HHS Public Access

Author manuscript

J Investig Med. Author manuscript; available in PMC 2015 April 30.

Published in final edited form as:

J Investig Med. 2009 December ; 57(8): 818–824. doi:10.231/JIM.0b013e3181c2996c.

Bridging Clinical Investigators and Statisticians: Writing the Statistical Methodology for a Research Proposal

Beverley Adams-Huet, MS and **Chul Ahn, PhD**

The University of Texas Southwestern Medical Center at Dallas, Department of Clinical Sciences, Dallas, TX

Keywords

statistical methods; statistician; clinical research; research proposal; protocol; study design; biostatistics

INTRODUCTION

Clinical research is judged to be valid not by the results but how it is designed and conducted. The cliché of ‘do it right or do it over’ is particularly apt in clinical research.

One of the questions a clinical investigator frequently asks in planning clinical research is “Do I need a statistician as part of my clinical research team?” The answer is “Yes!” since a statistician can help to optimize design, analysis and interpretation of results, and drawing conclusions. When developing a clinical research proposal, how early in the process should the clinical investigator contact the statistician? Answer - it is never too early. Statistics cannot rescue a poorly designed protocol after the study has begun. A statistician can be a valuable member of a clinical research team and often serves as a co-investigator. Large multicenter projects such as Phase III randomized clinical trials for drug approval by a regulatory agency nearly always have a statistician (or several) on their team. However, smaller, typically single center studies may also require rigorous statistical methodology in design and analysis. These studies are often devised by young clinical investigators launching their clinical research career who may have not collaborated with a statistician. Many clinical investigators are familiar with the statistical role in the analysis of research data1, but researchers may not be as aware of the role of a statistician in designing clinical research and developing the study protocol. In this paper we discuss topics and situations that clinical investigators and statisticians commonly encounter while planning a research study and writing the statistical methods section. We stress the importance of having the statistical methodology planned well in advance of conducting the clinical research study. Working in conjunction with a statistician can also be a key training opportunity for the clinical investigator beginning a clinical research career.

Corresponding Author: Beverley Adams-Huet, MS, Department of Clinical Sciences, Division of Biostatistics University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd. Dallas, TX 75390-9066. Tel: 214-648-2788; fax 214-648-3934 Beverley.huet@utsouthwestern.edu.

GETTING STARTED ON THE STATISTICAL ANALYSIS PLAN

Why work with a statistician?

The study design, sample size, and statistical analysis must be able to properly evaluate the research hypothesis set forth by the clinical investigator. Otherwise, the consequences of a poorly developed statistical approach may result in a failure to obtain extramural funding and result in a flawed clinical study that cannot adequately test the desired hypotheses. Statisticians provide design advice and develop the statistical methods that best correspond to the research hypothesis. For the planning of a clinical study, a statistician can provide valuable information on key design points as summarized in Table 1. The statistician can discuss with the clinical investigators questions such as: Is the design valid? Overly ambitious? Will the data be analyzable?

Very early in the planning stages, it is important to send the statistician a draft of the proposal. Any protocol changes may affect the required sample size and analysis plans so it is important to meet with the statistician throughout the planning stages and later if modifications have been made to the study design. Before the statistical section can be developed, what information does the statistician need? Questions from a statistician concerning design, power and sample size, and analysis may include:

- What is the research hypothesis?
- What is the type of study design?
- What is the most important measurement (primary outcome variable)?
- What is the type of variable and unit of measurement?
- What is a clinically meaningful difference for the primary outcome?
- How many subjects can be recruited or observed within a study period? How many groups or treatment arms are to be included in your design?
- Will there be an equal number of participants or observations in each group? i.e., what is the allocation ratio?
- How many total evaluations and measurements?
- For repeated measurements, what is the measurement interval?

You are not expected to have all the answers at your first meeting and ongoing conversations with the statistician can serve to develop these ideas. Eventually, the answers to these questions comprise the justification for the design selected, provide the basis for the sample size estimate, and drive the choice of statistical analysis. A brief consultation with a statistician will not be adequate to address these issues. The interaction with a statistician to construct the statistical section is not usually one meeting, email, or phone call. It is a process that will help you think through the design of your study. This is also an excellent opportunity to ask questions and enhance your statistical education. Additionally, the exchange of ideas is beneficial to the statistician who will better appreciate the clinical research question. The discussions with a statistician could lead to changes in study design, such as proposing a smaller, more focused study design to collect preliminary data.

