Published in final edited form as:

Clin Trials. 2016 October ; 13(5): 478–483. doi:10.1177/1740774516654862.

Statistical issues in pragmatic trials of behavioral economic interventions

Andrea B Troxel¹, David A Asch^{2,3}, and Kevin G Volpp^{2,3}

¹Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

²Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

³US Department of Veterans Affairs, Philadelphia, PA, USA

Abstract

Background—Randomized clinical trials provide gold-standard evidence for the efficacy of interventions, but have limitations, including highly selected populations that make inference on effectiveness difficult and a lack of ability to adapt and change midstream.

Methods—We propose two innovations for pragmatic trial design.

Results—Evidence-based evolutionary testing, a framework that allows adaptation of interventions and rapid-cycle innovation, preserves the power of randomization while acknowledging the need for adaptation and learning. An opt-out consent framework increases the fraction of the target population who participate in trials, but may lead to dampening of effect sizes.

Conclusion—Pragmatic trials offer numerous advantages in the evaluation of behavioral interventions in health. Statistical innovations, including evidence-based evolutionary testing and opt-out framing of consent and enrollment processes, can enhance the power of pragmatic trials and lead to more rapid progress.

Keywords

Effectiveness; generalizability; innovation; opt-out consent

Introduction

Randomized controlled trials (RCTs) have long been the gold standard for evaluation of therapies and health interventions—randomization provides the only solid foundation for

Reprints and permissions: sagepub.co.uk/journalsPermissions.nav

Corresponding author: Andrea B Troxel, Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, 622 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, USA. atroxel@mail.med.upenn.edu.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: A.B.T. is on the Scientific Advisory Board of VAL Health, a behavioral economics consulting firm. D.A.A. and K.G.V. are principals and owners of VAL Health.

direct comparison of groups unconfounded by measurable or unmeasurable characteristics.^{1,2} The classic RCT has well-known limitations, however. RCTs can be time-consuming and expensive. Most worryingly, they often have restrictive eligibility criteria, resulting in a highly selected trial sample that may not be representative of the target population. While randomization is critical to the valid comparison of groups exposed to intervention or control, eligibility restrictions often yield very different groups enrolled or not enrolled. This means that after considerable effort, expense, and time, RCTs may yield conclusions of limited value and applicability—providing high-confidence answers to questions of less broad value than intended. Pragmatic clinical trials reflect an attempt to address these limitations while preserving the crucial benefits of randomization.³

Pragmatic trials include at least four features that offer advantages.⁴ First, they specify a largely unrestricted population with few eligibility criteria, recruited from diverse settings. This leads to both increased feasibility and simplified screening procedures and broader generalizability to the target population. Second, pragmatic trials are designed to evaluate simple, clinically relevant interventions that require no specialized expertise or capabilities on the part of providers or the health care system. This means that any successful interventions can be widely scaled. Importantly, those interventions are intended to be as relevant at the end of the trial as they are at the start. Third, pragmatic trials rely on simple, easily measured outcomes that can be obtained from standard sources, such as electronic health records or insurance claims data. This means that supplemental data collection efforts are unnecessary, and trial participants need not have extra study-specific visits or interviews. Fourth, pragmatic trials often involve non-standard randomization in order to accommodate practical constraints; cluster randomization and/or uneven randomization ratios are common.

While pragmatic trials can apply to many areas of health, trials involving health behavior are especially well suited to these designs. For example, in a randomized trial comparing the effects of different drugs, our theory about the mechanism of effect might be largely based on our understanding of physiology. We might expect that physiology to be similar between those who enroll in the trial or not. In this case, results obtained even from a highly selected set of research participants might plausibly be seen as relevant to the much larger population of patients to whom the results might later be applied.

