

READERS' TOOLBOX

Understanding Research Methods

Novel Clinical Trial Designs to Improve the Efficiency of Research

Daniel I. Sessler, M.D.,
Paul S. Myles, M.B., B.S., M.P.H., D.Sc., F.A.N.Z.C.A.

Anesthesiology 2020; 132:69–81

Novel Clinical Trial Designs to Improve the Efficiency of Research

Novel trials are efficient alternatives to conventional trials that still address bias and confounding

Novel trial designs can address heterogeneous populations, cost effectiveness, bundled interventions, and real-world scenarios.
Examples: trauma, cardiac arrest, surgical site infections

Cons of novel trials include

- Often require complex statistical methods
- Less control of delivery of the intervention

- Finite resources limit how many prospective randomized trials can be conducted.
- Novel clinical trials can facilitate enrollment of many patients at modest cost, but raise challenges for planning, conduct, ethical oversight, and statistical analysis.

Sessler and Myles. *ANESTHESIOLOGY*. November 2020.

ANESTHESIOLOGY
Trusted Evidence. Discovery to Practice.
Copyright © 2019 Wolters Kluwer Health. All rights reserved. Published by Lippincott, Williams & Wilkins.

SUMMARY

Large randomized trials provide the highest level of clinical evidence. However, enrolling large numbers of randomized patients across numerous study sites is expensive and often takes years. There will never be enough conventional clinical trials to address the important questions in medicine. Efficient alternatives to conventional randomized trials that preserve protections against bias and confounding are thus of considerable interest. A common feature of novel trial designs is that they are pragmatic and facilitate enrollment of large numbers of patients at modest cost. This article presents trial designs including cluster designs, real-time automated enrollment, and practitioner-preference approaches. Then various adaptive designs that improve trial efficiency are presented. And finally, the article discusses the advantages of embedding randomized trials within registries.

(*ANESTHESIOLOGY* 2020; 132:69–81)

Maintaining the health and well-being of populations is a universal priority. Healthcare is among the largest employers and consumers of resources in our global economy. However, it is also clear that the current growth in healthcare expenditures is unsustainable, already being approximately 10% of the economy in the United Kingdom, Australia, and Canada—and 17.8% in the United States.¹ Governments around the globe are therefore seeking greater efficiency and better outcomes.

The highest level of clinical evidence is generally thought to be large randomized trials or a systematic review of several large trials.² Large trials generally provide reliable results because they limit confounding and, when blinded, also limit measurement bias. Large multicenter trials typically enroll patients in diverse study settings and are therefore also generalizable (fig. 1).^{3–6} Trial results are more likely to be rapidly implemented when generalizability is high and important patient-centered outcomes were included.^{6,7}

The difficulty is that enrolling large numbers of randomized patients across numerous study sites is expensive and often takes years. There will never be sufficient resources dedicated to conventional clinical trials to address the important questions in medicine. The cost, time, and difficulties of conventional trials have led to questions about their viability and a search for alternative approaches.^{8,9} Consequently, there are calls for improved efficiencies in

Image: D. I. Sessler and P. S. Myles.

Submitted for publication January 10, 2019. Accepted for publication August 9, 2019. From the Department of Outcomes Research, Anesthesiology Institute, Cleveland Clinic (D.I.S.); and the Department of Anaesthesiology and Perioperative Medicine, Alfred Hospital and Monash University, Melbourne, Australia (P.S.M.).

Copyright © 2019, the American Society of Anesthesiologists, Inc. All Rights Reserved. *Anesthesiology* 2020; 132:69–81. DOI: 10.1097/ALN.0000000000002989

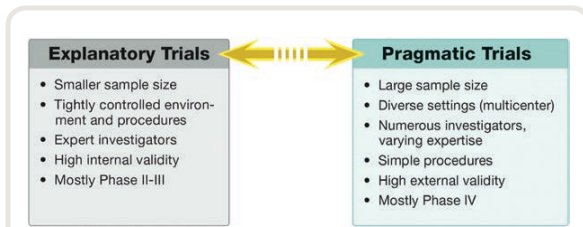


Fig. 1. Trade-offs between explanatory trials (can it work in ideal settings?) and pragmatic trials (does it work in the real world?).

Box 1. What to Look for in Research Using This Method

- Large number of participants, few exclusions
- Real-world setting(s)
- Patient-centered outcome(s)
- Practical results, applicable to your own practice

medical research¹⁰ that have led to growing interest in novel trial designs^{8,11,12} such as low-cost, large, pragmatic randomized trials.^{13,14}

Perhaps the most obvious alternatives to randomized trials are case-control or cohort studies based on “big data” registries with thousands to tens of millions of patient records. Large registries, combined with sophisticated analysis strategies such as propensity scoring, now provide valuable information, and for rare conditions or outcomes, they may be the only workable approach.^{15,16} However, even with the best analyses, observational data will always be subject to unknown degrees of bias and confounding that diminishes the reliability of the analyses and confidence in the conclusions.

Efficient alternatives to conventional randomized trials that preserve protections against bias and confounding are thus of considerable interest. Among the most useful are pragmatic trials conducted in “real-world” settings that

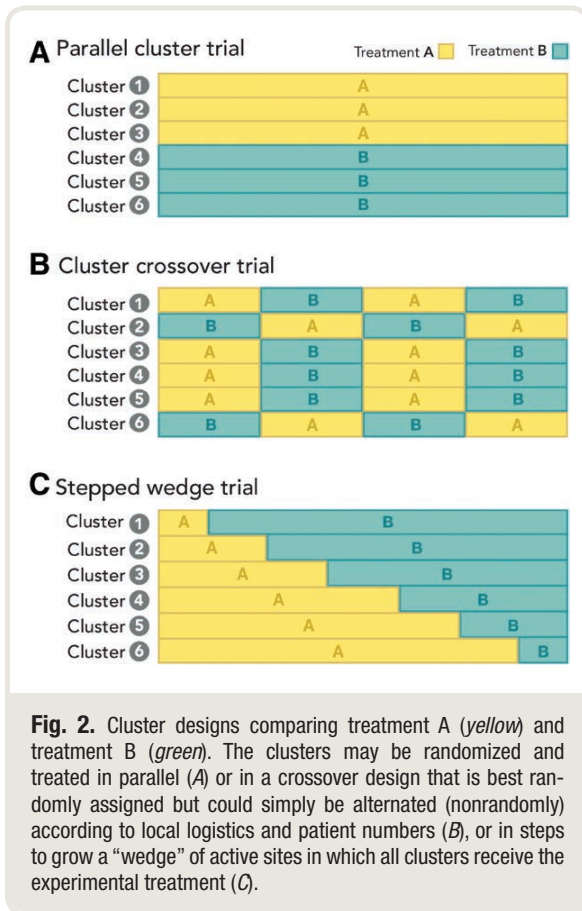
enhance their external validity and clinical value, sometimes referred to as effectiveness trials. A common feature of novel trial designs is that they are pragmatic and facilitate enrollment of large numbers of patients at modest cost. However, as with all aspects of trial design, there are important trade-offs to alternative designs that may or may not be worthwhile in various circumstances (table 1). Generally, parallel-group randomized and blinded trials should be considered the default, with alternative designs being adopted only when there are compelling reasons.

Cluster Trials

Systematic healthcare-related interventions such as implementation of clinical pathways or electronic records cannot be turned on or off for individual patients, which precludes conventional randomized trial designs. Such interventions are thus often evaluated using before-and-after designs. However, that approach is weak because it suffers from three major sources of unquantifiable error, each of which makes interventions appear beneficial even if they are not.