A general outline of the statistical methods section is shown in Table 2. There may be deviations from this format depending on the particular study design. The statistical write-up is rarely less than one page and may total several pages. Although some clinical investigators trained in statistics do prepare this section, more commonly the statistician constructs and writes up the statistical methods section for grants and protocols in close collaboration with the investigators. However, it is important that clinical investigators develop a conceptual understanding of the proposed statistical methodology. Take advantage of any study design and biostatistics classes offered at your institutions to make statistical collaborations more fruitful.

STUDY DESIGN

Type of Design

Before the statistical section can progress, the study design must be known. Study designs that are commonly used in clinical research include case-control, cohort, randomized controlled design, crossover, and factorial designs. A randomized controlled trial has many features but most commonly incorporates what is called a parallel group design where individuals are randomly assigned to a particular treatment or intervention group. In a crossover study, the subject participates in more than one study intervention phase, ideally studied in a random sequence, such as comparing triglyceride responses within the same individual on a low fat versus a high fat diet.

Sampling

How do we select participants for the study? There are many types of sampling procedures, the basis of which is to avoid or reduce bias. Bias can be defined as ”a systematic tendency to produce an outcome that differs from the underlying truth”.2 Although true randomness is the goal of a sampling, it is generally not achievable. The study subjects are not usually selected at random to participate in clinical research. Instead, in most clinical trials, the “random” element in randomization is that the consented subjects are assigned by chance to a particular treatment or intervention. The clinical inclusion and exclusion criteria coupled with informed consent will determine who will be the study participants and, ultimately, to what population the study results will be generalizable.

Sample size

With the study design and the make-up of the study sample determined, the sample size estimates can be obtained. Fundamental to estimating sample size are the concepts of statistical hypothesis testing, type I error, type II error, and power (Table 3). In planning clinical research it is necessary to determine the number of subjects to be required so that the study achieves sufficient statistical power to detect the hypothesized effect. If the reader is not familiar with the concept of statistical hypothesis testing, introductory biostatistics texts and many web sites cover this topic. Briefly, in trials to demonstrate improved efficacy of a new treatment over placebo/standard treatments, the null hypothesis is that there is no difference between treatments and the alternative hypothesis is that there is a treatment difference. The research hypothesis usually corresponds to the alternative hypothesis which represents a minimal meaningful difference in clinical outcomes. Statistically, we either 1)

reject the null hypothesis in favor of the alternative hypothesis or 2) we fail to reject the null hypothesis.

Typically, the sample size is computed to provide a fixed level of power under a specified alternative hypothesis. Power is an important consideration for several reasons. Low power can cause a true difference in clinical outcomes between study groups to go undetected. However, too much power may yield statistically significant results that are not meaningfully different to clinicians. The probability of Type I error (α) of 0.05 (two-sided) and power of 0.80 and 0.90 have been widely used for the sample size estimation in clinical trials. The sample size estimate will also allow the estimation of the total cost of the proposed study.

A clinical trial that is conducted without attention to sample size or power information takes the risks of either failing to detect clinically meaningful differences (Type II error) due to not enough subjects or taking an unnecessarily excessive number of samples for a study. Both cases fail to adhere to the Ethical Guidelines of the American Statistical Association which says “Avoid the use of excessive or inadequate number of research subjects by making informed recommendations for study size”.3

What information is needed to calculate power and sample size?

The components that most sample size programs require for input include:

- Choose Type I error (alpha)
- Choose Power
- Choose clinical outcome variable and effect size (difference between means, proportions, survival times, regression parameters)
- Variation estimate
- Allocation ratio

Clinical outcome measures—Clearly describe the clinical outcomes that will be analyzed to the statistician. The variable type (Table 4) and distribution of the primary outcome measurement must be defined before sample size and power calculations can proceed. The sample size estimates are mainly needed for the primary outcome. However, providing power estimates for secondary outcomes is often helpful to reviewers.

