

‘Pragmatic’ and ‘explanatory’ attitudes to randomised trials

Merrick Zwarenstein

Centre for Studies in Family Medicine, Department of Family Medicine, Western University, London, Ontario N6A 3K7, Canada

Corresponding author: Merrick Zwarenstein. Email: Merrick.Zwarenstein@ices.on.ca

Conceptualising two types of randomised control trials

In the early days of randomised control trials the ability to reduce confounding from known and unknown confounders rightly underpinned their rapid rise in methodological popularity. Ever since, however, concerns have been expressed by the clinician community and others about the difficulty of applying estimates of effects derived from randomised control trials to settings and patient populations outside those in which randomised control trials have often been done. Applicability, also known as generalisability or external validity, is important because systematic reviews of randomised control trials are used to inform decisions about whether patients in the wider world, that is, outside the trials contributing to systematic reviews, would benefit from receiving the treatment evaluated in those systematic reviews. These decisions take for granted that random allocation is required to minimise confounding in making treatment comparisons. Rather, they consider the similarity or differences between patients, practitioners, settings and approaches between the care of the patients in the trials included in systematic reviews, and those of patients, settings and care of patients outside the trials. This article provides an account of how, since the 1960s, an unfolding understanding of the issue of applicability led to the recognition that there might be different purposes for different kinds of randomised trials.

Following publication of the iconic report of the United Kingdom Medical Research Council’s randomised trial of streptomycin for pulmonary tuberculosis in 1948,¹ expository articles setting out the characteristics of and ground rules for randomised control trials began to appear, among which the best known are two articles by Austin Bradford Hill, a British statistician.^{2,3} By the end of the 1950s, the Council for International Organizations of Medical Sciences (established under the auspices of UNESCO and the WHO) had recognised that randomised control trials were of such importance that it convened an

international conference to discuss them in Vienna, and asked Bradford Hill to organise it. All the formal presentations were by British investigators, and their papers were published in a book edited by Hill,⁴ Bird⁵ and Chalmers.⁶ The proceedings were also translated into French and published in a book edited by Daniel Schwartz⁷ – a French statistician – and three French statistician colleagues.

At this early stage of its development, the purpose of a randomised trial was seen as singular and self-evident: to use randomisation to create groups comparable on known and unknown confounders, and thus obtain unbiased estimates of the effects of two or more interventions. In 1966, Marvin Schneiderman, a statistician at the US National Cancer Institute who had been involved in trials of cancer therapy in the early 1950s, summarised his thinking in a working paper prepared for a WHO Expert Committee on Cancer Treatment. His report noted that the singular purpose of randomised control trials had become dual:

In [cancer] chemotherapy at least, there appear to be two different kinds of trials, conducted for distinctly different purposes. There are the patient-orientated trial, and the drug-orientated trial. Patient-orientated trials are designed to give answers to the question ‘How shall I treat the next patient with cancer who comes into my care?’ The drug-orientated trials attempt to answer the questions ‘Has this drug enough promise that I can bring it into patient-orientated trials?’ and ‘If I were to bring it into a patient orientated trial, how is it best to give it?’ (p. 5)⁸

Schneiderman conceived of the two kinds of trials – both of which used random allocation to generate comparable groups – as evaluating treatment interventions for two different purposes. One was to inform the choice between alternative treatments for the current, real world patients seen by clinicians, while the alternative purpose was to understand whether the treatment had any promise. The wording ‘drug-oriented

trials' suggests something akin to what we would now think of as phase 2 trials, usually conducted during the early development and testing stages of a potential treatment. Under Schneiderman's later guidance, as Chief Statistician at the US National Cancer Institute, this division of the two types of trials into early and later stage trials led to the three phases of drug trial with which we are now familiar.

In an article published the following year, two French statisticians, Daniel Schwartz and his colleague, Joseph Lellouch, citing Schneiderman and others, and acknowledging that their viewpoint was not new, revisited this issue at greater length and depth,⁹ and went on to publish a book on the topic – essentially an extension of the paper.¹⁰ In the 1967 paper, Schwartz and Lellouch worked with a British statistician, Michael Healy, who although not listed as an author is acknowledged in the paper as contributing 'much more than a translator'. In this paper, Schwartz and Lellouch suggest, as had Schneiderman, that the goal of a clinical trial could be conceptualised as addressing one of two entirely different purposes. They designated the two types of trial as 'explanatory' and 'pragmatic'. As Schwartz and Lellouch state in the very first sentence of their paper:

It is the thesis of this paper that most therapeutic trials are inadequately formulated, and this from the earliest stages of their conception. Their inadequacy is basic, in that the trials may be aimed at the solution of one or other of two radically different kinds of problem; the resulting ambiguity affects the definition of the treatments, the assessment of the results, the choice of subjects and the way in which the treatments are compared.⁹

Schwartz and Lellouch proposed a different dichotomy from that of Schneiderman. While they agreed with him that each randomised control trial could be designed for one of two clearly defined, mutually exclusive and opposite purposes, one of the purposes they proposed was similar, and one different to Schneiderman's: (i) to confirm (or refute) a causal hypothesis about the mechanism of action of a treatment (different from Schneiderman's concept of the drug-related trial which evaluated the potential, or early stage promise, for that treatment) or (ii) to choose between alternative treatments (similar to Schneiderman). For Schwartz and Lellouch, the former kind of randomised control trial aimed to test hypotheses about the mechanisms of action of treatments, while the second was aimed at providing information to inform real world clinical choices. These two kinds of problem were each to be dealt with using a different approach to design in order to produce the information required. Schwartz

and Lellouch's rigorous way of thinking through the consequences of the duality in purpose of 'explanatory' and 'pragmatic' randomised control trials has had a major influence on our thinking about all randomised control trials, including those which are intended to be 'explanatory'. They summarise the two approaches as follows:

...(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.⁹

The goal of the current article is to explain the differences between 'explanatory' and 'pragmatic' attitudes to the design of randomised trials, as conceived by Schwartz and Lellouch. The text is organised in two parts: the first focuses on some of the less well-known conceptual and statistical distinctions between these two attitudes to trial design, as developed by Schwartz and Lellouch. The second part focuses on a number of specific aspects of randomised control trial design, such as inclusion criteria, setting, efforts to maintain adherence and measurement of outcomes, initially laid out by Schwartz and Lellouch, and subsequently popularised by David Sackett.¹¹

The less well-known ideas of Schwartz and Lellouch

Implications of 'pragmatic' and 'explanatory' objectives for choice of controls

Schwartz and Lellouch began their 1967 paper with a cleverly formulated hypothetical example of a randomised trial of two approaches to treating a particular but unspecified cancer: radiotherapy alone versus radiotherapy preceded by 30 days of sensitising chemotherapy. They argue that if the purpose of the trial is 'explanatory' (purely to test the hypothesis that sensitising chemotherapy improves the action of radiotherapy) then the radiotherapy in each arm of the trial should be given at the same time. This would mean that in the radiotherapy alone group, it would be delayed until day 31, to make it exactly comparable to the radiotherapy delivered postchemotherapy. This approach renders the two arms comparable in relation to the radiotherapy, the only difference being the presence or absence of

chemotherapy. It thus tests directly the hypothesis that chemotherapeutic presensitisation increases the effect of follow-on radiotherapy.

If, however, the purpose of the trial is ‘pragmatic’, that is, simply to identify which treatment is better, then it is clear that in the radiotherapy-only arm, the radiotherapy should be delivered immediately in the comparison group, at the same time as the sensitising chemotherapy is given to patients in the experimental intervention arm (in which the chemotherapy is started immediately and radiotherapy begins at day 31). This renders the start point of treatment in the two arms comparable, exactly as it would be delivered in the real world (without any delay in starting the treatment regimen for patients in either arm).

And so too with placebo comparison groups, whose purpose in ‘explanatory’ trials is to render the arms of the trial comparable by eliminating any difference between them that might arise from different subjective belief in the effectiveness of treatment, in other words, to eliminate one mechanism of action so as to leave the field clear to evaluate whether another – the hypothesised mechanism of action – exists. In ‘pragmatic’ trials, by contrast, the effect of belief and other contextual factors that may contribute to or detract from the biological mechanism of action of an intervention, whether they are due to a different ordering of treatments, or to differences in subjective belief, are simply absorbed as part of the intervention itself, with no attempt to eliminate their effects. In a ‘pragmatic’ trial there is little or no interest in confirming or refuting the hypothesis that a particular biological mechanism of action that might be leading to an effect. Instead there is a strong drive to measure the overall difference in effects in patients in the comparison groups, irrespective of the mechanism of those effects. A ‘pragmatic’ trial does not presume to comment on mechanism of effect, only on existence or not of these effects.

In fact, these two approaches to evaluating the effects of medical interventions had been distinguished in Paris more than 200 years previously, well before the formal description of randomised trials. There was intense public debate over the arrangements for assessing Franz Mesmer’s claims for the therapeutic value of ‘animal magnetism’. Mesmer¹² had proposed an unblinded randomised comparison of patients treated by him using ‘animal magnetism’, with similar patients treated by the methods used by orthodox medicine at the time. Mesmer was in no doubt that patients would feel better after he had administered ‘animal magnetism’, and that this would be confirmed in the (unblinded) trial he proposed. Members of the Royal Commission charged with assessing ‘animal magnetism’, by contrast, were concerned primarily with

whether there was a physical basis on which ‘animal magnetism’ could work.¹³ Their blinded trial found no evidence to support the belief that ‘animal magnetism’ existed. The unblinded design would not have been able to distinguish the placebo effect, arising from Mesmer’s powerful suggestive approach, from any effect due to some underlying physiological mechanism. So, where the issue at stake is the existence (or not) of a proposed mechanism of action, clearly the preferred attitude to trial design would be what came to be known, after Schwartz and Lellouch’s work, as ‘explanatory’. But where the issue is the choice between alternative interventions, the attitude which will result in a trial designed to mimic real world care to the greatest extent possible – the ‘pragmatic’ attitude – is likely to produce the most relevant and useful results.

Origins of the ‘intention-to-treat’ approach to analysis

The second ‘lost’ idea from Schwartz and Lellouch turns out to be a possible first mention of the now widely accepted ‘intention-to-treat’ approach to analysis (JLL explanatory essay).¹⁴ Schwartz and Lellouch developed this by thinking about how to handle withdrawals in each of their types of randomised trial. For ‘pragmatic’ trials, withdrawals or failures to take the prescribed treatment are simply part of the usual run of events during use of a treatment in the real world. The fact that patients withdraw is no reason to exclude them from the analysis in the arm to which they were randomised. Indeed, far from withdrawing them, retaining these patients in the trial reflects better what will happen when the treatment is applied to real patients in the real world, and because no patients in intervention or control arm are removed from the analysis, randomisation continues to ensure comparability.

