

Guideline for Writing a Study Protocol – RPG-04

Guideline

Purpose

Developing protocols for clinical research in collaboration with Children's Hospital investigators is one of the primary activities mandated by the mission of the Clinical Research Center (CRC). The written protocol is the embodiment of a clinical research plan and expresses its scientific agenda, clinical relevance, organizational thoroughness, and professional quality. A crucial interest of CRC is therefore to ensure that protocols developed under its auspices follow a high standard of scientific merit and methodological content. An additional aim of equal importance is that a protocol be clear and comprehensible to the variety of audiences that it addresses, including medical, statistical, and human-subjects reviewers as well as the co-investigators and staff who are ultimately charged with carrying out the research. A comprehensive, readily available Standard Operating Procedure for protocol development, comprising outlines, guidelines, templates, 'boilerplates,' comments, and examples, should enable, by its regular use, CRC staff to attain and maintain high standards in performing this central task, particularly when collaborating with inexperienced medical investigators.

The intent of this document is to furnish the reader with guidelines on the important scientific elements of a clinical research protocol; to describe their essential content and interrelationships; to indicate which elements are optional and which mandatory; and to illustrate their key features with examples. Guidelines will also be provided for handling the collaborative processes and collegial issues that often arise in protocol development. All investigators should follow the specific instructions provided by the CHB Committee on Clinical Investigation (CCI) when submitting a protocol for IRB approval, particularly regarding requirements for human subjects protection. Information and guidelines from CCI can be found on the hospital intranet under Clinical Research.

Definitions

Protocol

A protocol is a written document containing a comprehensive plan for clinical research. The protocol poses scientific questions and describes the methods by which they will be addressed.

It must be detailed enough to allow medical, statistical, and research-ethics reviewers to judge the merit of the plan, the qualifications and readiness of the investigators to carry it out, and the feasibility and general cost of performing it.

Principal Investigator (PI)

The Principal Investigator is the person proposing a clinical research project. A project can have only one PI. The PI is responsible for writing the protocol and must be prepared to

stand behind all its contents. The PI is also responsible for assembling the protocol and related materials; submitting it to review bodies and responding to their questions; and directing all scientific and administrative aspects should it be approved and support obtained.

Procedure

The Writing Process

Contents of a protocol cannot reach the highest standard unless the process of authorship is well controlled and professionally organized. Miscommunication is a serious threat to quality and is easy to fall into when the writing team is physically scattered, separated by discipline, and operating under pressures of time, all of which are commonly the case in clinical research. The following guidelines attempt to address the need for effective teamwork in protocol writing.

Role of the PI

The Principal Investigator is the author of the protocol and is responsible for all of its elements. The final document should read as if written in a single voice, even if portions originated from a number of different contributors.

Formative stages

Including CRC staff in the assembly of a protocol in its earliest stages, particularly when the investigator is relatively inexperienced, may save considerable effort later in the process. Participating in formative discussions or in the writing of a concept paper (see below) represents a chance for CRC collaborators to provide valuable input in the most fundamental and important areas of the research plan. These include formulation of the specific aims, choice of study design, specification of endpoints and treatment groups, and allocation of staff effort.

Concept paper

In some cases, the first object of protocol development is a brief concept paper or précis, summarizing the basic aims, design, and methods of the project. The concept paper may be required as a first step in the approval process by a potential funding agency, or it may simply serve as an early guide for further development of the research plan. Concept papers present an opportunity for CRC staff to shape a protocol, provide beneficial material early in the project when it is still flexible, and have a maximum of influence with a minimum of writing effort. This

‘short version’ of the research plan is also extremely useful as reference material when efforts begin on writing the full version, usually after some interval.

Written contributions

CRC staff may frequently be called upon to write full sections of a protocol, particularly those of a technical nature, ranging in length anywhere from a paragraph to several pages. Examples include the statistical analysis plan; database design; or randomization procedures. Such writing should be preceded by thorough discussions with the PI, to the extent that the CRC staff member is comfortable with his/her understanding of the aims, objectives, and design of the study. Adequate prior discussion will ensure that the contribution is apt and accurate and will not have to be revised multiple times owing to incomplete or faulty communication.

Contributed protocol sections should be written as if in the PI’s voice rather than attributed to the CRC staff. Such attributions, even if intended as giving due credit, are not appropriate in a protocol; they imply ‘proof by authority’ and appear to distance the PI from important elements of the research plan. In particular, power calculations and analytic methods, like all other medical and technical matters, must be presented and vouched for by the PI.

Example (don’t):

- The following power calculations were provided by Dr. Ying-Hua Jiang, a biostatistician in the CRC.

Editing and version control

CRC staff contributing written segments should expect that, after such sections have been incorporated into the developing protocol, the PI will allow the CRC staff member to review the full document to verify their correctness in context and to approve any editing performed by the PI. This expectation should be made clear to the PI early in the process. Such exchanges also give CRC a fuller view of the developing aims, design, logistics, measures, and organization of the study and may suggest corrections or improvements in the portion contributed by CRC.

All draft protocol versions should be dated in the footer of the document while under editing and revision. Only IRB approved versions of the protocol used during the study should be archived with the approval dates and version numbers in the footer.

