

This is the authors' final, peer reviewed manuscript published in  
**Journal of Clinical Epidemiology** 2009, 62 (5):495-8 with the same title. Accepted 31 August 2008.  
For publishers version: [http://www.jclinepi.com/article/S0895-4356\(09\)00045-6/abstract](http://www.jclinepi.com/article/S0895-4356(09)00045-6/abstract)

## A pragmatic resolution

*Andrew D Oxman, Carl Lombard, Shaun Treweek, Joel J Gagnier, Malcolm Maclure, Merrick Zwarenstein*

In this response to the mechanists<sup>1</sup> we will first address what the mechanists have labelled as areas of possible disagreement and areas of clear disagreement. We will then describe what we believe is at the heart of our disagreement and conclude with what we consider to be a pragmatic resolution.

### **1. Is the pragmatic-explanatory framework useful to decision makers?**

We agree that it is useful for decision makers to understand the difference between pragmatic and explanatory trials, limitations of explanatory trials for informing decisions about practice, and the rationale for pragmatic trials. We disagree that the mechanistic-practical framework is useful for decision makers. Because it confounds the applicability of the results of a trial with whether the purpose of the trial at its inception was pragmatic or explanatory, it is more likely to confuse decision makers than enlighten them.

### **2. Will introducing two new terms (and definitions) cause more confusion than elucidation?**

In their response the mechanists argue that their new terms and definitions are, in fact, not new, but are essentially the same as the original purpose-based definitions of the terms “pragmatic” and “explanatory” as defined by Schwartz and Lellouch.<sup>2</sup> They further argue that the problem was not with the original terms or definitions, but with how they have been interpreted subsequently. We agree that we should stick to the original terms and definitions. We do not agree with the mechanists' new definitions for the reasons we gave in our previous comment.<sup>3</sup> We also do not agree that our colleagues who have interpreted the pragmatic-explanatory framework subsequent to Schwartz and Lellouch have confused purpose with structure. Indeed, just as Schwartz and Lellouch related the purpose of a trial to decisions about its structure, subsequent descriptions of the pragmatic-explanatory framework have also begun with essentially the same definitions and described how those relate to how trials are designed. We do, however, agree that explanations of the pragmatic-explanatory framework have not always adequately taken account of the fact that what is pragmatic under “usual” circumstances, depends on what is “usual”.

We are delighted that the mechanists are prepared to withdraw their “new” terms and we hope that they will agree in the end to continue with the same definitions of pragmatic (or practical)<sup>4</sup> and explanatory that were used by Schwartz and Lellouch and have been used by others subsequently, which we agree refer to the purpose of a trial.

### **3. Is the mechanistic framework more open to abuse than the pragmatic-explanatory framework?**

The mechanists rightly point out that any framework is open to abuse. We agree.

#### 4. Should trials always be broad?

Of course not. The mechanists state: “While there is great merit in keeping pragmatic trials broad, when serious doubts about subgroup effects exist, the initial pragmatic trial may appropriately be more narrow.” We agree entirely with this; we do not think there is anything new in this statement;<sup>2</sup> and we do not think our colleagues who have interpreted the pragmatic-explanatory framework over the past 40 years have suggested anything different. There are good reasons for both explanatory and pragmatic trials, and there are risks with both.<sup>2,5</sup> There are also good reasons for trials that are neither completely pragmatic nor completely explanatory, as is usually the case.<sup>6</sup> There are also good reasons for sometimes having narrow inclusion criteria.

We do not, however, agree with the mechanists’ subsequent assertion that: “While pragmatists aim to ask a question that is applicable to many, trials that naively enrol a wide breadth of participants provide an answer that is applicable to nobody.” Firstly, this assumes that those designing a trial with a wide breadth of participants were “naïve”. While this may occasionally be the case, we would argue that overly narrow inclusion criteria are more common and that rather than being “naïve” it is generally wise for pragmatic trials to start out with broad inclusion criteria, unless there are sound reasons for narrowing them. As noted by Schwartz, Lellouch and others, the results of a trial with broad inclusion criteria are applicable to the collective breadth of those included. If the results suggested that an intervention had little or no effect, this would mean that it did not work in everybody included in the study, not that it does not work in anybody. This is nothing new. The dilemma of whether a pragmatic or explanatory trial should come first has been around for at least 40 years,<sup>2</sup> and has been summarised previously in a 2x2 table (Table).<sup>5</sup>