Effect size—As an example, suppose a parallel group study is being designed to compare systolic blood pressure between two treatments and the investigators want to be able to detect a mean 10 mm Hg difference between groups. This 10 mm Hg difference is referred to as the effect size, detectable difference, or minimal expected difference.

How is the effect size determined?—Choose an effect size that is based on clinical knowledge of the primary endpoint. A sample size that ‘worked’ in a published paper is no guarantee of success in a different setting. The selected effect size is unique to your study intervention, the specific type of participants in your study sample, and perhaps an aspect of the outcome measurement that is unique to your clinic or laboratory.4

The investigator and statistician examine the literature, the investigator's own past research, or a combination of the above to determine a study effect size. To investigate the difference in mean blood pressure between two treatments, the effect size options might be 2 mm Hg, 6 mm Hg, 10 mm Hg or 20 mm Hg. Which of these differences do you need to have the ability to detect? This is a clinical question, not a statistical question. Effect size is a measure of the magnitude of the treatment effect and represents a clinically or biologically important difference. Choosing a 20 mm Hg effect size yields a smaller sample size than a 10 mm Hg effect size since it is easier to statistically detect the larger difference. However, an effect size of 10 mm Hg or smaller magnitude may be more a realistic treatment effect and less likely to result in a flawed or wasted study.

Variation estimates for sample size calculations—In addition to effect size, we may need to estimate how much the outcome varies from person to person. For continuous variables, the variation estimate is often in the form of a standard deviation. If the hypothesized difference in systolic blood pressure is an effect size of 10 mm Hg, a study with a blood pressure standard deviation of 22 mm Hg will have lower power than a study where the standard deviation is 14 mm Hg. For a continuous outcome such as blood pressure, a measure of the variation is another part of the formula needed to compute the sample size. An estimate of variation can be derived from a literature search or from the investigator's preliminary data. Obtaining this information can be a challenge for both the clinical investigator and statistician.

Table 5 shows sample sizes scenarios for detecting differences in blood pressure when comparing two treatments based on a t-test. A standard deviation of 14 mm Hg is chosen to estimate the variation. Sample sizes are calculated for power of 0.80 and 0.90 at the two-sided 0.05 significance level. Notice that the smaller effect sizes require a larger sample size and that the sample size increases as the power increases from 0.80 to 0.90.

Determining a reasonable and affordable sample size estimate is a team effort. There are practical issues such as budgets or recruitment limitations that may come into play. A too large sample size could preclude the ability to conduct the research. The research team will assess scenarios with varying detectable differences and power as seen in Table 5 (calculations performed using PS power5 available at the website <http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize>). Typically a scenario can be worked out which is both clinically and statistically viable.

The elements of sample size calculations presented here pertain to relatively simple designs. Cluster samples or family data need special statistical adjustments. For a longitudinal or repeated measures design, the correlation between the repeated measurements is incorporated into the sample size calculations.6, 7

Do all studies need sample size and power estimates?

Pilot studies—Pilot studies may not need a power analysis since they are more about testing the protocol than testing a hypothesis.8 Sometimes there are no preliminary data and thus pilot data must be obtained to provide estimates for designing for a more definitive

study. However, calling a study a pilot study to avoid power analyses and to keep the sample small is misrepresentation.8

Precision—Sample size calculations are necessary when the study goal is precision instead of power. The goal may be to describe the precision of a proportion or mean or other statistic that is to be estimated from our sample. Precision in this context is based upon finding a suitably narrow confidence interval for the statistic of interest, such that the lower and upper limits of the confidence interval include a clinically meaningful range of values. We may want to know how many subjects are required to be 95% confident that an interval contains the true, but unknown, value. For example, how many subjects are needed for 10% precision if we expect a 30% allele prevalence in a genetic study? Instead of power, we estimate the sample size for the desired precision based on a single proportion of 0.30 and summarize by stating "With 80 subjects, the precision for a 30% allele prevalence rate is approximately 10% (95% confidence interval: 21% to 40%)." If greater precision is desirable then the sample size is increased accordingly.

Accounting for attrition

Withdrawal and dropout are unwelcome realities of clinical research. Missing data in clinical trials or repeated measurement studies are inevitable. Consider missing data issues when designing, planning and conducting studies to minimize missing data impact. Sample size estimates are finalized by adjusting for attrition based upon the anticipated number of dropouts.