Reference copy

Regardless of the extent of contributed material, CRC staff collaborating on a protocol should expect to receive a copy of the submitted protocol in its final form and should make that expectation clear to the PI. Besides representing collegial courtesy from the PI, a final copy is essential for CRC staff in answering questions that arise later in the protocol review process, which are often transmitted by phone or email under time pressure, and for making revisions and preparing resubmissions. CRC staff also benefit from having a copy of the completed protocol as

reference material for later work that is similar in design or content.

Administrative information

In its early drafts a protocol will commonly lack a good deal of information that is ‘obvious’ to the PI and therefore not supplied until the final revisions, such as the title of the study or key personnel. Also missing from early drafts may be items as yet undecided, such as budget, dates of conduct, or funding source. CRC staff should request this type of information as early as possible in order to establish the identity of the project in CRC records and prevent misunderstandings concerning personnel, budget, and other pertinent administrative arrangements.

Specific Aims

Specific Aims are the ‘heart and soul’ of every research proposal. They should be brief, explicit, and logically ordered. Every reviewer reads them first and relies on them as a framework for understanding the entire protocol. The primary scientific reviewers expect all details of design, conduct, measurement, and analysis to relate to the Specific Aims. Reviewers assigned to a specialty area, such as statistics or research ethics, may read little else than the Specific Aims in addition to their own assignment. A clear, sharp Specific Aims section is therefore vital to the construction of a successful protocol.

Specific Aims should occupy at most one page and may often take less. The section should begin with an introductory paragraph ‘setting the stage’ scientifically, including the most important terminology, abbreviations, background, and rationale for the study. These should be extremely brief; more detail should be reserved for the Background and Significance section.

The body of the section should be a short, numbered list of aims, typically 2-4, in logical order. Each aim should express a scientific question to be addressed by the research project. The sequence of aims should represent an unfolding investigation into the topic, with central questions first and incidentals last. Ordinarily, one aim should be identified as primary and should involve a primary endpoint (outcome variable).

Example:

- Aim 1 (primary). To describe the level and range of concentration of venous allosteric ligand binding protein (VALbp1) in serum of children with cystic fibrosis and compare it to that in normal children.
- Aim 2. To determine whether serum level of VALbp1 is associated with the presence of the CiMB4u allele in children with CF.
- Aim 3. To determine whether VALbp1 level is correlated with the severity of respiratory impairment as measured by forced vital capacity in children with CF.

The specific aims should represent questions that can be answered definitively and quantitatively and should be couched in terms of particular well-defined criteria or measured variables — in a word, they should be *specific*. The aims should be intellectual rather than operational and should

not be so general or vague that a well-informed reader cannot imagine how they would be addressed.

Example (don't):

- Aim 4. To recruit a cohort of patients with CF. (Too operational)
- Aim 5. To improve the sensitivity of our enzyme-linked immunospecific assay for VALbp1. (Too technical)
- Aim 6. To characterize V-cell response in the population of CF patients. (Too vague)
- Aim 7. To gain an understanding of the interaction between V-cell response, genotype, and phenotype in the CF population. (*Too broad*)

The listing of Specific Aims is commonly followed, or in some cases interleaved with, a statement of corresponding Hypotheses. Where each Aim expresses a scientific question, each Hypothesis represents the investigator's best guess at the answer. The list of hypotheses should run in strict parallel to the list of aims.

Example:

- Hypothesis 1. Serum VALbp1 is in the normal range (0–10 pmol/L) in at least 90% of children with CF but markedly elevated in a small subset.
- Hypothesis 2. Children homozygous for the CiMB4u allele will have extremely high serum levels of VALbp1, exceeding 10 times the normal maximum.
- Hypothesis 3. Pulmonary function in children with CF, as measured by forced vital capacity, is correlated poorly if at all with VALbp1 level (Spearman $r < 0.20$).

Note that each hypothesis is stated in terms of a specific variable and is testable by a statistical procedure. The structured statement of aims and hypotheses can thus feed naturally into a similarly structured listing of measurement techniques, statistical procedures, and power calculations later on in the protocol.

Background, Significance, and Rationale

In this section the investigator reviews the state of knowledge in the field, drawing on published literature, and makes a case that the aims addressed by this protocol will advance scientific understanding, clinical practice, or both.

The presentation of prior research should not be addressed solely to fellow medical specialists but should at least begin with essential definitions, abbreviations, and explications comprehensible to a reader with basic medical knowledge. The literature review should be thorough but not exhaustive, including only what is necessary to provide context and rationale for the proposed work.

The Background section often ends with a short paragraph making explicit the need for the present study: how it follows logically from prior studies, and what the anticipated findings would imply for scientific evidence and clinical practice.

Example:

Because prior research as just described has not established that elevated VALbp1 is a primary manifestation of genetic correlates of CF, as opposed to a secondary reaction to the multiple morbidities attending CF, our anticipated demonstration of a strong association with a recessive gene known to predominate in the Armenian-American population will have profound implications for public-health screening, acute treatment, and chronic care.

Preliminary Studies

The Preliminary Studies section should efficiently review the investigators' own prior work in the subject matter of the protocol, including published or unpublished research by the PI and key co-investigators. Pilot data, progress reports on ongoing studies, and related methodological studies should be featured if possible, presented item by item in the concise style of a published abstract. Small tables and graphs are appropriate and efficient for presenting work to date.

If the investigators have not recently performed any directly pertinent work, then a short synopsis of their professional preparation for the proposed study is appropriate in this spot.