**Table. The conclusions that can be drawn from explanatory and management trials\***

	Conclusion from this trial	
	Benefit clearly greater than harm (“Positive Result”)	Benefit clearly no greater than harm (“minimally important improvement” ruled out”)
<b>Explanatory Trial</b>	<u>Ambiguous</u> ; it “works” but will patients and clinicians jump through the hoops necessary for its success?	<u>Clearly sensible to abandon</u> this treatment for this condition
<b>Management Trial</b>	<u>Clearly worthwhile</u> to adopt this treatment	<u>Ambiguous</u> ; did it fail because it was worthless or because too few patients and clinicians followed directions and took their medicine?

\*From Sackett DL. Explanatory versus Management Trials. In: Haynes RB, Sackett DL, Guyatt GH, Tugwell P. Clinical Epidemiology: How to Do Clinical Practice Research. Philadelphia: Lippincott, Williams and Wilkins, 2005.

In the example of endarterectomy for carotid stenosis that the mechanists use, they assume prior knowledge of what the results would show. The problem is we do not know the results before doing the trial. For example, if someone had done a trial of patients with 50-70% stenosis first, there would still be a question about the effects of endarterectomy in patients with less stenosis.

All questions are always being answered by subgroup analyses (implicitly, if not explicitly). If a narrower trial was done first, the decision would be based on whatever information was available to suggest that that particular subgroup was likely to benefit and other groups would be less likely or unlikely to benefit, and the post-hoc interpretation would be based on direct evidence for that particular subgroup and whatever indirect evidence there was for other subgroups. Once there is direct evidence for different subgroups, such as people with different degrees of stenosis or different surgical procedures, a subgroup analysis needs to be used to answer the question of differential effects, whether it is done based on within study comparisons, between study comparisons or both.<sup>7</sup>

### **5. Is perspective more important than context?**

We would define context as the circumstances where and when an intervention is implemented that modify (or might modify) the effectiveness of the intervention. This clearly is closely related to the extent to which an intervention works under “usual” circumstances and there is not disagreement that this is therefore important to consider when specifying the pragmatic extremes of each relevant element of a trial’s design.<sup>6</sup>

On the other hand, it is not clear what the mechanists mean by “perspective”. The standard definition in economics and healthcare evaluation is the viewpoint from which a decision is taken. This has limited relevance to the design of a trial, but is important in determining which consequences to consider and which ones to ignore when making a decision, particularly which economic consequences to consider.<sup>8</sup> Thus, from the perspective of an individual patient relatively little weight might be given to the consequences of a decision on healthcare expenditures that are not paid out of pocket, whereas from the perspective of a policymaker with a societal perspective all expenditures would be included, regardless of who paid.

There are four problems with the mechanists focus on perspective, in addition to a lack of clarity as to what they mean by “perspective”. Firstly, the mechanists confound context and perspective when they describe someone’s “context” as “assessing the impact of introducing a new drug into the community (and ultimately whether to publicly fund that drug)”.<sup>1</sup>

Secondly, they assume that a “public health perspective” will favour broad inclusion criteria and a clinician-patient dyad’s or guideline developer’s perspective will favour narrower inclusion criteria, particularly with respect to compliance. This is a misunderstanding of the perspective of policymakers and others with a “public health perspective” and a misrepresentation of what guideline developers do or should do. Policymakers and others with a societal perspective are also interested in subgroup effects and, whenever possible, would want to have policies that take account of important differences in effects across different groups of patients, including differences in compliance, when making decisions such as whether to introduce or provide coverage for new drugs or other interventions. Clinical guideline developers, on the other hand, may take a broad perspective (including all important consequences) or a narrow perspective (at the extreme, only including consequences that are directly important to an individual patient). In other contexts some of the mechanists have argued that guideline developers should generally take a broad perspective.<sup>8</sup> Thus we can only assume that they mean something else by “perspective” in this context or that they are confused about what they mean as we are.