Randomization plan

Random allocation of subjects to study groups is fundamental to the clinical trial design. Randomization, which is a way to reduce bias, involves random allocation of the participants to the treatment groups. If investigators compare a new treatment against a standard treatment, the study subjects are allocated to one of these treatments by a random process. A general description of the randomization approach may be introduced in the clinical methods section of the proposal, for example, "Treatment assignment will be determined using stratified, blocked randomization". Specific randomization details will need to be elaborated upon in the statistical methods section, including how the allocation procedure will be implemented, e.g., via computer programs, web site, lists, or sealed envelopes. If stratification is deemed necessary, include in the proposal a description of each stratification variable and the number of levels for each stratum, for instance, gender (male, female), diabetes (type 1, type 2). However, keep the number of strata and stratum levels minimal.9 Discuss the advantages and disadvantages of the various allocation approaches with the study statistician.

Blinding

Knowledge of the treatment assignment might influence how much of a dosage change is made to a study treatment or how an adverse event is assessed. Blinding or masking is another component of study design used to try to eliminate such bias.10 In a double-blind randomized trial, neither the study subjects nor the clinical investigators know the treatment assignment.

Describe the planned blinding scheme. For example, “This is a double-blind randomized study to investigate the effect of propranolol versus no propranolol on the incidences of total mortality and of total mortality plus nonfatal myocardial infarction in 158 older patients with CHF and prior myocardial infarction.” Specify who is to be blinded and the steps that will be taken to maintain the blind. It is important that evaluators such as a radiologists, pathologists, or lab personnel who have no direct contact with the study subjects remain blinded to treatment assignments.

It may be impossible or difficult to use the double-blind procedures in some clinical trials. For example, it is not feasible to design a double-blind clinical trial for the comparison of surgical and non-surgical interventions. Or, blinding might not be completely successful; study personnel may be inadvertently alerted as to the probable treatment assignment due to the occurrence of a specific adverse event. If blinding is not feasible, offer an explanation for lack of blinding procedures in the research proposal.

STATISTICAL ANALYSIS METHODOLOGY

The statistical analysis methods for analyzing the study outcomes are to be carefully detailed. Specifying these methods in advance is another way to minimize bias and maintain the integrity of the analysis. Any changes to the statistical methods must be justified and decided upon before the blind is broken.11 In the statistical analysis plan not only must the statistical hypotheses to be tested be described and justified but we also detail which subjects and observations will be included or excluded in each analysis.

Analysis data sets

Intention-to-treat analysis—It is crucial to define the primary sample of subjects analyzed in the reporting of clinical trial results. Defined in Table 6, intention-to-treat (ITT) and per-protocol analyses are commonly reported in medical literature result sections. For a randomized study, intention-to-treat analysis is the gold standard for the primary analysis and the intention-to-treat principle is regarded as the most appropriate criteria for the assessment of a new therapy by the Food and Drug Administration and the National Institute of Health.12 An intention-to-treat data set includes all randomized subjects, whether or not they were compliant or completed the study. Adhering to the ITT principle mirrors what occurs in clinical practice where a patient may discontinue a medication or miss a clinic appointment. This avoids biases that can result from dropouts and missing data. However, the missing data must not bias the treatment comparisons13, otherwise the statistics may not be valid. This type of bias could occur if the dropouts or missed study visits are related to a particular treatment group and are not observed equally across all of the treatments.

A true intention-to-treat data set may not be attainable in all clinical trials. There might be no post-randomization or post-treatment data for a study subject who withdraws from the study at the initial study visit. Then the primary analysis might consist of all subjects who took at least one treatment dose or had at least one follow-up visit.11 Anticipate these possibilities as the study is designed and specify in the statistical analysis plan which subjects and observations will comprise the “full analysis set”. Pre-specification of these data sets prior to statistical analysis is imperative.

Per-protocol analysis—It may be of clinical interest to plan an analysis set which consists of only ‘completers’ or ‘compliers’. A per-protocol analysis, defined in Table 6, is more likely to be planned as secondary analyses. If the per-protocol analysis results are not consistent with the intention-to-treat analysis results, then closely examine the reasons behind any discrepancy.

