The University of North Texas Health Science Center
Texas College of Osteopathic Medicine
Department Of Psychiatry, Behavioral Health And Neuroscience
Alan L. Podawiltz, D.O.,M.S.,FAPA
Chair and Assistant Professor

RESEARCH MANUAL
A PRIMER FOR BASIC RESEARCH COMPETENCIES
AND RESEARCH PROJECTS

des Anges Cruser, Ph.D., MPA
Associate Professor
Research and Education Director

Updated April 29, 2008

RESEARCH MANUAL
DEPARTMENT OF PSYCHIATRY, BEHAVIORAL HEALTH, AND NEUROSCIENCE
UNTHSC - TCOM

TABLE OF CONTENTS

FORWARD

This Research Manual for the Department of Psychiatry, Behavioral Health, and Neuroscience is provided as a guide for resident physicians, faculty, and staff of the department.

This manual has been reviewed and approved by the Education Committee comprised as follows.

Scott Winter, M.D., DFAPA, Associate Professor & Vice Chair for Education UNTHSC, Director, Psychiatric Residency Program, JPS Health Network

Linda Holding, Psychiatric Residency Program Coordinator

Alan L. Podawiltz, D.O., M.S., FAPA, Chair & Assistant Professor, UNTHSC, Deputy Director, Psychiatric Residency Program, JPS Health Network

James Witschy, M.D., Associate Professor & Clerkship Director, UNTHSC

Helene Alphonso, D.O., Chief Resident for Academic Programs

Lori Shahidi, D.O., Chief Resident for Clinical Services

des Anges Cruser, Ph.D., MPA, Associate Professor, TCOM Director of Education and Research and Psychology, Department of Psychiatry UNTHSC

Lily Wong, - BSN, MHSM, CNAA, Director, Psychiatry Services, JPS Health Network

Dr. Cruser is the author of this manual. She develops and monitors the research education program, and oversees all research in the department.

Resident physicians in the Department of Psychiatry, Behavioral Health, and Neuroscience are assigned research mentors. All residents, beginning with the incoming trainees in the 2006-07 academic year are expected to conduct an independent mentored research project. In addition to opportunities to collaborate with psychiatry and psychology faculty in the department, a resident may collaborate with a doctoral health psychology student, a graduate student in the physicians' assistants (PA) program, or with faculty in the basic sciences departments or the school of public health at the university.

Research didactics are provided throughout the academic year.

The Guidelines section of this manual provides information about on-going individual and collaborative research in the department. All residents are expected to present the findings of their research at both the Research Appreciation Day at UNTHSC, and at the Research Day at the JPS Health Network. Guidelines for posters, for IRB protocol proposals, and for the IRB or IACUC procedures are provided in the Guidelines section of this manual.

Research design and methods texts are maintained in the library of the JPS Health Network and the UNTHSC Department of Psychiatry offices available for check-out to aid in the development and conduct of their research projects.

The Department of Psychiatry has existing data bases. The Division of Psychology has existing data that can be analyzed or examined to develop a research question. Make use of existing sources of information, but do not overlook, the possibility of conducting a longitudinal study of change in a group of patients to which you have access and in whom you have interest.

FEASIBILITY CHECKLIST

CHECKLIST FOR DETERMINING THE FEASIBILITY OF A RESEARCH PROJECT

The following questions will guide your thinking about the feasibility of a research project. These are not the only questions you should ask. These questions are a beginning. You should review these questions by yourself and again with your research mentor.

A. LITERATURE AND MENTORS
- ❑ Do I have some ideas from reading research articles or clinical case studies?
- ❑ Do I have some ideas from my clinical experience?
- ❑ Do I have a mentor who is conducting research in an area of interest to me?
- ❑ Has there been an idea generated by a journal club discussion/presentation?

B. WHAT KIND OF RESEARCH SHOULD I DO?
- ❑ Epidemiologic study?
- ❑ A clinical trial? (Use an experimental intervention and a control group?)
- ❑ A case control study?
- ❑ Retrospective cohort?
- ❑ Cross sectional?
- ❑ An arm of an existing research project?
- ❑ Replication of a study?
- ❑ Basic science bench research?
- ❑ Other? _____

C. TIME AND RESOURCES?
- ❑ Do I have easy access to data or patients?
- ❑ Who are the available experts or mentors with the time and experience to guide and support my research?
- ❑ What will I need and expect from that person or persons?
- ❑ How long do I have, or want, to take to conduct this research?
- ❑ Other: _____

D. DATA ACCESS
- ❑ Am I generating pilot data?
- ❑ Are pilot data available that can guide my research design development?
- ❑ Is there existing data that can be analyzed in a different way?
- ❑ Is there existing data in an electronic form that I can analyze such as a national survey or other national or local data base?
- ❑ Other _____

E. END POINTS AND PRODUCTS
- ❑ When will I publish or present the findings (Research Appreciation Day)?
- ❑ What is the benefit or value to the field of psychiatry, to our existing system of care, or to health care consumers?
- ❑ Is there a cost-benefit issue to address for this research project?
- ❑ What do I expect of myself during and after the project?
- ❑ Other _____

FEASIBILITY CHECKLIST

F. OTHER RESOURCES

❑ Do I have a biostatistician available to consult on data management and analysis?

❑ Am I trained in Human Subjects protection issues, IRB procedures, and/or have access to a clinical research coordinator to assist in data collection, and study reports/reviews?

❑ What expertise do I need to access?

❑ Do I need funding to conduct this research project?

❑ Other _____

G. MY OWN SKILLS

❑ What other knowledge do I need to complete this research project?

❑ Do I have or can I acquire the basic skill and knowledge to accomplish my goals for this research project?

❑ Am I committed to doing this research project?

❑ Other _____

The next step is the design phase and assistance with this can be found in the Worksheets at the end of Chapter I and in other source materials.

Note: The most often cited reason for not completing a clinical trial or epidemiologic study is lack of an adequate number of subjects or records. Rarely do researchers report that all subjects enrolled completed a study. Inclusion and exclusion criteria used to match subjects, narrow the focus, or control for sources of error, all may limit the total number of subjects or records one can include in a study, particularly a study done in a limited amount of time.

A research study is a significant investment of your time, energy, and intellectual capital. Try to develop a project that is realistic, but at the same time do not limit yourself by what you think might be problems or obstacles. Always look for a way to accomplish your goals.

Chapter I: Introduction to Research Designs

INTRODUCTION

This section describes several different types of research designs. For each design we have listed some of the benefits and limitations. We have also provided some of the key words essential to your understanding of these designs.

Knowing what types of research designs are most commonly used in Psychiatric and social & behavioral research, as well as other fields, is the first step in beginning to understand and conduct biomedical research. This will assist you in reading published research literature, and understanding how to craft a study of your own or with a mentor. Question, design, and methods are intricately linked with each other. The question drives the design and the design drives the methods section and data analysis.

Key Concepts in Research Design

Random *selection* (sampling) – in choosing participants for your study, each person in the population has an equal chance of being selected.

Random *assignment* – when assigning participants to the various groups, each person has an equal chance of being placed into one particular group.

Validity: There are many types of "validity" in research design.

Internal **Validity** – control for all outside influences that may influence the characteristics or status of the groups being studied, except of course the variable of interest. True internal validity has no extraneous factors affecting it. This is difficult in the real world. For example, can you tell a patient not to get any other treatment except the study treatment in a study of manual medicine treatments? Perhaps; but can you tell a patient not to take any other medications in a pharmaceutical study? Yes. And, for research that is in-vivo, in real life situations, there are ways to control for extraneous variables. Those variables can become part of the study such as medication use.

Random assignment helps control factors that might contaminate internal validity but nothing controls for it completely. Causal inference from a study finding is typically NOT DONE. Is there anything you can think of that causes something to happen and is accounted for as the ONLY thing that makes "it" happen?

External **Validity** – determines whether or not the research can be generalized to the larger population. External validity depends on the power of the study, the method, and criteria for selection of records or participants. It may also depend on the measurement instruments selected. Some instruments are designed to measure pain and others are designed to measure back pain specifically.

Reliability – Can your results be repeated by other researchers using your same methods?

Experimental Design

This would be a research project in which you apply a treatment or intervention or you observe a population under a certain number of controlled conditions. For an experimental design you can conduct pilot research with a small number of human subjects. This type of research is particularly valuable to discover and uncover whether a specific medicine, instructional material, therapy, or other manipulated condition changes outcomes of interest.

**Four Common
Research Designs**

1. Cohort Study
2. Correlational Study
3. Cross Sectional Study
4. Case Control Study

Each of these research designs can be used by a resident or a student. These studies are often exploratory or developmental, paving the way for another resident or student to replicate it or develop an arm of it, or modify it, thus improving the design based on the study's identified limitations. These studies are "publishable" as posters or short manuscripts.

A study involving human subjects requires Institutional Review Board approval, and some method of addressing consent to participate. In some cases research will require something called a "HIPPA" waiver. The Department has a Research Project Manager Angelita Trevino, M.S. who is the IRB and IACUC liaison, and will assist you and your mentor with these procedures. Research involving animal models will require "IACUC" review and approval. Ms.Trevino provides training in this and IRB processes, and consultation for your project.

---

1. Cohort Study

What is a cohort study? A cohort is a group of individuals studied over time.
A cohort study compares the incidence of a condition in a group at a defined baseline and at an endpoint.
Two groups or sub-groups are usually matched on some characteristic (gender, age, diagnosis, exposure, or possessing some other characteristic such as an event they have experienced in common or a genetic marker).

Methods and data management
The types of variables in a cohort study data may be nominal, ordinal, interval or ratio.
Simple studies may use t-tests, chi-squared analysis, or step-wise linear regression analysis
Complex studies may use survival analysis or odds ratios

Limits
Findings of these studies have limited applicability to other populations
There is risk of a Type II error
A cohort study may require a long period of time such as a year, or it may be a retrospective cohort study. It may require financial resources, and attrition may be a concern.

Strengths
Leads to larger research
Theory building
Exploratory and Developmental
Can provide insight to risk factors for conditions
May explain how to achieve desired clinical outcomes in specific groups

---

2. Correlational Studies

In a "correlational" study, you would compare two or more variables (factors or characteristics or qualities) in one group. For example, "What factors characterize persons who are readmitted to the emergency center compared to those admitted only once?" or "Is there a relationship between a history of jail incarceration and utilization of emergency services?" You could ask whether frequency or type of interpersonal violence that occurs in a person's life impacts clinical outcomes. Correlational studies can utilize existing data, or within an experimental study, the researcher may investigate the relationship between multiple conditions or outcomes within groups.

Multiple linear regression is one of the statistical tools that can be used to predict the values of one

variable based on its relationship to multiple other variables. The simplest correlational study would use Pearson's correlation coefficient.

For Correlational Study Designs:

Hypothesis or Question
    Helps to determine whether one condition is related to or associated with another
    Shoe size and earned income

Group
    Only one: Precise definitions of the characteristics of interest (may have sub-groups)

Methods and data management
    Scatter Plots are useful to explain the shape of the data
    Linear and curvilinear relationships can be found such as in the Circumplex model of family systems

Limits
    Difficult to generalize
    Type I error may be a risk in very large sample sizes

Strengths
    Leads to larger research
    Theory building
    Exploratory Developmental
    Relatively easy to conduct

---

3. Cross Sectional Study

A cross sectional study design is used to describe a population on several dimensions. It captures a condition at a specific moment in time.

You can determine or estimate the prevalence of a condition in a population. You can use this to generate questions for future research. You may have a clinic that began using Cognitive Behavioral Therapy last year and you want to look at the records to describe changes in certain test scores the therapist recorded. Or you might study the records of a group of adolescents who were admitted to the juvenile justice system for family violence to determine the case disposition based by academic performance in the record at the time of the admission. If there are adequate numbers to power a study you could use exploratory data analysis tools to richly describe the population and available clinical outcomes data by disease severity. This type of study also uses existing data or may take a sample of subjects in a given setting and take one measurement to describe a condition at a given point in time.

---

4. Case Control Study

A case-control study is basically a case report comparing those with and those without a particular condition. If you can get enough information about patients in a psychiatric clinic exposed to a particular event and patients in another clinic not exposed to that event you can examine use of crisis services prospectively from the time of the event. You can compare quality of life for a clinical population with a quality of life indicator for a non-clinical population. Headaches would be another possible topic if you can match cases of persons with and without specific headache complaints following a trauma, you could examine the use of pain relievers post-trauma or the number of visits for other health complaints.

These studies are typically of not much value unless the condition is rare or the treatment is experimental or novel so that you could match two clinical populations who did and did not take a particular medication and compare them on an outcome. If you had access to records from two clinics you could compare patients receiving a particular treatment with those not receiving that treatment.

Using/Selecting Questionnaires

When conducting a study using a questionnaire to measure a condition or outcome, consider existing instruments. It is better to use an existing instrument in a novel way, with a new population than to develop your own questionnaire. You can modify an existing survey, but not an instrument used to detect symptoms or diagnose a condition. If you are recording information never collected, consult with someone who knows how to construct interview questionnaires. The Department can provide consultation for you in this area, as can and should your mentor.

Seek out expert advice and guidance. There are multiple sources of information on designing valid and reliable questionnaires. Sampling is of critical importance and bias is another concern in the design and use of questionnaires.

Surveys

You can conduct a survey with a group. For example you can distribute a survey at a clinic and provide a box for deposit of the completed survey and still provide cash compensation.

Mail surveys can be inexpensive but labor intensive. You should always include a pre-paid return envelope. You will need to get consent procedures approved by an IRB.

On-line questionnaires are more efficient and the university has resources to assist with that through your research mentor.

Limitations: response rates are notoriously low. Ensuring understandability and readability is critical and so you must pre-test the questionnaire – any questionnaire on a group representative of the population of interest.

You can conduct a personal one-on-one interview which is time consuming but also inexpensive. You can use the telephone for this or recruit subjects – you can purchase lists of names and addresses.

Collaboration opportunities

There are opportunities to collaborate with a Health Psychology student, a Physicians Assistant Studies student, a specialty medicine physician such as a trauma physician, in addition to the on-going research in animal models of stress and behavior, dual-diagnosis research in neurosciences and community services, existing data analyses such as psychiatric emergency center services data base, Alzheimer's disease, child and adolescent mental health, and perinatal depression.

Texts

Motulsky, H. Intuitive Biostatistics (1995) Oxford University Press.

Gould, J.E. Experimental Methods for the Behavioral and Biological Sciences.

Peat, J., Barton, B. Medical Statistics: A Guide to Data Analysis and Critical Appraisal (2005) Blackwell Publishing.

Rouane, J. M.. Essentials of Research Methods (2005) Blackwell Publishing.

Mitchell, J.E., Crosby, R.D., Wonderlich, S.A., and Adson, D. E. Elements of Clinical Research in Psychiatry (2000) American Psychiatric Press.

Itulley

Neuteus

INTRODUCTION

What is the purpose of the Background and Significance section of a research proposal and how and why do I review the literature.

The purpose of the Background and Significance section is to describe the scientific knowledge, research recommendations, and gaps in that knowledge as portrayed in published peer reviewed or text-book materials relevant to the topic of interest.

This section provides the rationale for the proposed research and summarizes the expected contribution your research will make to the field.

In this section you will:

- Make a compelling case for your proposed research project. Why is the topic important? Why are the specific research questions important? How are the researchers qualified to address these?

- Establish that you are knowledgeable in the topic.

- Avoid outdated research.

- Use citations not only as support for specific statements but also to establish familiarity with all of the relevant publications and points of view, not just those that support your hypothesis.

- Make sure the citations are specifically related to the proposed research. Cite and paraphrase correctly and constructively.

- Highlight why research findings are important beyond the confines of a specific project, i.e. how the results might be applied to further research in this field or related areas.