Example:

Dr. Fleck has 15 years' experience in cluster-randomized trials related to heart disease, physical activity, and nutrition, including design and analysis of the Pawtucket Heart Health Program, the Child and Adolescent Trial for Cardiovascular Health, and the Rapid Early Action for Coronary Treatment trial.

The object of the Preliminary Studies section, whether featuring research synopses or long-term experience, is to demonstrate the qualifications, preparation, and professional readiness of key personnel to carry out the proposed studies.

Research Plan

The Research Plan forms the bulk of most protocols. It contains detailed specifications of how the study is to be conducted, including all procedures related to recruitment, treatment, measurement, and analysis. It should be written as a narrative (not merely a bullet-list) and addressed, like the Background, to an informed reader rather than a specialist. The level of detail should be sufficient for the reader to visualize the procedures that are described, assess their merit and appropriateness, and come away with a good idea of how they will be carried out.

Study design

The Research Plan should begin with a general characterization of the study design, using standard nomenclature. The reader should know at the very outset whether the study is

- Observational or experimental
- Prospective or retrospective

- Case series, case-control, or cohort design (if observational)
- Controlled, blinded, and/or randomized (if experimental)

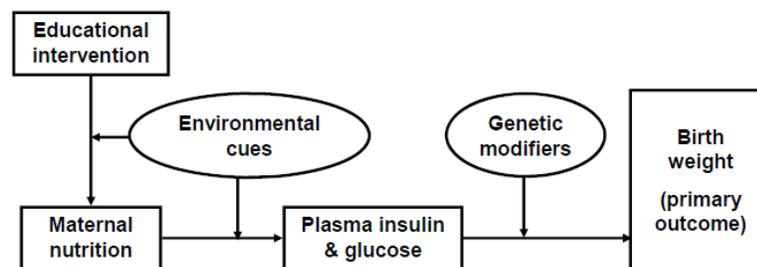
If the study is designed to be exploratory, pilot-scale, or hypothesis-generating rather than a definitive test of hypotheses, that intent should be made clear at the start of the Research Plan. A rationale should be provided for the choice of study design.

For clinical trials, one of the standard classes should be indicated:

- Phase I - Safety testing on a small sample. Open-label (unblinded) and uncontrolled.
- Phase II -Dose-finding, preliminary efficacy testing. Small sample, usually randomized and controlled, blinding optional. May use a surrogate outcome variable.
- Phase III - Definitive demonstration of efficacy, large sample, randomized, controlled, blinded. May include interim analysis and early-stopping rules.
- Other major features of the study design should be summarized at this point in the protocol, before plunging into detail. These include
- Number and nature of study arms (comparison, control, placebo)
- Planned sample size
- Designation of primary endpoint and study hypotheses
- Other variables involved in the major hypotheses (outcome, predictor, confounder, covariate, mediator, effect modifier)

A schematic research model is sometimes informative, showing the hypothesized causal, confounding, effect-modifying, and mediating relationships among study variables.

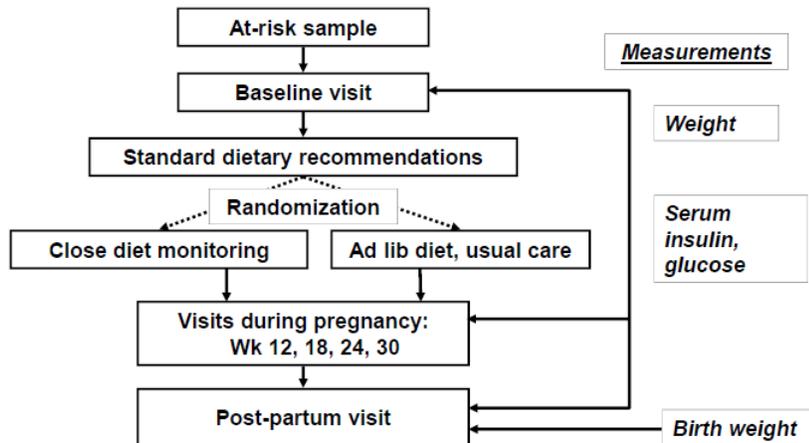
Example:



Nutrition education for pregnant prediabetic women

An operational diagram is likewise useful at the outset of the narrative, particularly for multistage recruitment schemes, longitudinal studies, trials with multiple arms or subgroups of interest, crossover experiments, and other complex designs.

Example:



Population

Clinical research by definition aims to answer questions about a human population, in most cases a particular subpopulation (e.g., children with asthma). This population should be specified in terms of inclusion and exclusion criteria that can be operationalized as a screening procedure for eligibility. Typical criteria include

- Age
- Sex
- Medical diagnosis, preferably with objective definition in terms of serum levels, signs and symptoms, and/or duration
- Availability for required visits
- Health status sufficiently good to allow the subject to give informed consent or assent and complete the study

Example:

Inclusion criteria

1. Male or female, age ≥ 1 yr and ≤ 12 yr at enrollment.
2. Diagnosis of hyperkoffatremia based on any two of the three criteria published in the 1998 HKT Consensus Conference: (a) plasma free keffatrine ≥ 60 mEq/L by pyracanthic gas chromatography; (b) genotype with two identifiable mutations consistent with HKT; or (c) abnormal transulnar compliance ratio (age ≥ 5 yr only).
3. Clinically stable with no evidence of abnormal muscle weakness or fainting spells.