Table 1. Strengths and Weaknesses of Novel Trial Designs

Design	Strengths	Weaknesses
Pragmatic trials	Generalizability: broad range of patients and practice settings Focus on outcomes relevant to patients, clinicians and policy makers Have greater statistical power to investigate subgroup effects Best to inform cost-effectiveness	Partially uncontrolled heterogeneity: lack of uniformity of treatment effect Varied expertise delivering the intervention Varied expertise in research procedures Limited information about treatment effect under optimal conditions
Cluster crossover trials	Test real-world effectiveness Ideal for complex or bundled interventions No need for individual consent and randomization Limits contamination bias Accounts for a learning effect Equipoise is less compromised	Need a larger sample size (more patients) because of intraclass correlation (can overcome with crossover, but this demands a longer trial duration) Reduced patient autonomy Blinding can be difficult Very specific reporting requirements
Stepped wedge trials	Have the strengths of cluster crossover trials All trial units eventually get the “new” treatment Deals with carryover effects	Ideally, should have rapid enrollment to control the duration of each period Possible temporal confounding More dropouts Difficult to plan and undertake interim analyses Greater complexity in the analysis of results
Adaptive, including platform, basket, and umbrella trials	Acknowledge the reality of new knowledge acquired during the conduct of the trial: reduced exposure to potential harm Inherent flexibility of varying interventions Flexible objectives Reduced trial duration and costs when testing numerous interventions or disease states Ideal for rare diseases (highly efficient)	Potential difficulties to identify all sources of bias Need to stringently control type I error Complex logistics Unexpected heterogeneity, effect size(s)



The first error is time-dependent confounding, which results because healthcare outcomes generally improve over time consequent to multiple small changes, many of which are unrecognized or poorly quantified. For example, concern about surgical site infections may prompt more frequent handwashing, better skin preparation, glove changes at various stages of surgery, restrictions on operating room traffic, air conditioning filter changes, etc. Because many factors change over time, there is no basis for attributing all improvement to a specific intervention. The second error is the Hawthorne effect, which refers to improvements that result from focusing attention on a particular outcome. The third bias is regression to the mean, which occurs when an intervention is prompted by high incidence of the outcome. However, the observed high incidence, say of infection, may simply be a random variation that will return to the mean level with or without intervention. It is consequently unreasonable to attribute observed benefit from one intervention, for example changing the warming system, because it may well have occurred without any intervention whatsoever. For additional details about these sources of error, see recent reviews.^{17–20}

Cluster designs eliminate many of the problems inherent in before-and-after studies. They are defined by exposure being allocated to *groups* of subjects. Typically, groups are defined by hospital or unit (fig. 2). Cluster approaches are especially useful for systematic changes such as implementation of rapid response teams that cannot be allocated individually or easily reversed. Cluster trials can be randomized or controlled without randomization and some involve crossover periods or sequential exposure. In this section, we will present just a few of many approaches to cluster trials.

Parallel Group Randomized Cluster Trials

Most conventional cluster trials simultaneously randomly assign groups of patients to one of two or more exposures. They thus resemble conventional parallel-group individually randomized trials, except that the randomization is on the basis of units rather than individuals. As with conventional trials, crossover designs can be incorporated into cluster designs. Clusters can be entire hospitals,²¹ units within a hospital,²² or even patients under the care of single physicians.²³ In each case, however, all patients in a particular group are given the same treatment (fig. 2A) and are thus normally conducted with waived consent.

An advantage of cluster randomized trials without crossover is that they reduce the Hawthorne effect and largely eliminate learning that is a common feature of other designs including conventional parallel-group randomized trials. For example, consider a conventional randomized trial of guided fluid management for prevention of intraoperative hypotension. Individual clinicians may quickly learn from cases in which fluid administration is guided that they normally give too much or too little fluid. To the extent that they adjust their practice in patients *not* randomized to guided fluid management, the difference between the treatment groups will diminish—perhaps to the point where it is no longer possible to demonstrate a real benefit of guided management. In contrast, there would be little opportunity for clinicians to learn from one treatment and apply the information to the other condition in a cluster trial because each site/clinician would be exposed only to one treatment.

Sample-size estimation for all types of cluster trials is complicated but depends importantly on the number of clusters. Enrolling many patients within a cluster, which is usually relatively easy, does not substitute for having many clusters.²⁴ The major challenge is thus usually establishing a sufficient number of clusters.

Randomized Stepped Wedge Trials

The term “stepped wedge” refers to the fact that at the beginning of the trial, the experimental treatment is used at none of the study sites. One randomly selected site then implements the experimental treatment. After a suitable interval, perhaps several months, another randomly selected site converts to the experimental treatment (fig. 2C). The

process continues until all study sites are active, thus producing a growing “wedge” of active sites.^{25–27} An advantage of the approach is that all trial units eventually convert to the experimental treatment (such as a specialized stroke service) that might even be nationally mandated. In contrast, the experimental treatment is only available to half the units during the trial period with a parallel-group cluster approach.

For example, the TRACE (routine postoperative Anesthesia visit to improve patient outcome) trial is a prospective, multicenter, stepped-wedge, cluster-randomized trial being conducted in The Netherlands to evaluate routine postoperative visits by an anesthesiologist to reduce the risk of postoperative complications.²⁵ All hospitals start simultaneously with a control phase in which standard care is provided. Sequentially, in a randomized order, hospitals cross over to the intervention phase (routine visits). The trial is currently underway and is recruiting 5,600 adult patients at high risk of postoperative complications (trial No. NTR5506).

With some loss of rigor, nonrandomized stepped wedge designs can also provide value. These might be used when investigators cannot reliably control when interventions are initiated at various sites. For example, switching to electronic records is a complex process that most hospitals will not initiate at a particular time to please an investigator. However, to the extent that electronic records are implemented on a pseudorandom basis (say based on funding availability), investigators could reasonably analyze their effects in a group of hospitals as they switch over time. Stepped wedge studies require complex statistical methodology, especially with nonrandomized designs that need to be adjusted for potential confounding factors.²⁸

Alternating or Randomized Multiple Crossover Cluster Controlled Trials

A recently developed pragmatic approach is an alternating cluster controlled trial. Effectively, these are cluster trials with interventions distributed in time rather than in space. In these trials, an intervention is implemented for a limited period, say 2 weeks, and then removed for a comparable amount of time. That 1-month block might be considered a before-and-after study with all the serious limitations inherent to that approach. However, the key element to alternating cluster trials is to repeat the study unit enough times to wash out time-dependent confounding from background improvements in healthcare and regression to the mean (fig. 2B). The Hawthorne effect is also limited because the observation period includes times with and without intervention, rather than just after an intervention. Furthermore, over a typical study period of a year or more, the trial becomes part of the clinical routine rather than a one-off focus on particular exposures and complications.

A defining aspect of alternating intervention cluster trials is that entire hospital units are studied and that the intervention of interest is applied to clusters of all or no

patients within a given unit during the designated periods. For example, an alternating cluster trial might include all patients in a set of operating rooms or a particular surgical ward. A corollary is that consent is usually waived because exposure within a given unit is based only on time period rather than patient characteristics or consent.

Alternating cluster trials are controlled because exposure allocation is not determined by patient or physician preference. However, they are not randomized. (The order of the exposure periods could be randomized, although doing so adds little value to the trial.) Unless surgeons specifically schedule patients to particular time periods, exposure will effectively be random. In practice, this approach provides most benefits of randomization but without the cost and difficulty.²⁹

Limitations of the approach include a lack of concealed allocation and lack of blinding. Additional limitations include clinician learning during intervention periods that improves care during nonintervention (or alternate intervention) periods. How likely such “contamination” is depends on the intervention(s) and the extent to which they can be controlled by investigators. Alternating cluster trials are thus best for interventions that truly are considered comparable by clinical teams and where the outcomes are objective and not exposed to detection bias—preferably recorded electronically.