- Stress any innovations in experimental methods (e.g., new strategies, research methods used, interventions proposed.

The Art of the Literature Search

The literature search is the process of finding research reports related to your research topic. Much information can be gained by doing this preliminary work and there are various ways to perform it. This section reviews some important reasons for conducting a literature search and outlines methods for four possible approaches to a literature search.

Why Perform a Literature Search?

1. To review the existing published research on the topic
2. Discover how and why the topic has been studied
3. Determine what has changed between the first and most recent published information
4. Focus your research by reviewing the methods and definitions used in prior studies
5. To learn the names of the prominent researchers in the area of interest
6. To help shape study design by learning from and improving on prior studies
7. Identify sources of valid, reliable measurements and procedures and invalid and irrelevant measures as well.
8. Determine if the study is important and useful

What is a review of the literature?

"A literature review is an account of what has been published on a topic by accredited scholars and researchers. The purpose is to convey to your reader what knowledge and ideas have been established on a topic, and what their strengths and weaknesses are. As a piece of writing, the literature review must be defined by a guiding concept "

Skills you need

- Ability to scan the literature efficiently, using manual or computerized methods, to identify the materials you will use and cite

- Ability to apply principles of analysis and critical thinking to identify valid, reliable research reports and identify and describe the gaps in the literature pertaining to your topic

- Organize your material around and related directly to your topic

- Synthesize results into a summary of what is and is not known

- Cogently discuss areas of controversy in the literature

- Formulate questions that need further research

Checklist Of Questions To Ask Yourself

☐ What key words have I found in the first three articles I have read?

☐ If these are the most recent articles on the topic, what references have they used?

☐ Do I need to include theory?

☐ Are there any existing "reviews of the literature" on my topic?

☐ Do I need to include Text Books in addition to published peer review articles?

☐ What unpublished material may help my review?

☐ Have I checked the literature cited by authors to ensure I have covered a broad scope? (when you begin to see citations repeated in the articles you review, you may have exhausted the field.)

☐ I have over 100 articles, how can I address this large scope?

☐ Have I used a written list of the specific outcome measures or population characteristics, or other criteria to cull out articles that are extraneous?

☐ Do I have a checklist to assess strengths and weaknesses of the materials?

☐ Have I cited and discussed studies contrary to my perspective?

☐ Will the reader find my literature review relevant, appropriate, and useful?

Questions To Use In Reviewing Articles And Texts

- Has the author formulated a problem/issue?
- Is it clearly defined? Is its significance (scope, severity, relevance) clearly established?
- Could the problem have been approached more effectively from another perspective?
- What is the author's research orientation (e.g., interpretive, critical science, combination)?
- What is the author's theoretical framework (e.g., psychological, developmental, feminist)?
- What is the relationship between the theoretical and research perspectives?
- Has the author evaluated the literature relevant to the problem/issue?
- Does the author include literature taking positions she or he does not agree with?
- Does the author provide complete information or how to access more information on the basic components of the study design?
- Are the measurements valid and reliable for this topic?
- Is the analysis of the data accurate and relevant to the research question?
- Are the conclusions validly based upon the data and analysis?
- Is there scientific evidence underlying the reasoning, or is the author promoting a perspective that is theoretical only?
- How does the author structure the argument? Can you "deconstruct" the flow of the argument to see whether or where it breaks down logically?
- In what ways does this book or article
  - contribute to our understanding of the problem under study,
  - provide information for practice?
- What are the strengths and limitations?
- How does this book or article relate to the specific thesis or question I am developing?

How to Conduct a Literature Search

A well-organized, goal-oriented approach to a literature search will save a lot of time in the library. Draft a literature search outline with the questions pertinent to your research. Faculty members, biostatisticians, and librarians can be of great assistance in this area. There are many different ways to conduct a literature search. A combination of the following four approaches can be used to avoid potential limitations in the search.

Ancestry approach
Begin with the most recent reference and trace its bibliography
Pros: describes evolution of the idea or technique, places study in an historical context
Cons: unveils many older references vs. more recent references

Descendency approach
Locate the classical reference and identify each subsequently published study that refers to the classical reference (via Science Citation Index or Social Science Citation Index)
Pros: identifies a large number of articles in sometimes unlikely journals
Cons: unveils articles of peripheral issues, may bypass opposing arguments in search

Database search (e.g. MEDLINE)
Databases cross-indexed by topic, content, author, year of publication
Librarian-assisted
Pros: intelligent control of study with manageable number of references
Cons: you could miss out on an original reference or rare, specialty literature and text book sources

Contact an author directly
Contact an association to learn of contacts
Pros: primary source, state-of-the-art information, may provide additional contacts, may identify students willing to assist with the study
Cons: can be restricted to a single point of view, information may not have been subjected to peer review

STEPS These are recommended steps (you may find you need some sub steps) write them in.

Select your source

Use multiple different key word variations (previously mapped out with your mentor or a librarian – based on your topic – or use words from articles or texts that you already have).

Read the abstract on line (print abstracts if needed but don't print a lot of articles until you are sure you want them printed – expensive – and wasteful in the end).

If needed, identify more precise, specific keywords and language or phrases.

Enter new keywords or new authors or follow links provided on the abstract pages.

Again, Read the abstracts critically before you "take" the article.

Select articles carefully, limiting yourself at first to only those you know you can read and understand.

Read the articles – Mark them with highlighters and note your questions or reactions on the paper – margins or on the back using as much paper as you can.

Put the ones you will keep in a three ring notebook and put tabs with the last name of the first author and the year of publication and key words if there is room.

Always make a list by last name of first author to put in the front of the binder for quick reference.

If you organize them electronically label them in the same way – last name of first author, date of pub, brief title.

Look at the references cited by the authors – you may want to read SOME of them also –.

Make your trail backward in time to trace the history and development of the topic.

Trace back to the seminal research in the area, if appropriate.

Look for meta analyses and focused or critical reviews of the research you are doing.

Include foreign articles – you can get translations if you need – and software is available to translate most languages .

INCLUDE research that you do not agree with or if you are studying women look at the same literature for men as appropriate -.

Begin writing your background and significance right away = it WILL change over time but your argument for your research project will develop, emerge, and become more logical and stronger with practice.

Final Notes

There is value in the old "index card" system of keeping notes on articles or use excel files so you can search later by key words from your own literature data base.

EXAMPLE

Topic

Author

Title

Journal, year, volume, and pages

Aim of article/research

Primary question/hypotheses

Results

Methods

Future research questions recommended

Comments: unique features, stimulating facts

A literature review is a piece of **discursive prose**, not a list describing or summarizing one piece of literature after another. It's usually a bad sign to see every paragraph beginning with the name of a researcher.   Organize the literature review into sections that present themes or identify trends, including relevant theory. You are not trying to list all the material published, but to synthesize and evaluate it according to the guiding concept of your thesis or research question.

If you are writing an **annotated bibliography**, you may need to summarize each item briefly, but should still follow through themes and concepts and do some critical assessment of material. Use an overall introduction and conclusion to state the scope of your coverage and to formulate the question, problem, or concept your chosen material illuminates. Attempt to group the literature under important medical or clinical sub-sections. This process will help you indicate comparisons and relationships.

CHAPTER III: THE RESEARCH PROPOSAL

INTRODUCTION

This Chapter "The Research Proposal" includes one section for each part of a traditional research proposal. It is based on the NIH-Gold Standard for research proposals of all kinds.

No matter what type of research you chose to conduct you will need a research proposal. This is the document used to communicate your plan for a research project to all relevant audiences.

The main sections always include at least the Specific Aims, the Background and Significance, (your own, if any, Preliminary Studies), and Methods.

For research using any information from or about human subjects or about animal use and care policies, check with the resources provided in the Department first.

Multiple sources of guidance on the content, format, and the process of developing each section of a research proposal for small and large studies are available in textbooks, on the world-wide-web, and at your institution. As a student you should learn to seek out these sources and use them to their maximum benefit. Research project development is not a solitary process. Although writing a research proposal is ultimately the responsibility of one person, it always benefits from team, consultant, mentor, and reviewer input to shape a well crafted proposal.

This principle applies also to the logistics of the research. Great research ideas and proposals must be feasible. Research must be "doable" in the real world regardless of the topic or the research design. Every stage and step in any research project, no matter how small or how large, must be carefully developed, considered, and communicated within the proposal so that it can be translated into an operational plan that will ultimately serve the goal of the project from beginning to end.

Objectives of a Research Proposal

A research proposal has two primary objectives.

1. To define and describe the research project
2. To establish the plan for the conduct of the project

Research with humans cannot be conducted without IRB approval, exempt, expedited, or full board review. Your research proposal will either be attached to your IRB forms or proposal, or it will be used to copy and paste into relevant sections of an IRB proposal. This depends on what your institutional guidelines require.

Organization of this Chapter

This chapter is organized around the major components used by the National Institutes of Health (NIH) for research proposals.

For more information, refer to the instructions for PHS 398 forms at
http://grants1.nih.gov/grants/funding/phs398/phs398.html   You may also look at a specific program announcement by any potential sponsor, including your own school. This chapter does not cover all possible information you may want to develop a research proposal. It includes the essentials. You may have unique needs. Your university or a potential grant agency may have unique requirements. Follow the guidelines most appropriate for your research/application.

This chapter provides
- A description of the purpose of each section of a research proposal,
- Prompts to guide the writer and
- Examples of strong and weak material.

Full proposal samples are available from the Department of Psychiatry Research Division with the permission of the Principal Investigator.

GOOD HABITS AND TIPS

You should be able to create a research proposal for pilot research in no more than about five or six pages. A research **grant** proposal would be longer.

Appendices are discouraged for several reasons, one of which is that if a researcher cannot say in the allotted pages what she wishes to communicate to the audience, revisions are needed. Always ask about unnecessary or necessary appendices, especially if the instructions are unclear or the appendix is required or requested by the sponsor.

All data collection forms (called DCFs – this acronym will be used again in this manual) or questionnaires you plan to use in the study must be thoroughly described, cited, justified, and when permitted or required, the DCFs will be an appendix to the Research Plan.

If a data collection form (DCF) is described in the body of the proposal, the form itself may need to be in an Appendix. For an IRB proposal you will normally include all DCFs. This is the way the IRB will determine the level of review required and will see what information you plan to collect from records or from individual subjects.

Some tips for conserving space in your proposal include such strategies as:
- Use a single space rather than double spaces between a period and the beginning of a sentence (yes this IS permissible),
- Use tables instead of vertical lists of items,
- Limit modifiers, and
- Choose words and sentence structures carefully.

Parsimony is important in scientific writing. Parsimonious means to be unusually or excessively frugal; and is herein applied to the use of words. Attempt to utilize the simplest assumption in the formulation of a theory or in the interpretation of data, especially in accordance with the rule of Occam's razor. Occam's razor is also called the *principle of parsimony*. These days it is usually interpreted to mean something like "the simpler the explanation, the better" or "don't multiply hypotheses unnecessarily". Select cautiously from among theories with equal explanatory power, and when giving explanatory reasons for something, avoid positing more than is necessary.

Good Habits For High Quality Scientific Writing

- Read other scientific material or proposals with a critical eye.
- Use active voice vs. passive.
- Use "this study" or "this proposed study" in reference to your proposal, and "that study" or "their study" in reference to another you are citing.
- Do <u>not</u> use the term "etc." Be precise. If there is too much to list in a sentence use a table or refer to a list elsewhere. On occasion you can say "such as" or "including but not limited to". Do not make the reader guess what you mean or have the impression that you are vague.
- Use the terms "valid" and "reliable" wisely. There are many different types of validity and reliability. Refer to a glossary of research terms. Ensure you are correctly using these and other research terms.
- A finding is either significant or not significant. Some authors erroneously use modifiers such as "highly" or "greatly". This is not sound research language.
- Always use correct grammar. Check and have others check your spelling. Do not rely solely on the computer for spelling or grammar checks.
- "Account for" means that a variable or a factor explains a portion of the effects of the independent variables on the dependent variable in a correlation analysis. For example, "age" may account for a portion of the difference between the time men and women take to fall asleep in a study of sleep aids.
- You may indeed make good use of the first person such as I or we.
- Keep it simple and logical.
- Get at least two detail-oriented individuals to read the proposal for you.

THE RESEARCH PROPOSAL
AND THE INSTITUTIONAL REVIEW BOARD (IRB) PROPOSAL

There is a difference between the Research Proposal and an IRB proposal. The IRB functions to protect the subjects in a study, and to protect the organization (the university, the hospital, and such) by ensuring ethical and legal practices in all research conducted by its faculty. The research proposal usually accompanies the IRB proposal that focuses on inclusion/exclusion of protected groups of persons, risks, and protections for subjects in a study.

The Institutional Review Board (IRB) Proposal

The Department of Psychiatry has a designated liaison to the local IRBs and communicates with IRBs at other institutions. Contact the Department's liaison to help develop and review your IRB proposal BEFORE you go to the IRB staff.

One reason for this is to have the luxury of an informal pre-review performed so that you experience as few delays as possible in satisfying the IRB requirements. The typical IRB WANTS research to proceed, and has the responsibility to protect the institution and the research subjects.

THE OUTLINE

The main components of a research proposal are listed below. These are major section headings that reflect the gold standard endorsed by the NIH and most published texts on writing research proposals. These sections are to be followed in the order presented. Although each major section may vary in the order in which sub-sections are presented, there are essential, commonly accepted elements that should be included. If a sponsor or institution has no requirements, you cannot go wrong with this formula. It is endorsed by the finest texts in clinical and social science research in the world.

RESEARCH PLAN STANDARD OUTLINE BASED ON NIH FORMS AND FORMAT
    A. Specific Aims
    B. Background And Significance
    C. Preliminary Studies Or Research (Only Your Own)
    D. Research Design And Methods
    E. Human Subjects (followed by vertebrate animals)
    F. Literature Cited
        Appendices

Each of the above components is restated in the text of this material.

- A description of the component
- Key questions to guide your thinking
- Prompts to guide your writing
- Examples of strong and weak material

(NOTE: some examples are paraphrased to give a brief sense of the tone of the component and for brevity. Complete examples are available upon request.)

THE ANNOTATED OUTLINE

Margins and Fonts

You may use any font that is easily readable. The preferred fonts are Helvetica 12 or Arial 11. You may use either 1″ margins all around, or the NIH guidelines. NIH margins are ½″ all around and font is specifically recommended in the instructions. CHECK YOUR MARGINS!

A good font to use for the NIH proposal is Arial 11. You do not need to keep that size in tables or figures, but nothing is to be smaller than 10 point font. Also, using a consistent font type/style is helpful to the reader. You may bold or italicize for emphasis or separation of key points.

This is Arial 11. This is Arial 12. This is Times Roman 12. This is Bookman Old Style 12.

The Research Plan

A. Specific Aims

Use this section to establish and briefly describe the rationale for the study. For a Master's Thesis Research Proposal, keep it short – about two paragraphs. For NIH and other research funding proposals the standard is one page.

This section is where you present in a logical, organized fashion the thought process that leads the reader to the same conclusion that you make in your statement of specific aims. This must "grab" the reader, make the reader interested in reading further, and make the most parsimonious statement of your intent that you can possibly make.

Always present your specific aim(s) and question(s) or hypotheses on the first page. You will probably begin with a longer version of this section and then edit and reshape it to fit.

This section must always contain:
1. First a statement of the broad, long range goal of the proposal
2. A few statements that justify that goal and describe the focus or topic of the research
3. One or two primary aims or the primary and secondary aims of the research project that will lead directly into the primary research question or hypotheses for the proposed study

The hypotheses may be linked specifically with each aim, and you can restate each aim with its associated hypotheses in the Methods section.

Above all: KEEP THIS SECTION SIMPLE, STRAIGHT FORWARD AND LOGICAL. This is your chance to take the reader by the hand and lead him from nothing to the end of the page "Ah-Ha". The reader should be able to determine the likelihood of approving your proposal this early in the reading. All else will be measured and linked back to this page.

