NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Early Stopping in Pragmatic Clinical Trials: Workshop Summary

Introduction and Approach

There is general agreement that data monitoring is needed for pragmatic clinical trials. However, the design and operational features of pragmatic trials may influence how these obligations may best be met. One key issue relates to decisions regarding the possibility of early stopping. While 2 publications have described many issues related to data monitoring in pragmatic trials and have suggested that decision-making for early stopping may be different in some pragmatic trials as compared with traditional explanatory trials (Ellenberg et al 2015; Simon et al 2019), these proposals regarding early stopping merit further scrutiny. In addition, much like some implementation research, a data monitoring committee (DMC) in a pragmatic trial may not have the data needed to make accurate determinations about stopping a trial (Fiscella et al 2020).

Identification and evaluation of scientifically and ethically relevant considerations for early stopping in pragmatic clinical trials requires an interdisciplinary approach, involving those with expertise in pragmatic trials, trial design, ethics, and DMCs. A virtual workshop on January 13, 2021, engaged key stakeholders to deliberate these issues, organized around the following topics:

- Early stopping for futility
- Early stopping for safety
- Early stopping for efficacy
- Recommendations for policy, practice, and future research

This meeting summary synthesizes key themes and challenges arising from the discussion and suggests areas for future exploration.

Overall Key Themes

Several key questions emerged from the discussion that cut across the individual topics and do not seem to have been addressed in the existing literature:

- How is the heterogeneity of intervention delivery in many pragmatic clinical trials relevant for data monitoring?
- Given that many pragmatic trials do not involve obtaining individual consent from those included within them, if, when, and how should subjects be informed about issues related to data monitoring?

Prepared by: Stephanie Morain, Judith Carrithers, Steven Goodman, and Jeremy Sugarman for the NIH Pragmatic Trials Collaboratory Ethics and Regulatory Core
Version: March 29, 2021 (updated July 29, 2025)

- When is a DMC needed for a pragmatic trial, compared with another means of data monitoring? For example, a pragmatic trial that compares the effectiveness of alternative implementation approaches to delivering a standard treatment or care process may not need a DMC.

Early Stopping for Futility

Early termination of a trial for futility is intended to avoid unnecessary burden on those involved and prevent needless use of resources. Part of the DMC's role is to monitor whether the trial is being conducted in such a way that it will be informative. The discussion about early stopping for futility was complicated by questions about how to define futility in pragmatic clinical trials.

Meeting participants stressed the importance of distinguishing between intervention futility, enrollment failure, and intervention fidelity. In addition, it is important to consider how to weigh these assessments in pragmatic trials, given the differences in the types of research questions posed by pragmatic trials and explanatory trials and the embedded nature of pragmatic trials and their integration into ongoing clinical care.

Fidelity

To date, many, perhaps most, pragmatic clinical trials find little difference in effectiveness between the intervention and control groups. Workshop participants discussed the importance of monitoring intervention fidelity, using pilot testing and implementation science methods to improve fidelity, and setting conditions for determinations of futility. A challenge for DMCs is determining whether an apparent absence of difference in outcomes between study arms is a result of the intervention not being implemented correctly or a result of the intervention itself being ineffective. Therefore, a 2-stage approach may be appropriate in evaluating the possibility of futility: taking into account the fidelity of delivery (and variation in that fidelity across sites) and the actual effectiveness of the intervention itself.

Intervention fidelity alone can be an important outcome in pragmatic clinical trials, but it is complicated by the flexibility often allowed in pragmatic trials, where some degree of heterogeneity in intervention delivery is assumed as part of the design. While measurement and documentation of intervention fidelity or heterogeneity can make a pragmatic trial less pragmatic, adherence monitoring can play a key role in providing context for the difference (or absence of such a difference) between study arms. However, in some situations, it may be difficult or infeasible to collect and analyze data on implementation outcomes in ways that would enable a DMC to assess interim data with sufficient certainty to make a determination about early termination on the grounds of futility.

Defining Futility

Futility is often based on primary trial outcomes. However, a pragmatic clinical trial that finds no difference in the primary outcome between study arms but includes a rigorous

process evaluation can be informative if it provides information about, for example, feasibility or buy-in for implementation of a particular intervention. Consequently, even if there is futility for the primary outcome, it might be appropriate to continue the trial for secondary outcomes. The criteria for futility assessment should be agreed on by the trial sponsor, investigators, and DMC prior to study initiation.