Statistical analysis

The statistical analysis plan is driven by the research questions, the study design, and the type of the outcome measurements. The analysis plan includes a detailed description of statistical testing for each of the variables in the Specific Aim(s). If several Specific Aims are proposed, we write an analysis plan for each Specific Aim. Plan descriptive analyses for each group or planned subgroup. If subjects were randomly assigned to groups, it is expected that there will be a description of subject characteristics that include demographic information as well as baseline measurements or co-morbid conditions. Specify anticipated data transformations that may be needed to meet analysis assumptions and describe derived variables to be created such as area under the curve. Incorporate confidence intervals in the analysis plan for reporting treatment effects. Confidence limits are much more informative to the reader than are p-values alone.14

Statistical details and terminology are not intended to be an obstacle for a young investigator. Instead this is where the statistical expert can be a valuable resource to help the investigators use the appropriate statistical methods and language that address the research hypotheses. Brief statistical analysis descriptions are shown in Table 7 for a randomized study and a longitudinal cohort study. In addition to the general methodology of Table 7, we explain in the statistical methods section how statistical assumptions or model diagnostics will be evaluated. Describe the hypotheses to be tested with the corresponding statistical tests for the primary, secondary, and exploratory analyses. In the medical literature, statistical analyses such as chi-square and t-tests, analysis of variance, regression modeling, and various nonparametric tests are common. However, the statistician is happy to advise whether these traditional methods are appropriate for the research question at hand or if other approaches would be more suitable.

Statistics, like medicine, is a large and diverse field; hence statisticians have specific areas of expertise. Some proposals may require one statistician for the design and analysis of medical imaging studies and another statistician for design and analysis of a microarray study. Often a proposal specifies one statistician as the study statistician and another statistician to serve on a Data and Safety Monitoring Board.

Interim analysis

Conducting a planned interim analysis in an ongoing clinical trial can be beneficial for scientific, economic, and ethical reasons.15 Formal interim analyses include stopping rules for terminating the study early if a treatment shows futility or clear benefit or harm. The termination of the estrogen plus progestin treatment arm of the Women’s Health Initiative clinical trial in 200216 when the treatment risks exceeded benefits demonstrates the strong clinical impact of interim analyses. However, interim analyses are not to be undertaken

lightly. Taking unplanned repeated looks at accumulating data is problematic. First, it raises the multiple testing issue so that adjustments to control the overall Type I error rate are necessary. Second, the results can interfere with the conduct of the remainder of the study, creating bias. Pocock17 and O'Brien & Fleming18 authored the classic approaches for defining statistical stopping rules. The alpha spending function described by DeMets and Lan19 provides some flexibility for the timing of interim analyses as well as controlling the Type I error rate. Clinical investigators must seriously consider what decisions might have to be made based upon interim analysis results and how this will affect an ongoing study.

OVERLOOKED OR INADEQUATELY DESCRIBED AREAS

Matching in case-control studies

A weakness that often surfaces in sessions reviewing research proposals is an inadequate description of matching. Matching is commonly used in case-control studies by selecting for each case a control with the same value of the confounding variable. However, in our experience, the term "matching" is used too loosely. To a reviewer matching implies the recruitment of matched pairs. This may not be the intention of the investigators or the planned statistical analysis approach. A proposal that states that the participants will be matched according the gender, race/ethnicity, age, and body mass index would raise quite a few questions because 'matching' on all these variables would be quite difficult to achieve in practice. Often what the investigator really would like to insure is that the study groups will be balanced with respect to these characteristics. This is described as "frequency matching". For continuous variables, such as age, the range that is considered a "match" needs to be specified. Indicate the target age range that is clinically comparable for your study, e.g., within 2 years or 5 years. Avoid matching on variables that are not known confounders as this may lead to loss of power.20

Missing data prevention

It is well known that dropouts and certain missing data patterns can impact a study's validity. Since statistical analyses cannot cure all problems associated with missing data, prevention is the best policy. To minimize dropouts and missed study visits, verify that the proposal has included a retention plan. Incorporate study procedures that may help to reduce the amount missing data, such as making regular calls to participants to better maintain contact as the study is underway. Every member of the research team must appreciate the need to reschedule or repeat key study visits or labs to the extent possible if the primary outcome measurement was not obtained. In order to obtain an analysis set that is consistent with the intention-to-treat principle, continue to schedule follow-up visits and collect primary outcome measurements for subjects who have discontinued their assigned treatment.

