Clinical and Translational Science

TUTORIAL OPEN ACCESS

Protocol Development for Investigator-Sponsored Clinical Studies

Lyn Frumkin

School of Medicine, SPARK at Stanford, Stanford University, Stanford, California, USA

Correspondence: Lyn Frumkin (lfrumkin@stanford.edu)

Received: 25 February 2025 | **Revised:** 4 April 2025 | **Accepted:** 11 April 2025

Funding: The author received no specific funding for this work.

Keywords: clinical protocol | clinical study | clinical trial | clinical trial protocol

ABSTRACT

Clinical trials with investigator sponsors at academic sites have increased, in part due to studies involving drug repurposing, the process of identifying new uses for existing drugs that are initially conducted in patients rather than healthy participants. In contrast to industry- or government-sponsored trials, investigator-sponsored clinical studies, also known as investigator-initiated trials, are typically conducted at one or several academic centers and are resource-limited by finances and patient numbers. These studies can serve as crucial pilot studies to inform the design of larger, more definitive clinical trials. Drawing from the experience of working with clinical researchers in academic settings, this tutorial presents guidelines for writing clinical protocols for resource-limited investigator-sponsored studies that meet international standards and optimize the detection of meaningful signals or outcomes that can lead to investigation in larger well-controlled trials.

1 | Introduction

The clinical trial protocol is a roadmap that provides sufficient detail to understand the background, rationale, objectives, study population, interventions, methods, statistical analyses, ethical considerations, and conduct of a clinical trial [1]. It allows others to confirm and extend key aspects of the trial and permits appraisal of its scientific and ethical rigor. The importance of a well-written protocol has been emphasized by journal editors, reviewers, researchers, and public advocates [1].

The classical phased model of drug development, outside of oncology, usually entails an initial Phase 1 study in 40–60 healthy participants, often involving sequential cohorts of 8 participants (6 active, 2 placebo) in an ascending dose escalation, initially with cohorts receiving single doses and then multiple doses for treatment up to 14 days. These studies may determine safety, pharmacokinetics (PK), pharmacodynamics, and maximum tolerated dose (the highest dose of treatment that does not cause unacceptable side effects). With nontoxic therapies, the highest initial dose is often determined by preclinical safety and PK results or halting of dose escalation based on a physiological effect in human cohorts. Subsequent doses in multiple-dose cohorts are based in part on PK and safety from the single-ascending dose phase. Phase 2 estimates initial clinical activity and safety in patients to determine the best endpoints, doses, and signal strength. Phase 3 confirms efficacy, usually in two well-controlled trials (or single trials with more stringent criteria, usually for rare conditions with high unmet medical needs), and establishes the safety profile in a larger number of patients [2, 3].

Early-stage clinical studies in academic settings that are sponsored by investigators rather than organizations or companies are often limited in resources (operational, financial, patient number), which challenges the ability to detect clinically relevant effects. This tutorial for writing clinical protocols for investigator-sponsored pilot studies is intended to highlight approaches that can better maximize the ability to detect meaningful signals or outcomes.

2 | Investigator-Sponsored Trials

In an investigator-sponsored trial, the investigator rather than a company, institution, or organization files the Investigational New Drug (IND) application and is the responsible party for the design, conduct, and analysis of the clinical investigation. Such trials often involve drug repurposing, the process of identifying new uses for existing drugs or active substances. As a result, the initial human clinical studies are conducted in patients rather than healthy participants because safety and PK data are available from studies of the previously approved indication. These initial phase 1b or 2a studies can employ dose-escalation or parallel-group designs involving various controls (e.g., placebo, historical, standard of care) to determine dose ranging, PK, safety, and preliminary efficacy. In contrast to industry- or government-sponsored trials, investigator-sponsored clinical studies are typically conducted at one or several academic centers, with fewer finances and operational resources, including the number of study participants. Despite these concerns, investigator-sponsored clinical studies usually have sufficient ability to evaluate for preliminary evidence of safety, efficacy, and variability, including the presence of biological activity predictive of clinical outcomes by close attention to protocol development.

3 | Protocol Development for Investigator-Sponsored Clinical Trials

3.1 | Impact of Protocol Deficiencies and CONSORT/SPIRIT Initiatives

Clinical trial protocols are read by a wide range of individuals, including investigators and study staff (e.g., coordinators, pharmacists, biostatisticians), Institutional Review Board (IRB) or Ethics Committee (EC) members, regulatory agency physicians/scientists, scientific reviewers, public advocates, the general public (through a clinical trial registry or if the journal publishing the study requires or allows the protocol to be included as a supplemental appendix), and industry or venture capital firms doing due diligence for funding/licensing.

The impact of protocol deficiencies can be profound and lead to poor trial conduct, protocol amendments, unreliable results, jeopardized publication, and an inability to judge the reliability and validity of findings important for systematic reviews [4]. This can lead to problems varying from misinterpretation of entry criteria, varying conduct of the trial across the study participants, misapplication of methods by others trying to replicate or extend trial findings, or other aspects of the trial that impact the integrity of the data. Furthermore, studies that enroll and expose participants to risk but are not designed properly for accurate evaluation of the study objectives are unethical.

The CONSORT (Consolidated Standards of Reporting Trials) Statement was an initial and important initiative to improve the reporting results of parallel-group randomized clinical trials [5, 6]. Subsequently, the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) Initiative was an international project involving 115 key stakeholders, including trial investigators, healthcare professionals, methodologists, statisticians, trial coordinators, journal editors, and representatives from the research ethics community, industry, and regulatory agencies to identify protocol deficiencies and draft guidelines for the minimum content of a clinical trial protocol [1]. This initiative identified major deficiencies, of which the most common were in describing the primary outcome, treatment allocation methods, use of blinding, methods for reporting adverse events, components of sample size calculations, and prespecified data analyses. As a result, SPIRIT provides a checklist of the minimum set of items to be addressed in protocols, consistent with the International Conference on Harmonisation (ICH) Good Clinical Practice E6 guidance, trial registration requirements from the World Health Organization, the International Committee of Medical Journal Editors, ClinicalTrials.gov Registry, the European Commission, and others [1]. The SPIRIT-Outcomes 2022 [7] and SPIRIT-Surrogate [8] extensions expanded the SPIRIT 2013 statement with additional recommendations. This article will now review the key guidelines to consider during trial design and protocol development of investigator-sponsored clinical studies.