Alternating cluster trials work best when the exposure can be tolerated by nearly all patients because it is impractical to enforce inclusion and exclusion criteria. (Enrolling nearly everyone does not preclude restricting analysis to an *a priori* defined population, nor does it prevent clinicians from using alternative approaches when an exposure may be suboptimal for a particular patient.) A consequence of broad enrollment is excellent generalizability. It is also possible to perform two or more simultaneous alternating cohort studies in the same unit(s) if the interventions and outcomes do not conflict. Similarly, the approach works perfectly well for factorial trials; for example, see trial No. NCT03657368, which uses an alternating cluster approach to factorially evaluate two different tidal volumes and two different positive end-expiratory pressure settings.

Because alternating cluster trials enroll large numbers of patients, baseline factors are inevitably well balanced. Statistical analysis can thus be as simple as chi-square or *t* tests. However, because the trials occur over a fairly long period, it is generally prudent to include time in a multivariable model and then estimate the independent effect of the intervention.

In recent years, alternating cluster designs have been used to compare isoflurane *versus* sevoflurane on the duration of hospitalization ($n = 1,584$),³⁰ 30% *versus* 80% intraoperative oxygen on infection and healing-related complications ($n = 5,749$),³¹ and normal saline *versus* balanced salt solutions in emergency departments ($n = 13,347$),³² operating rooms (NCT02565420, $n = 8,600$), and critical care units ($n = 15,802$).³³ Each was characterized by rapid enrollment,

and the trials were far less expensive than a conventional approach with individual patient randomization.

Multiple crossover cluster trials can also be randomized. That is, instead of simply alternating treatment allocations, to assign them randomly. Because it is usually unlikely that patients get into clusters because of trial allocation, random assignment is probably less important than for conventional trials. However, when practical, it is the preferred approach. Perhaps more importantly, allocation should be blinded when possible to avoid selection and measurement biases. Blinding can only be maintained with random allocation and is a compelling reason to randomize rather than alternate treatment interventions.

For example, consider a cluster trial of supplemental operating room air filtration and sterilization on surgical site infections. The treatment can be blinded by internally deactivating the supplemental filters. Combined with randomized allocation periods in each operating room (the clusters), blinding will prevent measurement bias.

Real-time Automated Enrollment and Randomization

There are conditions in which immediate treatment is necessary, and it is not possible or practical to obtain consent. The most obvious examples are sudden and unpredictable emergencies such as cardiac arrest or major trauma. For a time, concerns about subject autonomy prevented most such research. However, it was obvious that avoiding research in highly lethal emergency situations was in no one's best interest. Nearly all regulatory systems now therefore permit research in such conditions with more or less scrutiny and oversight. For example, many critical care trials have relied upon a deferred consent model, whereby next-of-kin may be informed of the research (but usually cannot legally consent) and given an opportunity to opt-out, and included surviving patients are later asked to provide consent for their data to be included in the trial.³⁴⁻³⁶

There are other, perhaps less obvious, situations that preclude obtaining prior consent. Consider, for example, most any unexpected intraoperative event such as anaphylaxis (1 event per 677 cases³⁷), severe airway problems, or serious hypotension. Because it would be impractical or impossible to obtain consent from all surgical patients in expectation of randomizing rare qualifying patients, waived consent or modified consent is necessary to study such conditions. Outright waiver is often appropriate when the test interventions are low risk or perhaps likely to be helpful compared with routine care. Alternatively, an institutional review board might request a modified consent, such as providing information in advance and the ability to opt out or requesting *a posteriori* consent from qualifying subjects and including their data in analysis only with approval.

A recent example of real-time automated enrollment and randomization was a trial that evaluated an alert for clinicians about intraoperative hypotension.³⁸ More than

14,500 operations were screened in real time to identify 3,955 surgeries during which systolic blood pressure was less than 80 mmHg for 3 consecutive min. Thus, nearly four patients would have been asked for consent with a conventional approach to identify each who had a qualifying hypotensive episode. When episodes were identified, patients were randomized by a decision-support computer to either no alert or to alerts warning of hypotension that were posted to the electronic record screen and to the pagers of the in-room clinician and attending anesthesiologist. There was no required response to the alerts. The primary outcome was the duration of hypotension.

In another example, 36,670 patients were electronically screened in real time to identify 7,569 who experienced triple low events (a combination of mean arterial pressure of 75 mmHg or less, minimum alveolar fraction of less than 0.8, and Bispectral Index of less than 45).³⁹ In this case, nearly five patients would have been asked to consent to identify each who had a qualifying episode, which would have made a conventional randomized trial impossible. Patients who experienced qualifying episodes were enrolled and randomized in real time by a decision-support computer to no alert or to alerts warning: "A triple low (MAP [mean arterial pressure], MAC [minimum alveolar concentration], and BIS [Bispectral Index]) condition has been detected. Consider hemodynamic support." Again, there was no required response to the alerts. The primary outcome was 90-day mortality.⁴⁰

Both trials were considered ethical, and consent was waived by the Cleveland Clinic (Cleveland, Ohio) institutional review board because alerts were unlikely to have been harmful, were possibly helpful, and because the trial would be impractical without waived consent. In both cases, the alerts proved to be unhelpful but also harmless. Real-time automated enrollment and randomization is a useful trial design for evaluating responses to relatively uncommon intraoperative events where it would be impractical to obtain consent from enough patients in advance.

Practice Preference Randomization

Variations in "standard" clinical care are widespread. For example, some sites routinely use midazolam to prevent intraoperative awareness during cardiac surgery, whereas others rarely do. Variations are most apparent across countries but can also be found within a single country or city. Natural clinical variation can be harnessed in a novel trial design termed practice-preference randomization.⁴¹ Practice-preference trials are a variation of "play-the-winner" rules used in some adaptive designs^{42,43} but without a dynamic component.^{44,45} In this case, the "winner" is an unequal allocation favoring existing practice (fig. 3). The approach is being used in a trial evaluating high-dose dexamethasone in cardiac surgery, which includes sites in Australia (that rarely use high-dose dexamethasone) and