Database

The integrity of the statistical analysis depends on the quality of the data. Obviously a study must contain high quality data (garbage in, garbage out), but steps to ensure this are frequently overlooked. Describe in the research proposal how data will be collected, de-identified, stored, and protected. It is vital that the clinical research team becomes skilled at

data management. Meet with a database expert early to discuss the design of a database and related forms and involve the statistician in the review of the forms. Development of the proper data forms and database prior to study activation is essential.

DISCUSSION

We have presented guidance to be considered when developing the statistical plan in proposals for clinical and translational research. All these approaches have the common theme of eliminating or reducing bias and improving study quality. Planning the statistical methodology IN ADVANCE is crucial for maintaining the integrity of clinical research. We hope we have conveyed that developing the statistical methods for a research proposal is a collaborative effort between statistical and clinical research professionals.

Writing the statistical plan is a multidisciplinary effort. Both the clinical investigator and statistician on the research team need to carefully review the final product and ensure that the science and statistics correspond correctly. Just as a statistician who can understand the clinical aspect of the research is particularly advantageous, endeavor to learn all you can from the statistical expert. Ask the statistician to explain the rationale of the statistical methodology so you can defend the statistical plans without the statistician at your side. The clinical investigator may not have to know how to perform complex analyses but does need to understand the general statistical reasoning behind the proposed statistical design and analysis. When clinical investigators have a basic proficiency in statistical methodology, not only are collaborations with statisticians more dynamic and fruitful, but the potential to develop into a strong, independent clinical investigator and mentor increases. This leads to the design and execution of more efficient and advanced research, increasing the productivity of the entire research team.

Statistical Resources and Education

What if the researcher does not have funding to support a biostatistician? One option is to include a biostatistician as a co-investigator in your grant proposal to cover salary and supplies needed to implement the statistical methods described in the grant. Hopefully there is a department or division of Biostatistics or related field at your or a nearby institution. If not, long distance collaborations can succeed via conference calls and email. The American Statistical Association (ASA) has an ASA consulting section <http://www.amstat.org/sections/cnsl/> where a clinical investigator can get assistance in finding a statistical consultant.

Some useful statistical websites for general statistical information and definitions include “The Little Handbook of Statistical Practice” <http://www.tufts.edu/~gdallal/LHSP.HTM>; “HyperStat Online Statistics Textbook” <http://davidmlane.com/hyperstat/index.html>; WISE Web Interface for Statistical Education <http://wise.cgu.edu/index.html>. Clinical trial statistical guidelines are documented in the International Conference on Harmonisation (ICH) Guidance for industry: E9 Statistical principles for clinical trials < http://www.fda.gov/>.11

As of September 2009, 46 medical research institutions in the United States have been granted a Clinical and Translational Science Award (CTSA, <www.ncrr.nih.gov/crctsa>). When the CTSA program is fully implemented, it will support approximately 60 centers across the nation. Some CTSA awardees offer biostatistical collaboration or institutional pilot grants for early career clinical investigators in need of statistical expertise. Many of these research centers offer Biostatistics courses or seminar series that are specifically designed for clinical researchers. This paper evolved from a CTSA course, "Clinical Research from Proposal to Implementation", taught at the University of Texas Southwestern at Dallas. Take advantage of any such course offerings and resources.

Conclusion

A successful research proposal requires solid statistical methodology. The written statistical methods section is the result of teamwork between the clinical investigators and statisticians. Collaborating with a statistician early and often will help the study proposal evolve into a strong application that increases opportunities for scientific acceptance and funding for conducting important clinical research studies.