Lastly, you will usually find some "redundancy" in the proposal. This means you will repeat what is stated here later in the paper.

Prompts:
- Tell the reader what you are going to do.
- Tell the reader why.
- Tell the reader how you will do this.
- Tell the reader what you told them.

Key Question: What is the focus of the study?

Example

Strong: The specific aim for this pharmacological treatment intervention trial is to test the efficacy of quetiapine versus risperidone to reduce <u>drug use</u> in psychiatric outpatients with co-occurring bipolar disorder and cocaine dependence.

Weak: *This proposal seeks to describe how one drug reduces drug use better than another drug in a special population of people who use cocaine.*

Example

Strong: The goals of this research are:
1. Determine the prevalence of co-occurring psychiatric and substance use disorders among a sample of Caucasian, African American and Hispanic persons seeking health care treatment.
2. Examine and compare the prevalence of health and healthcare disparities and treatment efficacy among ethnic groups in persons with co-occurring disorders.
3. Assess the need for specialized health care and utilization among ethnic groups.

Weak: *The focus of this study is how certain drug therapies work to reduce cocaine abuse in people with a mental illness.*

Thinking Exercise: What are the strengths of this example?

How can you improve this example?

Does it say what you think the person really means?

Are there any errors in thinking in this example?

The field of research into treatment efficacy for persons with co-occurring bipolar and substance use disorders is not yet well established. Persons with co-occurring disorders are challenging to study due to the complexity of their co-morbid illnesses. However, by carefully taking into consideration a number of variables that can influence mood and drug use outcomes, unbiased data can be obtained. As the subjects studied in this kind treatment intervention trial can be severely ill, it may not be wise to design a placebo study as recruitment may very difficult and you put the study at risk for an increase in serious adverse events. Instead, a double-blind randomized comparison trial of two FDA approved medications – one approved to treat bipolar depression and the other approved to treat acute bipolar mania can be utilized to test the effectiveness that these two medications may have in treating bipolar depression, mania, <u>*and*</u> drug use.

Key Question: Why is this area the focus?

Example

Strong: Over 60% of persons with bipolar disorder have had a diagnosis of substance abuse in their lifetime. Up to 30% of bipolar patients with a co-occurring substance use disorder have cocaine abuse or dependence. Higher rates of hospitalizations and poorer psychiatric recovery are found in those with co-occurring bipolar and substance use disorders than in those with bipolar disorder alone. Treatments to improve mood while reducing drug use have clinical relevance in treating this patient population. However, data examining the efficacy of FDA approved medications to treat both mood and drug use in persons with co-occurring bipolar disorder with stimulant dependence are not available. Thus, we propose a double-blind, randomized study to compare the effectiveness that quetiapine and risperidone may have in treating both mood and drug use in this patient population.

Weak: About 1.6 million people in the U.S.A have a substance use problem that no one sees. Not getting treatment might include others with depression or other mood problems. No one knows whether atypical antipsychotics can work in this co-occurring patient population, so we propose a study to examine treatment efficacy.

Area of focus continued

Example

Strong: In 2003, the Texas infant mortality rate was 6.6 deaths per 1,000 births – higher than the Healthy People 2010 goal of 4.5 per 1,000 live births.[17] Among Texas counties, **Tarrant County** had a 2003 infant mortality rate of **7.5 per 1,000 births** – higher than that year's United States rate of 6.9.[17] We have not found any evidence to suggest that Tarrant County rates have even slightly improved over the last 4 years. Under- and uninsured women are more likely to be affiliated with a low income minority group with frequent exposure to stress, anxiety, and depression resulting in poor maternal and infant health. To avert these negative maternal, birth, and infant health outcomes, women need to be educated about the connection between the mind and the body. We propose to use Promotoras-Community Health Workers to educate under- and uninsured pregnant women about mind and body relationships. Promotoras-Community Health Workers are state certified promoters of health education who support and strengthen minority and underserved communities.[9,14] Promotoras-Community Health Workers (CHW) are culturally sensitive, using native languages and value systems unique to their communities – non-Hispanic Black and Hispanic communities alike.[9,14]

Key Question: **What therefore will I do?**

Example

Strong: The primary aim is to compare the efficacy of quetiapine versus risperidone to improve mood and reduce drug craving in persons with co-occurring bipolar disorders and cocaine dependence.

The secondary aim is to examine relationships between mood improvement and drug use in relation to quetiapine or risperidone use.

A third aim is to determine the length of time for expected treatment efficacy in relation to dosing.

Weak: This study will collect mood and drug use data from subjects in both medication groups and determine if and when improvements occur in mood and drug use.

Example

Strong: In ***Phase I*** we will provide mental health and substance abuse (MH/SA) certification training developed by the Texas Department of State and Health Services to Tarrant County Promotoras-CHW. Promotoras-CHW will then educate, and disseminate information about the mind-body connection (i.e. mental *and* physical health) to disadvantaged child-bearing women in the community. Finally, we will evaluate how well Promotoras-CHW disseminate their knowledge about the mind-body connection in their communities.

Weak: In this study our promotoras will train female subjects to use better health behaviors and we will measure whether this happens and favorably improves health outcomes for the mother and child.

Example

Strong: In ***Phase I*** we will provide mental health and substance abuse (MH/SA) certification training developed by the Texas Department of State and Health Services to Tarrant County Promotoras-CHW. Promotoras-CHW will then educate, and disseminate information about the mind-body connection (i.e. mental *and* physical health) to disadvantaged child-bearing women in the community. Finally, we will evaluate how well Promotoras-CHW disseminate their knowledge about the mind-body connection in their communities.

Weak: In this study our promotoras will train female subjects to use better health behaviors and we will measure whether this happens and favorably improves health outcomes for the mother and child.

Example

Strong: We will conduct a 20-week double-blind, randomized trial comparing quetiapine to risperidone to improve mood symptoms and reduce drug use in outpatients with co-occurring bipolar disorder and cocaine dependence.

Hypothesis 1. Outpatients receiving quetiapine will have a greater improvement in manic and depression symptoms and have less frequent cocaine use than those receiving risperidone.

Hypothesis 2. Outpatients receiving quetiapine will fewer adverse events than those receiving risperidone.

Weak: *This study will help determine if these hypotheses are true, that patients using illegal drugs will improve their mania and depression and use less cocaine, and that patients with less cocaine use will use quetiapine.*

B. Background and Significance

In the Background and Significance section of the proposal you need to describe the existing research literature or scientific body of knowledge that builds the rationale for your proposal. This is one of the most difficult sections to write succinctly, coherently, and powerfully. You will write and re-write it, and re-write it again. Frustration is normal. Your mentor may see something on the third edition she did not see on the first – that is how it evolves – that is why you will have milestones, product deadlines and regular meetings with key individuals to write your proposal – or later your report.

Discuss how the current scientific knowledge in your specific area of focus supports your hypothesis or reveals important gaps in the scientific knowledge that you will address. The Background and Significance section should be written in such a way that it leads the reader to agree that there is either evidence from previous research that supports the direction of your proposal or that there is insufficient relevant research on the subject although the area is an important one to study. This is the section to present the theoretical framework in which your research is grounded.

If you are replicating a study, this is the place to say whose study, why, and what the original study did, and how you will improve knowledge by replicating it. If you are developing a new instrument for measuring outcomes of an intervention or chemical changes from a laboratory experiment, make your case for its value in this section. You can divide this section into each major aim, hypothesis, outcome, or strategy that you will use by using separate headings such as: *Subsyndromal PTSD as a Clinical Construct*, or *Resilience as a Neurochemical Process*. Also don't forget to consider using existing data no one has mined or explored based on your hypothesis or question.

The Background and Significance section should logically lead the reader into your Methods Section and have them say "now I see" why or "oh wow," or "yes!"

Key Questions to guide your thinking are as follow.

- What is the scientific or clinical theoretical framework for this research?
- What do we currently know about this area?
- What are the strengths of the current body of knowledge?
- What are the gaps in the current body of knowledge?
- How will this proposed study build on the strengths, correct the limitations and fill the gaps (or begin to do this)?

Your Background and Significance section may be organized many different ways. A good rule-of-thumb is to use it to logically support each of your aims and hypotheses. Include information that refutes your hypotheses also, to show that you know about these and will address them. If you are writing about non-surgical interventions, discuss what is pertinent about surgery. If you are writing about women's health, discuss similar conditions in the male population and how women are different. Use comparison groups and related subjects to provide a full and rich picture to the reader.

This section is normally three to five pages in a funding proposal of 25 pages. For your thesis proposal it may be only two to three pages at most.

Background and Significance Examples

Example

On the next two pages we provide an entire Background and Significance section of a final draft proposal. It is quite strong but too long (how could you strengthen it and be briefer?)

B. Background and Significance

This is a proposal for a two-phase services and education demonstration project to a) train community Promotoras-Community Health Workers to educate disadvantaged child-bearing women about the impact of depression has on maternal, birth, and infant health outcomes; b) evaluate the effectiveness of the education program in a group of 100 perinatal women. This study will occur in an area plagued with disastrously high infant mortality rates. The target population is under- and uninsured women and their infants who live in this area of Texas.

While this proposal will not directly study infant mortality, it is the ultimate targeted problem for a services demonstration project to reduce the effects of depression on maternal and infant health. The devastation wrought by perinatal depression must be considered within the context of the "worst case scenario" that has existed in Tarrant County for over 20 years. In this county women between the ages of 18 and 39 experience the highest infant mortality rate of 92.2% with 13.9% of the deaths occurring in non-Hispanic Blacks, 6.0% in non-Hispanic Whites and 5.5% in Hispanics.[17] Between 2001-2003 disease during gestation and poor fetal nutrition were the two leading causes of infant mortality; at 55.7% for non-Hispanic Black women, 52.6% for Hispanic women, and 49.8% for non-Hispanic White women living in Tarrant County.[17] The women most at risk for poor birth outcomes are those who are under- or uninsured. Healthcare coverage has been linked to income status. This suggests that income may predict whether a pregnant woman acquires perinatal or other health care.

Low income minority mothers are at risk for poor health and diseases such as obesity, diabetes, and hyper-tension which are linked to poor nutrition.[3,7] Poor nutrition *may not be solely* related to income, but may also be related to depression. Depression commonly overlaps with anxiety and stress which often results in poor eating behaviors accompanied with either a loss of appetite or overeating.[11] In addition, other depression symptoms such as difficulty concentrating, trouble sleeping, restlessness, feeling fatigued, or feeling hopeless and apathetic can negatively affect diet and nutrition including under- and overeating.[5,6]

Depression, anxiety, and stress facilitate increased release of the stress hormone, cortisol.[11,13] High cortisol levels due to psychosocial stress and depression are related to stunted fetal growth, preterm labor, low birth weight, and maternal and fetal infections.[8,11,15] Higher levels of stress found in minority mothers are often due to community segregation and are associated with low birth weight.[8] High levels of cortisol are significantly related to high levels of perceived stress.[13] High levels of stress-induced cortisol predict weight gain, obesity and metabolic syndrome[11] which are most prevalent in minority women.[7] Poor maternal nutrition is an apparent mediator to fetal nutrition and poor birth outcomes. In 2003, the Texas infant mortality rate was 6.6 deaths per 1,000 births – higher than the Healthy People 2010 goal of 4.5 per 1,000 live births.[17] Among Texas counties, Tarrant County had a 2003 infant mortality rate of 7.5 per 1,000 births – higher than that year's United States rate of 6.9.[17] We have not found any evidence to suggest that Tarrant County rates have even slightly improved over the last 4 years. Under- and uninsured women are more likely to be affiliated with a low income minority group with frequent exposure to stress, anxiety, and depression resulting in poor maternal and infant health. To avert these negative maternal, birth, and infant health outcomes, women need to be educated about the connection between the mind and the body. We propose to use Promotoras-Community Health Workers to educate under- and uninsured pregnant women about mind and body relationships. Promotoras-Community Health Workers are state certified promoters of health education who support and strengthen minority and underserved communities.[9,14] Promotoras-Community Health Workers (CHW) are culturally sensitive, using native languages and value systems unique to their communities – non-Hispanic Black and Hispanic communities alike.[9,14]

In reaching out to their communities, Promotoras-CHW successfully help eliminate ethnic and cultural barriers to accessing healthcare services.[9,14] Studies show that the use of Promotoras-CHW to educate patients about cardiovascular disease, diabetes, and obesity is very effective in reducing medical non-compliance and changing poor lifestyle habits such as increasing physical activity.

Medina et al. (2007)[12] examined the effectiveness of Promotores-CHW in a cardiovascular intervention / prevention study in Hispanic patients. The Promotores-CHW used the curriculum entitled *Salud para su Corazon* which is endorsed by the National Heart, Lung and Blood Institute of the National Institutes of Health. Pre- and post-intervention healthy lifestyle behavior scores measuring salt and sodium consumption, cholesterol and fat consumption, weight control practices,

and physical activity were collected. The baseline to exit change scores showed that Promotores-CHW intervention significantly improved heart-healthy scores and facilitated positive lifestyle changes by reducing salt, cholesterol and fat consumption, reducing weight, and increasing physical activity (p<0.01).[12]

A recent randomized, control study explored the effectiveness of using Promotores-CHWs to identify the risks and prevention for Type 2 diabetes and cardiovascular disease in 258 Hispanic families with overweight or obese children (N=1000). In contrast to control families who did not receive Promotora-CHW intervention, those who were exposed to Promotora-CHW were able to either maintain or lose significant amounts of weight. Further, the control group gained weight.[18] Balcázar, et al. (2005)[1] studied the effectiveness of promotores de salud (n=33) in improving heart-healthy behaviors among Latino families (n=190) participating in the model's *Salud para su Corazon* pilot program in seven United States cities. Results showed significant differences between pre- and post-test effectiveness of the Promotora-CHW model in improving heart-healthy behaviors, promoting community referrals, and screenings by 59% (p<.001).[1]

The evidence suggests that utilization of Promotora-CHW is an effective way to increase access to health services, increase knowledge, and promote positive behavioral changes. However, there have been no studies of Promotora-CHW intervention in perinatal women or in the area of mental health. Therefore, this proposed service and research project is critically needed and is highly relevant to addressing the needs and evaluating the outcomes in under and uninsured pregnant women.

Very Weak Example

In today's society some violent events that may seem extraordinary to some, occur frequently for others. But clinical assessments leave this out of the interview (Boody et al., 1998). Even things like being shot or shot at aren't asked even in a jail population. Patients are not usually asked about being hit hard and going unconscious (Peets and Jones, 2005). Even so, it would seem important to know about this for clinical treatment and how this can effect women and children in their families.

Being exposed to one violent event such as being assaulted by someone may not cause PTSD, but if they don't know they can't address it. This keeps the provider from helping the women to cope and avoid being hit (Sophia, Georges, and Sams, 2009). We found no research on whether violence and trauma are factors in perinatal depression. There is some indication, however, that repeated exposure to violence and trauma can cause negative effects. So we want to interview some of the patients in the OB clinic and see if they have depression.

C. Preliminary Studies

This is the section to describe your own preliminary data from research on this topic. It may not apply to your project. If your mentor has preliminary data and is the PI, this section should be written using his or her data.

There is another possible use for this section, even if there is no preliminary data from your own independent or collaborative work.

You may use this section to discuss your other experiences or knowledge to show your competence to conduct this study. Or you may want to demonstrate competence among the research team in the clinical topic or to show how the institution or research team has special unique properties making it most likely to succeed.

D. Research Design and Methods

The Research Design and Methods Section is usually the largest section of the proposal. It can be organized in many different ways according to the nature of your research. This section describes the following in carefully detailed steps listed on the next page.