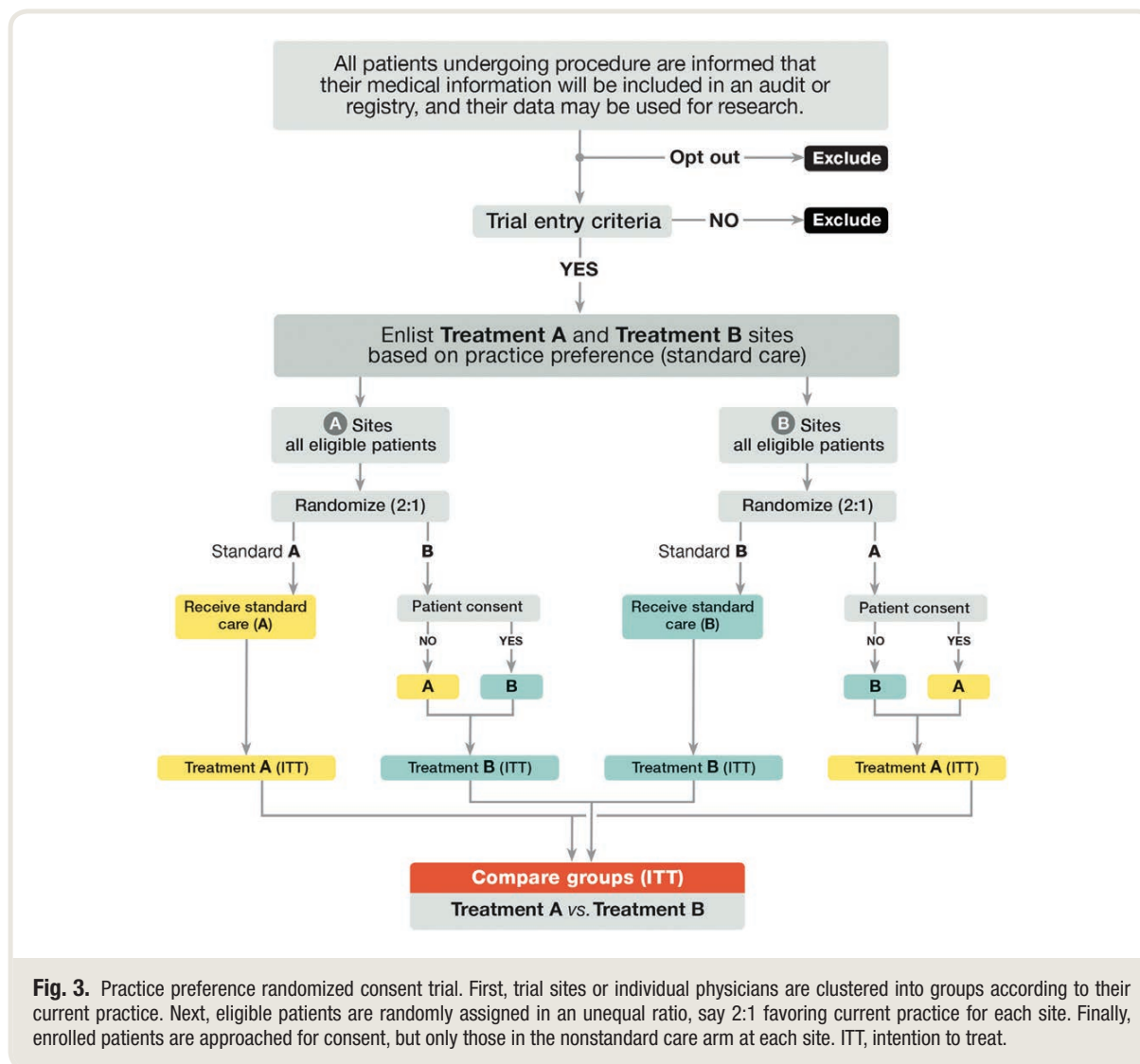


Fig. 3. Practice preference randomized consent trial. First, trial sites or individual physicians are clustered into groups according to their current practice. Next, eligible patients are randomly assigned in an unequal ratio, say 2:1 favoring current practice for each site. Finally, enrolled patients are approached for consent, but only those in the nonstandard care arm at each site. ITT, intention to treat.

The Netherlands (where high-dose dexamethasone is used routinely).⁴¹

There are three steps in practice-preference trials. Initially, trial sites or individual physicians are clustered into groups according to their current practice routine. The next step is to randomly assign eligible patients in an unequal ratio, say 2:1 favoring current practice for each site. Therefore most patients at a given site receive standard treatment for that site. The third step is to enroll patients, obtaining consent only for patients in the nonstandard arm at each site.

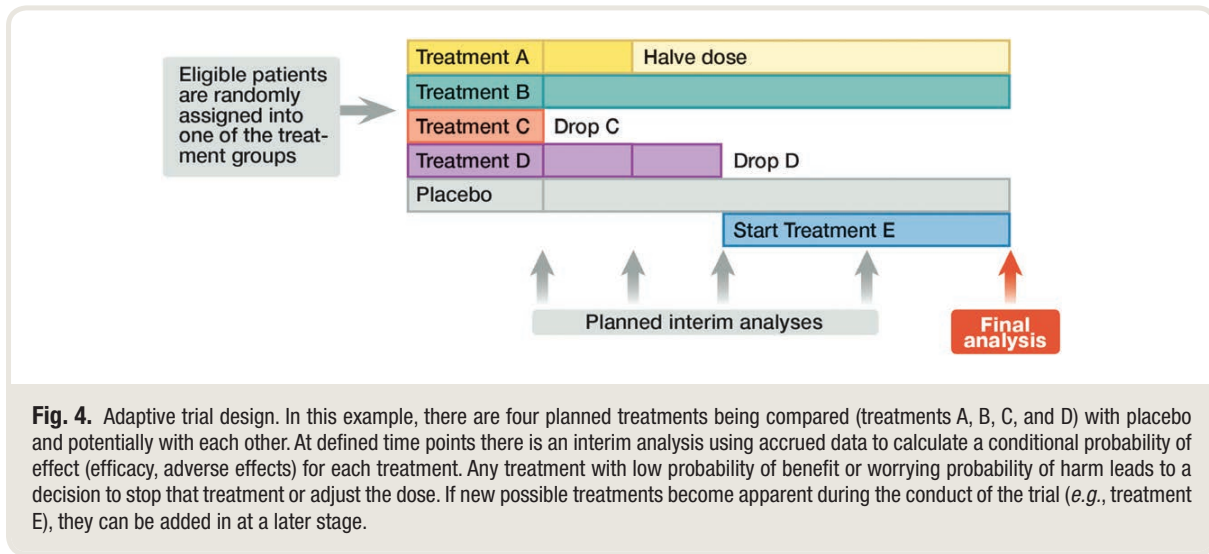
These trials are efficient because only a fraction, say a third, of the patients are approached for consent at each site. However, because at least several sites with differing definitions of routine care participate, overall enrollment is roughly balanced across the entire trial. Because clinician preference is respected, trial engagement is likely to be enhanced. A requirement for waived consent in the

standard-care group is trial-related measurements that are routine or at least minimally disturb participants.

Baseline imbalance and confounding are more likely with practice-preference than conventional randomized designs.⁴⁶ Selection bias is also possible because patients are approached for consent after random assignment to nonroutine care, whereas in a conventional trial consent is obtained from all patients before randomization. Statistical methods should be used to evaluate and, as necessary, adjust for both types of error.

Adaptive Designs

Because large clinical trials usually take years to complete, it is common for new and relevant information to be published during the conduct of a trial. Similarly, important information may accrue during a trial and be detected at



interim analyses. New knowledge sometimes suggests that aspects of an ongoing trial should be modified. For example, it may be necessary or advisable to alter enrollment criteria, drug dose, and which data are collected. It may also be necessary to recalculate the required sample size if the original assumptions prove incorrect.

Conventional clinical trials often include prespecified thresholds that determine whether a trial should be stopped or continued, which is perhaps the simplest type of adaptive design. Better, though, is the ability to systematically review accruing data and alter the protocol as necessary to reduce participant risk and enhance the clinical value of the results. Adaptive trial designs may include preplanned decision rules that permit changes to study population, assignment ratio, sample size, or study drug administration or dose (fig. 4).¹² Adaptive designs are likely safer for participants because patients at special risk can be excluded based on new information, and ineffective or excessive doses can be corrected. Furthermore, funding agencies and investigators are more likely to get return on their investments because adaptive trials are likely to achieve meaningful results with fewer patients in a shorter timeframe than conventional approaches.

Although generally efficient, adaptive designs introduce considerable complexities. For example, statistical analyses of adaptive trials need to account for multiple testing because of the frequent interim analyses and confounding caused by the baseline imbalance consequent to small numbers of participants and temporal trends.⁴⁷ Statistical simulations are often necessary, and Bayesian approaches (combining prior probability with observed results to estimate effect size) are sometimes used because prior information informs interpretation of sparse accruing data.⁴⁸ Another important factor is that adaptive options should be preplanned and incorporated into an *a priori* protocol with

appropriate statistical considerations. Unplanned protocol or analysis changes are often necessary, but they should be accurately presented as *post hoc* decisions rather than masked as “adaptive designs.”

