movement. *Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine* 7(3): 267–282.
- Rosengarten M and Savransky M (2019) A careful biomedicine? Generalization and abstraction in RCTs. *Critical Public Health* 29(2): 181–191.
- Sanders D and Haines A (2006) Implementation research is needed to achieve international health goals. *PLoS Medicine* 3(6): e186.

- Schwartz D and Lellouch J (1967) Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Chronic Diseases* 20(8): 637–648.
- Simpson B and Sariola S (2012) Blinding authority: Randomized clinical trials and the production of global scientific knowledge in contemporary Sri Lanka. *Science, Technology, & Human Values* 37(5): 555–575.
- Thorpe KE, Zwarenstein M, Oxman AD, et al. (2009) A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *Journal of Clinical Epidemiology* 62(5): 464–475.
- Timmermans S and Berg M (2003) *The Gold Standard. The Challenge of Evidence-Based Medicine and Standardization in Health Care*. Philadelphia: Temple University Press.
- Treweek S and Zwarenstein M (2009) Making trials matter: pragmatic and explanatory trials and the problem of applicability. *Trials* 10(37).
- Tunis SR, Stryer DB and Clancy CM (2003) Practical clinical trials: Increasing the value of clinical research for decision making in clinical and health policy. *JAMA* 290(12): 1624–1632.
- Ware JH and Hamel MB (2011) Pragmatic trials — guides to better patient care? *New England Journal of Medicine* 364(18): 1685–1687.
- Will CM (2007) The alchemy of clinical trials. *BioSocieties* 2(1): 85–99.
- Zuidgeest MGP, Goetz I, Groenwold RHH, et al. (2017) Series: Pragmatic trials and real world evidence: Paper 1. Introduction. *Journal of Clinical Epidemiology* 88: 7–13.
- Zvonareva O, Engel N, Kutishenko N, et al. (2017) (Re)configuring research value: international commercial clinical trials in the Russian Federation. *BioSocieties* 12(3): 392–414.
- Zwarenstein M (n.d.) Merrick Zwarenstein, MBBCh, MSc, PhD, homepage. Available at: https://www.schulich.uwo.ca/familymedicine/research/csfm/people/bios/zwarenstein_merrick.html (accessed 10 May 2019).
- Zwarenstein M and Treweek S (2009) What kind of randomized trials do we need? *CMAJ*. 180(10): 998–1000.
- Zwarenstein M, Treweek S, Gagnier JJ, et al. (2008) Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 337: a2390.

Notes

- 1 The terms 'pragmatic' and 'pragmatism' are associated with American Pragmatists, including, first of all, William James and John Dewey, and school of philosophy developed by them. When these terms are applied currently to clinical trials, this is done in a manner quite distinct from what was originally proposed by Pragmatist philosophers.
- 2 For instance, in 1926, Fisher wrote about evaluating new crops and fertilisers: "One way of making sure that a valid estimate of error will be obtained is to arrange the plots deliberately at random, so that no distinction can creep in between pairs of plots treated alike and pairs treated differently; in such a case an estimate of error, derived in the usual way from the variation of sets of plots treated alike, may be applied to test the significance of the observed difference between the averages of plots treated differently" (Fisher, 1926: 506-507).