Exclusion criteria

1. History of hypersensitivity to chelating agents.
2. Abnormal liver function tests at baseline visit.

3. Administration of any investigational drug within 30 days preceding baseline visit.
4. Chronic steroid use within 1 yr preceding baseline visit.
5. Residence more than 200 mi from study center.
6. Presence of any condition or abnormality that in the opinion of the Investigator would compromise the safety of the patient, quality or completeness of the data, or competence of the patient to provide informed consent or assent.

Recruitment

Concrete procedures for recruitment and obtaining informed consent should be described; including how and where subjects will be approached (newspaper, flyer, bedside, well-baby clinic), and by whom (primary care physician, nurse, research coordinator). The sampling frame should permit a wide range of variation for important study variables and should be free of selection bias.

Example (don't):

Subjects for the sunlight exposure calibration substudy will be recruited from children attending advanced swim classes at the YMCA outdoor pool in Waltham.

The process of recruitment and consent must ensure privacy and confidentiality, avoid coercion, provide adequate time and privacy for decision-making (in consultation with family if appropriate), and emphasize the subject's right to decline without compromise.

The recruitment section should also include an estimate of the anticipated rate of enrollment, considering typical case-flow or available pools, and specify the projected sample size and time required to assemble the full sample. (Sample size is specified here for the sake of operational context. Statistical justification comes later; see Sample size, precision, and power below.)

Randomization and blinding

For a blinded study, careful specification should be made as to who is aware of treatment assignments and who is not, including

- Subjects and family
- Principal Investigator
- Co-investigators
- Care providers
- Pharmacy
- Data collectors

Justification should be provided for any unblinding, particularly if it has any potential for affecting data. Clinical situations requiring unblinding should be anticipated and appropriate procedures devised.

For a randomized trial, the scheme and mechanics of assigning patients to treatment arms should be described thoroughly. Important details include

- How many treatment arms
- Balance or allocation ratio among treatment arms
- Strata, if any, within which balance or allocation ratio is to be maintained
- Method of generating random assignment lists (typically a mixture of permuted blocks, to avoid a predictable sequence)
- Where the assignment list will reside (pharmacy, sealed envelopes at nursing station, notebook in the study coordinator's locked drawer)
- Who will be privy to treatment assignments (if study is blinded)
- The sequence of events leading to a treatment assignment should be carefully described, including who is responsible for each stage of the process and how appropriate blinding is ensured.

Example:

•After informed consent is obtained, the study coordinator will record the subject's name in a sequential enrollment log, thus irrevocably assigning a study ID number as listed in the log. The study ID number will be communicated to the pharmacy (or study statistician, or other custodian of the randomization list), who will record it in the next place on the randomization list and respond with the assigned treatment arm.

Visit schedules

By enrolling in the study, subjects commit themselves to some number of encounters with study staff. These expectations should be outlined as clearly in the protocol as they will be to the subject. The schedule, location, and duration of visits should be specified, along with measures that will be taken to facilitate adherence to the protocol. Important details may include

- Instruction packets for subjects
- Telephone and mail reminders
- Help with transportation or reimbursement for parking expenses
- Incentives for participation

Procedures should be described for tracking subjects and maintaining contact information; implementing and documenting withdrawal or termination when needed; and providing subjects with information about the study findings afterward if they wish to receive it.

Treatment

Complete information should be provided on study medications and procedures, including formulation, route of administration, dosage, schedule, and dosage adjustment if indicated. For educational interventions, the content and purpose should be described and materials provided in an appendix if available.

Criteria for compliance and a method for monitoring it should be specified. In field trials, it is often important to specify a set of process and impact measures, which detail how, when, and how fully the intervention was delivered and received.

Measurements

Beginning with the primary endpoint, all study measurements should be listed, described, and justified, including

- Medical history
- Anthropometrics
- Clinical examinations and physiologic parameters (e.g., blood pressure)
- Blood tests (drug levels, lipids, hormones, immunology)
- Imaging procedures (CT, DXA)
- Psychological questionnaires
- Diet, physical activity, and other behavioral inventories
- Laboratory assays
- Genetic determinations

A complete schedule of measurements should be provided in tabular form, showing how many times each measurement is to be made and at what time points.

Important issues and details pertinent to a particular measure should be covered thoroughly, with reference to literature as appropriate.

Example:

Each adolescent's dietary fat intake will be assessed with three 24-hr diet recalls, administered over the telephone on randomly selected days during the one-month follow-up period. The telephone interview will be conducted by the study coordinator, who is well known to each participant. Data will be recorded on the Kasebier diet inventory, which has been validated for telephone administration and widely applied in epidemiologic studies (Brown, 1993). Although 24-hr recalls are not considered precise enough for clinical use, their application for large-sample studies is well established (Perry, 1991).

Data collection, entry, and management

The mechanics of data collection should be thoroughly described, including specifications of how the data will be recorded (paper forms, electronic transmissions), by whom, and at what time.

The case report forms in draft form, or similar forms from earlier studies, should be appended. A well-designed form is one that

- Is sequentially numbered

- Is easy to read and follow down the page
- Consists almost entirely of numerical, yes/no, or multiple-choice items
- Contains few if any blanks for free text
- Specifies units, number of digits, and number of decimal places for numerical items
- Provides check-boxes for multiple-choice items
- Permits blocs of non-applicable items to be skipped

Expert advice, guidance, and assistance with design of case report forms is available from CRC.