3.2 | Where to Begin: Clinical Trial Templates and Complete Protocols

There are many protocol templates available online, typically associated with a university's research hub to guide faculty in writing investigator-sponsored protocols. However, many publicly available templates of clinical trial protocols are written as outlines and may lack relevant detail. Excellent resources for publicly available clinical trial protocol templates are SPARK at Stanford [9] with guidance language preceding sections that contain full protocol language, University of California, San Francisco [10], University of Pennsylvania [11], and the US National Institutes of Health–Food and Drug Administration [12]. While complete protocols are increasingly available in appendices of journals reporting late-stage clinical trial results, they may contain sections and information that are not necessary for early investigator-sponsored trials. Continued efforts to improve and standardize protocol writing have led the International Council for Harmonisation of Technical Requirements of Pharmaceuticals for Human Use to release a draft guidance document for a clinical trial protocol template [13].

4 | Components of a Clinical Protocol

There are differences in the format and level of detail between protocols depending upon their phase, design, template being used, and styles. Key categories that should be in protocols, with subdivisions in some categories, are summarized in the ICH Good Clinical Practice E6 guidance for clinical trials whose data are intended for submission to regulatory authorities. These categories are also consistent with the SPIRIT Initiative, trial registration requirements from the World Health Organization, the International Committee of Medical Journal Editors, the European Commission, and other regulatory bodies. Table 1 shows the categories to be discussed in the current article.

TABLE 1 | Categories for an investigator-sponsored clinical study protocol.

Category	Subcategory
Introduction	
Objectives	
Endpoints	
Investigational Plan	
Outline of Visit Schedule	
Schedule of Events	
Discussion of Design	Rationale for design and controls Rationale for doses and dosing Rationale for assessments Rationale for endpoints
Data Analysis Methods	Determination of Sample Size Efficacy and Safety Variables Statistical and Analytic Plans Interim Analyses/Data Monitoring Committees
Study Population	
Treatments	
Adverse Events	Serious Adverse Events Reporting Requirements
Quality Control and Quality Assurance	
Administrative, Ethical, and Regulatory Considerations	Ethical Review Regulatory Considerations Study Documentation, Privacy, and Records Retention Declaration of Interests Data Availability Statement Study Finances Publication Statement

4.1 | Introduction

The four-step approach is most helpful: Disease, Investigational Treatment, Relationship of Treatment to Disease (i.e., rationale for use of the drug for a disease that includes relevant nonclinical, preclinical, and clinical data), and Risk/Benefit. Introductions of academic protocols can often be long and do not need to extensively review the literature or contain all the nonclinical and preclinical information. This is best summarized in the Investigator's Brochure, which is a compilation of clinical and nonclinical data about the investigational product that is the focus of the study; for repurposed indications, this may be the package insert for the already marketed drug. Instead, the introduction to an investigator-sponsored study should be a succinct summary of the four areas, but be informative and include a review of the standard of care (approved or unapproved) when describing the disease. Omitting the current standard of care may result from a belief that discussing approved drugs may decrease the perception of a need for new ones. However, unmet medical need is driven by suboptimal treatment, not the number of approved drugs. Finally, adequate citation of the initial approved indication(s) and safety profile of the repurposed drug is important, as the adverse event profile from other indications may be highly relevant to the new proposed use due to class effects of the drug.

4.2 | Objectives

There are different styles for this section, but simplicity is often best: "The primary objective of this study is to evaluate safety and tolerability of oral SA100 given twice daily for 8 weeks, assessed by clinical and laboratory adverse events. The secondary objective is to evaluate preliminary efficacy (clinical response) based on (a) accepted disease activity and lab measures, including histopathologic assessment of the colon by endoscopy, and (b) changes in serum biomarkers." Common problems are mixing aims, hypotheses, outcome measures, and endpoints so that the reader is unclear about the overriding objectives. This section is often unnecessarily complicated by the standard use of "aims" (specific, measurable goals of the research project) and "hypotheses" (explanation for the expected outcome based on existing knowledge) common in grants written by academic investigators.

4.3 | Endpoints

An endpoint is an event or outcome that can be measured objectively to determine whether the intervention being studied is active. The distinction between primary, secondary, and exploratory endpoints is less important in investigator-sponsored trials than in later-stage trials designed to obtain sufficient data and results to support the filing of an application for regulatory approval. However, it is still important to identify a primary endpoint to communicate what is considered the most relevant measure and to justify sample size estimations for the number of participants studied. Common problems in the discussion of endpoints are defining an endpoint measure, but not time (e.g., mean change from baseline in number of Grade 2 bleeding events rather than mean change from baseline to week 20 in number of Grade 2 bleeding events), not including responder endpoints (when appropriate), and not recognizing (or less commonly, overemphasizing) the value of surrogate endpoints.

4.3.1 | Responder Endpoint

A responder endpoint refers to a specific outcome measure that categorizes a patient as either a responder (they achieved an improvement based on set criteria) or a non-responder. The advantage of responder endpoints, when available, is that they build in clinically meaningfulness [14]. Examples of responder endpoints are glycated hemoglobin A1c concentration $\leq 6.5\%$ in Type 2 diabetes remission [15], the proportion of patients who achieve continuous abstinence from Weeks 9 through 12 in nicotine cessation studies [16], and the proportion of patients with pain and at least a 30% reduction (moderately important improvement) or at least a 50% reduction (substantial improvement) in pain intensity at a posttreatment time [17]. Proportion endpoints that include the number of participants with clinical

remission or clinical response (by predetermined criteria) are additional examples of responder endpoints. For some diseases, criteria for clinically meaningfulness are less defined or not agreed upon. The use of responder endpoints should be added to other endpoints as limitations can include less data than a continuous endpoint (e.g., mean change from baseline), cutoff values used to determine a responder that may be an arbitrary point on a continuous scale, and the potential requirement of more participants to detect a treatment effect [18]. As a result, using continuous variables in addition to responder endpoints to examine mean differences between groups remains important for determining clinical relevance.

4.3.2 | Surrogate Endpoint

Conducting a clinical trial with enough participants to detect a clinically meaningful outcome is ideal. Because investigator-initiated trials are usually limited in patient number to detect the most relevant clinical outcomes, consideration of the use of biomarkers and surrogate endpoints is important. A biomarker, usually a laboratory measure, is an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention [19, 20]. A biomarker can also be a surrogate endpoint, which is a substitute for an endpoint that is considered clinically meaningful. In studies limited in patient number to detect a clinically meaningful outcome, validated surrogate endpoints that reliably correlate with clinically meaningful endpoints may occur earlier, more frequently, and be better able to detect a true signal of effect. Alternately, surrogate endpoints that are not yet validated, yet reasonably likely to predict clinical benefit, may also be important to identify a signal that predicts benefit with clinical outcomes in a later-stage trial with an increased number of participants. Fleming and Powers [20] have provided examples of different categories of outcome measures, according to the level of evidence. Levels 2, 3, and 4 are considered indirect endpoints. Some examples are:

Level 1: A true clinical efficacy measure (when evidence establishing risk is acceptable in the context of evidence of benefit) such as death or hospitalization in heart failure.