Acknowledgments

Grant support: NIH CTSA grant UL1 RR024982

References

1. Altman DG, Goodman SN, Schroter S. How statistical expertise is used in medical research. JAMA. 2002; 287:2817–20. [PubMed: 12038922]
2. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA. 1993; 270:2598–601. [PubMed: 8230645]
3. American Statistical Association. Ethical guidelines for statistical practice: Executive Summary. Amstat News. Apr.1999 :12–5.
4. Lenth RV. Some Practical Guidelines for Effective Sample Size Determination. The American Statistician. 2001; 55:187–93.
5. Dupont WD, Plummer WD Jr. Power and sample size calculations. A review and computer program. Control Clin Trials. 1990; 11:116–28. Website: <http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize>. [PubMed: 2161310]
6. Ahn C, Jung SH. Effect of dropouts on sample size estimates for test on trends across repeated measurements. J Biopharm Stat. 2005; 15:33–41. [PubMed: 15702603]
7. Ahn C. Sample size requirement for clinical trials with repeated binary outcomes. Drug Information Journal. 2008; 42:107–13.
8. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract. 2004; 10:307–12. [PubMed: 15189396]
9. Pocock, S. Clinical Trials: a Practical Approach. New York: John Wiley and Sons; 1983.
10. Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. BMJ. 2000; 321:504. [PubMed: 10948038]
11. International Conference on Harmonisation (ICH). Guidance for industry: E9 statistical principles for clinical trials. Rockville, MD: Food and Drug Administration; Sep. 1998
12. Friedman, LM.; F, C.; DeMets, DL. Fundamentals of Clinical Trials. 3. New York: Springer; 1998.
13. Lachin JM. Statistical considerations in the intent-to-treat principle. Control Clin Trials. 2000; 21:167–89. [PubMed: 10822117]

14. Altman, D. Practical Statistics for Medical Research. New York: Chapman and Hall/CRC; 1991.
15. Chow, S-C.; Jen-Pei, L. Design and Analysis of Clinical Trials: Concepts and Methodologies. New York: John Wiley and Sons; 1998.
16. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA. 2002; 288:321–33. [PubMed: 12117397]
17. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. Biometrika. 1977; 64:191–200.
18. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics. 1979; 35:549–56. [PubMed: 497341]
19. DeMets DL, Lan KK. Interim analysis: the alpha spending function approach. Stat Med. 1994; 13:1341–52. discussion 53-6. [PubMed: 7973215]
20. Daya S. Strategies to control confounding in causation studies. Evidence-based Obstetrics and Gynecology. 2002; 4:1–2.
21. Toto RD, Adams-Huet B, Fenves AZ, Mitchell HC, Mulcahy W, Smith RD. Effect of ramipril on blood pressure and protein excretion rate in normotensive nondiabetic patients with proteinuria. Am J Kidney Dis. 1996; 28:832–40. [PubMed: 8957034]

Table 1

The role of the statistician in developing the statistical plan

- Clarify the research questions or hypothesis. Are the primary hypotheses clearly stated, adequate and realistic?
- Identify the outcome variables related to the research questions. Are the primary or secondary endpoints clearly defined?
- Does the study design appropriately and adequately address the proposed hypothesis?
- Are the issues of the bias, blinding or stratification properly handled in the study?
- Are the assumptions used for sample size estimation clearly elaborated and supported by proper preliminary data and/or references?
- Is there a clearly specified, appropriate data analysis plan?
- Is there an appropriate data and safety monitoring plan, interim analysis plan, or pre-established early stopping rule?

Table 2

Outline of the statistical methods section

- **I.** Study design
  - Type of design
  - Sampling
  - Power and sample size
  - Randomization and blinding
- **II.** Statistical analysis methodology
  - Define data analysis set
  - Statistical analysis
    - **i.** Primary analysis
    - **ii.** Secondary analysis
    - **iii.** Exploratory
  - Missing data
  - Multiplicity of testing
  - Subgroup analyses, covariates
  - Interim analysis

Table 3

Definitions for statistical hypothesis testing

- A Type I error (alpha) is the risk of inferring a difference between study groups when there is no such difference.
- A Type II error (beta) is the risk of inferring no difference between study groups when there is such a difference.
- Power is the statistical test's ability to detect a true difference. Power = 1 - Type II error.