In the ‘explanatory’ view, by contrast, withdrawal implies that the patient concerned may not have been suited for the treatment; in other words, that the initial definition of eligible patients was incorrect, suggesting that these patients should be excluded from the analysis. The proportion of withdrawals or the reasons for withdrawal may differ between the arms of the trial. Schwartz and Lellouch point out that there is no reliable and valid way to impute the impact of these withdrawals on the estimates of effects, given that the reasons for withdrawals are often unknowable.

The problems arising from loss of patients after randomisation became clear in the anturane reinfarction trial. At the request of the Food and Drug Administration, a reanalysis of the original data by David Demetz, an American statistician, revealed

that a number of patients had been excluded from the results submitted for consideration by the FDA.¹⁵ Their re-inclusion moved the result outside of statistical significance and established a US and then global regulatory demand that all trials be analysed on an intention-to-treat basis. This is now almost always viewed as the primary analysis, even in trials which are ‘explanatory’ in their intent. So, irrespective of whether a trial takes a ‘pragmatic’ or an ‘explanatory’ attitude, we now consider that the primary analyses should be based on all the participants randomly allocated to interventions in the trial – regardless of their adherence to interventions allocated – thus observing the so-called intention-to-treat principle (DeMetz, 2016, personal communication).

Choice of statistical error on which to focus analysis

The third subtlety in thinking that Schwartz and Lellouch brought to the randomised trial relates to the choice of statistical error on which to focus the analysis. There are two possible outcomes for any two-arm randomised trial: either differential effects of the treatments compared with statistical confidence, or no such differences will be detected. ‘Explanatory’ randomised trial usually focus on deciding which treatment is superior. They try to minimise both the chances of falsely declaring a positive or negative difference when there is in truth no difference (an alpha, or type 1, error) and the probability of concluding that there is no difference when in fact there is a real difference (a beta, or type 2, error). In ‘pragmatic’ trials, if the alternative treatments truly have a similar effect, we don’t care which one we recommend, since this error has no negative consequences for patients: either treatment is acceptable when the treatments do have similar effects. This means we can set the value of type 1 error at 100%, i.e. simply declare that type 1 error is always present (and so we will never declare superiority). And since we will always make a type 1 error, we can essentially ignore type 2 error and set it at zero, and thus declare that our statistical power to detect a difference between treatments is 100%.

Schwartz and Lellouch then point to an often ignored third error term, gamma, which is the probability of erroneously recommending the wrong treatment as superior, when it is actually inferior. This is precisely the goal of ‘pragmatic’ randomised trial – to avoid recommending the wrong treatment, a goal which is distinctly different from the traditional one of recommending the superior treatment with a known and low degree of error. Gamma error is usually ignored because it is miniscule when alpha and beta

are at commonly chosen levels. At 0.05 and 0.2, respectively, gamma error is only 10–7. In a ‘pragmatic’ trial, Schwartz and Lellouch propose that the only error we need to avoid is the one in which we would incorrectly recommend as superior the inferior treatment. With alpha error set at 100% and beta at 0 the value for gamma error becomes potentially important, and we can then use gamma with a known power and probability, as the statistic on which to declare that we have not selected the inferior treatment, i.e. the selected treatment is as good or better than the alternative.

In pragmatic randomised trial comparing treatments, trial efficiency is paramount, and so sample size should not exceed the minimum needed. Choosing gamma as the error term of interest will have a substantial effect on required sample size, reducing it by more than half. Despite its efficiency, this Schwartz and Lellouch innovation appears not to have been adopted in the world of clinical trials. This may be because the most widespread role of randomised trial (until the recent advent of comparative effectiveness research) has been in industry-funded regulatory submissions. Sponsoring companies’ interests depend on setting alpha and beta error at traditional values (such as alpha at 0.05 and beta at 0.2) in order to identify small superiorities over competing drugs or placebos.

The better-known ideas of Schwartz and Lellouch

In the conclusion of their celebrated 1967 article, Schwartz and Lellouch.⁹ stated:

Most trials done hitherto have adopted the explanatory approach without question; the pragmatic approach would often have been more justifiable. It is thus not surprising if these [explanatory] trials, difficult enough in themselves, raise still further difficulties at every stage and finish by satisfying neither doctor nor statistician.

The main work of translating Schwartz and Lellouch into terms accessible to clinician scientists designing randomised control trials was done by David Sackett and his colleagues in the third edition of their book *Clinical Epidemiology* (2005). Sackett was primarily a clinical trialist, having co-authored or supported the design and analysis of over 200 randomised trial. His extract of Schwartz and Lellouch selected a number of key design features and offered a clinically oriented description of the distinction between the two kinds of trials. Sackett¹⁶ defined ten attributes of randomised trial design and analysis, focusing on patient inclusion, adherence, setting, outcomes, follow-up and inclusions

Table 1. Sackett's selected attributes of 'pragmatic' and 'explanatory' approaches to randomised control trials.