Altering the Study Population

New information from other similar trials or the results of interim analyses often suggest that enrollment criteria should be modified. (A common enrollment change is to broaden criteria when trial enrollment is slower than expected, but this is a trivial case of adaptive design.) For example, new information might identify a subpopulation at especially high risk of toxicity. Simply prudence would suggest that patients at special risk be excluded—or at least be monitored especially carefully if the treatment is likely to overall beneficial.

Similarly, subpopulations might be identified that apparently receive little or no benefit from treatment or require a higher-than-typical drug dose to achieve comparable efficacy. Optimally designed protocols allow modifications that incorporate such obviously relevant information. It might also be necessary to increase trial size to provide sufficient power within clinically important subgroups. Adaptive designs increase the likelihood of the treatment being shown to be effective overall and of identifying subpopulations with substantively different responses.

Changing Treatment Group Assignment Ratios

Response-adaptive randomization refers to modifying random assignment ratios of multigroup trials when interim analyses suggest differential benefit. Consider, for example, a three-arm trial has an initial assignment ratio of 1:1:1 for treatments X, Y, and Z. At the first interim analysis, X and Y appear to provide more benefit than Z, even if the

assessment is hardly definitive given that enrollment is so far modest.

Nonetheless, a reasonable approach, when permitted by the original protocol, is to change the assignment ratio to “play the winner.” In this case a higher proportion of patients are assigned to an apparently more successful treatment. For example, the randomization might be altered to assign X, Y, and Z in a 2:2:1 ratio. Subsequent enrollment will therefore favor X and Y, thus providing extra precision about those treatments. Treatment Z may turn out to be comparable, inferior, or superior, but there is logic in directing resources to the treatments that look best. A similar approach can be used for trials that test multiple doses of a single drug. Random assignment favoring apparently more effective treatments is ethically appealing because it helps maintain equipoise throughout the trial as information accumulates.

Changing Sample Size

Sample size for formal trials is based on estimates of population variance (for continuous outcomes) or baseline incidence (for dichotomous outcomes), along with the treatment effect. Treatment effect, that is the true difference between the treatment groups, is the most important determinant of sample size, with the number of participants increasing exponentially as treatment effect diminishes. The difficulty, of course, is that the purpose of proposed trials is to determine the treatment effect, so investigators perforce do not know what it will be during the design phase.

Investigators should also consider what treatment effect is likely to be clinically meaningful, and that threshold should be specified in advance. In some cases, such as mortality, most any improvement would be considered meaningful. However, for trials evaluating process measures and mediators, statistically significant differences may not be clinically important. In such cases, statistically significant differences should be considered negative results when their magnitude does not reach a previously specified threshold.

Skilled investigators consider various strategies for increasing treatment effect such as enriching the population, using large exposures (e.g., high drug doses), and selecting continuous or composite outcomes. However, within a given trial design, an almost overwhelming temptation is to anticipate a treatment effect large enough to make the trial practical with respect to available subjects, time constraints, and especially available funding. The difficulty, of course, is that the true treatment effect is determined by biology, not the investigators' various constraints. It is thus common to complete planned enrollment and find a point estimate for treatment effect that is physiologically important but not quite statistically significant. Arguably, this is one of the worst outcomes for a trial because the investigators cannot claim a statistically significant benefit, but nor can they claim that there is no clinically important effect.

One cause for insufficient power is baseline variance or incidence being smaller than anticipated. That is relatively

easy to deal with because sample size can be reestimated without statistical penalty on the basis of variance or incidence across the entire study population.^{49,50} Prudent investigators thus monitor outcome incidence and variance across the study population (without distinguishing among groups) and reestimate the sample size if necessary.^{51,52} A more complicated situation arises when trial results are uninterpretable because treatment effect is smaller than anticipated. There are at least two ways to reduce that risk, both being variations of group-sequential designs.

One way to reduce the risk of insufficient power consequent to a smaller-than-anticipated treatment effect is to overpower the trial while including many interim analyses with stopping rules that allow enrollment to stop early if benefit or futility are clearly demonstrated. Typically, results of interim analyses are evaluated by an independent executive committee or data and safety monitoring board, and even then usually blinded to allocation (that is, on a “group A” vs. “group B” basis). If benefit is as anticipated or stronger, the trial will stop early at lower or comparable cost compared with a trial without interim analyses. However, if the treatment effect proves to be smaller than anticipated, the original design allows enrollment to continue—possibly to the point of clearly identifying benefit or futility. A second way to reduce risk of underpowering trials is to specify in the original protocol that sample size will be recalculated at given point or specified intervals based on (usually blinded) predefined outcomes of interest. Reestimation of the maximum sample based on observed treatment effects is complex, and it can be challenging to avoid both type 1 and type 2 statistical errors.⁵³

Changing Study Interventions: Platform Trials

Evaluation of new drugs is a costly process. Phase II drug development studies, for example, often evaluate several doses in an effort to determine an optimal one for testing in a phase III trial. It is hardly unusual for interim analyses to identify unexpected intolerance or toxicity or differential efficacy in subgroups. Many promising drugs fail early-phase testing because initial studies were unreliable; other drugs are found to be ineffective or harmful only because the wrong study population or dose was used or the sample size provided inadequate statistical power. Adaptive designs reduce risk by introducing flexibility into the evaluation process while maintaining statistical rigor.⁵⁴

An extension of the adaptive design concept is to establish an overarching platform protocol that permits introduction of additional treatments, for some treatments to be dropped, and for additional therapeutic questions to be tested without a set end date.^{55,56} The master protocol defines one or more target populations with a common system for patient selection, study procedures, process and outcome templates, and data management.

The approach is efficient because platform trials allow for concurrent or sequential evaluations of multiple treatments

within a single drug, condition, or surgical population.^{55,56} These have a common control arm, and many different treatment arms that are included or removed from the trial as futility or efficacy are demonstrated, often according to Bayesian decision rules.¹ Platform trials therefore avoid the need to design separate trials for each drug/treatment evaluation and to separately obtain ethical and regulatory approvals. An additional advantage is that accumulated infrastructure, procedures, and experience and knowledge is valuable and can roll over from one treatment to another within an overarching platform trial. The same general approach can be extended to cohort designs.⁵⁷

Subtypes of the platform design, also using master protocols and currently most often used in oncology, include basket and umbrella trials.^{58,59} A basket trial evaluates a specific treatment in a variety of diseases (*e.g.*, different types of cancer) defined by a particular molecular marker. An umbrella trial evaluates multiple treatments in one or more diseases, defined by an expectation that each included disease state could have a beneficial and similar response to any particular treatment. Here the set of diseases (*e.g.*, tumor types) is the “umbrella.”