Thirdly, the mechanists ignore the extent to which it is possible to accurately judge how compliant a patient is likely to be. When patients report that they are noncompliant, there is a high likelihood that they are. However, self-reported compliance is less useful because patients may still have been noncompliant.<sup>9</sup> Physicians’ sensitivity for identifying noncompliant patients may be as low as 10%.<sup>9</sup> Thus, patients’ and clinicians’ understanding of what effects to expect based on an assumption of compliance may overestimate the effects that could be expected. The

average effect from the results of a trial that included typical patients with a typical range of compliance would then, in fact, provide a better estimate of the benefits that most patients could expect. Similarly, clinicians' have a limited ability to assess their own competence<sup>10</sup> Thus their understanding of what effects to expect based on an assumption of competence may also overestimate the effects that could be expected. Again, the average effect from the results of a trial that included typical clinicians with a typical range of competence would likely provide a better estimate of the benefits that most patients could expect.

Finally, the mechanists' assert that those arguing for more pragmatic trials "blindly advocate for broad trials" whereas individual patients and clinicians would often be more interested in the results of trials with narrower inclusion criteria. Patients and clinicians would undoubtedly like trials that reflect their exact circumstances as closely as possible (as would policymakers), but this ignores what is possible (see Box) and the extent to which the results of narrowly focused trials and subgroup analyses can be misleading.<sup>11</sup> We agree that there are sometimes good reasons for not having broad inclusion criteria, but the results of trials are always average results and clinician-patient dyads will never have information that is specific for a single patient, unless they do an N of 1 trial.

**Box. What is it reasonable for clinicians (and patients) to expect from a trial?**

*"It is right for each physician to **want** to know about the behavior to be expected from an intervention or therapy when applied to his particular individual patient (to whom the physician has the strongest ethical obligation). It is not right, however, for a physician to **expect** to know this – except, possibly, for the most dramatically effective and time-tested interventions or therapies. Most useful interventions or therapies change, for the better; the chance of a favorable outcome from a smaller chance to a larger chance. Most physicians and surgeons recognize this and do not demand (though they may rightfully ask for) detailed and reliable forecasts for individual patients."* JW Tukey<sup>12</sup>

**6. Should the term "mechanistic" be used instead of "explanatory" with corresponding changes in the definition?**

Having gone from wanting to introduce a "new" framework, then reverting to wanting to claim that their new framework was, in fact, the same as the original pragmatic-explanatory framework, the mechanists now again want to introduce what is, in fact, a new term and a new definition. We do not disagree with using the term "proof of concept" in describing the purpose of explanatory trials. We do disagree with the suggestion that the use of a surrogate outcome is the defining feature of explanatory trials.

There are ten ways in which the design of pragmatic and explanatory trials can differ, for each of these there is a continuum from extremely explanatory to extremely pragmatic, and trials can be pragmatic in some ways and explanatory in others.<sup>6</sup> The nature of the primary outcome is one way in which pragmatic and explanatory trials can differ. The use of an outcome, possibly a surrogate outcome, that the experimental intervention is expected to have a direct effect on is characteristic of explanatory trials, whereas pragmatic trials would characteristically measure patient-important outcomes, which would directly inform practice decisions. Although explanatory trials may help to understand mechanisms of treatment effects, they are primarily designed to test whether interventions have hypothesised effects under optimal circumstances, not necessarily to investigate "possible mechanisms of effect". The strength of explanatory trials is that a "negative" result can directly inform practice, since an intervention that does not work under optimal circumstances is unlikely to work under usual circumstances. The weakness of explanatory trials is that "positive" results do not directly inform practice, although they may directly inform practice under a narrow set of optimal circumstances and they can inform decisions about future research. Pragmatic trials, on the other hand have the opposite strengths

and weaknesses (Table). The weakness of pragmatic trials is that with “negative” results it is unclear whether the intervention is “worthless” or whether it might, in fact, be worthwhile under some (more optimal) circumstances or for a subgroup of patients. The strength of a pragmatic trial is that “positive” results can directly inform decisions under the “usual” conditions for which the trial was intended to be applicable.