Level 2: A validated surrogate (for a specific disease setting and class of interventions and when interventions are safe, with strong evidence that risks from off-target effects are acceptable) such as hemoglobin A1c for clinical effects on long-term risk of microvascular complications in Type 2 diabetes mellitus.

Level 3: A non-validated surrogate, yet one established to be 'reasonably likely to predict clinical benefit' (for a specific disease setting and class of interventions and when interventions are safe, with evidence that risks from off-target effects are acceptable) such as large and durable effects on viral load in some human immunodeficiency virus settings.

Level 4: A correlate that measures biological activity but has not been established to be at levels above, such as fever for the treatment of community-acquired bacterial pneumonia.

Surrogate endpoints, including use as secondary endpoints to support other primary outcome measures, are important considerations for initial resource-limited clinical trials. Less commonly, investigator-initiated protocols that use surrogate endpoints as the primary outcome measure may overemphasize their importance by not recognizing the need for later translation into more clinically meaningful endpoints. While the use of surrogate endpoints, especially as primary outcome measures, provides efficiency in early-stage studies (i.e., reducing sample size and duration), limitations include an inability to predict the intervention effect on the target outcome and overestimating treatment effects [21]. In investigator-sponsored trials with limited resources, the advantage of considering the use of surrogate endpoints to identify important signals that can result in subsequent investigation usually outweighs the disadvantages when coupled with recognition of limitations. Identifying and reporting surrogate endpoints that include justification for use as a primary outcome measure in randomized controlled trials (SPIRIT-Surrogate, CONSORT-Surrogate) was recently added as extension items to the main SPIRIT and CONSORT protocol checklists [22].

4.4 | Investigational Plan

This section should be a brief and well-written summary paragraph of the study: "This is a 12-week, multicenter, open-label, dose-escalation Phase 1b study of SA100 given orally to patients with ulcerative colitis. Eight patients will be sequentially enrolled in 1 of 4 cohorts. Treatment will be given for 8 weeks followed by a 4-week nontreatment follow-up period (study duration of 12 weeks). The study plans to enroll 24 patients at one site."

Brief paragraphs describing the type and timing of assessments should follow: "Safety evaluations will consist of physical examination, vital signs, and laboratory evaluations at screening or baseline, and Weeks 2, 4, 8, and 12 visits." "Efficacy evaluations will consist of..."

4.5 | Outline of Visit Schedule

This section should include a visit-by-visit description that includes the visit date (e.g., Week 2) and a bulleted list of assessments or procedures to be obtained. Common problems include describing a visit only by number and not the specific time (e.g., Visit 3 rather than Week 16 Visit), adding endpoints to assessments "flexible sigmoidoscopy to evaluate whether...", and not defining whether screening (entry) procedures can also be used as baseline values if both times are close together. Instead, certain procedures are often incorrectly listed in both screening and baseline when the intention is that a screening lab or procedure can also be used as a baseline assessment. The assessments must support the calculation of the endpoints.

4.6 | Schedule of Events—Table

A Schedule of Events as a table with time (columns) and procedures (rows) can be added to the body of the protocol or as

an appendix and helps summarize study events for study staff and others. Similar to the outline of the visit schedule, a common error is not defining the difference between screening and baseline assessments, the former required for evaluation of inclusion/exclusion criteria (screening) rather than when they are taken. Both screening and baseline assessments (generally defined as the last value before beginning study treatment) can be taken on the same day or at different times. A common problem is that a screening lab also intended to be used as a baseline lab is listed in both places without a footnote, which implies two tests. In some situations, a screening lab should not be used as a baseline lab due to differences in collection times that make screening values impractical for change from baseline analyses; in this case, listing a lab or procedure at both screening and baseline without further clarification is appropriate.

4.7 | Discussion of Design

This section should be brief and succinct, and explain the rationale for the design (including use of controls), doses, assessments (including duration of assessment), and endpoints. Common problems include dose-escalating using safety criteria that are more appropriate for oncology (toxic) therapies, and not mentioning a regulatory precedent or guidelines as a basis for the use of endpoints.

4.7.1 | Discussion of Design (Use of Control)

Options for control groups include: open-label, unblinded groups (with or without control groups), placebo group (open-label or blinded), active treatment with a specific (usually approved) drug, standard of care (optimal treatment available), and standard care (clinical practice in a real-world setting). If an active control treatment is used, discussion of what would be considered a similar effect to the investigational drug should occur, as discussed in ICH E10 Guidance [23]. In addition, crossover designs, more common in PK studies, are increasingly used in early-stage trials with a limited number of participants. A crossover design is a repeated measurement design where each participant receives different treatments during different periods; they cross over from one treatment to another during the trial after a washout period. This type of design has two potential advantages: (1) the influence of confounding covariates is reduced because each crossover patient serves as their own control, and (2) the design is statistically efficient in that in a resource-limited study it requires fewer participants than non-crossover designs. Limitations of crossover designs include longer study duration, more study visits (burden on patients), and, most importantly, the risk of residual effects from the prior therapy. This most often occurs when the washout period is based on pharmacological indices (e.g., biological half-life of the drug) without taking into account that the clinical effects may last considerably longer.

Description of the rationale for use of a placebo can often be something as succinct as: placebo effects exist with this disease, and the present design allows the most direct assessment of the safety and preliminary efficacy of the investigational agent in this patient population.

4.7.2 | Rationale for Use of Placebo Control

The use of a double-blind, placebo control is optimal for demonstrating a drug effect and identifying underlying disease-related patterns. Enrollment in a placebo arm is acceptable when there is no effective intervention or when added to an established intervention [24, 25]. Use of placebo is also ethical if withholding effective treatment will not cause more than minor risk, patients are fully informed about alternatives, and effective risk mitigation plans exist. The use of placebo in the absence of extensive and consistent historical control data is methodologically important if ethically allowed.

In some oncology settings, placebo is rarely used alone due to practical (toxicity is unblinding) and ethical (irreversible harm) reasons. Alternatively, with most diseases, withholding a drug in a placebo arm does not affect long-term health. It is also common to use a placebo arm in participants on stable background therapy, with standard-of-care treatment receiving placebo or active drug. In some situations, three treatment arms that include a placebo (investigational drug, active control, placebo) may help distinguish between a study that does not work (results of all treatment arms are similar) versus a drug that does not work (only active arm shows efficacy).