- How you will achieve the specific aims
- What activities you will undertake to test the hypotheses for your study
- What you will do in your experiment or what you will do to the human or animal subjects
- How you will do it
- When
- Where
- How
- How often
- Why
- What will you use to determine if the action you propose will generate the reaction or the result/outcome you hypothesize that it will generate – your measurements.

In this section you must address:

1. Subjects: Who will be in the study? How will you select them by inclusion and exclusion criteria, and how will you recruit them? (NOTE: Human subjects (risk and exclusion justification) issues are covered in a separate section E and do not go here)

2. The independent variables: What you are observing or what you expect to change – the outcomes or measures of change

3. The 'Power' of the study: See guidelines on when and how to obtain a power analysis to estimate sample size

4. The dependent variables: What you are using as an intervention or treatment

5. Other data such as categorical variables

6. The hypotheses to be tested, question to be explored, or the product to be developed

7. The design of the study: randomized block, repeated measures, longitudinal, retrospective, prospective, cross-sectional, existing data, controlled, blinded, and the like.

8. The instrumentation to collect data, and to measure the experimental change, the clinical outcomes, the validity of the theory, or the utility/validity of the product

9. The statistical tests you plan to use to test the hypotheses or question

Examples follow.

(NOTE: Examples are divided by plausible but not required sub-sections)

D.1. Overview

Example

Strong: This proposed on-going study will use existing data to provide information useful for planning services, determining utilization patterns, and educating providers, referral sources and policy makers. It will use <u>existing data</u> from the electronic database already created and maintained by the JPS Department of Psychiatry.

Methods and Procedures

There is no treatment involved, and no experimental procedures involving human subjects. Only existing data will be analyzed.

Two data files will be created and retained:

1. An Original Master Data File

2. A De-Identified Research Study File

Criteria for inclusion and exclusion: source of data

Since there is no age restriction on admissions to these three service programs, data will be gathered on individuals of all ages, including persons under the age of 18.

Weak: This experimental health education project will measure changes in women's attitudes toward healthy behavior and their knowledge of food values to improve how they prepare meals for their families. We will see the subjects every week in each group who will receive three experimental interventions and one with no intervention but will complete the questionnaires.

```
                        Research Design
            Ξ   Double blind
            Ξ   Controlled
            Ξ   128 subjects randomized equally to four groups:
                1. Health Education video and Group Discussion
                2. Health Education video only
                3. Book to read
                4. No intervention
            Ξ   Six weekly visits to the research clinic
```

```
                      Outcome Measures
```

```
        PRIMARY OUTCOMES                    PHYSICAL
                                          PERFORMANCE
                                       SECONDARY OUTCOMES

 Attitudes toward health
 SF36 at baseline and endpoint            ✠  Strength
 Knowledge of food values                 ✠  Energy
 Dietary Integration Health Questionnaire ✠  BMI
 Depressive symptoms                      ✠  VAS
 Health behaviors                         ✠  Cholesterol
 Health diary
```

D.2. Subjects

D.2.1. Power Analysis

Example 1

Strong

We calculated the number of subjects needed for each treatment group by using the mean crowley's intellectual perception test (CIP) as the primary outcome measure. The mean score for a psychiatric population in the CIP is 74 with a standard deviation of 12. Studies indicate that unmanaged stress and poor diet contribute to abnormal scores above 140 (Eisen[12], Ramey[11] and Kilmer[24]). Using PASS 2000 (Hintze, 2001[26]), we estimated that with 78 subjects in each of our three experimental groups (for a total of 234 subjects), we will be able to detect a change in the CIP of at least 20% at end-point following health education and counseling. We set the power (beta) at 85% power with a significance level (alpha) of 0.05. These estimates are based on the findings of Ramey and colleagues who found changes in health behaviors following illustrative education improved between 15 and 30% in women with young children and less than a college education.[11] To allow for an anticipated attrition rate of 20% based on previous studies, we will recruit 100 subjects per group for a total of 300 subjects. This calculation of power, estimated effect size, and confidence intervals, protects against a Type II error, and requires a manageable number of subjects. Similarly, we have calculated that 78 subjects per group will enable us to detect moderate changes in other physiological and clinical outcome measures, daily functioning and pain while maintaining optimum statistical power.

Example 2

Strong

A power analysis restraining the 95% confidence interval to +/- 5% indicates that an adequate sample size to compute the prevalence of depressive symptoms in the perinatal population at UNTHSC and JPS is 410. Thus, a sample of 450 perinatal women will be adequate to power this pilot study.

D.2.2. Subject Selection and Recruitment

Example 1

Strong

Women between the ages of 21 and 45 who are pregnant or who have delivered a child within the past six months who are under the care of a physician at either UNTHSC or the JPS Health Network OBGYN department are eligible for this study.

We are interested in women at any gestational age with the pregnancy confirmed by the physician, because it is unclear from the literature at what point depressive symptoms may occur during pregnancy. With an adequate sample of women at many different stages of pregnancy, we will be able to perform a sub-group analysis by gestational age to better understand observed phenomena.

There may be a brief window of time for us to engage a representative sample of postpartum mothers in this study because there is an important group of women, who are thought to be Mexican Spanish, who deliver their babies and do not return for postpartum or pediatric care. In order to maintain the power of the study and obtain an adequately representative sample, we have bracketed our sample by the inclusion and exclusion criteria listed below.

Example 2

Strong

The Volunteer Promotoras-CHW in *Phase I* and research staff in *Phase II* will refer all under- and uninsured pregnant women in need of primary healthcare to The Tarrant County Public Health Department, the Fort Worth Public Health Department, the Mission Fort Worth Clinic, the Albert Galvan Clinic, the John Peter Smith (JPS) Health Network and/or the University of North Texas Health Science Center (UNTHSC) patient care departments. In particular, we have the support of the UNTHSC Chairman of the OB/GYN departments, Dr. Ralph, to provide subjects with referrals to residents and nurse midwives for perinatal and postnatal care through the federally-funded perinatal Child Health Insurance Program. The Mental Health Association of Tarrant County (MHATC) will serve as the primary referral access resource for mental healthcare.

D.3. Methods

Begin with a statement such as "This section describes how we will collect the data pertinent to this proposed study. It is divided into (your number) sections: 1) study protocol, 2) experimental groups and interventions 3)......."

D.3.1. Study Protocol

Example

Strong: This is an 18-month prospective service demonstration program of a "Promotoras Mental and physical Health Program" (PMPHP) that contains a 2-day training component that will be implemented and then evaluated over a 2-year period. We will provide menal health and substance abuse (MH/SA) certification training for Tarrant County Promotoras-CHW. Research Promotoras-CHW (n=3) serving in the role of Research Assistants, will perform study assessments and educate study subjects (n=100) using the 18-module curriculum.

Weak: Promotoras trained as study personnel will conduct interviews in phase II of this longitudinal study. They will train about 100 study subjects, and measure different outcomes by conducting interview.

D.3.2. Experimental Groups and Interventions
Example

Strong: The Volunteer Promotoras-CHW in *Phase I* and research staff in *Phase II* will refer all under- and uninsured pregnant women in need of primary healthcare to The Tarrant County Public Health Department, the Fort Worth Public Health Department, the Mission Fort Worth Clinic, the Albert Galvan Clinic, the John Peter Smith (JPS) Health Network and/or the University of North Texas Health Science Center (UNTHSC) patient care departments. We intend to enroll 100 eligible women to complete the health education program. Institutional Review Board approvals will be obtained prior to beginning the study.
Inclusion Criteria: Women who are: (1) non-Hispanic Black, Hispanic, and non-Hispanic White, (2) under- or uninsured, (3) currently pregnant, (4) between the ages of 18-40, and (5) English or Spanish speaking. Subjects will be stratified by trimester (1st, 2nd, 3rd). Trimester will be determined using the date of the last menstrual period.
Exclusion Criteria: Women who have: (1) Life threatening illnesses (i.e. HIV/AIDS, metastatic cancer, brain tumor, etc.) or (2) a recent history of attempted suicide (i.e. within the last 12-months) will be ineligible to participate in the study.

Weak: We will select 10 promotoras from community trained women to recruit 100 subjects to educate. These women will be referred by clinics at the university who are eligible for treatment. They will get a consent before they participate. This demonstration project has only one group, all women get all treatments. The women have to be pregnant and older than 18 years old. They can't have any life threatening diseases.

D.4. Methods

D.4.1. Outcome Measures

Example

Strong: The primary outcome for this study will be the quantity of cocaine used each week, measured in grams. A biological measure will be taken twice a week to help validate cocaine use – urine drug screens measuring the cocaine metabolite benzoylecgonine (300ng/mL). Each urine toxicology drug screen has the capability to test for drug use within the past 72-hours. A self-report measure will also be taken to evaluate how many grams are reported as used, and how many days per week drugs were used. Therefore, there are three quantifiable ways to determine cocaine use that can be reported in number of days and grams. The reductions in grams and days of use can be used to determine the efficacy between study medication groups.

Thinking through selecting your outcome measures

For a study of two drugs to reduce cocaine use, a self-reported measure of cocaine use represented as a total dollar amount spent on cocaine each week is not valid because substance abusers can get free drugs in exchange for 'services rendered' (i.e. prostitution or drug running). If the researcher does not use an objective/biological measure, cocaine use cannot be accurately identified using the National Institute of Drug Abuse cut-off of 300ng/mL. For self-report measures remember that these subjects may want to stay in a study, and may try and please the researcher; thus introducing a bias known as the 'halo effect'...telling the researcher what they think they want to hear.

OR using a clinician rated Global Assessment of Functioning Scale based on a measure from 0-100 ranging from "a danger of seriously hurting self or others" to "superior functioning – no symptoms," requires caution. The GAF is reliable and valid for Emergency Department patients and for some inpatient populations' but the GAF may not be valid for outpatients. Also, the GAF is not specific to the research goal of examining incremental improvement of manic and depression symptoms over time as a function of study medication.

Strong: Figure 2 illustrates the evaluable factors that we hypothesize will influence maternal, birth, and infant health outcomes. The hypotheses will be tested with repeated monthly assessments of depression, anxiety, perceived stress, quality of life, and physical health status using well known, validated instruments. Outcome measures include objective and self-reported variables. We will measure hypertension (e.g. blood pressure), and hyperglycemia (e.g. urine glucose and ketones) as these biomarkers have been associated with depression and anxiety symptoms as well as complicating gestational weight retention. We will also measure basal metabolic rate (BMR) and percentage of fat mass as obesity and being overweight have been associated with depression symptom severity.
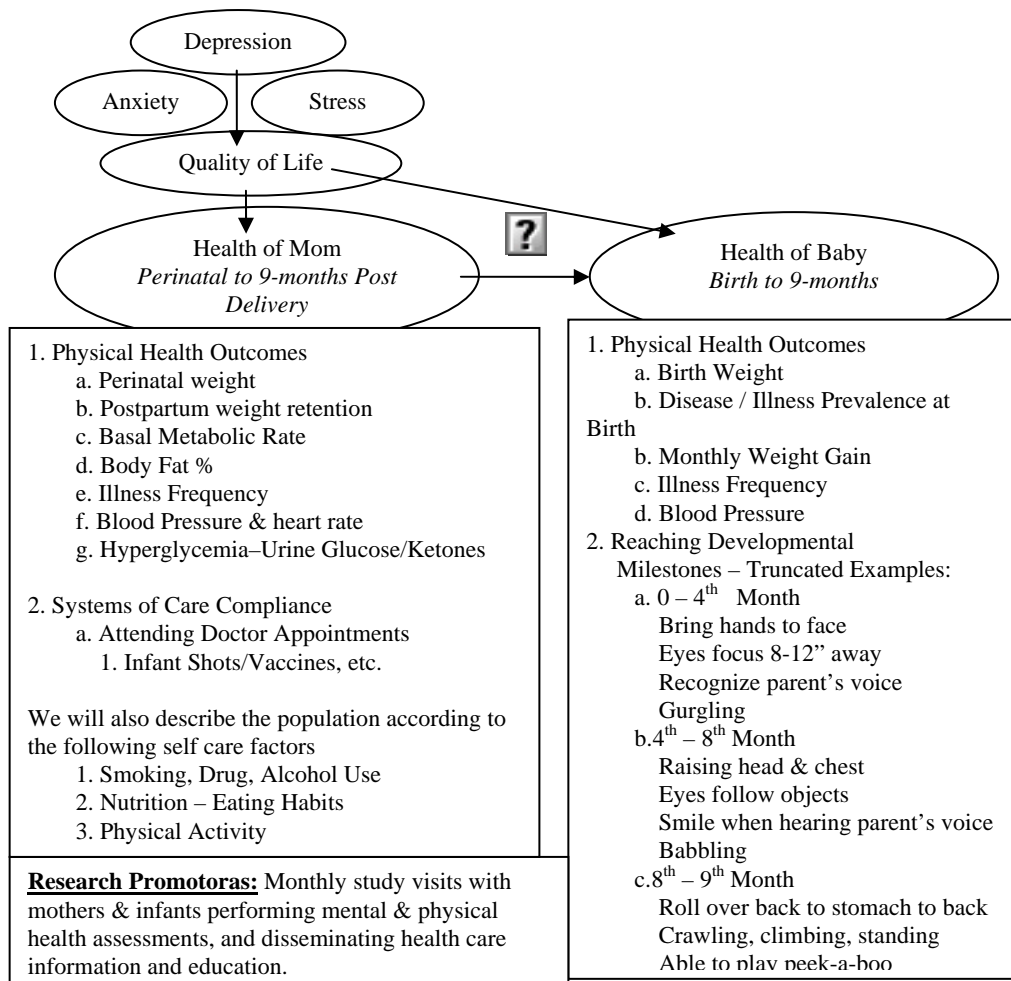
D.5. Data Analysis Plan
<u>Example</u>

<u>Strong:</u> **Aim one:** Estimate the prevalence of depressive symptoms among perinatal women in a sample of 450 patients of the UNTHSC and JPS Health Network OBGYN departments. The Edinburg Depression Scale will provide data that will be used to compute point and interval estimations of the prevalence of the depressive symptoms likely to be found in these clinic populations. SPSS version 14 will be used to estimate prevalence using a non-linear hierarchical model. A more conservative estimate will also be computed using Cox regression analysis.

Aim two: Analyze relationships between perinatal depressive symptoms, maternal health, and exposure to traumatic life events. Trauma loads will be computed using orthogonal varimax rotated factor analysis scores. Analysis of variance (ANOVA) will be computed for sub-groups of subjects stratified by high, medium and low EDS scores. Pearson correlation coefficients will be computed to ascertain the strengths of the relationships between depressive symptoms and trauma loads. Logistic regression will be used to determine the extent to which trauma loads may predict EDS scores, controlling for key demographic and socioeconomic variables.

<u>Weak:</u> We will analyze the data using the most appropriate parametric and non-parametric tests. Our approach to the data will be descriptive mainly, with two-tailed hypothesis testing.

D.6. Study Timetable and Management Plan

A research proposal should always contain a timetable and management plan, however brief it may be. It lets the reader know that you have considered the logistics of carrying out your research plan. The shape of this depends on your research, particularly if it is sequential, or multi-center, or involves a data safety and management board (DSMB) or other advisory groups. It lists the key steps involved in the project from start-up to final report and marks the weeks or months during which these will occur. It may be similar in format to the protocol time schedule and assures the reviewer that you have a plan in mind even though it may be modified later.

A simple time table may be like this:

	Grant Year 1												Grant Year 2											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Complete study Manual	X	X																						
Training on intervention reliability		X	X					X			X			X										
Subject Recruitment			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Enrollment & screening/baseline measures			X		X		X		X		X	X		X										
Interventions			X	X	X	X	X		X	X	X		X	X	X		X	X	X		X	X	X	
Evaluation of progress and protocol								X			X			X			X			X				X
Six month follow-up											X	X	X	X		X	X		X	X				
Study analysis & final report																								
Publications & possible R-01 application																								

On the next page is presented a two phase longitudinal demonstration project with a research component measuring intervention outcomes and program effectiveness.