In contrast, in a randomized trial comparing the effects of different techniques to encourage medication adherence, diet, or fitness, our theory about the mechanism of effect might be largely based on our understanding of human behavior and motivation. Those underlying processes might differ considerably between those willing or not willing to enroll in a clinical trial. If only the most motivated people are willing to enter a trial that aims to increase medication adherence, then we ought to have less confidence that the differences we observe in the highly selected enrolled group will be relevant to a largely unselected group likely to have considerably less motivation. In this case, the selection process for trial enrollment establishes predictable external confounding that defeats or at least undermines the reduction of internal confounding provided by randomization. To the extent that pragmatic designs facilitate studies in less selected populations, they reduce this challenge to external validity.

Behavioral economics

These considerations are of growing importance because trials of behavioral interventions are increasingly generated from the promise of behavioral economics. Behavioral economics seeks to combine insights from both economics and psychology and recognize that as humans we often fail to do things that are in our best interests. Traditional economic theory views these departures as irrational. The primary insight of behavioral economics is not just that we tend to be irrational, but that we tend to be irrational in predictable ways. Indeed, it is the predictability of our common decision errors that can help us design tools to overcome them.⁵ For example, we eat the tempting dessert before us right now even though we know it competes with our long-term goal of losing weight; we fail to take our medications today even though we know they will improve our future health. These are examples of present bias, in which events occurring in the present or near future are weighted disproportionately to events in the more distant future. Other biases identified by behavioral economics include our difficulties understanding low probability events—leading people to equate risks of 1% with risks of 0.01%; our sensitivity to social norms and comparisons—leading people to do what they believe others do, regardless of whether that is the right decision; loss aversion—in which we are more motivated by the prospect of losing US\$100 than the prospect of gaining US\$100; and status quo bias—in which whatever choice is presented as the default is substantially favored, whether it was a good choice or not. None of these tendencies is rational in the sense that each of them is as likely to lead us away from our goals as toward them.^{6–8} Recent advances in behavioral economics have harnessed these common decision errors to redirect behaviors toward personal goals, like weight loss or fitness, rather than away from them. For those reasons, behavioral economics offers considerable innovation and promise for improving the outcomes of chronic conditions like heart disease and diabetes that often require considerable engagement of patients toward their management.

Chronic diseases are often managed through long-term use of medications, and yet consistent use of medications is extremely poor. A study of 4783 patients with hypertension prescribed a once-a-day antihypertensive (arguably the simplest regimen) showed that only about half were still taking the prescribed medication after a year.⁹ A similar pattern emerges for patients with high cholesterol who are prescribed statins.^{10–12} In patients with coronary artery disease, medications such as aspirin, beta blockers, statins, and antiplatelet agents significantly reduce the rate of repeat cardiovascular events and repeat treatment procedures in these patients. Yet, adherence to these medications is frustratingly low;^{12–19} for example, only about 39% of patients prescribed statins, aspirin, beta blockers, and platelet blockers after a heart attack are still taking those medications a year later.²⁰ Existing approaches to improving medication adherence, including patient education, encouragement of self-monitoring, and simple financial incentives such as copayment reductions, have had only modest effects.^{21,22}

Statistical issues in pragmatic trials

Two recent trials provide useful examples of the application of interventions drawn from behavioral economics aimed at improving medication adherence, conducted in a way that incorporates pragmatic trial features. The first was a four-arm cluster-randomized trial of

patient incentives, provider incentives, and shared incentives in 1503 patients with high cardiac risk and elevated high-density lipoprotein cholesterol (LDL-C) (the “shared incentives” trial).²³ The second was a two-arm trial of a comprehensive automated hovering intervention in 1509 patients discharged from the hospital after a myocardial infarction the HeartStrong Program²⁴. “Automated hovering”²⁵ is a concept developed in recognition of the fact that most patients, even those with multiple chronic diseases, spend only a fraction of their waking hours each year with health care providers; effective interventions must engage patients not only in the physician’s office, but in their daily lives, and in ways that are low-cost and scalable. The HeartStrong program used a suite of interventions including a daily lottery for medication adherence, automated feedback to patients and their designated medication partners, and access to program advisers and social workers. Each of these trials involved a number of pragmatic aspects raising important statistical issues and challenges. In the next sections, we describe these challenges, discuss two broad approaches to overcoming them, and discuss their implementation in the two trials described.