The ability to recruit patients with rare diseases into investigator-initiated trials is a special challenge given the resources already imposed by limited sites and operational abilities. In many situations, the use of a placebo arm will lessen enrollment. As a result, it is critical to evaluate whether such a control is necessary for trials where no or few placebo effects exist, such as with hemodynamic measurements in pulmonary arterial hypertension [26, 27]. When the use of a placebo arm is ethical, the availability of an open-label extension trial results in increased receptivity of participants to enroll in a placebo-controlled trial. Extension phases can assess the durability of a benefit and collect longer-term safety data and can also be attractive to patients considering participation in placebo-controlled trials. These studies allow those who complete the initial trial without deterioration or discontinuation due to worsening and are not receiving the highest dose of active drug to enroll in an active treatment arm (pending results of the initial study); however, such extensions are usually not practical due to the budget constraints of investigator-initiated trials. Finally, early-stage studies that use a 2:1 (active: placebo) randomization may increase the ability to accrue more study participants despite an increased sample size needed to achieve the same level of precision or statistical power as equal randomization; however, significant limitations warrant caution when considering higher ratios.

4.7.3 | Rationale for Doses and Dosing

Determining the appropriate dose and regimen is critical during the development of new drugs and has contributed to a significant number of failures of the first regulatory review cycle because of uncertainties in the dose selection rationale [28]. With repurposed drugs evaluated in investigator-initiated clinical trials, dosing information is available from an approved drug for a different therapeutic indication and may reduce these uncertainties. Small phase 1b or 2a trials often require judgment

rather than prespecified criteria when deciding if it is safe and appropriate to advance to the next dose.

4.7.4 | Rationale for Assessments and Endpoints

In an investigator-sponsored clinical trial, the rationale for selecting assessments and endpoints is related to maximizing the ability to measure clinically relevant outcomes that are specific, measurable, achievable, relevant to the patient population, and sensitive enough to detect a meaningful and true treatment effect. The rationale for assessments and endpoints can often rely on expert and regulatory consensus, including the design of studies for drugs in the proposed indication that have achieved regulatory approval.

4.8 | Data Analysis Methods

This section summarizes the determination of sample size, efficacy and safety variables with a summary of assessments and when each assessment is obtained (rationale is in earlier section), the statistical and analytical plans that include handling of missing data, how safety, efficacy, and other endpoints are analyzed, and interim analyses/Data Monitoring Committee.

An emphasis is placed here on the determination of sample size, which is a critical component of investigator-sponsored clinical trials that are challenged by limited operational and financial resources. Greenland et al. [29] have noted in discussing the misapplication of statistical principles that a key challenge is that there are no interpretations of these concepts that are at once simple, intuitive, correct, and foolproof. A biostatistician is critical to the trial design, sample size estimations, and data analysis methods.

4.8.1 | Determination of Sample Size

Sample size calculations convey the degree of confidence that a given number of participants will be adequate to detect a predetermined or greater clinically relevant treatment effect. The most common approach to sample size determination is to identify the number of study participants that will provide an adequate level of confidence to conclude at study completion that a treatment has efficacy when it truly does (power, usually 80% or 90%) and a low probability to falsely conclude a treatment has efficacy when it truly does not (alpha, usually 5%). Pilot studies with small numbers of participants may have lower power than desired (i.e., decreased ability to detect a true treatment effect) at a standard significance level. Aside from wasted resources (time and money), studies that cannot assess the safety and efficacy of a treatment because of an inadequate number of participants should not be conducted on ethical grounds. It is wrong to expose patients to a drug in a new setting if the study cannot support meaningful conclusions.

4.8.2 | Power and Sample Size Calculation

Power is the probability that, given a prespecified true difference between two groups, the quantitative results of a study will be deemed statistically significant [30]. Power is a pretrial concept, which is of limited use to apply to observed differences after the study. Because the statistical power of a study is the probability that a significance test will detect an effect that truly exists, the focus of a trial that fails to reach significance should be the observed treatment effect and the confidence interval.

A failed trial may be partially the result of a true effect that is smaller than is clinically meaningful. Alternately, a common error is estimating treatment effects known to be unrealistically favorable to justify both a small number of evaluable patients and a desire for 80% power. Goodman and Berlin [30] have emphasized that results of this practice are journals filled with reports of possibly clinically important but statistically nonsignificant effects (i.e., a smaller clinically relevant true effect could not be detected) and research that is inefficient and wasteful. Assumptions underlying sample size estimations should be carefully considered.

4.8.3 | Calculating Sample Size

Calculation of sample size (number of study participants) relies on the following choices:

- Estimated treatment effect (assumed to be an effect size that is clinically relevant).
- Assignment ratio of treatments (e.g., 1:1).
- Power, which is usually set at 80% or 90%, is the probability of reaching statistical significance for a given treatment effect or greater if one truly exists ("true positive"). Low power can lead to falsely rejecting an effective therapy (Type II error).
- Significance level, or alpha, is the probability of falsely detecting a treatment effect when none truly exists ("false positive" or Type I error). This can lead to falsely accepting ineffective therapy. As a result, alpha is typically set to be small (e.g., 5%). Type 1 error is more often set to be two-sided, which examines whether a result is significantly different from a reference point in either direction (greater or less than), while a one-sided test (that offers more statistical power) evaluates if the result is significantly different in one specific direction.
- The standard deviation (variability) associated with measuring the primary endpoint for continuous variables is often the most difficult parameter to set and should be based on previous trials or other historical information.

Determining sample size based on a level of precision in the estimation of treatment efficacy by proposing the desired width of the confidence interval is an alternate approach [31]. Planning sample size based on precision lessens the problem of dichotomizing results into statistically significant vs. nonsignificant with resulting misinterpretation of nonsignificant but meaningful findings as well as small significant findings that are not clinically relevant [31, 32].

4.8.4 | Differences in Sample Size for Trials of Approved Treatments

Sample sizes of both early and late-stage trials of approved therapies can differ dramatically, influenced by the magnitude of treatment effects based on approved therapies or other historical data.

When placebo-adjusted treatment effects are modest, placebo responses are significant, and event rates are low, a relatively large number of participants are required to detect a clinically meaningful effect at a statistically significant level. In contrast, historically large effects from efficacious drugs and a low or absent placebo response contribute to a relatively low number of patients needed to show a predetermined clinically meaningful effect.

Investigator-sponsored studies are also conducted in settings such as rare or incurable cancers, where even one or two responders may be enough to justify further study. These initial studies typically have no control group, small numbers of patients, and no sample size justification or formal statistical analysis plan.

4.8.5 | Sensitivity Analyses: Sample Size Calculations

Small variations in estimated treatment effects can greatly affect the estimated number of participants/treatment arm. Knowledge of what is considered the minimal clinically important difference (MCID) for treatment effects on a disease, which is the smallest change in a treatment outcome that a patient would identify as important, can greatly aid in sample size calculations. Sample size determination for overly optimistic treatment effects can result in an inability to detect a lower but clinically meaningful effect. Conversely, estimations for treatment effects that are too small can expose participants to unnecessary risk to detect an effect at a significance level that can be clinically insignificant.