For laboratory measures, important logistical details should be provided, including sample handling, processing, storage, shipping, and if analysis is to be performed at an outside laboratory, the mode by which results will be transmitted and entered or imported electronically into the study database.

Data must be entered from the paper forms into a database in a form amenable to summary and analysis. In most cases, general-purpose spreadsheets (e.g., MS Excel) are not adequate for this purpose, because they allow too much latitude in form and content and result in disorganized or “unclean” data. Software specifically designed for scientific data collection, such as SPSS Collector or REDCap, is available from several sources including the Research Computing group at CHB. A good database program can be programmed to reflect closely the sequence and structure of the study’s case report forms, thus facilitating data entry and quality control. Among the features that can be programmed in are numerical range checks, logic checks, skip patterns, and double data entry. The result is well-structured, “clean” data that can be imported directly into statistical packages for summary, reporting, monitoring, and analysis.

Procedures should be described for ensuring confidentiality of study data. Personal identifiers should be kept separate from the body of study data, linkable only by a study ID number. This should be a unique code, not a set of initials or medical record number.

Security features of the database should be detailed, including password protection and back-up archiving of computer files. Personal information and randomization lists should be password-protected on the computer and hard copy kept in a locked cabinet, accessible only to appropriate staff.

Example:

- Data collected during the study visit will be recorded on paper case report forms, faxed to the study coordinator from each site (see draft, Appendix B).
- Subjects’ confidentiality will be maintained by utilizing a numeric study code in lieu of identifiable subject information. The study identification log will be kept in a locked file accessible only to study staff for the purposes of enrolling subjects.
- From case report forms, the data will be entered into a REDCap database. Data capture screens will be designed to mimic the case report forms, promoting accurate and efficient

data entry. Validation features, including range checks, logic checks, and double data entry, will be programmed to facilitate data entry and reduce error.

- System security will be maintained by requiring user authentication to gain access to the file on the hospital's private secure network. Daily server back-up activities will be conducted to permit data recovery in case of loss of files.

Quality control

Producing high-quality data requires active attention at every step of the study. This section should include plans for proactive measures to ensure accuracy and reliability in advance of data collection, as well as ongoing quality checks (e.g., on a random 10% sample) and data-cleaning at the end of the study. Typical measures include

- Calibration of equipment
- Training of data collectors (in some cases formal certification)
- Site visits to inspect facilities, interview staff, and verify data forms and source documents
- Systematic collection of blind duplicate samples for lab analysis
- Substudies to assess inter-rater reliability
- Double-checking of filled-in data forms by a senior staff member
- Double data entry
- Data cleaning by study statisticians

Data and safety monitoring

The purpose of this section is to specify procedures for

1. Monitoring data; and
2. Monitoring safety;

thus the term **Data AND Safety Monitoring**.

In a multisite trial, the monitoring function is commonly performed by an independent Data and Safety Monitoring Board, appointed by the sponsoring agency (often National Institutes of Health) and reporting to that agency. The DSMB typically meets once or twice a year to receive reports on the progress of the study and advises the sponsor on whether to continue or terminate the study. Data-monitoring aspects of the periodic report include

- Progress of enrollment
- Completeness and timeliness of data collection
- Subjects' compliance with the protocol
- Sample attrition (for cohort studies)
- Methodology and reliability substudies
- Baseline and interim data analysis
- Progress of publications

Safety monitoring functions include

- Adverse event reports
- Safety lab data (e.g., liver function tests)
- Implementation of stopping rules (e.g., early stopping of a drug trial for demonstrated elevation of adverse event rate)
- Formal recommendation to continue or terminate the study

Smaller studies, with a single site or no external sponsorship, may take a more streamlined approach to data and safety monitoring; but at the very least a data and safety monitoring plan is required by most IRBs including that of CHB. An independent board is not a necessity; the Principal Investigator may fill this function, preferably with the assistance of one or two consulting colleagues. The plan may simply consist of a procedure and schedule for periodically examining the status of enrollment, data collection, compliance, and adverse events. The purpose is to monitor, audit, judge, and correct as needed.

Primary and secondary endpoints

Although the primary and secondary outcome variables are presumably mentioned at several earlier points in the protocol, it is beneficial to catalogue and define them formally at this point in the Research Plan — after covering the operational details of data collection, but before entering into the intricacies of statistical analysis and inference. The primary endpoint should be highlighted and carefully defined, followed by secondary endpoints roughly in order of their role in the Specific Aims.

In some cases, an endpoint is not a simple measurement or dichotomous classification, but the end product of a complex calculation. Such data reduction procedures should be specified in detail, with reference to literature if part of a standard instrument-scoring algorithm.

Examples:

- Area under the insulin concentration curve following oral glucose administration
- Composite clinical outcome, defined by death or rehospitalization or respiratory function below 70% of normal for age
- Domain score on a psychological questionnaire

Statistical analysis plan

Whereas the Specific Aims ask a set of scientific questions, this section should specify in thorough technical detail how the investigator plans to answer those questions, using the data collected.

Before going into details, it is often useful to describe the organizational setting for this phase of the study. An important question is who will conduct the analysis; possibilities include

- The investigator alone
- The investigator with consultation from CRC or other statisticians
- Statisticians in collaboration with the investigator

The intended software tools should be mentioned, particularly if specialized techniques are to be used.