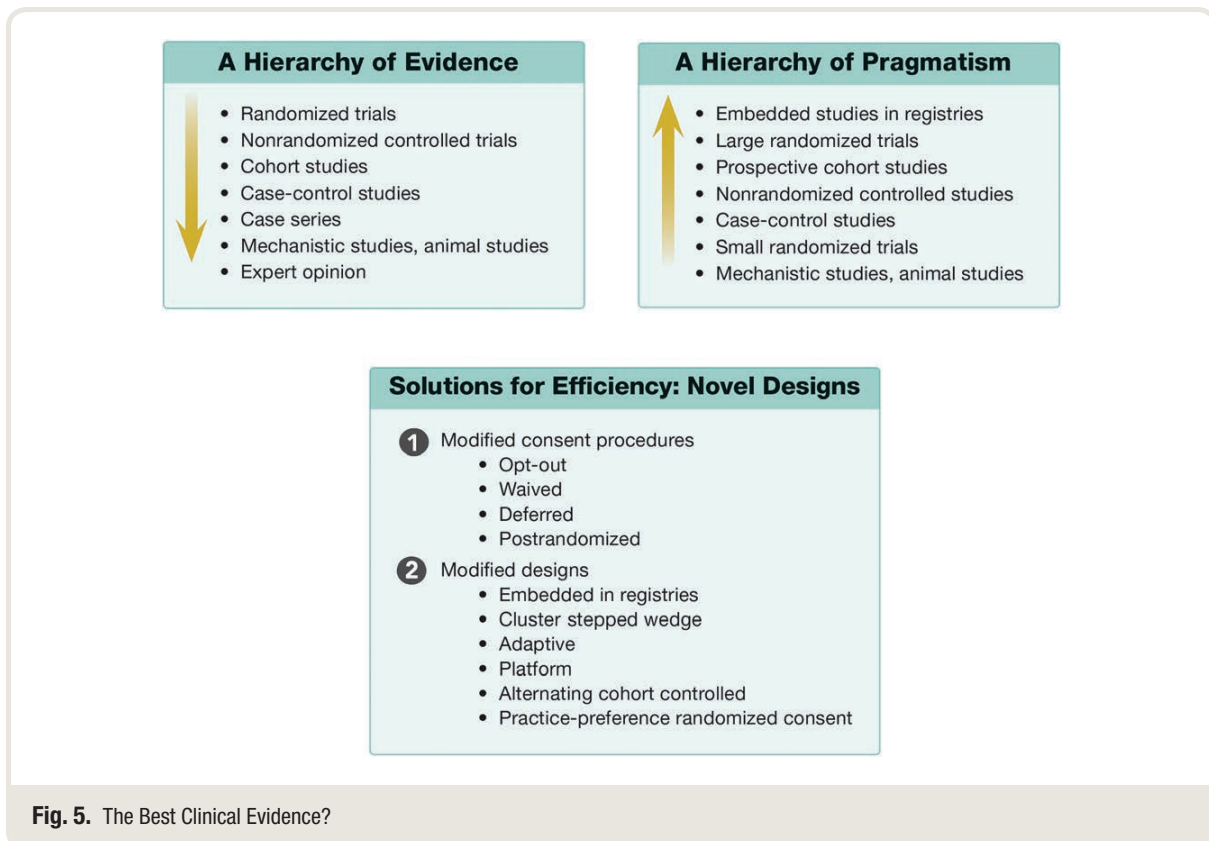
Platform, basket, and umbrella trials share the common characteristic of addressing multiple questions in a single study program, maximizing efficiency by obtaining more information in a shorter time. Adaptive designs in general,

and especially platform, basket, and umbrella trials, are statistically challenging and require complex sample-size estimates and analyses. Professional statistical collaboration is essential.

Embedding Randomized Trials within Registries

Randomized trials are rightly considered to provide the most reliable evidence of treatment effects, but trial entry criteria and stringent methodologies can limit generalizability. Pragmatic randomized trials include a broader range of patients and often in diverse healthcare settings, thereby offering greater generalizability. Even so, there are often groups of patients (*e.g.*, those with cognitive impairment or language difficulty) or healthcare settings (*e.g.*, low resource areas) that are not representative. Large amounts of data kept in electronic medical records and various registries are typically available for research. Clinical information systems and electronic health records can facilitate routine collection of process and outcome data to support low-cost clinical research.⁶⁰ To the extent that necessary outcome data are available from registries, trials can be conducted simply by controlling exposure allocation.⁶¹

Introduction of random assignment into a clinical registry combines the strengths of a large pragmatic trial while costing much less than a typical trial, which requires individual



Box 2. Where to Find More Information on This Topic

- Basic Health and Human Services Policy for Protection of Human Research Subjects (Available at: <https://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID=83cd09e1c0f5c6937cd9d7513160fc3f&pitd=20180719&n=pt45.1.46&r=PART&ty=HTML>; accessed October 7, 2019)
- Consort 2010 statement: Extension to cluster randomised trials. *BMJ* 2012; 345:e5661
- U.S. Food and Drug Administration. Informed Consent: Guidance for IRBs, Clinical Investigators, and Sponsors. July 2014 (Available at: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/informed-consent>; accessed October 7, 2019)
- U.S. Food and Drug Administration. Adaptive Designs for Clinical Trials of Drugs and Biologics: Guidance for Industry (Available at: <https://www.fdanews.com/ext/resources/files/2018/2/09-28-18-AdaptiveTrials.pdf?1538162421>; accessed October 7, 2019)
- An Introduction and Overview of Innovative Trial Design (Available at: <https://www.nhlbi.nih.gov/events/2018/introduction-and-overview-innovative-trial-design>; accessed October 7, 2019)

participant evaluation, completion of case-report forms, and subsequent transfer of information to the trial database (fig. 5).⁶² Registry-based randomized trials represent the epitome of comparative effectiveness research: generalizability of the findings, rapid enrollment, high completion, and often longer-term follow-up. The mechanics of randomization will depend on the context but will most often be similar to conventional randomized trials such as a central web-based system. However, alternative types of randomization or pseudorandomization can also be used when appropriate.

Registry-based randomized trials nonetheless present important challenges including broad capture to avoid exclusion bias, variable quality data in registries, and ethical concerns about lack of patient autonomy (if explicit consent is not obtained) and privacy (keeping research data confidential). For example, data quality in registries virtually never approaches the accuracy level of audited and monitored case-report forms. As in any other study using registry data, the validity of the underlying data is key. Investigators thus need to have a good understanding of the extent to which measurement bias, random error, and missing data might falsify analyses. Registry-based trials are nonetheless a disruptive technology in clinical research⁶³ because they simplify and speed enrollment while much reducing cost. An additional advantage of trials conducted within registries is that it is easy to subsequently evaluate the extent to which trial results are adopted into everyday practice.

Alternating cohort and other cluster trials often use data from electronic records to evaluate outcomes,^{30,31} but it is also possible to conduct otherwise conventional clinical

trials in which all or nearly all outcomes are obtained from electronic records rather than case reports. At the Cleveland Clinic, for example, we routinely conduct trials in which all substantive outcomes are obtained from various electronic registries.⁶⁴ The efficiency of this approach is obvious, and it much reduces the cost of conducting trials.

Conclusions

There will never be enough conventional clinical trials to address even a small fraction of important clinical questions. Novel trial designs, some with waived or modified patient consent, are increasingly being used to answer research questions more efficiently. Modifications to conventional trial designs that introduce flexibility and efficiency are also becoming more common. Many of the enhancements we review are attractive because they speed enrollment and lower the cost of research. They also raise new challenges in terms of planning, conduct, ethical oversight, and statistical analysis.

Research Support

Support was provided solely from institutional and/or departmental sources.

Competing Interests

The authors declare no competing interests.

Correspondence

Address correspondence to Dr. Myles: Alfred Hospital, Commercial Road, Melbourne, Victoria 3004, Australia. p.myles@alfred.org.au. Information on purchasing reprints may be found at www.anesthesiology.org or on the masthead page at the beginning of this issue. ANESTHESIOLOGY's articles are made freely accessible to all readers, for personal use only, 6 months from the cover date of the issue.