TIMELINE – The Mind-Body Connection in Perinatal Depression: A Role for Promotoras-CHW	

PROJECT YEAR ONE

AUG 2008	SEP 2008	OCT 2008	NOV 2008	DEC 2008	JAN 2009	FEB 2009	MAR 2009	APR 2009	MAY 2009	JUN 2009	JUL 2009
1st Year Funding & Start-up	Notify Partners & Promotoras-CHW of Upcoming Training		Create PHASE I & PHASE II Databases		Conduct MH/SA Training	Conduct Monthly PHASE I Effectiveness Evaluation Surveys with Partners & Volunteer Promotoras-CHW						
						On-going PHASE I Survey & PHASE II Study Data Entry					
Launch PHASE I Arrange Training Dates & Travel for Sue Gallego	Assemble & Copy All PHASE I Mental Health & Substance Abuse (MH/SA) Training Packets				Select Volunteer Promotoras-CHW (N=15)	Researchers Have Monthly Meetings with Partners & Promotoras-CHW to Engage Them in PHASE I & PHASE II Study Program Goals and Objectives					
Launch PHASE II Begin IRB Approval Process Hire & Train Res. Asst. Promotoras-CHW (N=3)	Compile & Copy All PHASE II 18-Module Curricula & Monthly Assessment Packets for Study Subjects (N=100)		Receive IRB Approval			Begin Enrollment of 1st 20 PHASE II Study SUBJECTS (N=20)		Enroll 15 New Study Subjects (N=35)		Enroll 15 New Study Subjects (N=50)	
			Distribute Recruitment Flyers to Partners & Promotoras-CHW		PHASE II Continuous Monthly Study Visits Feb. - July 1 Visit Per Month Per Subject						

PROJECT YEAR TWO

AUG 2009	SEP 2009	OCT 2009	NOV 2009	DEC 2010	JAN 2010	FEB 2010	MAR 2010	APR 2010	MAY 2010	JUN 2010	JUL 2010
2nd Year Funding	Continue to Engage Partners & Volunteer Promotoras-CHW			PI, Co-PI, Co-Invs. Conduct Refresher MH/SA CEU Lectures for Res. Asst. & Volunteer Promotoras-CHW in January					Prepare for the Tarrant County Wellness Fair in August Presenting the PHASE I Mind-Body Connection 9-Module Curricula		
IRB Cont.. Review	Subject Chart Review Continuous Data Cleaning & Entry										
Continue Distribution of PHASE I 9-Module Curricula – Effectiveness Evaluation Surveys Continue											
Continue Distribution of PHASE II Recruitment Flyers to Partners & Promotoras-CHW Through December				Enroll Last 20 Subjects (N=100)					Researchers Update Partners & Volunteer Promotoras-CHW re: PHASE II Closing Enrollment Continuing PHASE I 9-Module Dissemination		
Enroll 30 New Subjects (N=80) "Limit Enrollment to 1ST & 2nd Trimester Subjects Only"				"Limit Enrollment to 7th & 8th Month Term Subjects Only"		"Limit Enrollment to 9th Month Term Subjects Only"					
PHASE II Continuous Monthly Study Visits August – July: 1 Visit Per Month Per Subject - Enrollment Ends April 30											

PROJECT YEAR THREE

AUG 2010	SEP 2010	OCT 2010	NOV 2010	DEC 2011	JAN 2011	FEB 2011	MAR 2011	APR 2011	MAY 2011	JUN 2011	JUL 2011
3rd Year Funding	Continue Study Data Cleaning & Entry for PHASE I & PHASE II Data						Study Data Analyses		Presenting & Submitting Study Results for Publication & HRSA		
IRB Cont. Review	PHASE II Final Study Visits for 3rd Trimester Enrollees: All Subjects Complete Study by Feb. 28						Meet with Partners to Review Data from PHASE I & PHASE II Study Results		Promotoras-CHW Transition to Community Workforce		
PHASE II Continuous Monthly Study Visits with Remaining Subjects August – Feb.											

E. Human Subjects

This section must be thorough. Tutoring is always needed from someone with training and current, up-to-date knowledge and experience in issues of ethics in human subjects' research. There is a separate tutorial chapter and training provided by the Department of Psychiatry that includes CITI training and NIH sponsored training.

G. Literature Cited

You may use number references in the text and number the literature list, OR use names in parentheses and alpha-order the literature list.

Example

Abate C, Patel L, Rauscher FJ III, Curran T (1990) Redox regulation of Fos and Jun DNA-binding activity in vitro. *Science* **249:** 1157-1161.

Andrews NC, Faller DV (1991) A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells*. Nucleic Acids Res* **19:** 2499.

Chapter IV Research Design and Methods – The Research Plan

<div style="text-align:center">

INTRODUCTION

</div>

<div style="text-align:center">

Section D of the Research Proposal

Research Design and Methods

is the Heart of the Proposal

</div>

In this section of the research proposal, you are going to describe exactly what you will do and how you will do it to obtain the answer to the research question or test the hypothesis. Exactly HOW are you going to conduct this study?

As you write this section, you will begin to understand that you may need to revise your background and significance or that you want to "tweak" your question. This section should flow naturally from the specific aims page, even though it is separated by Section B, Background and Significance, and Section C, Preliminary Studies.

Keep it simple.

How will you answer your question? A clearly organized Method section is critical. Many researchers are confused about the distinction between the design discussion, and the methods description, leading them to lump the two together.

If you decide to dismiss this as "just an issue of semantics", you will find that its contents can be quantified by grant reviewers. This is the DESIGN **AND** METHODS section. In fact, the NIH changed the name of this section to include the word <u>design</u> because many proposals omitted that as a sub-section.

<u>Present the design and methods sub-sections separately</u>.

Design is the way in which you conceptualize your research project, and addresses structure. Your research question or your hypothesis drives the design. The design drives the methods.

Methods are detailed tasks, steps, stages, and procedures you will use to conduct the research.

The design discussion is relatively brief, particularly when compared to the methods narrative. However, it should be inherently interesting. It can be very creative. For example, you could have designed your project in any number of ways, but you chose this one. Why?

Methods, by contrast, are straightforward. It is critical to describe exactly what you will do to conduct the study from beginning to end. Provide sufficient detail for a complete evaluation of your work. This is particularly important if you are proposing to develop a new methodology or a new technique.

In the methods section you will include a discussion of the strengths and limitations of your study. Discuss any technical problems that may arise and what alternate plans you may implement. This is a sophisticated part of the proposal, and it should be carefully and thoroughly reviewed repeatedly with a mentor.

End the research design and methods section with a timeline. It is important to convey that your project can be conducted within the proposed time and that you have a logical, well conceived plan to implement.

There is no one standard way to organize the Research Design and Methods section.

Recommended Outline

D. Research Design and Methods

    D.1. Introduction

        D.1.1. Restate the research aims, question/s or hypotheses

        D.1.2. Describe how this section is organized

    D.2. Power Analysis

    D.3. Inclusion and Exclusion/Selection Criteria

    D.4. Outcome Measures – variables of interest

    D.5. Data management and analysis

    D.6. Project management plan and time lines

ANNOTATED OUTLINE

Tasks Associated With Research Design and Methods Section

Introduction

The reader has just finished either the background and significance in which you made your case that this project is supported by the literature and/or fills a gap in the literature, or the preliminary studies section in which you discuss your or your mentor's previous work, or present evidence of the ability to conduct this study.

Begin the Research Design and Methods section with something like:

"The purpose of this study is to test whether a community education model using Promotoras, improves birth outcomes for the mother and the newborn."

Then tell the reader how you have organized the material.

Example

The following material includes first a discussion of the research project design, next a discussion of the power of the study and the inclusion/exclusion (selection) criteria. Next we provide information about the outcome variables of interest and how we propose to manage and analyze the data to address the study questions/hypotheses. Last, we provide a project management plan and time-lines for completing the project.

Key Questions to guide your thinking are as follow.

What is the scientific or clinical theoretical framework for this research?

What do we currently know about this area?

What are the strengths of the current body of knowledge?

What are the gaps in the current body of knowledge?

How will this proposed study build on the strengths, correct the limitations and fill the gaps (or begin to do this)?

Design

Example 1

We propose to use a retrospective case-control design to observe whether psychiatric outpatients with a dual diagnosis of a thought disorder and a substance abuse disorder utilize crisis care more frequently than those with a single primary diagnosis of a thought disorder, and no substance abuse disorder. Subjects' records will be randomly pulled from the admissions to the Psychiatric Emergency Center (PEC) for calendar year 2003, clustered by diagnosis. Records will be searched for crisis services utilization for 12 months following the person's first admission to the PEC in 2003.

Example 2

We will utilize a cross sectional, random sample of patients admitted to the Trinity Springs Pavilion Inpatient Unit over five years, from January 2001 through December 31, 2005 to examine whether length of stay is associated with the mediation administered at admission to the unit.

Example 3

Existing data from the Alzheimer's Disease screening profiles from 180 patients evaluated in the internal medicine clinics between July 2003 and July 2005 will be used to randomly select 50 patients to interview for health care utilization patterns in the three to five years since the assessment was provided. The hypothesis is that patients who are diagnosed before age 70 will have utilized more and varied health care services than those diagnosed after age 70 regardless of the severity of the disease at the time of the assessment.

Power Analysis

After the research question(s) or hypotheses are final you should determine what result(s) you anticipate. Typically, the smaller the sample size, the larger any difference between group scores will have to be in order to achieve statistical significance.

Statistical power analysis is a set of procedures and formulas that allow us to determine how likely we would achieve statistical significance with a particular sample size (given an assumed true difference between groups).

If the likelihood is good (e.g. greater than or equal to an 80% chance), then the sample size is considered adequate.

Power Analysis Examples

You expect elderly patients with chronic illnesses to have a certain amount of cognitive decline that is greater by about 25% compared to those with no chronic illnesses at the same age.

You can use data from the literature or a theoretical clinical foundation to estimate the anticipated effect size for your study. For example, if the literature reports a 10% difference in neurocognitive performance on a particular test of executive functioning between patients with and without multiple sclerosis, and there is little else available for other types of chronic health diseases regarding executive functioning, you can use this data to estimate your anticipated effect size.

You need to have as much valid and reliable data as possible to estimate a sample size for your study at a specified power. This applies to ANY design. This applies if you are taking all your data from medical records. This applies if you are using existing data from another source. Some research designs eat power, such as using more than two groups, or using repeated measures and attempting to assess changes at time intervals.

Don't try to do a power analysis or estimate a sample size by yourself. Use an expert. For example, don't not lick your finger and hold it up to the wind, or put your thumb up to the portrait or MRI and say "oh, about a hundred should do it" or "Seems like 25 would be okay".

Regardless of whether you intend to, or can afford to adhere to the sample size indicated by a power analysis, it is <u>always</u> important to include a discussion of this effort.

If you are willing to accept a small change or difference you may not need so many subjects. If you are conducting an exploratory study or a case-control study you do not need a specific sample size but you always should guard against both Type I and Type II errors (see glossary of terms).

Outcome Measures

Be clear, precise, and specific about how you are going to measure your outcomes of interest. Define them exactly. Usually you cannot provide too much detail in this aspect of your research. The definition of outcomes will determine what information/data you are going to collect, what you are going to use to collect/record it, and how you will acquire it.

Data Management and Analysis

Reference the statistical software that will be used for the analysis.

Meet with a biostatistician who will help you THINK critically about your design and methods, not merely crunch numbers. This will require an investment of time on your and their part.

Think through the way you will collect the research and how your data base will be set-up.

Invest the time to understand why a configuration of a data set can be your best flying dream or your worst nightmare.

Correlational vs. Experimental Research

Most empirical research belongs clearly to one of those two general categories. In correlational research we do not (or at least try not to) influence any variables but only measure them and look for relations (correlations) between some set of variables, such as blood pressure and cholesterol level. In experimental research, we manipulate some variables and then measure the effects of this manipulation on other variables; for example, a researcher might artificially increase blood pressure and then record cholesterol level. Data analysis in experimental research also comes down to calculating "correlations" between variables, specifically, those manipulated and those affected by the manipulation. However, experimental data may potentially provide qualitatively better information: Only experimental data can conclusively demonstrate causal relations between variables. For example, if we found that whenever we change variable A, variable B changes, then we can conclude that "A influences B." Data from correlational research can only be "interpreted" in causal terms based on some theories that we have, but correlational data cannot conclusively prove causality.
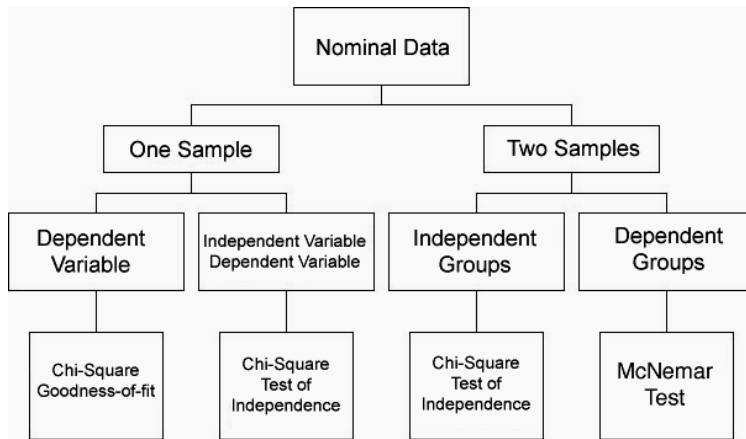
DECISION TREE FOR NOMINAL DATA ANALYSIS

DECISION TREE FOR ORDINAL DATA ANALYSIS
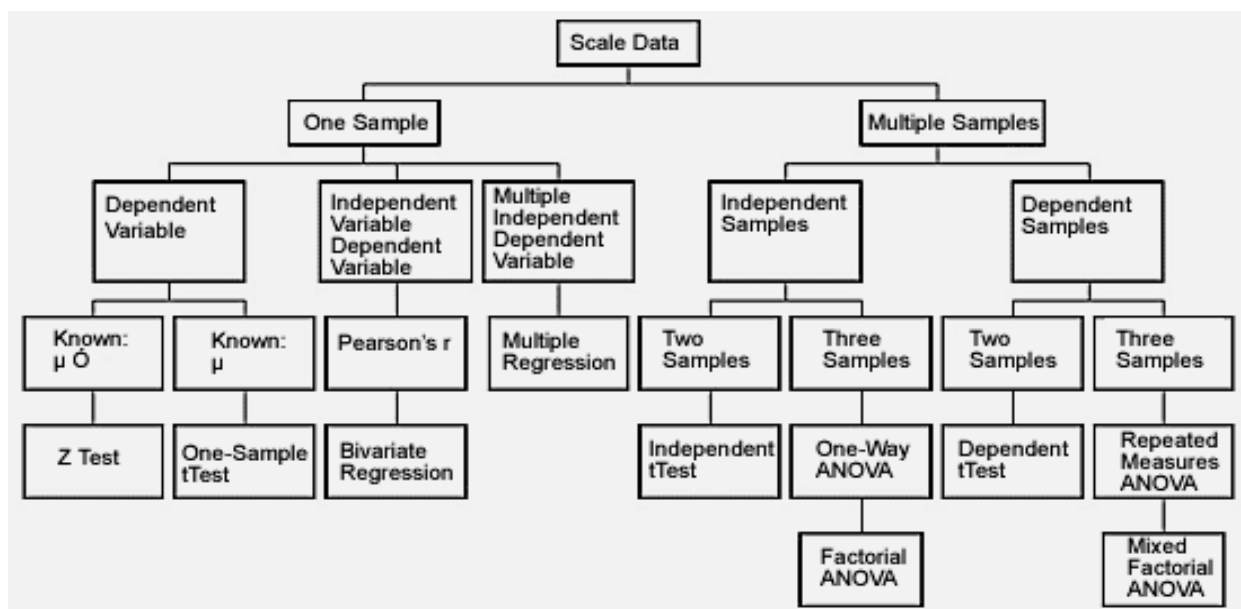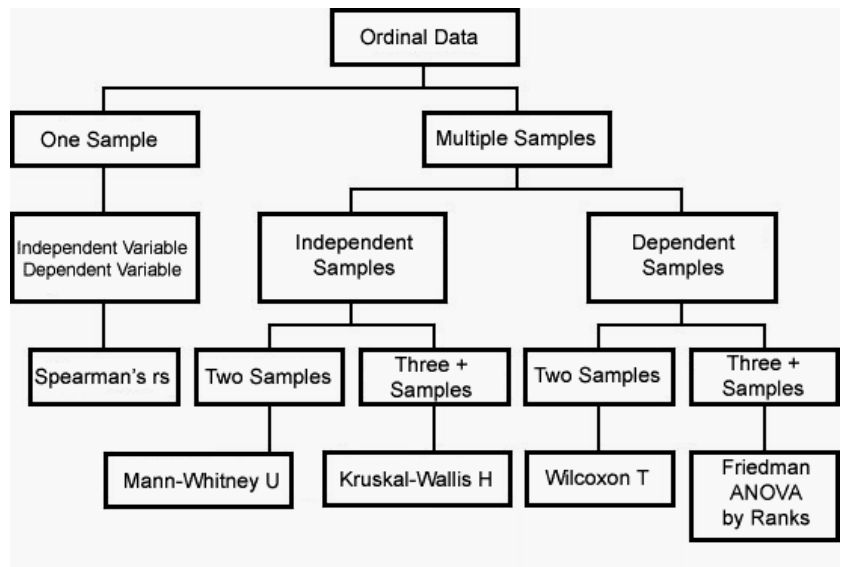
DECISION TREE FOR CONTINUOUS-SCALED DATA ANALYSIS