Example:

Statistical analysis will be performed by the PI, Dr. Fahey, with guidance from biostatistical consultants at CRC. SPSS software will be used for simple descriptive statistics and the primary analysis, a 2-sample t-test. LogExact software will be used for the large contingency-table analysis required for Aim 3.

An introductory paragraph should summarize a set of simple, exploratory statistical analyses for purposes of description and data-cleaning, which constitute a vital preliminary step in any analytic plan. These typically include

- Descriptive statistics for continuous measures (mean, standard deviation, quartiles, extreme values), overall and by important subgroups
- Tabulation and cross-tabulation of discrete variables (sex, diagnosis, measurement times, qualitative outcomes)
- Graphical summaries (distributions, scatterplots, boxplots, individual time courses)
- Correlation tables, principal components, or factor analysis

The main body of the analytic plan should be an itemized set of data-analytic procedures, keyed one by one to the Specific Aims and designed to address the questions that they pose. For descriptive Aims, the analysis may consist of reporting a simple percentage with 95% confidence limits. For most other Aims the analysis requires a statistical test. A cursory list of statistical tests, resembling the table of contents of a biostatistics text, is not adequate.

Example (don't):

We will analyze study data with Student t-tests, chi-square statistics, Spearman and Pearson correlations, and multiple logistic regression.

Plans for statistical modeling and testing should be described in full, specifying the variables involved and their role in the procedure.

Example:

To address Specific Aim 1, we will use the Student t-test to compare mean birth weight between normal neonates and those with evidence of fetal hypoproteinemia.

If the choice of analytic method is contingent in part on the nature of the data, then alternatives should be listed and the rationale given for the ultimate choice.

Example:

If the distribution of fetal lung volumes is highly skewed, then we will replace the t-test with the Wilcoxon rank-sum test, an analogous nonparametric procedure.

The p-value that will be considered significant for statistical testing should be specified. For the primary endpoint, 5% is the conventional value. If multiple comparisons are planned because of a large number of important secondary endpoints, predictor variables, or interactions of interest, then a rationale should be provided for adjusting the critical p-value to avoid Type I error.

Example:

In diagnostic screening studies or ROC analysis, the “gold standard” for calculating sensitivity and specificity should be clearly indicated. Similarly, in describing planned regression analyses, it is essential to specify explicitly which is the dependent variable and which is the independent variable.

Multiple regression is a particular challenge to describe clearly. The standard nomenclature of elementary biostatistics (ANOVA, ANCOVA, mixed-model) is not always adequate to characterize a complex mixture of discrete and continuous predictors or fixed, random, and repeated-measures effects. In such cases, the best method for specifying the analytic plan is to lay out the regression model term by term, starting with the dependent variable and proceeding through the list of predictors, covariates, and interaction terms with a careful description of the nature and role of each variable. A model equation or quasi-equation is often useful.

Example:

- To address Specific Aim 3, we will determine the joint effects of TR4 allele, gender, birth weight, and severity of initial diagnosis on motor function as measured by the Boswell scale at the 1-yr, 2-yr, and 3-yr follow-ups, adjusted for random effects of test site and hospital. The regression model may be summarized as
- $Y = \text{Allele} + \text{Time} + \text{Gender} + \text{Birth weight} + \text{Severity} + \text{random effects} + \text{interactions} + \text{error}$.
- The dependent variable Y is the Boswell score at 1 yr, ranging 1-20. The primary independent variable is TR4 allele, which has 4 variants. Time is the repeated-measures fixed effect, representing normal motor development over 3 yr. We will assume a first-order autoregressive covariance structure for Time.

- Gender, birth weight, and severity are included as a covariate adjustment because they are known to influence the Boswell scale. Random variation among birth hospitals and testing sites is represented by Gaussian deviates.
- The interaction of primary interest is Allele \times Time, allowing for delayed expression of the TR4 variants. A second important interaction is Allele \times Gender, addressing the hypothesis of Mack et al. (1997) that the activation of TR4 gene products is more marked in males than in females.

If multiple regression is a central tool in the analytic plan, a model-building strategy should be outlined. Commonly employed options include

- Automatic selection procedures (backward, forward, best-subset)
- A fixed core of obligatory predictors and covariates
- Alternative coding of covariates (continuous, dichotomous, quantiles)
- Strategies for identifying and choosing among collinear covariates
- Strategies for identifying predictors, confounders, and mediating variables

Following specification of statistical tests and model-building procedures, a number of general issues may need addressing. These include

- Missing data and methods for assessing the magnitude of any resulting bias
- Adjustment of critical p-values for multiple comparisons, from a study-wide perspective
- Methods for examining collinearity of predictors and distinguishing their joint and independent effects
- Interim analyses, stopping rules, and their implementation
- Alternatives for secondary analysis, such as data transformations, two-stage analyses, semi-parametric or non-parametric techniques, or ordinal logistic regression
- Choice of software

Sample size, precision, and power

Clinical research, being in essence experimentation on humans, can be justified only if it has some chance of producing interpretable results. This issue applies to the entire protocol, including formulation of hypotheses and choice of measures, but it is most sharply embodied in the necessity of adequate sample size. One of the principal reasons for using statistical methods to decide between alternative hypotheses is that, despite the fact that statistical conclusions are necessarily uncertain, the degree of uncertainty can be quantitated. The statistical parameters of precision and power, both critically dependent on sample size, provide a quantitative estimate of the likelihood of obtaining meaningful results. A well thought-out analysis of precision and power is therefore a necessary component of any clinical research protocol.