Table 4

Variable types and derivations to be described in the statistical analysis plan

Describe each variable and type to be collected

- Categorical
  - Two categories (binary or dichotomous)
    - Sex (male, female), Diabetes (Type I, Type 2)
  - More than two categories
    - Blood type (O, A, B, AB)
- Ordinal
  - Attitudes (strongly disagree, disagree, neutral, agree, strongly agree)
  - Cancer stage (I, II, III, IV)
- Survival (time to event)
  - Transplant free survival time
  - Time free from infection
- Numerical
  - Discrete
    - Number of abnormal cells
  - Continuous
    - Age
    - Serum creatinine
    - Log_e triglycerides (non-normally distributed, log transformed due to skewness)
- Derived variables
  - Absolute and percent change
  - Area under the curve (AUC)
    - Insulin AUC from an oral glucose tolerance test
    - Receiver operating characteristic curve (ROC)
  - Pharmacokinetic parameters

Table 5

Scenarios for choosing sample size

Primary Outcome Variable	Effect size Mean detectable difference between groups	Estimated standard deviation[a]	Sample size per group[b] α = 0.05 Power = 0.80	Sample size per group[b] α = 0.05 Power = 0.90
Blood pressure				
Systolic blood pressure, mm Hg	6	14	86	115
	8	14	49	65
	10	14	32	42
	20	14	9	11

[a] cite a reference where this estimate was derived (e.g. Toto et. al. 199621) or cite own preliminary data

[b] name the software used for the computations; PS Power program by Dupont and Plummer5

Table 6

Analysis data sets

Intention-to-treat principle — the principle that asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a subject (i.e., the planned treatment regimen) rather than the actual treatment given. It has the consequence that subjects allocated to a treatment group should be followed up, assessed, and analyzed as members of that group irrespective of their compliance with the planned course of treatment.
Full analysis data set — the set of subjects that is as close as possible to the ideal implied by the intention-to-treat principle. It is derived from the set of all randomized subjects by minimal and justified elimination of subjects.
Per protocol data set (valid cases, efficacy sample, evaluable subjects sample) — the set of data generated by the subset of subjects who complied with the protocol sufficiently to ensure that these data would be likely to exhibit the effects of treatment according to the underlying scientific model. Compliance covers such considerations as exposure to treatment, availability of measurements, and absence of major protocol violations.

From ICH E9: Guidance for Industry - E9 Statistical Principles for Clinical Trials, U.S. Department of Health and Human Services, Food and Drug Administration, September 1998

Table 7

Statistical analysis plans

A. Statistical analysis example for a randomized study.
Statistical analysis: The full analysis set will include patients who have received at least one dose of medication or had one or more post-randomization, follow-up evaluation. Descriptive statistics will be computed for each treatment group, Medians and percentiles will be reported for skewed continuous variables. For primary and secondary outcomes, descriptive statistics and 95 percent confidence intervals will used to summarize the differences between groups. The primary outcome of systolic blood pressure and other continuous variables will be assessed with a repeated measures analysis using a mixed linear model approach. Since many of the inflammatory markers are positively skewed, IL-6 and CRP will be log transformed prior to analysis. The Wilcoxon Rank Sum test will be used to compare pill counts between groups. Hypothesis tests will be two-sided using the 0.05 significance level. Bonferroni type adjustments for multiple testing will be implemented to control type I errors. Statistical analysis will be performed with SAS software (SAS Institute, Cary, NC, USA).
B. Statistical analysis example for a longitudinal cohort study
Descriptive/comparative statistics defining the biomarker levels in the different disease activity classes. We will compute and compare the mean/median, and inter-quartile range of urine biomarker levels in different disease activity groups, after partitioning patients in various ways: patients who attain any of the primary disease outcomes, i.e., WHO Class IIII-or-IV glomerulonephritis, patients with nephritic or nephrotic flares, or end stage renal disease. Additionally, we will define the biomarker levels in patients with the following disease features: anemia, leucopenia, or thrombocytopenia. For comparing multiple patient groups, analysis of variance (ANOVA) or the Kruskal-Wallis test will be used, depending on whether the biomarker values are normality distributed. Data transformations will be performed if necessary. If the omnibus ANOVA or Kruskal-Wallis test yields $p<0.05$, we will conduct pairwise group comparisons using either t-tests or Wilcoxon rank sum tests with Bonferroni corrections. The generalized estimating equations (GEE) approach will be used to evaluate if urinary biomarkers vary significantly over time among different disease activity classes.