	Type of Data		
Goal	Measurement (normal)	Rank, Score, or Measurement (non-normal)	Binomial (Two Possible Outcomes)
Describe one group	Mean, SD	Median, interquartile range	Proportion
Compare one group to a hypothetical value	One-sample t test	Wilcoxon test	Chi-square
Compare two unpaired groups	Unpaired t test	Mann-Whitney test	Fisher's test (chi-square for large samples)
Compare two paired groups	Paired t test	Wilcoxon test	McNemar's test
Compare three or more unmatched groups	One-way ANOVA	Kruskal-Wallis test	Chi-square test
Compare three or more matched groups	Repeated-measures ANOVA	Friedman test	Cochrane Q**
Quantify association between two variables	Pearson correlation	Spearman correlation	Contingency coefficients**

IN THE FOLLOWING PAGES WE OFFER A GLOSSARY OF THE MOST COMMONLY USED RESEARCH AND STATISTICAL TERMS.

THESE ARE **NOT** ALPHABETICAL BUT INSTEAD THEY ARE GROUPED BY LIKE CONCEPTS.

Elementary Concepts in Statistics:

http://www.statsoft.com/textbook/esc.html

http://www.bath.ac.uk/e-learning/gold/glossary.html#N1862

http://www.ecs.org/html/educationIssues/Research/primer/foreword.asp
A Policymaker's primer on education research
Mid-continent research for education and learning

How to determine that a result is "really" significant: There is no way to avoid arbitrariness in the final decision as to what level of significance will be treated as really "significant." The selection of a level of significance, up to which level you would consider the results as valid, is both arbitrary and depends on the type of study and the knowledge of the researcher in the topic. In practice, the final decision depends on whether the outcome was predicted a priori, or found post hoc in the course of analysis, on the total amount of consistent supportive evidence in the entire data set, or on "traditions" existing in the particular area of research. Typically, in many sciences, results that yield $p \leq .05$ are considered borderline statistically significant, but this level of significance still involves a somewhat high probability of error (5%). Results that are significant at the $p \leq .01$ level are commonly considered statistically significant, and $p \leq .005$ or $p \leq .001$ levels are often called "highly" significant. But remember that those classifications represent nothing but arbitrary conventions that are only informally based on general research experience.

In exploratory and developmental studies, you can address clinical significance at ≥ 0.05 but ≤ 0.10 to 0.12. Read more about interpreting P values in the resources.

Dependent vs. independent variables: The terms dependent and independent variable apply most often to experimental research in which there is an intervention or a manipulation of a condition. A manipulated variable is an "independent" variable, independent of the behavior of the subjects. Some outcome variable is expected to be "dependent" on the manipulation or experimental conditions. That is to say, it depends on "what the subject will do" in response to it.

Independent variables are those that are manipulated, whereas dependent variables are measured or registered. This distinction appears terminologically confusing to many because, as some students say, "all variables depend on something." However, once you get used to this distinction, it becomes indispensable.

Somewhat contrary to the nature of this distinction, these terms are also used in studies where we do not literally manipulate a condition/independent variable, but we assign subjects to "experimental groups" based on some pre-existing properties of the subjects. For example in an experiment comparing males with females on white cell count (WCC), gender could be the independent variable and WCC the dependent variable.

P value: In a statistical hypothesis test, the P value is the probability of observing a test statistic at least as extreme as the value actually observed, assuming that the null hypothesis is true. This probability is then compared to the pre-selected significance level of the test. If the P value is smaller than the significance level, the null hypothesis is rejected, and the test result is termed significant.

The P value depends on both the null hypothesis and the alternative hypothesis. In particular, a test with a one-sided alternative hypothesis will generally have a lower P value (and thus be more likely to be significant) than a test with a two-sided alternative hypothesis. However, one-sided tests require more stringent assumptions than two-sided tests. They should only be used when the data meet those assumptions.

Measurement scales: Variables differ in "how well" or how accurately they measure the outcome or condition. There is always some measurement error involved in every measurement. The "type of measurement" determines the amount of information a variable provides. Generally speaking, variables are classified as (a) nominal, (b) ordinal, (c) interval or (d) ratio.

Nominal, ordinal, and interval scales are ordinal variables. A nominal (classification) measurement provides less information than an ordinal (ranking of a measure such as high, medium and low) measurement, but we cannot say "how much less" or how this difference compares to the difference between ordinal and interval scales.

a. *Nominal variables* are qualitative measures of group membership. For example, two individuals are different in terms of variable A (e.g., they are of different race), but we cannot say which one "has more" of the quality represented by the variable. Typical examples of nominal variables are gender, race, ethnic group, a color or hue, and city.

b. *Ordinal variables* allow us to rank order items we measure according to how much that variable has of something, but they do not allow us to say "how much more," in the sense of double or triple or half of something. A typical example of an ordinal variable is the socioeconomic status of families. For example, we know that upper-middle is higher than middle but we cannot say that it is, for example, 18% higher.

c. *Interval variables* allow us not only to rank order the items that are measured, but also to quantify and compare the sizes of differences between them. For example, temperature, as measured in degrees Fahrenheit or Celsius, constitutes an interval scale. We can say that a temperature of 40 degrees is higher than a temperature of 30 degrees, and that an increase from 20 to 40 degrees is twice as much as an increase from 30 to 40 degrees. In the use of the visual analog scale, a pain of 4 is not twice as much as a pain of 2, nor is my pain level of 5 the same as your pain level of 5.

d. *Ratio variables* are very similar to interval variables. In addition to all the properties of interval variables, ratio variables have an identifiable absolute zero point, thus they allow for statements such as x is two times more than y. Typical examples of ratio scales are measures of time or space. For example, as the Kelvin temperature scale is a ratio scale, we can say not only that a temperature of 200 degrees is higher than one of 100 degrees, but also that it is twice as high. Interval scales do not have the ratio property.

Most statistical data analysis procedures do not distinguish between the interval and ratio properties of the measurement scales.

Null hypothesis: A hypothesis test involves calculating the probability of seeing a result at least as extreme as the observed data, given some initial assumption about the underlying probability distribution. We base the initial assumption on some general concept that nothing noteworthy is happening, such as "the means for all groups are the same regardless of treatment," or "the survival rate remains the same for subjects given drug X or drug Y." We call the initial assumption the null hypothesis.

You then compare the calculated probability with your pre-selected significance level. If the P value is smaller than the significance level, the null hypothesis is rejected, and the test result is termed significant. For the two-sample unpaired t test, the null hypothesis is that the two population means are equal, and the t test involves finding the probability of observing a t statistic at least as extreme as the one calculated from the data, assuming the null hypothesis is true.

Sensitivity: The sensitivity of a test is the probability that the test will declare the condition of interest present when the condition is in fact present.

Specificity: The specificity of a test is the probability that the test will declare the condition of interest absent when the condition is in fact absent.

Normal probability plot: A normal probability plot, also known as a normal Q-Q plot or normal quantile-quantile plot, is the plot of the ordered data values (as Y) against the associated quantiles of the normal distribution (as X). For data from a normal distribution, the points of the plot should lie close to a straight line.

Normal (Gaussian) distribution: The normal or Gaussian distribution is a continuous symmetric distribution that follows the familiar bell-shaped curve. The distribution is uniquely determined by its mean and variance. It has been noted empirically that many measurement variables have distributions that are at least approximately normal. Even when a distribution is non-normal, the distribution of the mean of many independent observations from the same distribution becomes arbitrarily close to a normal distribution as the number of observations grows large.

Skewness: Skewness is a lack of symmetry in a distribution. Data from a positively skewed (skewed to the right) distribution have values that are bunched together below the mean, but have a long tail above the mean. (Distributions that are forced to be positive, such as annual income, tend to be skewed to the right.) Data from a negatively skewed (skewed to the left) distribution have values that are bunched together above the mean, but have a long tail below the mean. Boxplots may be useful in detecting skewness to the right or to the left; normal probability plots may also be useful in detecting skewness to the right or to the left.

Type I and Type II error: When testing a null hypothesis, there are two ways to draw a mistaken conclusion from the test.

Type I error is to incorrectly conclude that the null hypothesis is false when it is in fact true. You find a significant effect/change/difference when in fact there is not one. The probability of this error is usually denoted by the Greek letter alpha. By selecting the significance level (alpha-level) for a hypothesis test, you specify the value of the Type I error you are willing to tolerate if the null hypothesis is true.

A Type II error occurs when the test fails to reject the null hypothesis when it is in fact false. You miss an effect/change/difference when there actually is one, thus failing to support your alternative hypothesis.

The probability of this error is usually denoted by the Greek letter beta, and is equal to 1 minus the power of the test. The probability of a Type II error depends on the significance level (alpha-level) of the test, the components of the calculation of the test statistic, and on the specific alternative hypothesis under consideration.

In general, for a particular hypothesis test, significance level, and alternative hypothesis, the probability of a Type II error decreases as the amount of data collected increases. Because the probability of a Type II error depends on the alternative hypothesis, it can only be calculated with reference to a specific alternative hypothesis. A power curve shows (1 - beta) plotted for different possible alternatives, such as the possible differences between the two population means for a two-sample unpaired t test.

Violation of assumptions: Statistical hypothesis tests generally make assumptions about the actual or hypothesized population(s) from which the data were sampled. For example, many normal theory-based tests such as the t test and ANOVA assume that the data are sampled from one or more normal distributions, as well as that the variances of the different populations are the same (homoscedasticity). If test assumptions are violated, the test results may not be valid.

Robust: Robust statistical tests are tests that operate well across a wide variety of distributions. A test can be robust for validity, meaning that it provides P values close to the true ones in the presence of (slight) departures from its assumptions. It may also be robust for efficiency, meaning that it maintains its statistical power (the probability that a true violation of the null hypothesis will be detected by the test) in the presence of those departures.

Chi-square test for goodness of fit: The chi-square test for goodness of fit tests the hypothesis that the distribution of the population from which nominal data are drawn agrees with a posited distribution. The chi-square goodness-of-fit test compares observed and expected frequencies (counts). The chi-square test statistic is basically the sum of the squares of the differences between the observed and expected frequencies, with each squared difference divided by the corresponding expected frequency.

Chi-square test for independence (Pearson's): Pearson's chi-square test for independence for a contingency table tests the null hypothesis that the row classification factor and the column classification factor are independent. Like the chi-square goodness-of-fit test, the chi-square test for independence compares observed and expected frequencies (counts). The expected frequencies are calculated by assuming the null hypothesis is true. The chi-square test statistic is basically the sum of the squares of the differences between the observed and expected frequencies, with each squared difference divided by the corresponding expected frequency. Note that the chi-square statistic is always calculated using the counted frequencies. It cannot be calculated using the observed proportions, unless the total number of subjects (and thus the frequencies) is also known.

Inappropriate use of chi-square test: Pearson's chi-square test for independence for a contingency table involves using a normal approximation to the actual distribution of the frequencies in the contingency table. This approximation becomes less reliable when the expected frequencies for the contingency table are very small. A standard (and conservative) rule of thumb (due to Cochran) is to avoid using the chi-square test for contingency tables with expected cell frequencies less than 1, or when more than 20% of the contingency table cells have expected cell frequencies less than 5. In such cases, an alternate test like Fisher's exact test for a 2x2 contingency table should be considered for a more accurate evaluation of the data.

Correlation: The most commonly used correlation statistic is the Pearson correlation coefficient (r). This statistic measures both the **strength** and **direction** of the linear relationship between two variables.

Correlation Example

Suppose we want to look at the relationship between age and height in children. We select a group of children for study, and for each child we record their age in years and their height in inches. We could plot these values on a graph so that the child's age would be on the horizontal axis and the child's height would be on the vertical axis. Each dot on the plot represents a single child's age and height. This is called a **scatter plot**.

Since older children are generally taller than younger children, we would expect the dots on the plot to roughly approximate a straight line (a linear relationship between the variables) and that the line will slope upward (since age and height tend to increase at the same time).

Correlation Coefficient: The Pearson correlation coefficient (r) is a number between -1 and +1 that measures both the strength and direction of the linear relationship between two variables.

The magnitude of the number represents the strength of the correlation. The larger r is in absolute value, the stronger the linear association between X and Y. If r is 0, X and Y are said to be uncorrelated, with no linear association between X and Y. Independent variables are always uncorrelated, but uncorrelated variables need not be independent.

A correlation coefficient of zero represents no linear relationship (the scatter plot does not resemble a straight line at all), while a correlation coefficient of -1 or +1 means that the relationship is perfectly linear (all of the dots fall exactly on a straight line).

The sign (+/-) of the correlation coefficient indicates the direction of the correlation. A positive (+) correlation coefficient means that as values on one variable increase, values on the other variable tend to also increase; a negative (-) correlation coefficient means that as values on one variable increase, values on the other tend to decrease, that is, they tend to go in opposite directions.

T-Test: There are several kinds of t-tests, but the most common is the "two-sample t-test" also known as the "Student's t-test" or the "independent samples t-test". The t-test compares means.

T-Test Example

The two sample t-test computes differences in means on some measure between two independent populations.

For example, we might have a research hypothesis that rich people have a different quality of life than poor people. We give a questionnaire that measures quality of life to a random sample of rich people and a random sample of poor people. The null hypothesis, which is assumed to be true until proven false, is that there is no difference between these two populations.

We gather sample data and observe that the two groups have different average scores. But does this represent a real difference between the two populations, or just a chance difference in our samples?

T-Test Statistic

The statistics t-test allows us to answer this question by using the t-test statistic to determine a p-value that indicates how likely we could have gotten these results by chance. By convention, if there is less than a 5% chance of getting the observed differences by chance, we reject the null hypothesis and say we found a statistically significant difference between the two groups. See Statistical Data Analysis for more information about hypothesis testing.

Always check for assumptions about the data before used a statistical test.