We continue to find the terms pragmatic and explanatory and their definitions, which have been used over the past 40 years, useful. We do not find the mechanists’ arguments for using new terms, new definitions or a new framework compelling. We also find their arguments that their framework and concepts are at the same time “new” and consistent with the old terms and framework confusing.

### **At the heart of our disagreement**

Having agreed on the importance of context in determining what “usual circumstances” are, our approaches to addressing this are different in two important ways. First, the mechanists have selected new words and definitions. We think that is a mistake and will confuse people. It has certainly confused us. Second, having rightly pointed out the importance of context, the mechanists have attempted to address this by equating “practicality” with applicability. In doing so they have confounded how a trial is designed in relationship to its purpose with the applicability of its results. Our approach is to recognise the specific settings or conditions for which a trial is intended to be applicable prior to specifying the explanatory and pragmatic extremes of each domain and then determining how explanatory or pragmatic a trial is in relationship to those extremes for each of the ten ways in which explanatory and pragmatic trials can differ.<sup>6</sup>

### **A pragmatic resolution**

As the mechanists point out,<sup>1</sup> we agree about a lot, including the importance of being clear over the purpose of a trial and the importance of context in relationship to what is “normal” or usual. Of course this is not new. As noted 40 years ago by Schwartz and Lellouch:<sup>2</sup> “The distinction between “normal” and “laboratory” conditions clearly depends upon the level of current practice and would tend to vanish if this level were to rise”. We also agree that trialists should be clear about their purpose and context, that they should design their trial in relationship to its purpose, and that trials should be as broad as possible.

The starting point for the mechanists’ “new” framework was two perceived limitations in “the current interpretation of the pragmatic-explanatory framework”.<sup>13</sup> The first was that it confounds purpose with structure. We are not convinced this was a problem and we agree that the purpose of a trial should determine its design. The second limitation was that it ignores “the varying perspective of those using RCT results”. Having agreed that the primary usefulness of the pragmatic-explanatory framework is for those designing trials, not for users of trials, this no longer appears to be a problem.

Thus it seems the mechanists are left without an argument for a “new” framework and that we should continue to develop and use the pragmatic-explanatory framework, which has proven to be useful, although under-utilised, over the past 40 years.

**References**

- <sup>1</sup> Karanicolas PJ, Montori VM, Devereaux PJ, Schünemann H, Guyatt GH. The practicalists' resosne. *J Clin Epidemiol* 2008;
- <sup>2</sup> Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967; 20:637-48.
- <sup>3</sup> Oxman AD. Why I will remain a pragmatist: four problems with the impractical mechanistic framework and a better solution. *J Clin Epidemiol* 2008;
- <sup>4</sup> Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA* 2003; 290:1624-32.
- <sup>5</sup> Sackett DL. Explanatory versus Management Trials. In: Haynes RB, Sackett DL, Guyatt GH, Tugwell P. *Clinical Epidemiology: How to Do Clinical Practice Research*. Philadelphia: Lippincott, Williams and Wilkins, 2005.
- <sup>6</sup> Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, Tunis S, Bergel E, Harvey I, Magid DJ, Chalkidou K. A proposal for graphing randomized controlled trials within the pragmatic-explanatory continuum: PRECIS. *J Clin Epidemiol* 2008;
- <sup>7</sup> Caldwell DM, ADES AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331:897-900.
- <sup>8</sup> Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Vist GE, Schunemann HJ, and the GRADE Working Group. Incorporating considerations of resource use. *BMJ* 2008; 336:1170-3.
- <sup>9</sup> Stephenson BJ, Rowe BH, Haynes RB, Macharia WM, Leon G. Is this patient taking the treatment as prescribed. *JAMA* 1993; 269:2779-81.
- <sup>10</sup> Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA*. 2006; 296:1094-102.
- <sup>11</sup> Oxman AD, Guyatt GH. A consumers' guide to subgroup analyses. *Annals of Internal Medicine* 1992; 116:78-84.
- <sup>12</sup> Tukey JW. Some thoughts on clinical trials, especially problems of multiplicity. *Science* 1977; 198:679-84.
- <sup>13</sup> Karanicolas PJ, Montori VM, Devereaux PJ, Schünemann H, Guyatt GH. A new framework for designing and interpreting randomized trials. *J Clin Epidemiol* 2008;