Two primary advantages of pragmatic trials are simplicity and generalizability. Simplicity, in both the design of the intervention and the conduct of the trial and data collection, makes trial participation possible for a wider range of participants (both individual participants and health care organizations that might eventually implement the intervention under study). This, in turn, leads to greater generalizability, with the ultimate goal of assessing effectiveness and not simply efficacy. These same features, however, can lead to an attenuation of the effect of the intervention because underlying motivation and activation likely have important moderating effects on response. As a result, pragmatic trials of behavioral interventions may require more participants in order to achieve adequate statistical power to detect effect sizes that are both realistic and clinically meaningful.

Two innovations

Evidence-based evolutionary testing

We have adapted the traditional RCT using the framework of evidence-based evolutionary testing (EBET)²⁶ to better incorporate ongoing improvement in care processes but retain the appealing RCT features that allow definitive program evaluation. Rather than deploy a single intervention during the proposed study period, we begin with our current state-of-the-art standard (call this Version 1). After conducting side experiments designed to optimize certain features of the intervention, as well as gaining operational experience from the period of initial deployment, we propose deployment of an enhanced intervention (called Version 2) about midway through the trial period. Randomization to active treatment at a 2:1 ratio relative to control provides equally sized groups receiving usual care, Version 1, and Version 2, thus optimizing statistical power across a range of hypotheses of interest. The primary hypothesis test involves comparing the combined intervention groups to the control group; the unbalanced randomization makes this comparison marginally less efficient than a standard balanced approach. It allows additional tests of interest, however, such as comparison of Version 1 and Version 2 to the control group. Note that a comparison of each version alone to its contemporaneous control group preserves direct randomizationbased comparability, but has lower efficiency because of the smaller sample size; comparison of a

single version to the combined control group results in the opposite effects. It may eventually be of interest to compare Version 1 and Version 2 directly to isolate the effect of improvements made during the course of the trial; because this comparison is not protected by randomization and is confounded by time, careful consideration must be given to the length of the study period and any time-cohort changes in population or related treatments that may have occurred. We recommend allowing the Type I error to be considerably higher than the usual level for this comparison, since it will only proceed if both Versions 1 and 2 are significantly better than control. In that setting the negative impact of a Type I error is markedly less important; that is, a false conclusion that Version 2 is superior to Version 1 has fewer adverse consequences when both versions are effective compared to control.

We implemented the EBET approach in the HeartStrong trial. The initial intervention consisted of provision of electronic pill bottles to patients, combined with daily feedback and a daily lottery for medication adherence; patients also named medication support partners and had contact as needed with study engagement advisers and social workers. Side experiments during the first year of the study focused on the effect of support partner nomination and optimizing the recruitment process. Using these results and incorporating our operational experience, we deployed an enhanced Version 2 about midway through the trial period. Since our leading indicator of success (medication adherence) was high, we did not make extensive changes to the protocol; however, we implemented improvements to the recruitment process and the online patient interface. The primary outcome was time to first fatal or nonfatal acute vascular event or revascularization, including acute myocardial infarction, unstable angina, stroke, acute coronary syndrome admission, or death. We anticipated the 1-year event rate in the usual care group to be approximately 23%²⁰ and designed the trial to provide 80% power to detect a hazard ratio of 0.7; this corresponds to a 6% decrease in the event rate in the intervention group. Thus, we planned to accrue 1500 patients over the 24-month period in which Versions 1 and 2 were implemented; the 2:1 randomization ratio yielded 500 control and 1000 intervention subjects, approximately 500 each receiving Versions 1 and 2. When comparing the individual versions to the contemporaneous usual care group, 500 subjects per arm provide 80% power to detect a hazard ratio of about 0.5, which corresponds to a reduction in the 1-year event rate of 11% points. In direct comparisons of Versions 1 and 2, we allowed a more generous Type I error of 0.25, since this comparison will proceed only if both versions are more effective than usual care, and in that scenario the implications of committing a Type I error are considerably reduced. With 500 patients receiving each version, we have 80% power to detect a hazard ratio of 0.62; assuming a baseline event rate of 17% in Version 1, this corresponds to a reduction of about 6% in Version 2.