Determining the MCID is critical as other critical values for sample size calculation are conventionally set (e.g., alpha, 5%, power, 80% or 90%). Criteria for MCIDs are usually determined by consensus of experts, meta-analyses showing correlations with functional improvement (including quality of life), patient-reported outcomes, or regulatory precedent. Additional methods used to determine the difference to be used in sample size calculations include subjective assessment by patients of their change in health status or a larger difference than the minimum that would need to be observed to justify a change in practice [33, 34].

When the MCID is established, small changes in assumptions can affect sample size estimations. It is important to recognize that estimated treatment effects can be based on different criteria ranging from practitioner judgment, patient reports, consensus from studies correlating an outcome with quality of life scales, regulatory approval, and pivotal studies showing placebo-adjusted treatment effects. If known, using the MCID as the basis for estimated treatment effects for sample size calculation is optimal. Novel approaches to determine the clinical meaningfulness of a treatment by combining multiple endpoints in a prespecified analysis have also been proposed [35].

4.8.6 | Guidelines for Power and Sample Size Calculations

- Power of 80%–90% and significance level (alpha) of 0.05 (one-sided or two-sided) are traditionally used for Phase 1b or 2a studies.
- Clinical meaningfulness or scientific conclusion should not be based on a P value being "significant" per se [36, 37]. The actual size and practical relevance of the observed effect, often measured by effect size and confidence interval, should also be taken into account when interpreting clinical significance. In addition, small but statistically significant effects may be misinterpreted as clinically important.
- *p* value should be interpreted as a continuum, where a lower value gives increased strength of evidence of a real effect [36, 37].
- The use of confidence interval in place of or in addition to P values can provide critical information about the probability that a true value will fall within a set of values, and the direction and strength of the demonstrated effect. This information can help the evaluation of the clinical relevance of the study findings.
- In some cases, increased alpha (e.g., 10%) may be appropriate in pilot studies with small sample sizes where the risk of not identifying ineffective therapy (Type I error) may be less important for future study than identifying a potentially effective therapy (e.g., power, 90%).
- Estimated treatment effects and variability can be based on drugs approved for the proposed indication or other rational approaches (observed differences from studies and use of predicted CI).

4.8.7 | Resources for Sample Size Estimations

There are many excellent free online programs for sample size calculations, including University of California, San Francisco Sample Size Net [38] and Southwest Oncology Group Statistical Tools [39]. In addition, sample size calculators can also be found on Apple and Android (Google Play) app stores.

4.8.8 | Power and Detectable Differences/ Final Precision

Goodman and Berlin [30] have emphasized problems with studies that have low power or an unrealistically high effect size used in sample size calculations. The inability after the trial to distinguish between clinically important and unimportant results, which will be expressed in the form of wide confidence intervals, can be the result. A confidence interval can be a valuable way to describe probability as a set of true but unknown differences that are statistically compatible with the observed difference [29]. Determining realistic and meaningful estimates of the treatment effect and confidence interval can serve as an optimal basis for determining sample size.

Goodman and Berlin [30] note that precision (defined as the width of the 95% confidence interval) and power are linked to sample size and are mathematically related. As a result, the approximate size of the confidence interval after the experiment can be predicted before the experiment, and this prediction should supplement traditional sample size calculations and ideally be reported in protocols.

4.8.9 | Statistical and Analytical Plans That Include Handling of Missing Data, Efficacy Analyses, Safety Analyses, and Interim Analyses/Data Monitoring Committee

This section can be either straightforward or challenging depending upon the basic approach of how safety, efficacy, and PK data will be analyzed, including the approach for missing data. The goal for any study is to have all samples drawn and evaluations conducted at the designated time points with no missing data. Methods for dealing with missing data vary by therapeutic area and endpoint and require careful consideration of whether the missing data is a function of efficacy or safety. In early phase clinical research, there are often multiple endpoints that are important for data to be repeatedly examined. These conditions may require a discussion of adjustments to statistical testing or confidence intervals, or an explanation of why such adjustments are not needed. The use of a Data Monitoring Committee (also known as a Data and Safety Monitoring Board, Data and Safety Monitoring Committee, or an Independent Data Monitoring Committee) to conduct a statistical review of accumulating data from the ongoing study is less common in small, early-stage investigator-sponsored trials [40].

4.9 | Investigator-Sponsor Trial Recommendations

As noted above, investigator-sponsored clinical trials with a limited number of participants may not be able to detect a true and clinically relevant treatment effect with high probability (are "underpowered"). At a given sample size, factors that contribute to lowered power to detect a given clinically relevant effect, and can lead to overestimations or underestimations of the true effects of an intervention, include high placebo rates, low event rates, modest treatment effects, and benefit from slowing of worsening (as opposed to improvement per se). Despite this concern, most investigator-sponsored studies usually have sufficient ability to evaluate for preliminary evidence of safety, efficacy, and variability, including the presence of a biological signal.

In resource-limited investigator-sponsored trials, identifying clinically important effects at study completion can be helped prestudy [41] by:

- Incorporating multiple biologically-related endpoints.
- Use of biomarkers that are biologically plausible.
- Use of a valid or reasonably likely valid surrogate endpoint that may be highly affected by treatment.
- Using relevant assessments with low placebo response rates.
- Specifying endpoints that occur frequently.
- Enrolling patients at higher risk.
- Prolonging follow-up.

Sample size calculations can also often be powered at times based on the precision of estimates and should involve consultations with a biostatistician and clinical trialist.

4.10 | Study Population

This section includes:

- Brief summary of population (healthy vs. disease, severity).
- Study setting (e.g., community clinic, academic center) and countries
- Entry criteria for study participation.
- specify both inclusion and exclusion criteria.
- Criteria for discontinuing study drug and study participation.
- Strategies for participant recruitment and retention

Entry criteria should attempt to maximize signal detection. This can be achieved by enrolling best responders, which is preferable in many cases, and excluding those who are too well and may have little room for improvement (ceiling effect) or may be too refractory to show benefit (floor effect).

Common problems include:

Vague entry criteria that may be interpreted differently by study staff:

- *Example*: exclusion of "those who smoke."
- *Better*: exclusion of "Current smokers or users of e-cigarettes or nicotine replacement products unable to avoid using these products from at least 48 h before check-in through the final study visit."

Entry criterion so close to the definition of success that it may be met by natural variation independent of active drug or placebo.

- *Example*: platelet count > 90 K allowed in a study of thrombocytopenia when a count of 100 K or greater represents the success of the study drug.

Unspecified duration of background stable therapy before study entry. Beginning already-approved treatments too close to the study entry onset date can result in a confound of benefits from this treatment inaccurately being attributed to placebo or investigational drug.