Element	Explanatory (or efficacy) trial	Pragmatic (or effectiveness or management) trial
The question	Can this Rx work under ideal circumstances?	Does this Rx benefit under usual circumstances?
Participant eligibility	Strict: restricted to high-risk, highly responsive, highly compliant.	Free: everyone with the condition of interest.
Experimental intervention	Inflexible, with strict instructions for every element. Both participants and practitioners are usually blind. Cross-overs are prohibited	Highly flexible, as it would be used in routine healthcare. Nobody is blind. Cross-overs are permitted.
Comparison intervention	Inflexible, with strict instructions (often employs a placebo). Both participants and practitioners are usually blind. Cross-overs are prohibited.	Usual care for this condition in this setting. Nobody is blind. Cross-overs are permitted.
Practitioner expertise	Only practitioners and settings with previously documented high expertise.	Full range of practitioners and settings in which a successful intervention would be applied.
Participant compliance with interventions	Closely monitored and may be a prerequisite for study entry. Both prophylactic strategies (to maintain) and 'rescue' strategies (to regain) high compliance are used.	Unobtrusive (or no) measurement of compliance. No special strategies to maintain or improve their compliance.
Practitioner adherence to protocols	Close monitoring into how well clinicians and centres are adhering to the trial protocol and 'manual of procedures', triggering vigorous interventions whenever deficient.	Unobtrusive (or no) measurement of practitioner adherence. No special strategies to maintain or improve their adherence.
Follow-up intensity	Frequent, highly intense, with extensive data collection.	Usual intensity for this condition and setting, or restricted to administrative databases on mortality and utilisation.
Primary outcome	A restricted set of events, composite outcomes or surrogate outcomes, often determined by blinded experts and adjudicators.	A broad set of events of importance to participants, determined in the routine course of healthcare.
Primary analysis	Might try to justify excluding non-compliers or non-responders.	Never deviates from 'intention-to-treat' analysis of all participants who entered the trial.

and exclusions from analysis, for each of which differences between 'pragmatic' and 'explanatory' approaches might exist (Table 1).

'Explanatory' attitude to trial design

Schwartz and Lellouch proposed that trials designed with an 'explanatory' attitude were aiming to confirm

or refute a hypothesis about the mechanism of action of a clinical treatment or intervention by assessing whether an intervention has the effects predicted from the hypothesised mechanism of action. 'Explanatory' trials are therefore designed in ways which maximise the contrast between the interventions being compared, increasing the strength of any signal of a real difference between the

intervention and control groups, and reducing ‘noise’ in the overall measurement. Efforts are made to minimise all sources of variation other than the interventions themselves. These design features ensure that the comparison is made under ‘optimum’ conditions, aimed at increasing the contrast between the intervention under study (usually new) and the comparator (old intervention, or placebo).

To increase the chances of treatment differences occurring and being detected, patients in ‘explanatory’ trials are often chosen to be at the more severe end of the spectrum of the target disease to maximise their ‘room to respond’ to the treatment under test. Patients may also be selected to have as few other sources of variation in response to treatment as possible so that differences in intervention effects are not masked or diluted. For example, ‘explanatory’ trials may exclude patients known after a pretrial testing period to adhere poorly to prescribed interventions. This ensures that patients who do participate in the trial are more likely to use the interventions as intended.

It is also important to use outcome measures that are likely to pick up any differences that exist between the comparison groups. Under intense scrutiny, multiple measurements of many parameters are taken to provide evidence on how the presumed mechanism of action is unfolding under treatment. To increase contrast between intervention and control groups, ‘explanatory’ trials usually choose short-term physiological, biochemical or other process measures which tap directly into the expected underlying mechanism of action, and which minimise the sample sizes required as the outcome can be measured on each patient, whereas more patient-relevant outcomes, such as mortality, may occur in only a small proportion of patients.

The highest safe dose of the new intervention under test is prescribed to increase the chances of detecting a difference, whereas the comparator may be a placebo or an average dose of a current treatment. Investigators monitor intervention use intensively, and remind participants and practitioners of protocols, so that the intervention is properly prescribed, administered and used.¹⁷ Since a specific mechanism of action is the principal target for understanding, the new intervention is most often compared to placebo, in order to eliminate perceptions as a cause of apparent changes in outcomes.¹⁸

Blinding also helps to reduce changes in outcome due to mechanisms other than those under study. When patients, clinicians delivering the intervention and those measuring the outcomes or analysing the data, or some combination of these, are blinded,

the impact of their beliefs on the interventions’ effectiveness on perceptions of outcome will be reduced. Of course, blinding is more important for subjective outcomes.

These ‘explanatory’ strategies reduce variability in trials, thus increasing the likelihood of detecting an effect arising from the suspected mechanism of action. Failure to detect any difference in outcomes between the interventions compared can then be attributed more confidently to a lack of differential intervention effects, rather than problems in delivery, dilution, responsiveness or failure to use the interventions properly. These kinds of restrictions mean that ‘explanatory’ trials tend to be done in ‘ideal’ settings, where trial staff have ‘control’, and adherence by practitioners and patients is measurable.

‘Pragmatic’ attitude to trial design

Schwartz and Lellouch contrasted ‘explanatory’ trials with ‘pragmatic’ trials. These are simpler and seek only data to inform direct comparisons between two or more alternative interventions under real world conditions. ‘Pragmatic’ trials may compare a new treatment with an existing treatment or several alternatives compared to each other. In ‘pragmatic’ trials the underlying mechanism of action is not of primary interest. All that is sought is an unbiased comparison of treatments as they are used in the real world.

The ‘pragmatic’ attitude aims to increase the range of healthcare settings, recipients and practitioners of the interventions in a trial, with a view to representing the range of those who will be prescribing and receiving those interventions or treatments shown to be useful and acceptably safe. Of course, there are many real worlds, and so judgement has to be applied to interpret the degree to which a given trial context, patients and situation applies to the specific patients or healthcare setting in which a decision maker (patient, clinician or policy maker) is operating. Trials at the ‘pragmatic’ end of the spectrum thus tend to have wide inclusion criteria, often based on ‘the uncertainty principle’, namely that patients are eligible to participate unless it is certain that they should not receive one of the interventions being compared. In other words, all patients who might be offered the intervention when widely applied should it be shown to be effective are eligible for inclusion in a ‘pragmatic’ trial.