Ultimately, DMCs should consider what constitutes futility for assessing the appropriateness of early stopping in a particular pragmatic trial. For example, when comparing treatments that are available and in widespread use, finding that one approach is very unlikely to be more effective than another could have tremendous value for clinicians, patients, payers, and other stakeholders. Given that both treatments are widely used, there is arguably no ethical duty to stop the trial, and there could be substantial social value in completing it as planned, so as to provide the strongest possible evidence for lack of difference in effectiveness between the treatments. When effectiveness is similar, practitioners and patients will likely be more comfortable making treatment decisions based on other factors such as types of side effects, convenience, and cost.

Early Stopping for Safety

Clinical trials are sometimes stopped early when the risks to those enrolled clearly outweigh the potential benefits. Nevertheless, making such a determination in pragmatic clinical trials may involve special considerations, including the risk of premature termination for safety when benefits and adverse events accrue at different rates, when the time of ascertainment of data on benefits and risks or the intensity of data collection differs between study arms, and when there are competing effects (whether positive or negative) in the same study arm. For example, contact with participants may be unequal across study arms, such as in trials in which data for an arm of a trial (for example, a usual care control group) are ascertained from healthcare system records rather than through additional interactions with study staff. In such trials, there could be a high risk of bias from direct comparison of adverse events between groups, which could mislead decision-makers trying to determine whether early stopping for safety is warranted.

When making decisions about early stopping for safety in pragmatic trials, DMCs may need to consider implementation challenges across institutions, implications of cluster randomization and other complex study designs that raise many other challenges, and potential risks to clinicians and institutions. Because many pragmatic trials use electronic health record data to identify adverse events, data are rarely available in real time to the DMC (or to investigators), making interim assessments about safety difficult or impossible.

Evidence supporting a decision to stop a trial must be robust and persuasive. In making this determination, DMCs should also consider the distinction between unanticipated events and unanticipated results. An example is the Women's Health Initiative hormone replacement therapy study, which examined the benefits and risks of hormone therapy for chronic disease prevention in predominately healthy postmenopausal women. While hormone replacement therapy was in widespread use because it was believed to prevent a

variety of conditions, including coronary heart disease, no large-scale randomized trial had compared the associated risks and benefits related to long-term health. While emerging safety information from the Women's Health Initiative suggested that hormone therapy was associated with net harm (rather than net benefit, as might have been expected), the DMC recognized the need for definitive data to change practice and thereby prevent potentially even greater harm for future patients. Consequently, the DMC waited until the evidence for harm was very strong; with a similar data pattern in a study of an investigational therapy not yet available, the DMC might have acted more quickly to recommend termination.

Finally, a pragmatic trial that compares the effectiveness of delivering a standard intervention "A" and an intervention "A-plus" (that is, the standard intervention plus an add-on strategy) may involve new and unknown risks. These A vs A-plus trials may require careful monitoring for potential adverse consequences of the "plus" that could result in the need to make decisions about early stopping on the grounds of safety. Here, the "plus" may not be something that an individual might have otherwise received as part of standard care, which can create different obligations to ensure they are not made worse than they might otherwise have been outside the pragmatic trial.

Early Stopping for Efficacy

Some clinical trials are stopped early because an interim analysis demonstrates efficacy. One principle involved with making such decisions in traditional explanatory trials is that research subjects should not be exposed to an experiment if there is not clinical equipoise. Panelists discussed whether clinical equipoise or other factors are the proper framework for determining whether to stop a trial for efficacy, or whether features of pragmatic clinical trials, such as the embedded nature of the trials or comparing treatments in widespread use, may shape the nature of obligations to those enrolled in pragmatic trials as compared to explanatory trials.

Pragmatic trials challenge assessments for whether and when to stop the trial early based on efficacy. For example, a pragmatic trial may test an intervention while addressing heterogeneous healthcare delivery across sites. If there are large differences in some sites and not others, should the trial be stopped in some sites or all sites? While site differences are also of interest in traditional trials, they may have particular relevance for pragmatic trials, which generally provide less specific direction to sites about implementation of interventions. Therefore, differences in sites may be more important in consideration for early stopping for efficacy in pragmatic trials. In this sense, a typical pragmatic trial has a double hypothesis: one about effectiveness, another about delivery, adherence, fidelity, and other operational outcomes.