There is an important balance between homogeneity vs. heterogeneity in determining study participant entry criteria. There is a substantial gain in statistical power by focusing intervention on a homogenous patient population most likely to respond. In contrast, there is also a need to generalize findings to a broader group of heterogeneous patients who might benefit from treatment. There is no perfect answer and the decision is always a tradeoff; however, protocols of pilot studies or investigator-sponsored studies that are limited in resources and oriented toward detection of a meaningful signal should err on the side of selecting a more homogenous population most likely to respond given the limited sample size and power. Even in late-stage trials, recruitment of responders allows for the best detection of a treatment effect [42].

All study participants have a right to discontinue study drug or withdraw from the study for any reason at any time. This section summarizes the criteria for both a participant's early discontinuation of study drug and withdrawal from the study, including whether replacement of the participant with another participant can occur. In many situations, it is advantageous for the integrity of the final data analysis for a participant with early discontinuation of study drug to complete remaining scheduled visits and procedures, provided consent to do so has not been withdrawn. Continued follow-up also aids safety assessment—for example, helps determine if an adverse event resolves following discontinuation of study drug.

Finally, general strategies for participant recruitment and retention may be included in this section, or a separate one. This may include how potential participants will be identified, the types of recruitment strategies planned, and the anticipated number of sites and participants to be enrolled if the trial involves different countries.

4.11 | Treatments

This section summarizes:

- Method of Assignment to Treatment (e.g., Randomization).
- Materials and Supplies.
  - Formulation, packaging, and labeling.
  - Storage and handling.
- Dosage and Administration.
- Blinding (including method of emergency unbinding).
- Concomitant Therapy.
- Study drug and study discontinuation criteria.

A common problem is not including language that the study drug should be discontinued if a prohibited medication that would have precluded study entry is taken during the study.

4.12 | Adverse Events, Serious Adverse Events, and Reporting Requests

The adverse event section of any clinical protocol should contain similar language, in part because definitions of an adverse event (AE), a serious adverse event (SAE), as well as reporting requirements to regulatory authorities are standardized by ICH and mandated by federal regulations [43].

It is helpful in investigator-sponsored protocols to include language that an AE includes a laboratory value that is judged to be clinically significant by the investigator, which aids in case report form construction and collection of safety information. In addition, important events that occur between the signing of the informed consent and initiation of study drug, and are judged to be related to a study procedure, should be recorded. Common problems in this section include, in studies with more than one site, contact information for reporting SAEs to the sponsor, typically within 24 h of becoming aware of the event, and criteria and timing for notifying the IRB or EC of both SAEs and "reportable events."

4.13 | Quality Control and Quality Assurance

The purpose of an IRB or EC, which is comprised of at least five members (including at least one not affiliated with the institution) with varying backgrounds, is to protect human participants in a research study from harm. Consistent with ICH Good Clinical Practice E6 guidance [44], this section should contain standardized language that states that the IRB or EC is formally designated to approve, monitor, and review biomedical research involving humans, protect human subjects from harm, and is empowered by federal regulations to approve, require modification before approval, or disapprove research based on scientific, ethical, and regulatory considerations. In addition, this section states that the investigator allows IRB or EC review and regulatory inspection of trial-related documents and procedures, and that the investigator is responsible for ensuring that data are generated, documented, and reported in compliance with the protocol, Good Clinical Practice, and all regulations.

4.14 | Administrative, Ethical, and Regulatory Considerations

The Declaration of Helsinki [45] is a set of ethical principles for medical research that involves human participants and is widely followed globally. This section of the protocol, in most cases, contains standardized language that each potential participant must be adequately informed of risks, and additional protections should follow the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule (termed differently in regulatory areas outside the United States). This section states that a participant has the right to discontinue a study at any time for any reason without prejudice to future medical care, specifies that appropriate compensation and treatment must be provided for those harmed as a result of study participation, and affirms that the rights, safety, and well-being of subjects prevail over the interests of science and society. This section also summarizes the responsibilities of the sponsor-investigator to supply information to the investigative site's IRB and conduct the study in accordance with the protocol and ethical principles in the Declaration of Helsinki and all applicable guidelines, laws, and regulations. Finally, aspects of data availability, clinical monitoring, study finances, declarations of interest, trial registration, and publication policy are typically summarized in this section.

4.14.1 | Ethical Review

This section summarizes the responsibility of the investigator to present the risks and benefits to the participant in simple terms using the informed consent document (or assent obtained from any minor participant, when applicable). A study participant has the right to discontinue a study at any time for any reason without prejudice to future medical care. In addition, appropriate compensation and treatment must occur for subjects harmed as a result of study participation.

The requirements of the sponsor-investigator to supply specific materials to the investigative site's IRB (e.g., protocol and amendments, required safety reports, etc.) and what documentation to retain, if applicable, are also summarized here.

4.14.2 | Regulatory Considerations

The sponsor-investigator must acknowledge that the conduct will be conducted following the protocol and ethical principles in the Declaration of Helsinki and all applicable guidelines, laws, and regulations. Further, the sponsor-investigator will initiate changes to the protocol as necessary (except for those to eliminate an immediate hazard to a study participant) and seek IRB approval before implementing. This section summarizes the responsibility of the sponsor-investigator to enroll participants who have met protocol eligibility and report violations to the local IRB following their policies. The sponsor-investigator may terminate the study at any time.

4.14.3 | Study Documentation, Privacy, and Records Retention

This section specifies the duration for which government agency regulations and directives require study data and related records to be kept after the study completion, with additional considerations that an IRB or EC might require longer retention periods. Additional language related to what type of data and where it will be stored, site monitoring to ensure data integrity and rights of participants, and records containing participant information must be handled following the requirements of privacy rules are also typically included. Furthermore, case report forms and other study documents should be completed following the instructions provided by the sponsor-investigator, including the instructions for the coding to protect participant identities.

4.14.4 | Declaration of Interests

This section refers to a statement where investigator-sponsors involved in a research study disclose any potential conflicts of interest, including financial or non-financial relationships that could influence their objectivity (or give the appearance of a conflict) in the study design, execution, or interpretation of results.

4.14.5 | Data Availability Statement

A data availability statement outlines how the data collected during the trial will be accessible to researchers and the public, including where the data will be stored, under what conditions it can be accessed, and any limitations regarding participant privacy or ethical considerations that may restrict full data sharing. It details the plan for making clinical trial data available after the study is completed. In place of this section, this information can also be specified in the Study Documentation, Privacy, and Records Retention or Publication Statements.

4.14.6 | Study Finances

Sources and types of financial, material, and other support are included here.