‘Pragmatic’ trials tend to be flexible in the way in which the interventions being compared are delivered, with few or no restrictions on clinicians. They are able to provide care in whatever ways they would have done without the trial, under typical conditions

of the healthcare system in which they work. Since this is how they provide care in the real world – variably over time, between patients and among themselves, with little in the way of rigid adherence to strict rules – this can be expected to improve the applicability of the trial results to the real world. Making the interventions as realistic and flexible as possible will help to make the results more likely to inform their use by decision makers, whether at the locale of the trial, or more widely.

‘Pragmatic’ trials seldom monitor or intervene to support adherence by patients or practitioners to some predesigned, detailed intervention protocols. Follow-up of patients, practitioners or outcomes is minimised and kept unobtrusive in ‘pragmatic’ trials, to avoid intrusions that may change the behaviour of patients or clinicians from their normal, unmonitored practice.

The primary outcome in ‘pragmatic’ trials is usually chosen because it is of unambiguous importance to the research users whose uncertainties are being addressed, for example, mortality data from a reliable registration system.

Placebos are rarely considered appropriate in ‘pragmatic’ trials despite their informal use in everyday practice. Likewise, blinding is not acceptable in real world clinical care, so ‘pragmatic’ trials tend to use unblinded approaches, at least for the patient and the clinician delivering the treatment, but sometimes also for the person assessing the effects. Since ‘pragmatic’ trials focus on real world usefulness of an intervention, patient-oriented outcomes are preferred. This has the result of including in the final assessment of effects any which might be due to perception and observer biases. This is not viewed as problematic (even though it distorts the effects of underlying mechanisms of effect), because patient perceptions will be part of the usefulness of the treatment under real world conditions.

Implications for trial design

The two attitudes, ‘pragmatic’ and ‘explanatory’, highlighted by Schwartz and Lellouch are very different from each other and result in different approaches to trial design. Two multicentre trials evaluating the same drug and done and reported concurrently, one more ‘pragmatic’, the other more ‘explanatory’, illustrate the differences in execution if not in intention. The two studies were the first trials of intravenous streptokinase used for thrombolysis in myocardial infarction with more than 1000 patients – the Indexed Sequential Access Method (ISAM) trial with 1700 patients reported in June 1986,¹⁹ and the Gruppo Italiano per lo Studio della Streptochinasi

nell’Infarto (GISSI) trial with 11,806 patients reported in February 1986.²⁰ Both trials tested 1.5 million units of streptokinase infused intravenously in patients who had had a myocardial infarction.

The ‘explanatory’ trial

The more ‘explanatory’ trial, conducted in the US, was a multicentre trial, with blinding of patients and clinicians achieved by using placebo infusions matching the drug infusions. The sites participating in the trial all had to be specialist hospitals, and many exclusion criteria were applied in the selection of patients for participation (e.g. patients over 75 years of age were excluded). This resulted in a longer intake/eligibility process, thus reducing the applicability of the trial to patients seen in usual care settings. In particular, the time window for inclusion in the study (time is of the essence in dissolving the clot in a coronary artery) was narrowed from 12 to 6 h. Although it has since been shown that an even smaller number of hours is optimal, at the time, the choice to narrow the window also narrowed the applicability of the findings. It failed to answer the then current questions about the usefulness of new treatment for patients presenting during all of the 12 h after onset that was typical of the delays in hospitalisation and treatment common at that time.

The ISAM study’s declared primary outcome was all-cause mortality within 21 days. No statistically significant difference was found, but the conclusion of the ISAM study describes a ‘trend towards reduced mortality’ and focuses on the subgroup treated within 3 h of the onset of symptoms, in which a statistically significant effect was found. Despite having declared mortality as the trial’s primary outcome, the main discussion in the report of the trial was on a supposedly secondary question, the drug’s mechanism of action, specifically, ‘What is the effect of streptokinase on myocardial muscle injury, as measured by serum creatine kinase of myocardial origin?’ A statistically significant benefit of treatment on this outcome was found, using a complicated measure: time to peak serum levels of myocardial creatine kinase, with the size of the infarct estimated by integrating the area under the time-to-peak levels curve. This was followed up 3–4 weeks later on an available (and likely less severe) group of 848 survivor patients, in whom angiography was used to measure systolic ejection fractions. In the conclusion, these mechanism findings were given prominence. And in the final sentence of the paper, attention was drawn to the five patients with or suspected of having intracranial haemorrhages, with the sentence ‘Intracranial bleeding partly outweighs its beneficial effects’. This large

and no doubt expensive trial contributed little to clinical decision making, and even its conclusions on mechanisms are of doubtful meaning given that the primary outcome of the trial showed no statistically significant effect.

The ‘pragmatic’ trial

The more ‘pragmatic’ trial was done in Italy, where a national collaborative research group tested streptokinase in early acute myocardial infarction in 11,806 patients in 176 coronary care units (over 90% of the coronary care units in Italy). All patients admitted within 12 h after the reported onset of symptoms, and with no contraindications to streptokinase, were randomised by central telephone randomisation to receive this drug in addition to whatever usual treatment the participating hospital provided. The patients allocated to the control group received this usual (unspecified) care but did not receive streptokinase.