Research vs Clinical Care

A major discussion point was whether there are obligations to patients that are different in research as opposed to clinical care, and, if so, how these obligations should be understood

in the context of pragmatic clinical trials, which deliberately integrate these two activities. This may affect whether there is an obligation to stop a pragmatic trial.

Pragmatic trials evaluate effectiveness, not efficacy. Therefore, should differences between outcomes be considered research risks or clinical risks? A related consideration involves whether considerations for stopping pragmatic trials based on evidence of overwhelming effectiveness might vary based on whether the superior intervention is in fact available to those receiving the presumed inferior intervention.

Factors to Consider

A fundamental ethical justification for DMCs is the need for independent monitoring of interim comparative data to ensure that the trial does not unfairly disadvantage individuals assigned to one study arm as compared to another. Pragmatic clinical trials often have multilevel heterogeneity, with a broad range of patients, communities, organizational factors (including healthcare system leaders, clinicians, and staff, their culture, their organizational systems, and their receptiveness). These factors militate against premature termination for benefit. One possible solution could be for all pragmatic trials to include some mixed methods or implementation science input to guide making such determinations, to provide additional context to help interpret study findings related to effectiveness.

Recommendations for Policy, Practice, and Future Research

In the final session, panelists recommended that it is helpful to return to original guidance regarding DMCs. The basic principles apply to all trials, whether pragmatic clinical trials or traditional explanatory clinical trials. However, what is required for DMCs to operationalize these basic principles in pragmatic trials may need to be distinct from that in explanatory trials, as previously described.

Although there is currently little available data regarding the deliberations of DMCs in pragmatic trials about the possibility of early stopping (and gathering such information would likely be worthwhile), there seems to be a need to clarify policies and practices for DMCs in pragmatic trials in this regard. For instance, the significance of different types of futility may be different in a pragmatic trial than in an explanatory trial. Considerations for early stopping should be carefully defined upfront, and data must be convincing to make a decision for early stopping. Furthermore, deliberation is warranted around trial integrity as a whole, including the role of the DMC in considering how to adjudicate the potential continued importance of primary and secondary outcomes given emerging data.

Finally, the workshop discussions brought to light the importance of several unsettled larger issues that affect the rather narrow remit of DMCs. These include concerns about properly allocating resources for research. A pragmatic trial that is unable to answer the research question because of futility or implementation challenges may not represent a good use of resources. Similarly, the fundamental approach to funding and operationalizing research may undermine the ability to rapidly obtain data. To meet the hopes and

expectations of having pragmatic clinical trial data available to inform healthcare decision-making, such challenges will need to be addressed.

References

Ellenberg SS, Culbertson R, Gillen DL, Goodman S, Schrandt S, Zirkle M. 2015. Data monitoring committees for pragmatic clinical trials. *Clin Trials*. 12(5):530-6. doi: 10.1177/1740774515597697. PMID: 26374679.

Fiscella K, Sanders M, Holder T, et al. 2020. The role of data and safety monitoring boards in implementation trials: When are they justified? *J Clin Transl Sci*. 4(3):229-232. doi: 10.1017/cts.2020.19. PMID: 32695494.

Simon GE, Shortreed SM, Rossom RC, Penfold RB, Sperl-Hillen JAM, O'Connor P. 2019. Principles and procedures for data and safety monitoring in pragmatic clinical trials. *Trials*. 20(1):690. doi: 10.1186/s13063-019-3869-3. PMID: 31815644.

Meeting Participants

Joe Ali
Jeremy Brown
Rob Califf
Judith Carrithers
Dave DeMets
Susan Ellenberg
Kevin Fiscella
Thomas Fleming
Andrew Garland
Steven Goodman
Adrian Hernandez
John Lantos
Susan Mendley
Stephanie Morain
Pearl O'Rourke
Frank Rockhold
Greg Simon
Kayte Spector-Bagdady
Catherine Stoney
Jeremy Sugarman
Ann Trontell
Wendy Weber
Dave Wendler
Ben Wilfond
Janet Wittes
Doug Zatzick