The simple fact that other studies of the topic have employed a particular sample size is not adequate justification for its use in a proposed protocol.

Preliminary data are useful as a basis for precision and power calculations because they provide provisional estimates of important parameters, such as

- General prevalence of a dichotomous outcome
- Variability (standard deviation) of individual observations of a continuous measure; and in some cases
- Rough magnitude of effect to be expected from a treatment

These quantities may be available from pilot studies, from published literature, or from reasonable speculation. If no single value can be cited with confidence, a range of plausible values may be considered, in order to demonstrate that adequate precision or power will be obtained regardless of the outcome. Another option, if preliminary data are not available for a measured endpoint, is to express the endpoint in multiples of 1 standard deviation (“effect size”) and refer the detectable effects in the proposed study to conventional benchmarks for “small” and “large” effect size.

Because these options are available, lack of preliminary data does not excuse the investigator from providing precision or power estimates.

Precision is the uncertainty attending a particular parameter estimate, usually expressed as a standard error or as a 95% confidence interval. The magnitude of the standard error is generally proportional to the inverse square-root of sample size. For a continuous parameter, such as a mean or regression coefficient, calculation of the standard error requires knowledge of the underlying variability of data points, in the form of the standard deviation of individual observations. The precision of a correlation coefficient or a binomial proportion depends on the magnitude of the true correlation or proportion. As noted, these quantities may be available from prior studies or reasonable speculation.

Power is the likelihood of obtaining a significant result in a statistical test. Power calculations are thus required only if the study involves some test of hypothesis; otherwise precision rather than power is the pertinent concept. Power calculations, like precision calculations, typically involve a standard error and therefore require some estimate of variability, correlation, or proportion, which should be obtained from prior data or speculation.

Power may be defined as the probability of detecting a true underlying effect (group difference, ratio, regression slope) of a particular magnitude. The power of a study is thus not a single number, but a function describing how sensitively the proposed inferential procedure responds to the presence of the true (but unknown) effect over the plausible range of magnitude for that effect. Conventionally, 80% or 90% is considered acceptable power. The investigator’s obligation is not to give a blanket promise of statistical significance; rather it is to demonstrate that the proposed study, taking into account all its features including experimental design, outcome measure, strength of effect, analytic method, and sample size, will have an 80% or 90% chance of attaining statistical significance if the true effect is large enough to be clinically important.

One option for power calculations is to display the entire power function (power vs. magnitude of effect), showing that effects of low clinical significance may not be detectable, but those large enough to be interesting will carry a high probability of detection (statistical significance). A second option is to designate a certain “minimum interesting difference,” i.e., a desired detectable effect or threshold for clinical significance, and demonstrate that the power is adequate to detect that magnitude of effect. A third option is to fix power at 80% or 90% and calculate the magnitude of effect that is detectable with that level of power. In each of these modes of presentation, the investigator implicitly acknowledges that effects smaller than a certain magnitude may go undetected without serious regret.

The power calculation should be provided for the primary analysis of the primary endpoint at a minimum, and optionally for the more important secondary analyses (subgroup effects, interactions) and secondary endpoints. The calculation should be based on formulas pertinent to the intended method of analysis, including the specified critical p-value, citing if possible the pertinent prior data and mathematical formulas.

Example:

- For the Student 2-sample t-test, the detectable difference in weight loss between N intervention subjects and M control subject is given by $s \times (1/N + 1/M)^{1/2} \times (t_{\alpha/2} + t_{\beta})$, where s is the standard deviation of weight loss among children and $t_{\alpha/2}$ and t_{β} are Student deviates for the Type I and Type II error rates respectively.
- The formula indicates that with a 2-sided p-value $\alpha=0.05$, our sample of 300 intervention and 200 control children will provide 80% power ($1-\beta$) to detect a difference of 2.8 kg.
- This detectable effect is smaller than the 3.2 kg achieved by Aaron et al. (1998) in obese children of comparable age, using a less intensive intervention than the one we propose. The JCIF Consensus Panel (2010) designated any short-term weight loss below 3 kg to be clinically insignificant. We believe therefore that the proposed study is sufficiently powered to demonstrate the anticipated result, or any result that is clinically important, with statistical significance.

Timeline

A calendar should be provided for the entire study showing the major phases of the protocol and illustrating their sequence, duration, and overlap. Typical blocs are the following:

- Development of final protocol and manual of operations
- Training
- Screening and enrollment
- Treatment period
- Follow-up period
- Analysis, publication, and close-out

Organization chart

For large studies particularly, an organization chart is a valuable tool. This chart should have the Principal Investigator at the top, representing the fact that the PI is responsible for all aspects of the study. Below the PI is the line staff responsible for study operations. Other branches might represent co-investigators, remote sites, clinical contributors, specialty teams (laboratory, diet assessment, pharmacy), a Steering Committee and subcommittees, NIH project office, and advisory boards. The chart serves to catalogue the study personnel and delineate their functions, responsibilities, relationships, and channels of communication.