The trial was not blinded; patients were included during the usual flow of care, specific consent was waived, which simplified this process, and data collection was minimal. Only two exclusion criteria were applied (contraindications to streptokinase, and more than 12 h since onset of pain), resulting in a high inclusion rate. Of 31,826 patients admitted to hospital for myocardial infarction, 11,806 were randomised to enter the trial. Of the remaining 20,020 patients, 52.8% were ineligible because they reached hospital more than 12 h after the onset of symptoms; 20.8% were excluded due to contraindications to the drug; and only 9.5% excluded for administrative or unspecified reasons. Over 99% of the randomised patients were included in the analysis. The primary outcome was mortality at 21 days after the onset of symptoms, an obviously patient-relevant outcome measure easy to ascertain with great certainty. Overall mortality at 21 days was 10.7% in streptokinase recipients versus 13% in controls, an 18% reduction ($p = 0.0002$, relative risk 0.81).

The differences

The differences between these two near simultaneously published studies by separate research teams of the same treatment are marked. The ‘pragmatic’ study takes all-comers, from almost all possible sites in an entire country where the treatment could be offered, allows patients to be treated as would be habitual at those sites of care. The trial shows a distinct benefit in an important, patient-relevant outcome (and because of its huge size is able to use subgroup analysis to show that the window of

effectiveness is narrower than the inclusion criterion). By simplifying the inclusion and data collection processes, the trial could be made enormous, with tight confidence intervals around the estimates of this important effect. The trial was prominently published in *The Lancet* and frequently cited in later discussions of intravenous streptokinase treatment. This trial may well have changed practice globally, as was its explicit intent: ‘to test in a formal prospective trial whether effective and safe thrombolysis could be achieved with intravenous streptokinase under routine conditions in the majority of patients...’

This contrasts with the confusing approach of the ISAM trial, which combined an explicit primary outcome of mortality, almost ignored in the discussion, with an intense and costly focus on mechanism of action, resulting in a trial that was too small, yielding a difference in mortality that was not statistically significant, and was of little use in informing future clinical practice. Because the trial failed to show a clear benefit on mortality, the findings on the drug’s possible mechanism are moot.

The contrast between the two trials can be illustrated graphically (Figures 1 and 2) using PRECIS – the Pragmatic Explanatory Continuum Indicator Summary²¹ and the updated tool, PRECIS-2.²²

The ‘pragmatic–explanatory continuum’

As implied in the title of the Pragmatic Explanatory Continuum Indicator Summary (PRECIS) tool, most randomised control trials cannot be readily deemed either ‘pragmatic’ or ‘explanatory’ – dichotomously – but rather fall somewhere along a ‘pragmatic–explanatory’ continuum. This continuum is likely to be multi-axial, as illustrated in the PRECIS-2 wheels shown above (Figures 1 and 2). An important consequence of the development of the PRECIS tool is that it has made it possible to visually depict the position of a trial on this multi-axial continuum and also examine associations between the position of a trial on this multi-axial continuum and important characteristics of randomised trial, for example, to test the idea that that pragmatic designs are less likely to identify treatment differences.²³

In this article I have described the novel understanding of randomised trial developed by Schwartz and Lellouch in 1967, and some of the antecedents, and the features of their characterisation on a continuum between explanatory trials (for testing hypotheses about mechanisms of action) and pragmatic trials (for choosing among intervention options). This description included randomised trial design features that are already widely adhered to but

Figure 1. PRECIS 2 wheel: GISSI trial.