Factor Analysis: A factor is a single discrete classification scheme for data, such that each item classified belongs to exactly one class (level) for that classification scheme. For example, in a drug experiment involving rats, sex (with levels male and female) or drug received could be factors. A one-way analysis of variance involves a single factor classifying the subjects (e.g., drug received); multi-factor analysis of variance involves multiple factors classifying the subjects (e.g., gender and drug received).

Nonparametric tests: Nonparametric tests are tests that do not make distributional assumptions, particularly the usual distributional assumptions of the normal-theory based tests. These include tests that do not involve population parameters at all (truly nonparametric tests such as the chi-square goodness of fit test), and distribution-free tests, whose validity does not depend on the population distribution(s) from which the data have been sampled. In particular, nonparametric tests usually drop the assumption that the data come from normally distributed populations. However, distribution-free tests generally do make some assumptions, such as equality of population variances.

When to Use Which Method

It is not easy to give simple advice concerning the use of nonparametric versus parametric procedures. Each nonparametric procedure has its peculiar sensitivities.

For example, the **Kolmogorov-Smirnov** two-sample test is not only sensitive to differences in the location of distributions (for example, differences in means) but is also greatly affected by differences in their shapes.

The **Wilcoxon matched pairs test** assumes that one can rank order the magnitude of differences in matched observations in a meaningful manner. If this is not the case, one should rather use the Sign test.

In general, if the result of a study is important (e.g., does a very expensive and painful drug therapy help people get better?), then it is always advisable to run different nonparametric tests. Should discrepancies in the results occur contingent upon which test is used, one should try to understand why some tests give different results. On the other hand, nonparametric statistics are less statistically powerful (sensitive) than their parametric counterparts, and if it is important to detect even small effects (changes/differences) (e.g., is this food additive harmful to people?) one should be very careful in the choice of a test statistic.

Large data sets and nonparametric methods: Nonparametric methods are most appropriate when the sample sizes are small. When the data set is large (e.g., *n* > 100) it often makes little sense to use nonparametric statistics at all. When the samples become very large, then the sample means will follow the normal distribution even if the respective variable is not normally distributed in the population, or is not measured very well.

Thus, parametric methods, which are usually much more sensitive (i.e., have more *statistical power*) are in most cases appropriate for large samples. However, the tests of significance of many of the nonparametric statistics described here are based on asymptotic (large sample) theory; therefore, meaningful tests can often not be performed if the sample sizes become too small. Please refer to the descriptions of the specific tests to learn more about their power and efficiency.

Nonparametric Correlations

The following are three types of commonly used nonparametric correlation coefficients

Spearman R. Spearman R (Siegel & Castellan, 1988) assumes that the variables under consideration were measured on at least an <u>ordinal</u> (rank order) scale. Individual observations should be amenable to ranking into two ordered series. Spearman R can be thought of as the regular <u>Pearson product moment correlation coefficient</u>, that is, in terms of proportion of variability accounted for, except that Spearman R is computed from ranks.

Kendall tau. Kendall tau is equivalent to Spearman R, requiring the same underlying assumptions. It is also comparable in terms of its statistical power. However, Spearman R and Kendall tau are usually not identical in magnitude because their underlying logic as well as their computational formulas are very different. Siegel and Castellan (1988) express the relationship of the two measures in terms of the inequality:

Kendall tau and Spearman R imply different interpretations: Spearman R can be thought of as the regular <u>Pearson product moment correlation coefficient</u>, that is, in terms of proportion of variability accounted for, except that Spearman R is computed from ranks. Kendall tau, on the other hand, represents a probability, that is the difference between the probability that in the observed data the two variables are in the same order versus the probability that the two variables are in different orders.

Gamma. The Gamma statistic (Siegel & Castellan, 1988) is preferable to Spearman R or Kendall tau <u>when the data contain many tied observations.</u> In terms of the underlying assumptions, Gamma is equivalent to Spearman R or Kendall tau. In interpretation and computation it is more similar to Kendall tau than Spearman R. Gamma is also a probability. It is the difference between the probability that the rank ordering of the two variables agree minus the probability that they disagree, divided by 1 minus the probability of ties. Thus, Gamma is basically equivalent to Kendall tau, except that ties are explicitly taken into account.

Wilcoxon rank sum, **Kendall's S and the Mann-Whitney U test** are exactly equivalent tests. In the presence of ties the Mann-Whitney test is also equivalent to a chi-square test for trend.

In most circumstances a two sided test is required; here the alternative hypothesis is that x values tend to be distributed differently to y values. For a lower side test the alternative hypothesis is that x values tend to be smaller than y values. For an upper side test the alternative hypothesis is that x values tend to be larger than y values.

The Mann Whitney U statistic is a method for comparing two independent random samples (x and y). Samples of size n1 and n2 are pooled and Ri are the assigned ranks. U expresses the number of times observations in one sample precede observations in the other sample in the ranking.

Assumptions of the Mann-Whitney test include random samples from populations; independence within samples and mutual independence between samples; measurement scale is at least ordinal.

A confidence interval for the difference between two measures of location is provided with the sample medians. The assumptions of this method are slightly different from the assumptions of the Mann-Whitney test: random samples from populations; independence within samples and mutual independence between samples; two population distribution functions are identical apart from a possible difference in location parameters

Kappa Statistics: an index that compares "agreement" against what might be expected by chance. Kappa can be thought of as the chance-corrected proportional agreement. Possible values range from +1 (perfect agreement) via 0 (no agreement above that expected by chance) to -1 (complete disagreement).

To assess the accuracy of any particular measuring 'instrument', it is usual to distinguish between the *reliability* of the data collected and their *validity*. **Reliability** is essentially the extent of the agreement between repeated measurements. **Validity** is the extent to which a method of measurement provides a true assessment of that which it purports to measure.

When studying the variability of observer *categorical* ratings, two components of possible lack of accuracy must be distinguished. The first is inter-observer bias, which is reflected in differences in the marginal distributions of the response variable for each of the observers (Cochran's Q-test is the appropriate test for the hypothesis of no inter-observer bias). The second is observer disagreement, which is indicated by how observers classify individual subjects into the same category on the measurement scale (Kappa coefficient is one of the most common approaches). In this part, we will focus on the Kappa coefficient (or Kappa statistics).

Example using Kappa: 29 patients examined by two independent raters (see Table). 'Yes' means the patient has a specific disease, whereas 'No' means the patient does not have that disease.

		Rater A		
		No	**Yes**	**Totals**
Rater B	No	10 (34.5%)	7 (24.1%)	17 (58.6%)
	Yes	0 (0.0%)	12 (41.4%)	12 (41.4%)
Total		10 (34.5%)	19 (65.5%)	29

Kappa = (Observed agreement - Chance agreement)/(1 - Chance agreement)
Observed agreement = (10 + 12)/29 = 0.76
Chance agreement = 0.586 * 0.345 + 0.655 * 0.414 = 0.474
Kappa = (0.76 - 0.474)/(1 - 0.474) = 0.54

Parametric statistics are statistics that estimate population parameters. A parameter is a constant in an equation that varies in other equations of the same general form, or one of a set of measurable factors such as temperature and pressure that define a system, determine its behavior, and may vary in an experiment.

Parametric inferential statistical methods are mathematical procedures for <u>statistical hypothesis testing</u> which assume that the distributions of the variables being assessed belong to known parametrized families of <u>probability distributions</u>. In that case we speak of <u>parametric model</u>.

All parametric tests involve (a) estimating at least one population parameter, (b) assumptions about the distribution of the population from which the data were randomly sampled, and (c) assumptions about the measurement of the <u>dependent variable</u>.

For example, <u>analysis of variance</u> **(ANOVA)** assumes that the underlying distributions are <u>normally</u> distributed and that the <u>variances</u> of the distributions being compared are similar. The <u>Pearson product-moment correlation coefficient</u> also assumes normality.

While parametric techniques are robust – that is, they often retain considerable <u>power</u> to detect differences or similarities even when these assumptions are violated – some distributions violate the assumptions so markedly that a <u>non-parametric</u> alternative is more likely to detect a difference or similarity.

ANOVA: While the t-test is usually used to compare means between two groups, ANOVA is more appropriately used to compare means between 3 or more groups.

There are several varieties of ANOVA, such as one-factor (or one-way) ANOVA, two-factor (or two-way) ANOVA, and so on, and also repeated measures ANOVA. Factors are independent variables, each of which must be measured on a categorical scale - that is, levels of the independent variable must define separate groups.

One-Way ANOVA Example

One-factor ANOVA, also called one-way ANOVA is used when the study involves 3 or more levels of a single independent variable. For example we might look at average test scores for students exposed to one of three different teaching techniques (three levels of a single independent variable).

ANOVA Statistics

The null hypothesis for ANOVA is that the mean (average value of the dependent variable) is the same for all groups. The alternative or research hypothesis is that the average is not the same for all groups.

The ANOVA test procedure produces an F-statistic, which is used to calculate the p-value. In the most rigid interpretation, $p < .05$ indicates that the null hypothesis if false. However, there may be clinically important information hidden with p values under .10.

With ANOVA, if the null hypothesis is rejected, then all we know is that at least 2 groups are different from each other. In order to determine which groups are different from which, post-hoc t-tests are performed using some form of correction (such as the Bonferroni correction) to adjust for an inflated probability of a Type I error.

Within effects: In a <u>repeated measures</u> <u>ANOVA</u>, there will be at least one factor that is measured at each level for every subject. This is a <u>within</u> (repeated measures) factor. For example, in an experiment in which each subject performs the same task twice, trial number is a within factor. There may also be one or more factors that are measured at only one level for each subject, such as gender. This type of factor is a <u>between</u> or grouping factor.

MANOVA Interaction: In <u>multi-factor analysis of variance</u>, factors A and B interact if the effect of factor A is not <u>independent</u> of the level of factor B. For example, in a drug experiment involving rats, there would be an interaction between the factors sex and treatment if the effect of treatment was not the same for males and females.

Repeated measures ANOVA: In a repeated measures <u>ANOVA</u>, there will be at least one factor that is measured at each level for every subject in the experiment. This is a <u>within</u> (repeated measures) factor. For example, in an experiment in which each subject performs the same task twice is a repeated measures design, with trial (or trial number) as the within factor. If every subject performed the same task twice under each of two conditions, for a total of 4 observations for each subject, then both trial and condition would be within factors. In a repeated measures design, there may also be one or more factors that are measured at only one level for each subject, such as gender. This type of factor is a <u>between</u> or grouping factor.

Linear logistic model: A linear logistic model assumes that for each possible set of values for the independent (X) variables, there is a probability p that an event (success) occurs. Then the model is that Y is a linear combination of the values of the X variables:

Linear regression: In a linear regression, the fitted (predicted) value of the response variable Y is a linear combination of the values of one or more predictor (X) variables:

fitted Y = b0 + b1*X1 + b2*X2 + ... + bk*Xk.

An X variable in the model equation could be a nonlinear function of an observed variable (e.g., one might observe distance, but use distance squared as an X variable in the model, or X2 might be the square of X1), as long as the fitted Y remains a sum of terms that are each an X variable multiplied by a coefficient. The most basic linear regression model is <u>simple linear regression</u>, which involves one X variable: fitted Y = b0 + b1*X.

Multiple linear regression refers to a linear regression with more than one X variable. **The purpose of linear regression is to "predict"**

It determines the extent to which there is a linear relationship between a dependent variable and one or more independent variables. There are two types of linear regression, simple linear regression and multiple linear regression.

In **simple linear regression** a single independent variable is used to predict the value of a dependent variable.

In **multiple linear regression** two or more independent variables are used to predict the value of a dependent variable. The difference between the two is the number of independent variables. In both cases there is only a single dependent variable.

Linear Regression - Data Considerations

The dependent variable must be measured on a continuous measurement scale (e.g. 0-100 test score) and the independent variable(s) can be measured on either a categorical (e.g. male versus female) or continuous measurement scale. There are several other assumptions that the data must satisfy in order to qualify for linear regression.

Correlation and Regression

Simple linear regression is similar to correlation in that the purpose is to measure to what extent there is a linear relationship between two variables. The major difference between the two is that **correlation makes no distinction between independent and dependent variables** while linear regression does. In particular, **the purpose of linear regression is to "predict"** the value of the dependent variable based upon the values of one or more independent variables.

Path Analysis is a sophisticated statistical mapping procedure for use with complex multi-factorial designs. This is often used in studies of the etiology or causes of behaviors. It could be used in a study of OMM outcomes if multiple factors are hypothesized to contribute to an outcome.

Survival Analysis This is also a sophisticated statistical tool. It is used to predict the end point for subjects in a study for whom the data is not collected. It may predict, for example what outcomes might occur in a sample of patients with pulmonary dysfunctions over time, given a set of initial parameters. Often it is referred to the prediction of when a subject will die.

Confidence Intervals: We generally use CIs to express the results of statistical tests because they convey more information than P values alone.

The confidence level sets the boundaries of a confidence interval. The confidence level is conventionally set at 95% to coincide with the 5% convention of statistical significance in hypothesis testing. In some studies wider (e.g. 90%) or narrower (e.g. 99%) confidence intervals will be more appropriate or required. This depends on the nature of the study.

Common **A 95% CI is the interval that you are 95% certain contains the true population value as it might be estimated from a much larger study**. The value in question can be a mean, the difference between two means, a proportion, etc. The CI is usually, but not necessarily, symmetrical about this value.

Pure Bayesian The Bayesian concept of a **credible interval** is sometimes put forward as a more practical concept than the confidence interval. For a 95% credible interval, the value of interest (e.g. size of treatment effect) lies with a 95% probability in the interval. This interval is then open to subjective molding of interpretation. Furthermore, the credible interval can only correspond exactly to the confidence interval if prior probability is so called "uninformative".

Pure frequentist Most pure frequentists say that it is not possible to make probability statements, such CI interpretation, about the study values of interest in hypothesis tests.

Odds Ratio: http://www.bmj.com/cgi/reprint/320/7247/1468?ck=nck
In recent years odds ratios have become widely used in medical reports. Odds ratios provide an estimate (with confidence interval) for the relationship between two binary ("yes or no") variables. Second, they enable us to examine the effects of other variables on that relationship, using logistic regression. Third, they have a special and very convenient interpretation in case control studies (dealt with in a future note). An odds ratio is a way of representing probability, especially familiar for betting. For example, the odds that a single throw of a die will produce a six are 1 to 5, or 1/5. The "odds" is the ratio of the probability that the event of interest occurs to the probability that it does not. "Odds" is often estimated by the ratio of the number of times that the event of interest occurs to the number of times that it does not.

Boxplot: A boxplot is a graph summarizing the distribution of a set of data values. The upper and lower ends of the center box indicate the 75th and 25th percentiles of the data, the center box indicates the median, and the center + indicates the mean. Suspected outliers appear in a boxplot as individual points o or x outside the box. The o outlier values are known as outside values, and the x outlier values as far outside values. If the difference (distance) between the 75th and 25th percentiles of the data is H, then the outside values are those values that are more than 1.5H but no more than 3H above the upper quartile, and those values that are more than 1.5H but no more than 3H below the lower quartile. The far outside values are values that are at least 3H above the upper quartile or 3H below the lower quartile.

Study Timetable and Management Plan

A research proposal should always contain a timetable and management plan, however brief it may be. It lets the reader know that you have considered the logistics of carrying out your research plan. The shape of this depends on your research, particularly if it is sequential, or multi-center, or involves a data safety and management board (DSMB) or other advisory groups. It lists the key steps involved in the project from start-up to final report and marks the weeks or months during which these will occur. It may be similar in format to the protocol time schedule and assures the reviewer that you have a plan in mind even though it may be modified at a later time. SEE THE EXAMPLE IN CHAPTER I.

You should be able to write a Research Design and Methods section using a non-experimental design in two to three pages. Flow charts are great to use.