References

1. Papanicolas I, Woskie LR, Jha AK: Health care spending in the United States and other high-income countries. *JAMA* 2018; 319:1024–39
2. Devereaux PJ, Yusuf S: The evolution of the randomized controlled trial and its role in evidence-based decision making. *J Intern Med* 2003; 254:105–13
3. Collins R, MacMahon S: Reliable assessment of the effects of treatment on mortality and major morbidity: I. Clinical trials. *Lancet* 2001; 357:373–80
4. Hennekens CH, Demets D: The need for large-scale randomized evidence without undue emphasis on small trials, meta-analyses, or subgroup analyses. *JAMA* 2009; 302:2361–2
5. Myles PS: Why we need large randomized studies in anaesthesia. *Br J Anaesth* 1999; 83:833–4
6. 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
7. Feldman AM: Bench-to-bedside: Clinical and translational research: Personalized medicine: Precision medicine: What's in a Name? *Clin Transl Sci* 2015; 8:171–3
 8. Reith C, Landray M, Devereaux PJ, Bosch J, Granger CB, Baigent C, Califf RM, Collins R, Yusuf S: Randomized clinical trials: Removing unnecessary obstacles. *N Engl J Med* 2013; 369:1061–5
 9. Jackson N, Atar D, Borentain M, Breithardt G, van Eickels M, Endres M, Fraass U, Friede T, Hannachi H, Janmohamed S, Kreuzer J, Landray M, Lautsch D, Le Floch C, Mol P, Naci H, Samani NJ, Svensson A, Thorstensen C, Tijssen J, Vandzhura V, Zalewski A, Kirchhof P: Improving clinical trials for cardiovascular diseases: A position paper from the Cardiovascular Round Table of the European Society of Cardiology. *Eur Heart J* 2016; 37:747–54
 10. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP, Al-Shahi Salman R, Chan AW, Glasziou P: Biomedical research: Increasing value, reducing waste. *Lancet* 2014; 383:101–4
 11. Hadorn D, Wilson N, Edwards R, Blakely T, Sarfati D: How to substantially increase recruitment in cancer trials in New Zealand. *N Z Med J* 2013; 126:57–68
 12. Berry DA: Adaptive clinical trials in oncology. *Nat Rev Clin Oncol* 2011; 9:199–207
 13. Ford I, Norrie J: Pragmatic trials. *N Engl J Med* 2016; 375:454–63
 14. Calvo G, McMurray JJ, Granger CB, Alonso-García Á, Armstrong P, Flather M, Gómez-Outes A, Pocock S, Stockbridge N, Svensson A, Van de Werf F: Large streamlined trials in cardiovascular disease. *Eur Heart J* 2014; 35:544–8
 15. Concato J: Is it time for medicine-based evidence? *JAMA* 2012; 307:1641–3
 16. Concato J: Observational *versus* experimental studies: What's the evidence for a hierarchy? *NeuroRx* 2004; 1:341–7
 17. Sessler DI, Imrey PB: Clinical Research Methodology 1: Study designs and methodologic sources of error. *Anesth Analg* 2015; 121:1034–42
 18. Sessler DI, Imrey PB: Clinical Research Methodology 2: Observational clinical research. *Anesth Analg* 2015; 121:1043–51
 19. Mascha EJ, Sessler DI: Segmented regression and difference-in-difference methods: Assessing the impact of systemic changes in healthcare. *Anesth Analg* 2019; 129:618–33
 20. Ho AMH, Phelan R, Mizubuti GB, Murdoch JAC, Wickett S, Ho AK, Shyam V, Gilron I: Bias in before–after studies: Narrative overview for anesthesiologists. *Anesth Analg* 2018; 126:1755–62
 21. Guidet B, Leblanc G, Simon T, Woimant M, Quenot JP, Ganansia O, Maignan M, Yordanov Y, Delerme S, Doumenc B, Fartoukh M, Charestan P, Trognon P, Galichon B, Javaud N, Patzak A, Garrouste-Orgeas M, Thomas C, Azerad S, Pateron D, Boumendil A; ICE-CUB 2 Study Network: Effect of systematic intensive care unit triage on long-term mortality among critically ill elderly patients in France: A randomized clinical trial. *JAMA* 2017; 318:1450–9
 22. Chen CC, Li HC, Liang JT, Lai IR, Purnomo JD, Yang YT, Lin BR, Huang J, Yang CY, Tien YW, Chen CN, Lin MT, Huang GH, Inouye SK: Effect of a modified hospital elder life program on delirium and length of hospital stay in patients undergoing abdominal surgery: A cluster randomized clinical trial. *JAMA Surg* 2017; 152:827–34
 23. Kappen TH, Moons KG, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, van Klei WA: Impact of risk assessments on prophylactic antiemetic prescription and the incidence of postoperative nausea and vomiting: A cluster-randomized trial. *ANESTHESIOLOGY* 2014; 120:343–54
 24. Brown AW, Li P, Bohan Brown MM, Kaiser KA, Keith SW, Oakes JM, Allison DB: Best (but oft-forgotten) practices: Designing, analyzing, and reporting cluster randomized controlled trials. *Am J Clin Nutr* 2015; 102:241–8
 25. Smit-Fun VM, de Korte-de Boer D, Posthuma LM, Stolze A, Dirksen CD, Hollmann MW, Buhre WF, Boer C: TRACE (routine postoperative Anesthesia visit to improve patient outcome): A prospective, multicenter, stepped-wedge, cluster-randomized interventional study. *Trials* 2018; 19:586
 26. Haugen AS, Søfteland E, Almeland SK, Sevdalis N, Vonen B, Eide GE, Nortvedt MW, Harthug S: Effect of the World Health Organization checklist on patient outcomes: A stepped wedge cluster randomized controlled trial. *Ann Surg* 2015; 261:821–8
 27. Fenger-Gron M, Vestergaard M, Pedersen HS, Frost L, Parner ET, Ribe AR, Davydov DS: Depression, antidepressants, and the risk of non-valvular atrial fibrillation: A nationwide Danish matched cohort study. *Eur J Prev Cardiol* 2019; 26:187–95
 28. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ: The stepped wedge cluster randomised trial: Rationale, design, analysis, and reporting. *BMJ* 2015; 350:h391
 29. Concato J: When to randomize, or “evidence-based medicine needs medicine-based evidence.” *Pharmacoepidemiol Drug Saf* 2012; 21:6–12
 30. Kopyeva T, Sessler DI, Weiss S, Dalton JE, Mascha EJ, Lee JH, Kiran RP, Udeh B, Kurz A: Effects of volatile anesthetic choice on hospital length-of-stay: A retrospective study and a prospective trial. *ANESTHESIOLOGY* 2013; 119:61–70
 31. Kurz A, Kopyeva T, Suliman I, Podolyak A, You J, Lewis B, Vlah C, Khatib R, Keebler A, Reigert R, Seuffert M, Muzie L, Drahuschak S, Gorgun E, Stocchi L, Turan A, Sessler DI: Supplemental oxygen and surgical-site