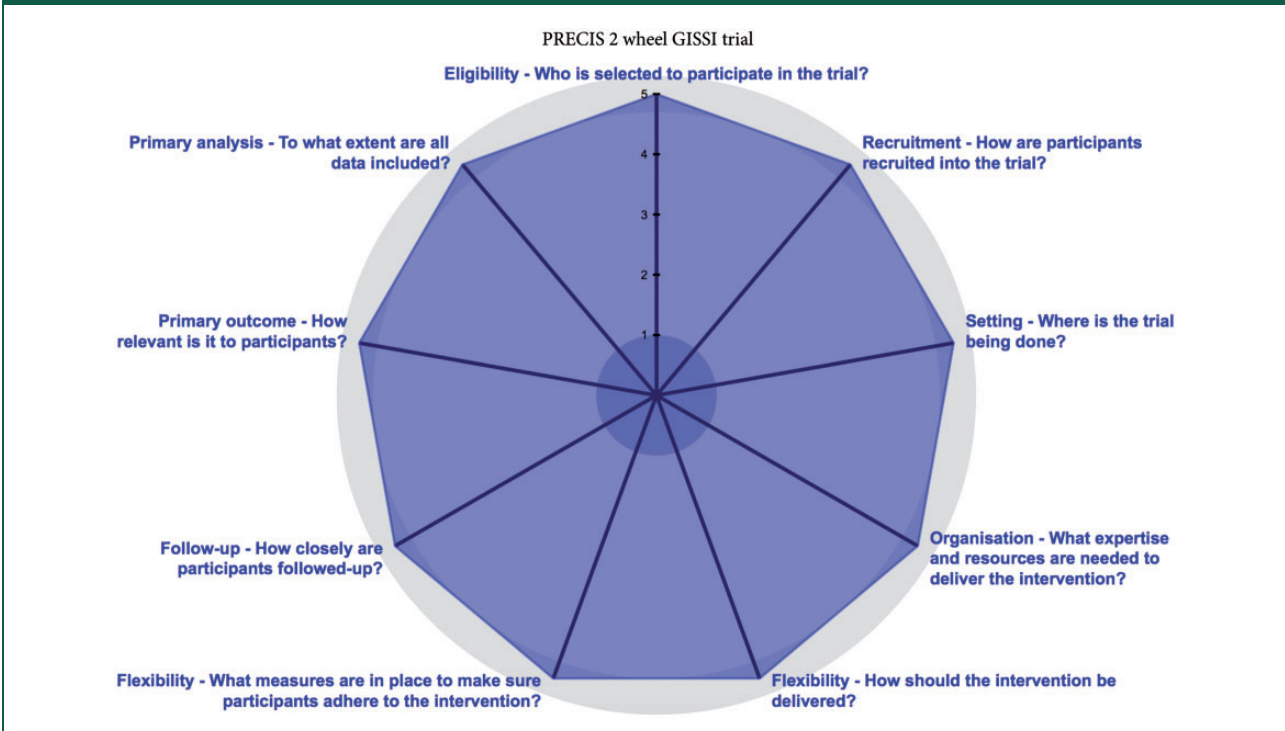
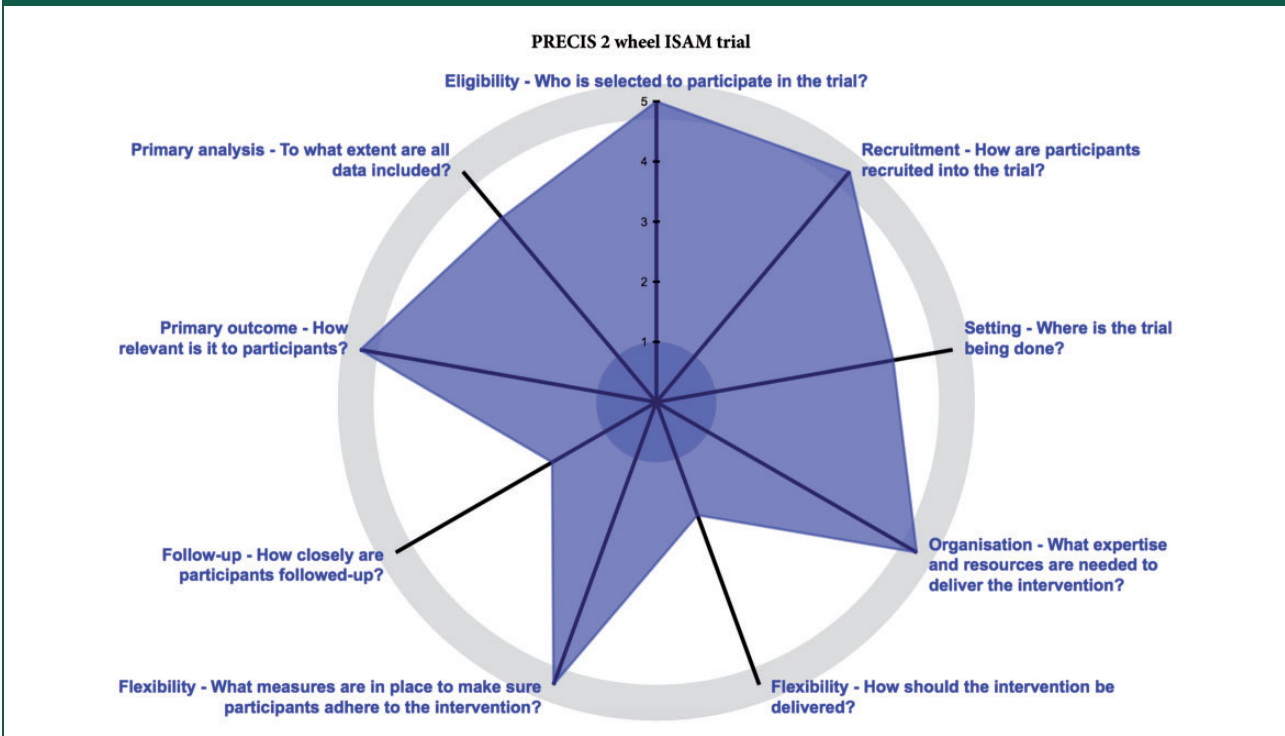


Figure 2. PRECIS 2 wheel: ISAM trial.



not attributed to Schwartz and Lellouch such as the intention-to-treat analysis, some that are less well understood, such as the alternative statistic for testing differences in outcomes, gamma and others that are increasingly widely cited. I look forward to the ever wider use of this set of concepts first clarified by Schwartz and Lellouch, and also of further development and use of the tools which have been built to make use of their work (PRECIS-2).²² Wider use of the Schwartz and Lellouch concepts should result in randomised trial with designs that align more closely with their intended purpose. I look forward to the further spread of Schwartz and Lellouch's ideas, and of the tools, such as PRECIS-2, built to help those who design randomised trial and use their results. I hope the resulting randomised trial will help us compare and select among the available clinical, service delivery or policy alternatives.

Declarations

Competing Interests: None declared.

Funding: None declared.

Ethical approval: Not applicable.

Guarantor: MZ.

Contributorship: Sole authorship.

Acknowledgements: I am grateful to Iain Chalmers for drawing my attention to the Mesmer experiments; to the 1960 publication by Schwartz, Flamant, Lellouch and Rouquette; and to the GISSI Trial as an example of a pragmatic trial.

Provenance: Invited article from the James Lind Library.