Opt-out consent framework

A second innovation in pragmatic trial design incorporates the use of opt-out framing for the consent process. Traditional RCTs use opt-in consent, in which potential participants are identified and contacted. After hearing the study description, they decide whether or not to participate. The default is non-participation. In traditional RCTs, barriers to enrollment are often high: the study may require extra clinic visits, procedures, and/or interviews. As a result of this process, consent rates are often low and only a small portion of the eligible

population is ultimately enrolled. In the shared incentives trial, for example, approximately 6% of the targeted population eventually enrolled.

One goal of pragmatic trials is to increase the proportion of the target population that ultimately enrolls in the trial. As mentioned above, streamlining eligibility criteria and study requirements can partially accomplish this goal. Changing the framing of the consent process to an opt-out scenario, in which the default is participation rather than non-participation, is another powerful approach. In a strict opt-out approach, potential participants are automatically enrolled and randomized in advance; they are then informed and given the opportunity to decline participation by opting out. Failure to opt out is deemed tacit consent. A middle ground approach uses opt-out *framing* to encourage the default of participation, by contacting potential participants and describing their enrollment in a new program; following this contact, consent is still obtained from participants, and choice is not restricted. Use of the opt-out framing is generally expected to result in a more diverse sample of trial enrollees, including many who would not enroll under the traditional opt-in paradigm. This approach enhances generalizability. It also changes, likely reduces, the differential effectiveness between intervention and control arms. Opt-out framing is suitable for low-risk interventions, such as those in the two trials described here. The interventions do not involve prescribing new or untested medical therapies, but rather provide reinforcement of choices already made by clinicians to use a therapy with a very favorable benefit/risk ratio; an example is adherence to medications that lower cholesterol among patients at high risk of cardiovascular disease.

In both the shared incentives trial and the HeartStrong trial, we conducted side experiments evaluating the use of opt-out framing of the consent process. In the shared incentives trial, we randomized a separate population of patients with uncontrolled blood sugar to either opt-in or opt-out framing for enrollment.²⁷ Subjects randomized to opt-in were sent the standard letter and then contacted to invite their participation. Subjects randomized to opt-out were sent a similar letter explaining the program but were informed they had been enrolled and provided a number to call to opt out. In both arms, subjects were required to give informed consent and were considered formally enrolled when they appeared in person at a baseline study visit. All subjects were then offered home monitoring devices with a lottery for daily monitoring for 3 months. Preliminary analysis indicates an approximate tripling of the enrollment rate in the opt-out arm compared to the opt-in arm. Adherence to daily monitoring was higher in the opt-in arm than in the opt-out arm. These results are consistent with the view that opt-out enrollment provides more generalizable information. These results are also consistent with greater difficulty to demonstrate efficacy, which is likely to be appropriate given that what we really want to know is how effective the intervention will be in a less selected population and that almost always requires a higher evidentiary standard.

In the HeartStrong trial, we conducted an opt-out experiment in patients discharged from the University of Pennsylvania Health System (UPHS) with myocardial infarction. We sent 50 patient study packets that included the letter describing the study along with the electronic pill bottle kits used to measure medication adherence and determine lottery eligibility. We provided opt-out information and followed up with telephone calls to obtain consent. We compared the consent rate in these 50 patients to that of UPHS patients with similar

insurance whom we recruited for the HeartStrong trial. The consent rate among the optout patients was more than doubled compared to that among UPHS HeartStrong patients. Preliminary information on adherence measured with the electronic pill bottle indicates that it was comparable between the two groups.