4.14.7 | Publication Statement

This section outlines the sponsor's plan for disseminating the study findings, including where and how the sponsor intends to publish the results. Language typically specifies that (in a multicenter trial) neither the complete nor any part of the results of the study, nor any of the information provided by the sponsor to perform the study, will be published or passed on to any third party without the consent of the study sponsor. In addition, any investigator involved with this study is obligated to provide the sponsor with complete test results and all data derived from the study. Optimally, this section also should state that design elements of the protocol will be posted in a publicly accessible database (such as ClinicalTrials.gov) and that, upon study completion and finalization of the study report, the results of the trial will either be published or posted in a publicly accessible database of clinical trial results. Similar to a Data Availability Statement, this section usually details the plan for making clinical trial data available after the study is completed, which is often that deidentified patient data that support the findings of this study will be available from the corresponding author upon reasonable request.

5 | Grammar, Spelling, References

Attention to writing is mandatory for clarity of communication. Besides the necessary critical review by others, there are many excellent sources for scientific writing guidance [46–48]. In addition, free online grammar and spelling programs, including ones that are built into the word processing program used, are available.

5.1 | Critical Review by Others Gives Different (and Valuable) Perspectives

Critical review and thoughtful criticism by others are necessary for both content and clarity of communication.

6 | Conclusion

There remains a critical need to develop therapies for many diseases that currently lack treatments [49]. This need is underscored by an appreciation that drug development, in general, has a high attrition rate, rising costs, and increased delays in bringing treatments to market. Investigator-sponsored clinical trials in academic environments can be part of known innovation that leads to significant impacts [50, 51]. These studies are also more likely to explore the possible efficacy of an intervention in a more severe group of patients with comorbidities and a higher risk of death, target populations with high unmet medical needs while less affected by commercial objectives, and involve drug repurposing that facilitates development because the initial safety and efficacy testing in the approved indication

has already been done [52, 53]. Such trials, however, are typically challenged by significant limitations in operational and financial resources, including the number of study participants. Despite these challenges, close attention to protocol development can greatly maximize the scientific and ethical rigor of these investigations and detect relevant outcomes that can lead to larger, well-controlled trials.

Acknowledgments

I thank Drs. Paul Flyer, David Goodkin, Curt Scribner, and Lee-Jen Wei for their helpful discussions and critical review. Linda MacKeen provided invaluable editorial support.

Conflicts of Interest

The author declares no conflicts of interest.

References

1. A.-W. Chan, J. M. Tetzlaff, D. G. Altman, et al., "SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials," *Annals of Internal Medicine* 158 (2013): 200–207.

2. U.S. Food and Drug Administration, Providing Clinical Evidence of Effectiveness for Human Drugs and Biological Products, 1998, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/providing-clinical-evidence-effectiveness-human-drug-and-biological-products.

3. A. V. Morant, V. Jagalski, and H. T. Vestergaard, "Characteristics of Single Pivotal Trials Supporting Regulatory Approvals of Novel Non-Orphan, Non-Oncology Drugs in the European Union and United States From 2012–2016," *Clinical and Translational Science* 12 (2019): 361–370.

4. S. Hopewell, I. Boutron, A. W. Chan, et al., "An Update to SPIRIT and CONSORT Reporting Guidelines to Enhance Transparency in Randomized Trials," *Nature Medicine* 28 (2022): 1740–1743.

5. C. Begg, M. Cho, S. Eastwood, et al., "Improving the Quality of Reporting of Randomized Controlled Trials. The CONSORT Statement," *JAMA* 276, no. 8 (1996): 637–639, https://doi.org/10.1001/jama.276.8.637.

6. D. Moher, K. F. Schulz, D. Altman, and CONSORT Group, "The CONSORT Statement: Revised Recommendations for Improving the Quality of Reports of Parallel-Group Randomized Trials," *JAMA* 285 (2001): 1987–1991.

7. N. J. Butcher, A. Monsour, E. J. Mew, et al., "Guidelines for Reporting Outcomes in Trial Protocols: The SPIRIT-Outcomes 2022 Extension," *JAMA* 328, no. 23 (2022): 2345–2356, https://doi.org/10.1001/jama.2022.21243.

8. A. M. Manyara, P. Davies, D. Stewart, et al., "Reporting of Surrogate Endpoints in Randomised Controlled Trial Protocols (SPIRIT-Surrogate): Extension Checklist With Explanation and Elaboration," *BMJ* 386 (2024): e078525.

9. Stanford SPARK, Stanford SPARK Clinical Trial Protocol Template, https://sparkmed.stanford.edu/spark-scholars/resources.

10. University of California, San Francisco, Clinical Trial Protocol Development, https://hub.ucsf.edu/protocol-development.

11. University of Pennsylvania, Protocol Template Clinical Trial, https://irb.upenn.edu/homepage/formsandtemplates.

12. National Institutes of Health-Food and Drug Administration, Phase 2 and 3 IND/IDE Clinical Trial Protocol Template, https://grants.nih.gov/policy-and-compliance/policy-topics/clinical-trials/protocol-template.

13. ICH, M11 Template: Clinical Electronic Structured Harmonised Protocol, https://www.ema.europa.eu/en/ich-m11-guideline-clinical-study-protocol-template-technical-specifications-scientific-guideline.

14. U.S. Food and Drug Administration, Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Guidance for Industry, 2009, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims.

15. C. N. Sandler and M. E. McDonnell, "The Role of Hemoglobin A1c in the Assessment of Diabetes and Cardiovascular Risk," *Cleveland Clinic Journal of Medicine* 83, no. 5 Suppl 1 (2016): S4–S10.

16. H. J. Warraich, B. A. King, W. M. Compton, et al., "Opportunities for Innovation in Smoking Cessation Therapies: A Perspective From the National Institutes of Health and U.S. Food and Drug Administration," *Annals of Internal Medicine* 178 (2025): 122–125.

17. R. H. Dworkin, S. R. Evans, O. Mbowe, and M. P. McDermott, "Essential Statistical Principles of Clinical Trials of Pain Treatments," *Pain Reports* 6, no. 1 (2020): e863, https://doi.org/10.1097/PR9.0000000000000863.

18. S. M. Snappin and Q. Jiang, "Responder Analyses and the Assessment of a Clinically Relevant Treatment Effect," *Trials* 8, no. 31 (2007).

19. Biomarkers Definitions Working Group, "Biomarkers and Surrogate Endpoints: Preferred Definitions and Conceptual Framework," *Clinical Pharmacology & Therapeutics* 69, no. 3 (2001): 89–95, https://doi.org/10.1067/mcp.2001.113989.

20. T. R. Fleming and J. H. Powers, "Biomarkers and Surrogate Endpoints in Clinical Trials," *Statistics in Medicine* 31 (2012): 2973–2984.