References

1. Medical Research Council. Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. *BMJ* 1948; 2: 769–782.
2. Hill AB. The clinical trial. *Br Med Bull* 1951; 7: 278–282.
3. Hill AB. The clinical trial. *N Engl J Med* 1952; 247: 113–119.
4. Hill AB. *Controlled Clinical Trials*. Oxford: Blackwell, 1960.
5. Bird SM. The 1959 meeting in Vienna on controlled clinical trials – a methodological landmark. *JLL Bulletin: Commentaries on the history of treatment evaluation*. See <http://www.jameslindlibrary.org/articles/the-1959-meeting-in-vienna-on-controlled-clinical-trials-a-methodological-landmark/> (2014, last checked 24 April 2017).
6. Chalmers I. UK Medical Research Council and multicentre clinical trials: from a damning report to international recognition. *JLL Bulletin: Commentaries on the history of treatment evaluation*. See <http://www.jameslindlibrary.org/articles/uk-medical-research-council-and-multicentre-clinical-trials-from-a-damning-report-to-international-recognition/> (2013, last checked 24 April 2017).
7. Schwartz D, Flamant R, Lellouch J and Rouquette C. *Les Essais Thérapeutiques Cliniques*. [Controlled Clinical Therapeutic Trials]. Paris: Masson, 1960.
8. Schneiderman MA. Therapeutic trials in cancer. In: *Working paper prepared for WHO Expert Committee on Cancer Treatment*, Geneva, Switzerland, 9–15 March 1965. WHO/CANC/66.66, 1966.
9. Schwartz D and Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967; 20: 637–648.
10. Schwartz D, Flamant R and Lellouch J. *L'essai Thérapeutique chez l'Homme*. Paris: Flammarion, 1970.
11. Haynes RB, Sackett DL, Guyatt GH and Tugwell P. *Clinical Epidemiology: How to Do Clinical Practice Research*. 3rd ed. Philadelphia, PA: Lippincott, Williams and Wilkins, 2005.
12. Mesmer FA. *Précis historique des faits relatifs au magnétisme animal jusques en avril 1781*. Par M. Mesmer, Docteur en Médecine de la Faculté de Vienne. Ouvrage traduit de l'Allemand [Historical account of facts relating to animal magnetism up to April 1781. By M. Mesmer, Doctor in Medicine of the Vienna Faculty. Work translated from German]. A Londres [false imprint, probably Paris.], 1781:111–114, 182.
13. Donaldson IML. Mesmer's 1780 proposal for a controlled trial to test his method of treatment using 'animal magnetism'. *JLL Bulletin: Commentaries on the history of treatment evaluation*. See <http://www.jameslindlibrary.org/articles/mesmers-1780-proposal-for-a-controlled-trial-to-test-his-method-of-treatment-using-animal-magnetism/> (2005, last checked 24 April 2017).
14. Furberg CD. How should one analyse and interpret clinical trials in which patients don't take the treatments assigned to them? *JLL Bulletin: Commentaries on the history of treatment evaluation*. See <http://www.jameslindlibrary.org/articles/how-should-one-analyse-and-interpret-clinical-trials-in-which-patients-dont-take-the-treatments-assigned-to-them/> (2009, last checked 24 April 2017).
15. Temple R and Pledger GW. The FDA's critique of the anturane reinfarction trial. *N Engl J Med* 1980; 303: 1488–1492.
16. Sackett DL. Explanatory and pragmatic clinical trials. *Arch Med Wewn* 2011; 121: 259–263.
17. Peto J. Reflections on the importance of strict adherence to treatment protocols: acute lymphoblastic leukaemia in children in the 1970s in the US and the UK. *JLL Bulletin: Commentaries on the history of treatment evaluation*. See <http://www.jameslindlibrary.org/articles/reflections-on-the-importance-of-strict-adherence-to-treatment-protocols-acute-lymphoblastic-leukaemia-in-children-in-the-1970s-in-the-us-and-the-uk/> (2016, last checked 24 April 2017).
18. Kaptchuk TJ. A brief history of the evolution of methods to control observer biases in tests of treatments. *JLL Bulletin: Commentaries on the history of treatment evaluation*. See <http://www.jameslindlibrary.org/articles/a-brief-history-of-the-evolution-of-methods-to>

control-of-observer-biases-in-tests-of-treatments/ (2011, last checked 24 April 2017).

19. Intravenous streptokinase in acute myocardial infarction (I.S.A.M.) Study group. A prospective trial of intravenous streptokinase in acute myocardial infarction (I.S.A.M.). Mortality, morbidity, and infarct size at 21 days. *N Engl J Med* 1986; 314: 1465–1471.
20. Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico (GISSI). Effectiveness of intravenous thrombolytic treatment in acute myocardial infarction. *Lancet* 1986; 1: 397–402.
21. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol* 2009; 62: 464–475.
22. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE and Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015; 350: h2147.
23. Young SL, Wolfenden L, Clinton-McHarg T, Waters E, Pettman TL, Steele E, et al. Exploring the pragmatic and explanatory study design on outcomes of systematic reviews of public health interventions: a case study on obesity prevention trials. *J Public Health* 2014; 36: 170–176.

How do you get more people
to read and cite your research?



EXPLAIN

it in plain language



SHARE

it via web, email
and social media



GET STARTED

www.growkudos.com

- ✓ Free service
- ✓ Quick and simple to use
- ✓ Proven to increase readership
- ✓ Keeps track of outreach wherever you do it
- ✓ Centralized reporting of full text downloads, altmetrics and citations
- ✓ Across all publications with a CrossRef DOI

KUDOS 
Greater Research Impact

SAGE
Publishing