Both of these studies provide compelling evidence that enrollment into pragmatic trials can be markedly increased by re-framing the consent process as opt-out rather than opt-in; we achieved nearly a threefold increase in enrollment with this approach. There is some evidence that this is accompanied by a dampening of the effectiveness of the intervention, at least on intermediate adherence outcomes. Even if that is true, however, it is likely to more accurately represent effectiveness, which is typically the main consideration.

A more pragmatic design eliminates consent entirely, even for randomized trials. Such trials would more accurately estimate expected effectiveness in unselected populations. Unconsented randomized trials must be carefully restricted to contexts where risks are very low and overseen by experienced institutional review boards. Many quality improvement projects conducted in naturalized clinical settings fit this pattern. For example, if either of two approaches to streamlining patient scheduling activity (such as a concierge service or an online portal) would be acceptable to roll out for all patients without consent as part of an ongoing quality initiative, randomization of participants (without consent) can determine which is better without exposing participants to risk.²⁸

Discussion

Pragmatic trials offer a number of improvements over traditional RCTs, including simplicity and broadened eligibility. These same features may conspire to reduce achievable effect sizes, however, highlighting the need for innovations in pragmatic trial design. We offer two such innovations.

The first, EBET, allows for rapid-cycle innovation while maintaining the comparability afforded only by randomization. This approach recognizes that typical clinical trials set an intervention in motion at the start, but rarely anticipate the expected advances in knowledge that occur in parallel, advances that could have allowed for ongoing adaptation and enhancement of interventions as might occur in the real world. Careful attention to hypothesis testing is needed within this framework, but an understanding of the implications of Type I and Type II errors for different kinds of hypotheses allows for efficient comparisons and ongoing optimization of interventions.

The second innovation, conversion to an opt-out framing of trial enrollment, derives from a recognition of the power of defaults. Choice architecture and judicious use of defaults have been shown in numerous settings to optimize desired outcomes without restricting freedom of choice. With continued interest in pragmatic trials and the call for comparative effectiveness trials to be situated in “real world” settings, using an opt-out default appears to be a promising way to increase the generalizability of findings. This approach may be particularly important for trials aiming to demonstrate and compare the effectiveness of interventions that require significant participant motivation and ongoing engagement.