21. A. M. Manyara, P. Davies, D. Stewart, et al., "Definitions, Acceptability, Limitations, and Guidance in the Use and Reporting of Surrogate End Points in Trials: A Scoping Review," *Journal of Clinical Epidemiology* 160 (2023): 83–99.

22. A. M. Manyara, O. Ciani, R. S. Taylor, SPIRIT-Surrogate and CONSORT-Surrogate Project team, and Consensus Group, "Transparent and Complete Reporting of Surrogate Endpoints in Trials: The Spirit-Surrogate and Consort-Surrogate Extensions," *Lancet* 404, no. 10450 (2024): 322–324, https://doi.org/10.1016/S0140-6736(24)01306-0.

23. ICH, Guidance for Industry. E10 Choice of Control Group and Related Issues in Clinical Trials, www.fda.gov/media/71349/download.

24. R. Temple and S. S. Ellenberg, "Placebo-Controlled Trials and Active-Control Trials in the Evaluation of New Treatments. Part 1: Ethical and Scientific Issues," *Annals of Internal Medicine* 133 (2000): 455–463.

25. Council for International Organizations of Medical Sciences (CIOMS), *International Ethical Guidelines for Health-Related Research Involving Humans*, 4th ed. (WHO Press, 2016).

26. L. Frumkin, "The Pharmacological Treatment of Pulmonary Arterial Hypertension," *Pharmacological Reviews* 64, no. 3 (2012): 583–620, https://doi.org/10.1124/pr.111.005587.

27. N. Galié, D. Badesch, R. Oudiz, et al., "Ambrisentan Therapy for Pulmonary Arterial Hypertension," *Journal of the American College of Cardiology* 46 (2005): 529–535.

28. O. Sander, B. Magnusson, I. Ludwig, et al., "A Framework to Guide Dose and Regimen Strategy for Clinical Drug Development," *CPT: Pharmacometrics & Systems Pharmacology* 10 (2021): 1276–1280.

29. S. Greenland, S. J. Senn, K. J. Rothman, et al., "Statistical Tests, P Values, Confidence Intervals, and Power: A Guide to Misinterpretations," *European Journal of Epidemiology* 31 (2016): 337–350.

30. S. N. Goodman and J. A. Berlin, "The Use of Predicted Confidence Intervals When Planning Experiments and the Misuse of Power When Interpreting Results," *Annals of Internal Medicine* 121 (1994): 200–206.

31. K. J. Rothman and S. Greenland, "Planning Study Size Based on Precision Rather Than Power," *Epidemiology* 29 (2018): 599–603.

32. E. van Zwet, A. Gelman, S. Greenland, G. Imbens, S. Schwab, and S. N. Goodman, "A New Look at P Values for Randomized Clinical Trials," *New England Journal of Medicine Evidence* 3 (2024): EVIDoa2300003.

33. M. L. Bell, "New Guidance to Improve Sample Size Calculations for Trials: Eliciting the Target Difference," *Trials* 19 (2018): 605.

34. J. A. Cook, S. A. Julious, W. Sones, et al., "DELTA2 Guidance on Choosing the Target Difference and Undertaking and Reporting the Sample Size Calculation for a Randomised Controlled Trial," *Trials* 19 (2018): 606.

35. D. Li, C. M. McDonald, G. L. Elfring, et al., "Assessment of Treatment Effect With Multiple Outcomes in 2 Clinical Trials of Patients With Duchenne Muscular Dystrophy," *JAMA Network Open* 3, no. 2 (2020): e1921306, https://doi.org/10.1001/jamanetworkopen.2019.21306.

36. S. Goodman, "A Dirty Dozen: Twelve P-Value Misconceptions," *Seminars in Hematology* 45 (2008): 135–140.

37. B. J. Harvey and T. A. Lang, "Hypothesis Testing, Study Power, and Sample Size," *Chest* 138 (2010): 734–737.

38. M. A. Kohn and J. Senyak, Sample Size Calculators, https://www.sample-size.net.

39. Southwest Oncology Group, CRAB SWOG Statistical Tools Calculators, https://stattools.crab.org.

40. U.S. Food and Drug Administration, Establishment and Operation of Clinical Trial Data Monitoring Committees. Guidance for Clinical Trial Sponsors, 2006, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/establishment-and-operation-clinical-trial-data-monitoring-committees.

41. L. Frumkin and T. McCluskey, "Clinical Trial Design," in *A Practical Guide to Drug Development in Academia: The SPARK Approach*, 2nd ed., ed. D. Mochly-Rosen and K. Grimes (Spinger, 2023), 164–183.

42. U.S. Food and Drug Administration, Enrichment Strategies for Clinical Trials to Support Determination of Effectiveness of Human Drugs and Biological Products. Guidance for Industry, 2019, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/enrichment-strategies-clinical-trials-support-approval-human-drugs-and-biological-products.

43. U.S. Food and Drug Administration, Safety Reporting Requirements for INDs and BA/BE Studies. Guidance for Industry and Investigators, 2012, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/safety-reporting-requirements-inds-investigational-new-drug-applications-and-babe.

44. ICH (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use), Guideline for Good Clinical Practice E6 (R3), 2025, https://www.ich.org/page/efficacy-guidelines#6-2.

45. World Medical Association, "World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Participants," *JAMA* 333, no. 1 (2025): 71–74, https://doi.org/10.1001/jama.2024.21972.

46. R. B. Daroff, L. P. Rowland, and C. Scism, "Suggestions to Authors," *Neurology* 38, no. 10 (1988): 1657–1658, https://doi.org/10.1212/wnl.38.10.1657.

47. JAMA Network Editors, *AMA Manual of Style: A Guide for Authors and Editors*, 11th ed. (Oxford University Press, 2020).

48. University of Chicago Press Editorial Staff, *The Chicago Manual of Style*, 18th ed. (University of Chicago Press, 2024).

49. C. J. L. Murray, "The Global Burden of Disease Study at 30 Years," *Nature Medicine* 28 (2022): 2019–2026.

50. T. Takebe, R. Imai, and S. Ono, "The Current Status of Drug Discovery and Development as Originated in United States Academia: The Influence of Industrial and Academic Collaboration on Drug Discovery and Development," *Clinical and Translational Science* 11 (2018): 597–606.

51. M. S. Kinch, C. Horn, Z. Kraft, and T. Schwartz, "Expanding Roles for Academic Entrepreneurship in Drug Discovery," *Drug Discovery Today* 25 (2020): 1905–1909.

52. T. I. Oprea, J. E. Bauman, C. G. Bologa, et al., "Drug Repurposing From an Academic Perspective," *Drug Discovery Today. Therapeutic Strategies* 8 (2011): 61–69.

53. P.-F. Laterre and B. François, "Strengths and Limitations of Industry vs. Academic Randomized Controlled Trials," *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases* 21 (2015): 906–909.