- infections: An alternating intervention controlled trial. *Br J Anaesth* 2018; 120:117–26
32. Self WH, Semler MW, Wanderer JP, Wang L, Byrne DW, Collins SP, Slovis CM, Lindsell CJ, Ehrenfeld JM, Siew ED, Shaw AD, Bernard GR, Rice TW, Investigators S-E: Balanced crystalloids *versus* saline in noncritically ill adults. *N Engl J Med* 2018; 378:819–28
 33. Semler MW, Self WH, Wanderer JP, Ehrenfeld JM, Wang L, Byrne DW, Stollings JL, Kumar AB, Hughes CG, Hernandez A, Guillaumondegui OD, May AK, Weavind L, Casey JD, Siew ED, Shaw AD, Bernard GR, Rice TW; SMART Investigators and the Pragmatic Critical Care Research Group: Balanced crystalloids *versus* saline in critically ill adults. *N Engl J Med* 2018; 378:829–39
 34. Honarmand K, Belley-Cote EP, Ulic D, Khalifa A, Gibson A, McClure G, Savija N, Alshamsi F, D'Aragon F, Rochweg B, Duan EH, Karachi T, Lamontagne F, Devereaux PJ, Whitlock RP, Cook DJ: The deferred consent model in a prospective observational study evaluating myocardial injury in the intensive care unit. *J Intensive Care Med* 2018; 33:475–80
 35. Cook DJ, Blythe D, Rischbieth A, Hebert PC, Zytaruk N, Menon K, Erikson S, Fowler R, Heels-Ansdell D, Meade MO: Enrollment of intensive care unit patients into clinical studies: A trinational survey of researchers' experiences, beliefs, and practices. *Crit Care Med* 2008; 36:2100–5
 36. Cooper DJ, McQuilten ZK, Nichol A, Ady B, Aubron C, Bailey M, Bellomo R, Gantner D, Irving DO, Kaukonen KM, McArthur C, Murray L, Pettilä V, French C; TRANSFUSE Investigators and the Australian and New Zealand Intensive Care Society Clinical Trials Group: Age of red cells for transfusion and outcomes in critically ill adults. *N Engl J Med* 2017; 377:1858–67
 37. Saager L, Turan A, Egan C, Mascha EJ, Kurz A, Bauer M, Besson H, Sessler DI, Hesler BD: Incidence of intraoperative hypersensitivity reactions: A registry analysis. *ANESTHESIOLOGY* 2015; 122:551–9
 38. Panjasawatwong K, Sessler DI, Stapelfeldt WH, Mayers DB, Mascha EJ, Yang D, Kurz A: A Randomized Trial of a Supplemental Alarm for Critically Low Systolic Blood Pressure. *Anesth Analg* 2015; 121:1500–7
 39. Sessler DI, Sigl JC, Kelley SD, Chamoun NG, Manberg PJ, Saager L, Kurz A, Greenwald S: Hospital stay and mortality are increased in patients having a “triple low” of low blood pressure, low bispectral index, and low minimum alveolar concentration of volatile anesthesia. *ANESTHESIOLOGY* 2012; 116:1195–203
 40. Sessler DI, Turan A, Stapelfeldt WH, Mascha EJ, Yang D, Farag E, Cywinski J, Vlah C, Kopyeva T, Keebler AL, Perilla M, Ramachandran M, Drahuschak S, Kaple K, Kurz A: Triple-low alerts do not reduce mortality: A real-time randomized trial. *ANESTHESIOLOGY* 2019; 130:72–82
 41. Myles PS, Dieleman JM, Forbes A, Heritier S, Smith JA: Dexamethasone for Cardiac Surgery trial (DECS-II): Rationale and a novel, practice preference-randomized consent design. *Am Heart J* 2018; 204:52–7
 42. Wang X: Play-the-winner rule and adaptive designs of clinical trials. *IJMMS* 2001; 27:229–36
 43. Brown CH, Ten Have TR, Jo B, Dagne G, Wyman PA, Muthén B, Gibbons RD: Adaptive designs for randomized trials in public health. *Annu Rev Public Health* 2009; 30:1–25
 44. Kuznetsova OM, Tymofyeyev Y: Preserving the allocation ratio at every allocation with biased coin randomization and minimization in studies with unequal allocation. *Stat Med* 2012; 31:701–23
 45. Heritier S, Lô SN, Morgan CC: An adaptive confirmatory trial with interim treatment selection: Practical experiences and unbalanced randomization. *Stat Med* 2011; 30:1541–54
 46. Taves DR: The use of minimization in clinical trials. *Contemp Clin Trials* 2010; 31:180–4
 47. Heritier S, Lloyd CJ, Lô SN: Accurate p-values for adaptive designs with binary endpoints. *Stat Med* 2017; 36:2643–55
 48. Meinzer C, Martin R, Suarez JI: Bayesian dose selection design for a binary outcome using restricted response adaptive randomization. *Trials* 2017; 18:420
 49. Gould AL: Interim analyses for monitoring clinical trials that do not materially affect the type I error rate. *Stat Med* 1992; 11:55–66
 50. Proschan MA: Two-stage sample size re-estimation based on a nuisance parameter: A review. *J Biopharm Stat* 2005; 15:559–74
 51. Kieser M, Friede T: Re-calculating the sample size in internal pilot study designs with control of the type I error rate. *Stat Med* 2000; 19:901–11
 52. Friede T, Kieser M: Sample size recalculation in internal pilot study designs: A review. *Biom J* 2006; 48:537–55
 53. Posch M, Bauer P: Interim analysis and sample size reassessment. *Biometrics* 2000; 56:1170–6
 54. Curtin F, Heritier S: The role of adaptive trial designs in drug development. *Expert Rev Clin Pharmacol* 2017; 10:727–36
 55. Lin J, Bunn V: Comparison of multi-arm multi-stage design and adaptive randomization in platform clinical trials. *Contemp Clin Trials* 2017; 54:48–59
 56. Parmar MK, Sydes MR, Cafferty FH, Choodari-Oskooei B, Langley RE, Brown L, Phillips PP, Spears MR, Rowley S, Kaplan R, James ND, Maughan T, Paton N, Royston PJ: Testing many treatments within a single protocol over 10 years at MRC Clinical Trials Unit at UCL: Multi-arm, multi-stage platform, umbrella and basket protocols. *Clin Trials* 2017; 14:451–61
 57. van der Velden JM, Verkooijen HM, Young-Afat DA, Burbach JP, van Vulpen M, Relton C, van Gils CH, May AM, Groenwold RH: The cohort multiple randomized

- controlled trial design: A valid and efficient alternative to pragmatic trials? *Int J Epidemiol* 2017; 46:96–102
58. West HJ: Novel precision medicine trial designs: Umbrellas and baskets. *JAMA Oncol* 2017; 3:423
 59. Renfro LA, Sargent DJ: Statistical controversies in clinical research: Basket trials, umbrella trials, and other master protocols: A review and examples. *Ann Oncol* 2017; 28:34–43
 60. Cowie MR, Blomster JI, Curtis LH, Duclaux S, Ford I, Fritz F, Goldman S, Janmohamed S, Kreuzer J, Leenay M, Michel A, Ong S, Pell JP, Southworth MR, Stough WG, Thoenes M, Zannad F, Zalewski A: Electronic health records to facilitate clinical research. *Clin Res Cardiol* 2017; 106:1–9
 61. Simon KC, Tideman S, Hillman L, Lai R, Jathar R, Ji Y, Bergman-Bock S, Castle J, Franada T, Freedom T, Marcus R, Mark A, Meyers S, Rubin S, Semenov I, Yucus C, Pham A, Garduno L, Szela M, Frigerio R, Maraganore DM: Design and implementation of pragmatic clinical trials using the electronic medical record and an adaptive design. *JAMIA Open* 2018; 1:99–106
 62. Li G, Sajobi TT, Menon BK, Korngut L, Lowerison M, James M, Wilton SB, Williamson T, Gill S, Drogos LL, Smith EE, Vohra S, Hill MD, Thabane L: 2016 Symposium on Registry-Based Randomized Controlled Trials in Calgary: Registry-based randomized controlled trials: What are the advantages, challenges, and areas for future research? *J Clin Epidemiol* 2016; 80:16–24
 63. Lauer MS, D'Agostino RB Sr: The randomized registry trial: The next disruptive technology in clinical research? *N Engl J Med* 2013; 369:1579–81
 64. Heddle NM, Cook RJ, Arnold DM, Liu Y, Barty R, Crowther MA, Devereaux PJ, Hirsh J, Warkentin TE, Webert KE, Roxby D, Sobieraj-Teague M, Kurz A, Sessler DI, Figueroa P, Ellis M, Eikelboom JW: Effect of short-term *vs.* long-term blood storage on mortality after transfusion. *N Engl J Med* 2016; 375:1937–45