Acknowledgments

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Piantadosi, S. Clinical trials: a methodologic perspective. 2nd. New York: John Wiley & Sons; 2005.
2. Friedman, LM., Furberg, CD., DeMets, DL. Fundamentals of clinical trial. 4th. New York: Springer; 2010.
3. Chalkidou K, Tunis S, Whicher D, et al. The role for pragmatic randomized controlled trials (pRCTs) in comparative effectiveness research. *Clin Trials*. 2012; 9:436–446. [PubMed: 22752634]
4. 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; 290:1624–1632. [PubMed: 14506122]
5. Loewenstein G, Asch DA, Volpp KG. Behavioral economics holds potential to deliver better results for patients, insurers, and employers. *Health Aff*. 2013; 32:1244–1250.
6. Johnson EJ, Goldstein D. Medicine: do defaults save lives? *Science*. 2003; 302:1338–1339. [PubMed: 14631022]
7. Madrian BC, Shea DF. The power of suggestion: inertia in 401(k) participation and savings behavior. *Q J Econ*. 2001; 116:1149–1187.
8. Cronqvist H, Thaler RH. Design choices in privatized social-security systems: learning from the Swedish experience. *Am Econ Rev*. 2004; 94:424–428.
9. Vrijens B, Vincze G, Kristanto P, et al. Adherence to prescribed antihypertensive drug treatments: longitudinal study of electronically compiled dosing histories. *BMJ*. 2008; 336:1114–1117. [PubMed: 18480115]
10. Jackevicius CA, Mamdani M, Tu JV. Adherence with statin therapy in elderly patients with and without acute coronary syndromes. *JAMA*. 2002; 288:462–467. [PubMed: 12132976]
11. Ho PM, Spertus JA, Masoudi FA, et al. Impact of medication therapy discontinuation on mortality after myocardial infarction. *Arch Intern Med*. 2006; 166:1842–1847. [PubMed: 17000940]
12. Hudson M, Richard H, Pilote L. Parabolas of medication use and discontinuation after myocardial infarction—are we closing the treatment gap? *Pharmacoepidemiol Drug Saf*. 2007; 16:773–785. [PubMed: 17486661]
13. Newby LK, LaPointe NM, Chen AY, et al. Long-term adherence to evidence-based secondary prevention therapies in coronary artery disease. *Circulation*. 2006; 113:203–212. [PubMed: 16401776]
14. Benner JS, Glynn RJ, Mogun H, et al. Long-term persistence in use of statin therapy in elderly patients. *JAMA*. 2002; 288:455–461. [PubMed: 12132975]
15. Gislason GH, Rasmussen JN, Abildstrøm SZ, et al. Long-term compliance with beta-blockers, angiotensin-converting enzyme inhibitors, and statins after acute myocardial infarction. *Eur Heart J*. 2006; 27:1153–1158. [PubMed: 16399775]
16. Kulkarni SP, Alexander KP, Lytle B, et al. Long-term adherence with cardiovascular drug regimens. *Am Heart J*. 2006; 151:185–191. [PubMed: 16368315]
17. Simpson E, Beck C, Richard H, et al. Drug prescriptions after acute myocardial infarction: dosage, compliance, and persistence. *Am Heart J*. 2003; 145:438–444. [PubMed: 12660666]
18. Eagle KA, Kline-Rogers E, Goodman SG, et al. Adherence to evidence-based therapies after discharge for acute coronary syndromes: an ongoing prospective, observational study. *Am J Med*. 2004; 117:73–81. [PubMed: 15234641]
19. Blackburn DF, Dobson RT, Blackburn JL, et al. Adherence to statins, beta-blockers and angiotensin-converting enzyme inhibitors following a first cardiovascular event: a retrospective cohort study. *Can J Cardiol*. 2005; 21:485–488. [PubMed: 15917876]
20. Choudhry NK, Avorn J, Glynn RJ, et al. Full coverage for preventive medications after myocardial infarction. *N Engl J Med*. 2011; 365:2088–2097. [PubMed: 22080794]

21. Bosworth HB, Powers BJ, Olsen MK, et al. Home blood pressure management and improved blood pressure control: results from a randomized controlled trial. *Arch Intern Med.* 2011; 171:1173–1180. [PubMed: 21747013]
22. Omhoni S, Guarda A. Impact of home blood pressure telemonitoring and blood pressure control: a meta-analysis of randomized controlled studies. *Am J Hypertens.* 2011; 24:989–998. [PubMed: 21654858]
23. Asch DA, Troxel AB, Stewart WF, et al. Effect of financial incentives to physicians, patients, or both on lipid levels: a randomized clinical trial. *JAMA.* 2015; 314:1926–1935. [PubMed: 26547464]
24. Troxel AB, Asch DA, Mehta SJ, et al. Rationale and design of a randomized trial of automated hovering for post-myocardial infarction patients: the HeartStrong program. *Am Heart J.* Accepted.
25. Asch DA, Muller RW, Volpp KG. Automated hovering in health care—watching over the other 5000 hours. *N Engl J Med.* 2012; 367:1–3. [PubMed: 22716935]
26. Volpp KG, Terwiesch C, Troxel AB, et al. Making the RCT more useful for innovation with evidence-based evolutionary testing. *Healthc.* 2013; 1:4–7.
27. Aysola, J., Troxel, AB., Asch, DA., et al. A randomized controlled trial of opt-in versus opt-out enrollment into a diabetes management intervention. Perelman School of Medicine, University of Pennsylvania; Philadelphia, PA: Jun 15. 2015 Technical Report
28. Selker, H., Grossmann, C., Adams, A., et al. The common rule and continuous improvement in health care: a learning health system perspective. Washington, DC: Institute of Medicine; 2011.