Human Subjects

Risks

Risks to human subjects vary over a wide range. It is never plausible to state that a clinical protocol carries no risk, no matter how seemingly benign; ‘negligible’ or ‘minimal’ risk is the minimum. Minimal risks include the slim possibility of infection from a routine blood draw; the stress involved in interrupting routine clinical care to answer a questionnaire; or the incidence of a rare side-effect of a common medication. Severe risks include toxicity of trial medication or adverse outcomes of surgery. A clinician PI will normally be able to list and characterize the pertinent risks for his/her protocol; if not, appropriate models can often be found in a colleague’s protocols or in CRC archives.

The protocol should detail how and why subjects are exposed to risk, whether the risks are associated with the hospital environment, the home, or the community at large, and what agents or materials if any represent a source of risk.

Protection against risk

A clinical investigator, having listed risks of various types, is responsible for anticipating and guarding against them to the greatest degree possible.

The first important aspect of protection is looking out for the subjects’ interest during their progress through recruitment and informed consent. Although recruitment may be covered thoroughly in the Research Plan, here the narrative should emphasize the human-subjects aspect rather than the logistics of the process. Important issues include

- The subject’s competence to consent or assent, as a function of age and mental status
- Assurance of the subject’s comprehension of any medical or technical language concerning the study
- Measures taken to avoid coercion in the context of medical care, including disproportionate compensation
- Mechanisms for assuring subjects that they are free to withdraw from the study at any time, regardless of the amount of data collected, and without compromising their medical care
- Formalities of documentation required in the granting of consent, such as witnessing the signature, keeping the signed documents on file, and providing copies to the subject

Additionally, this section should describe a variety of active measures designed to protect against risk. Examples include the following.

- Prevention of accidental exposure or injury;
- Confidentiality of personal information, records, and identifiers;
- Provision of contact information to the subject, including telephone numbers for questions about the protocol (typically the PI or project coordinator) or one's rights as a research subject (preferably an institutional representative not associated with the project);
- Procedures for reporting, recording, and reviewing adverse events, whether or not directly due to study activities;
- Membership, mandate, and meeting schedule for data and safety monitoring committee or board.

Many of the above items will necessarily recap material included earlier in the protocol. They are nevertheless worth repeating in this context for completeness and for ease of reference.

Benefits

Risk to human subjects can only be justified in light of potential benefits, either tangible or intangible.

This section should first detail direct benefits to research subjects. These may include

- Potential for cure or alleviation of disease
- Improvement in quality of life
- Enhanced knowledge or skill pertinent to health
- Compensation, as cash payment or equivalent (e.g., gift cards)
- Reimbursement for expenses of participation, such as travel, parking, or time lost from work

If no direct benefit to the subject is anticipated, this section may point out the social benefit that will be gained from advancing scientific knowledge or clinical practice as a result of the findings, possibly contributing to future treatment of others with similar conditions. If the potential advances are medically far-reaching or widely applicable in the public-health sphere, this aspect of benefit should be emphasized.

An effort may be made here to justify the ratio or balance of benefit against risk, particularly if the risk is more than minimal.

Inclusion of Women, Minorities, and Children

An explicit declaration should be made that women, racial and ethnic minorities, and children will be appropriately represented in the study sample, unless there is some scientific reason for

restricting the sample. Many Children's Hospital protocols are of course directed exclusively to children, for the reason that prior studies included only adults and medical understanding of the pediatric condition is therefore imperfect. Some studies concern female physiology alone. Some studies of genetics or medical services in underserved populations necessarily focus disproportionately on a particular racial or ethnic group. Any such reasons for a non-representative sample, whether implemented by design or incurred *de facto* and no matter how 'obvious,' should be thoroughly explained.

Literature Cited

The Literature Cited section should list, in a format similar to that used in journal articles, the sources cited throughout the protocol. Most citations will derive from the Background and Significance, where prior work in the field is summarized, and from Measures (under Research Plan), where assays, measurement tools, questionnaires, and other methodological sources are specified. Like the text itself, the reference list should be thorough but not exhaustive and serve only to document what is directly pertinent to this protocol. The PI's and co-investigators' own publications may be listed here if they fit the criteria of direct pertinence.

Related Content

- NONE

References

Guidelines for Writing Research Protocols. <http://ohsr.od.nih.gov/info/sheet5.html>. National Institutes of Health, Office of Human Subjects Research, 2002.

Piantidosi S. Clinical trials: a methodologic perspective. New York: John Wiley & Sons, 1997, pp. 92-100.

Hulley SB, Cummings SR, Browner WS, Grady D, Hearst N, Newman TB. Designing Clinical Research, Second Edition. Philadelphia: Lippincott Williams & Wilkins, 2001.

Inouye SK, Fiellin DA. An evidence-based guide to writing grant proposals for clinical research. *Ann Intern Med* 241:274-282, 2005.

Document Attributes

Title	Writing a Study Protocol – RPG-04		
Author	Henry Feldman, PhD	Date of Origin	August 2008
Reviewed/ Revised by	Ellis Neufeld, MD, PhD	Dates Reviewed/Revised	08/28/13
Copyright	©Boston Children’s Hospital, 2013	Last Modified	08/28/2013
Approved	<u>SIGNATURE ON FILE</u> Stavroula Osganian, MD, ScD, MPH Co-Chief, Clinical Research Center	<u>SIGNATURE ON FILE</u> Ellis Neufeld, MD, PhD Associate Chief, Division of Hematology/Oncology	

Disclaimer: Should Hospital and CRC policies conflict, Hospital policy will supersede CRC policy in all cases.