The guiding principle for writing the Method section is that it should contain sufficient information for the reader to determine whether the plan you have to collect information from specific sources is logical, well considered, free of gaps (or you can address any inherent gaps), free of leaps of logic, doable, appropriate to the question.

A good design and methods section should contain sufficient details for another qualified researcher to implement the study.

THE LIST BELOW is for you to check these aspects of your proposal or manuscript. These are the 10 most often cited reasons for proposals or manuscripts to be rejected. These reasons are taken from many different sources that repeat these reasons, thus they are presented here as the most often cited

TOP TEN REASONS FOR REJECTION OF MANUSCRIPTS AND PROPOSALS.

- ☐ Failure to provide the proper context to frame the research question
- ☐ Failure to cite landmark studies
- ☐ Failure to accurately present the theoretical and empirical contributions by other researchers
- ☐ Failure to stay focused on the research question
- ☐ Failure to develop a coherent and persuasive argument for the proposed research
- ☐ Too much detail on minor issues, but not enough detail on major issues
- ☐ Too much rambling -- going "all over the map" without a clear sense of direction.(The best proposals move forward with ease and grace like a seamless river)
- ☐ Too many citation lapses and incorrect references
- ☐ Too long or too short
- ☐ Sloppy writing

Can you decide which is number one and which is number 1 and which is number 10?

FINAL CHECKLIST

The Research Design and Methods section is the largest section of the proposal. It can be organized in many different ways according to the nature of your research. The Research Design and Methods section must describe the below listed aspects of the study. This list is presented here as a checklist for you to use as you develop this section of your proposal or scientific report.
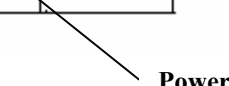
☐ How you will achieve you specific aims

☐ What activities you will undertake to test the hypotheses for your study

☐ What you will do in your experiment, or what you will do to the human or animal subjects

☐ How you will do it

☐ When

☐ Where

☐ How

☐ How often

☐ Why

☐ What you will use to determine if the action you propose will generate the reaction or the result/outcome you hypothesize that it will generate – your measurements

☐ Subjects: Who will be in the study? How will you select them by inclusion and exclusion criteria, and how will you recruitment them? (NOTE: Human subjects (risk and exclusion justification) issues are covered in a separate section E and do not go here)

☐ The independent variables: what you are observing or what you expect to change – the outcomes or measures of change

☐ The 'Power' of the study: see guidelines on when and how to obtain a power analysis to estimate sample size

☐ The dependent variables: what you are using as an intervention or treatment

☐ Other data such as categorical variables

☐ The hypotheses to be tested, question to be explored, or the product to be developed

☐ The design of the study: randomized block, repeated measures, longitudinal, retrospective, prospective, cross-sectional, existing data, controlled, blinded, and the like

☐ The instrumentation to collect data, and to measure the experimental change, the clinical outcomes, the validity of the theory, or the utility/validity of the product

☐ The statistical tests you plan to use to test the hypotheses or question

Power Analysis and Sample Size Estimation
Daisha J. Cipher, Ph.D.
Department of Biostatistics, SPH

The *power* of a statistical test is the probability of correctly rejecting a false null hypothesis. This probability is inversely related to the probability of making a Type II error. Recall also that we choose the probability of making a Type I error when we set Alpha and that if we decrease the probability of making a Type I error we increase the probability of making a Type II error. The relationships are defined in the table below:

Classical Hypothesis Testing

The Classical Hypothesis Testing approach involves stating a null hypothesis (H_0; statement of no effect), setting an alpha level (α; Type I error), a beta level (β; Type II error), power of the statistic ($1-\beta$), computing the relevant statistic(s) to address the null hypothesis, and finally, rejecting or accepting the null. In the classical approach, all of these steps are determined *a priori* to conducting the study.

Power and Alpha

The probability of correctly retaining a true null has the same relationship to Type I errors as the probability of correctly rejecting an untrue null does to Type II error. Yet, as I mentioned above, if we decrease the odds of making one type of error, we increase the odds of making the other type of error. What is the relationship between Type I and Type II errors? The probability of making a Type II error (and thus having low power) varies as a function of Alpha (Type I error). The lower our Alpha, the less likely we are to make a Type I error, but the more likely we are to make a Type II error.

references: Kirk (1988), Maxwell & Delaney (1990), Stevens (1996), Winer (1991)
Cipher 2001

Anytime we test whether a sample differs from a population or whether two sample come from 2 separate populations, there is the assumption that each of the populations we are comparing has its own mean and standard deviation (even if we do not know it). The distance between the two population means will affect the power of our test. In the following demonstration an increase in the variance (the spread of the distribution) shows a corresponding overlap in the two distributions and an increase in Beta.

Power as a Function of Sample Size, Variance, and Effect Size

You should notice in the last demonstration that what really made the difference in the size of Beta was how much overlap there was in the two distributions. When the means were close together the two distributions overlapped a great deal compared to when the means were farther apart. Thus, anything that effects the extent the two distributions share common values will increase Beta (the likelihood of making a Type II error).

Sample size has an indirect effect on power because it affects the measure of variance we use to calculate the t-test statistic. Since we are calculating the power of a test that involves the comparison of sample means, we will be more interested in the standard error (the average difference in sample values) than standard deviation or variance by itself. Thus, sample size is of interest because it modifies our estimate of the standard deviation. When N is large we will have a lower standard error than when N is small. In turn, when N is large well have a smaller Beta region than when N is small.

Effect size measures, such as Cohen's d, R^2, or η^2 are indications of the magnitude of the difference, variance shared, or variance accounted for (respectively) in your variables. The most common effect size measure for a two-sample design is the Cohen's *d*. The d is calculated as:

$$d = \frac{\overline{X}_1 - \overline{X}_2}{s}$$

where the standard deviation (*s*) would be that of either group, considering that both groups are assumed to have approximately equal SDs.

The resulting *d* value represents the difference between the means of Group 1 and 2, in standard deviation units. Thus, a value of 1.0 means that the two groups differed exactly one standard deviation unit from one another. Cohen (1965) defined effect size as "the degree to which a phenomenon exists." He described effect sizes (*d* values) as having the following qualities:

d	
.25	small
. 50	medium
1.0	large

One can still use Cohen's *d*, even if the design has more than two groups. The *d* is calculated for one pair at a time.

Estimating Sample Size and Power for Experimental Designs

How do you ensure that your experiment's sample size is adequate? Moreover, how can you know the extent of the power of your statistical test? The procedure called *power analysis* can answer both of these questions. The four factors involved in a power analysis are:

1. Level of significance (α, or alpha level) set by you, the researcher.
2. Probability of obtaining a significant result (power desired, or 1 - β).
3. The population effect size, or the hypothesized effect (or difference) between your groups.
4. Sample size.

Knowing any three of the above allows you to compute the fourth.

The most common power analysis actually entails estimating a sample size before the study begins. The following formula estimates adequate sample size for a two-sample design, two-sided test:

$$ n = \frac{2(z_{1-\alpha/2} + z_{1-\beta})^2}{d^2} $$

where:

$z_{1-\alpha/2}$ refers the z value that cuts off the middle 95% of a normal distribution

$z_{1-\beta}$ refers the desired power level

d   is the Cohen's d value (see formula above)

c:\…PowerAnalysis.doc   page 3
references: Kirk (1988), Maxwell & Delaney (1990), Stevens (1996), Winer (1991)
Cipher 2001

The UNTHSC-TCOM Department of Psychiatry   58   ©dAC 08

Typically, $z_{1-\alpha/2}$ is 1.96 and $z_{1-\beta}$ is .80 (which means that you have a 20% chance of making a Type II error).

If you don't want to do power analyses by hand, you can refer to the many tables in the literature that will estimate sample size, power, or both. Otherwise, you can look on the Internet for online power calculators. Some popular ones are:

http://calculators.stat.ucla.edu/powercalc/
http://www.math.yorku.ca/SCS/Demos/power/

When you finally compute the sample size that you need, you can interpret your findings as follows:

"I must have X amount of participants **in each group** to achieve a statistical power of .80, using a significance level of .05."

If you do NOT have X sample size per group above, if you run your analyses anyway and do not achieve significance, **you cannot differentiate between the possibilities that:**

1. There are no differences in the population; or
2. There ARE differences but you don't have the adequate sample size to indicate that.

Example of Sample Size Calculation For 2 Samples Using Online Power Calculator

Using the UCLA Department of Statistics website (calculators.stat.ucla.edu/powercalc), we can estimate our sample size for a two-sample design. Using the following hypothetical data reflecting IQ scores from two groups, we'll compare the calculator's results to our hand calculations.

Group 1 $\overline{X}$ = 100 $s = 15$ α= .05   β= .20   1 - β = .80
Group 2 $\overline{X}$ = 107.5
(this means you're estimating d to equal .50)

Power Calculator

Choose a Model and Push a Button. Disclaimer.

NORMAL	Power for a given Sample Size	Sample Size for a given Power
1 Sample	○	●
2 Sample, Equal Variances	○	●
2 Sample, Unequal Variances	○	●
Lognormal	○	●
EXPONENTIAL	**Power for a given Sample Size**	**Sample Size for a given Power**
1 Sample	○	●
2 Sample	○	●
BINOMIAL	**Power for a given Sample Size**	**Sample Size for a given Power**
1 Sample	○	●
1 Sample Arcsine	○	●
2 Sample Arcsine	○	●
2 Sample Median	○	●
Fisher's Exact Test	○	●
Proportion Responders	○	●
Case Control	○	●
POISSON	**Power for a given Sample Size**	**Sample Size for a given Power**
1 Sample	○	Not Available
2 Sample	○	Not Available
CORRELATION COEFFICIENT	**Power for a given Sample Size**	**Sample Size for a given Power**
1 Sample	○	●

Click this button to estimate sample size

UCLA Department of Statistics
Last updated: 22-May-2002
Access count is: 150755, since 18-Mar-2002
Maintained by: webstaff [webstaff@stat.ucla.edu]

Normal Power Calculations

Normal Distribution 2-Sample Equal Variances

μ_1 The Mean of Population 1	100
μ_2 The Mean of Population 2	107.5
Sigma Common Standard Deviations for both Populations	15

Number of Sides
Specifies Alternative Hypothesis.
One sided and $\mu_1 > \mu_2 \Rightarrow H_1 : \mu_1 > \mu_2$
One sided and $\mu_1 < \mu_2 \Rightarrow H_1 : \mu_1 < \mu_2$
Two sided $\Rightarrow H_1 : \mu_1$ not equal μ_2

○ 1 Side
◉ 2 Sides

Significance Level
The Significance Level of the test or Prob(reject null hypothesis ($H_0 : \mu_1 = \mu_2$) given it is true)   .05

Power
The Power desired for the test or Prob(reject H_0 given that H_a is true)   .80

Submit Query

c:\…PowerAnalysis.doc   page 5
references: Kirk (1988), Maxwell & Delaney (1990), Stevens (1996), Winer (1991)
Cipher 2001

The UNTHSC-TCOM Department of Psychiatry   60   ©dAC 08

Result:

N-1: 63.

N-2: 63.>
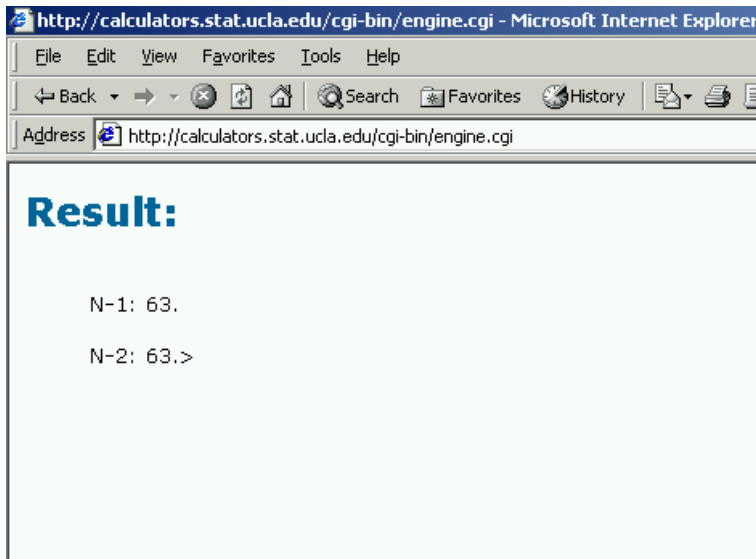
The online calculator estimated that we need 63 observations in EACH group to achieve a power of .8 at alpha=.05, estimated effect size, .50.

Now, using our equation, we can plug in our numbers to compare results:

$$n = \frac{2(z_{1-\alpha/2} + z_{1-\beta})^2}{d^2}$$

$$n = \frac{2(.80 + 1.96)^2}{.50^2}$$

$$n = \frac{2(.80 + 1.96)^2}{.50^2}$$

$$n = \frac{15.24}{.25}$$

$$n = 60.96$$

Note that our estimated sample size, 61, is a little smaller than the estimated sample size from the online calculator. The online calculator probably uses 2.0 to represent $z_{1-\beta}$, instead of 1.96. When we use 2.0 instead of 1.96, our formula yields 62.72.

c:\...PowerAnalysis.doc   page 6
references: Kirk (1988), Maxwell & Delaney (1990), Stevens (1996), Winer (1991)
Cipher 2001

The UNTHSC-TCOM Department of Psychiatry   61   ©dAC 08

APPENDIX IV - B
RESEARCH PROPOSAL WORKSHEETS

GUIDE FOR WORKSHEETS

The sections in this appendix will help you write statements and draw figures to think though specific parts of your research project. Completing these questions will help you make connections between various parts of your research plan. You will literally sketch out your ideas, steps, and activities you will need to guide you from idea through formulation of hypotheses and design to implementation. Your responses to these items and prompts should take you from the first step on day one to the last step on the day that you will write your report.

THE QUESTION

This section will help you formulate your primary research question or hypothesis. You should finish this section with one sentence. The problem can be stated as a question (e.g. Is there a relationship between the medication protocol for crisis management at intake to the emergency center and length of stay?) or as a hypothesis (e.g. Adherence to the treatment plan depends on the level of satisfaction the patient has with the case management services.) OR if you have two groups you might be interested in whether adolescents in an inpatient unit have different scores in a selected cognitive functioning test compared to adolescents in a juvenile detention facility.

If you are going to conduct a study with existing data you may have a research question such as "Does presenting problem at admission to the emergency center predict length of stay for those patients referred to the inpatient unit following crisis stabilization?" Or you might ask whether patients not referred to the inpatient unit following crisis stabilization relapse to another intensive level of care within the next 60 days.

1. First you need to craft your statement of the problem of interest Consider these points:
   Is your interest focused on a specific condition or diagnosis?
   Are you interested in a particular segment of the clinical populations?
   Will you want to compare two or more groups or focus on the characteristics and changes in one group?

2. Write one sentence for your main research question or hypothesis. It must withstand the test of being focused and answerable.

3. Is this problem or question of long-term interest such as relapse factors in substance abusers? Or will it address an immediate policy issue such as formulary or triage methods?

4. Next, consider why this research topic or question is important to study.
   Think this way: "I want to study "this" because- - - - - - .
   I think that the length of stay has increased in the inpatient unit since we have stopped using atypical antipsychotics".
   Here you are attending to the "so what" test. If the data has never been examined, why is it being collected, what can it tell us about our system of care or this group of patients, or

a community need? Will your study possibly generate new questions that need to be researched? Is there a study you have read that you want to replicate, or a gap in the research literature on a topic that pertains to a particular population?
Write why this question, topic or hypothesis is important:

5. To answer your question or test your hypothesis what research design will you use?

   Experimental _____ Number of groups _____, control group? _____

   Non-experimental ___
   Prospective or retrospective cohort
   Cross-sectional study
   Case-control
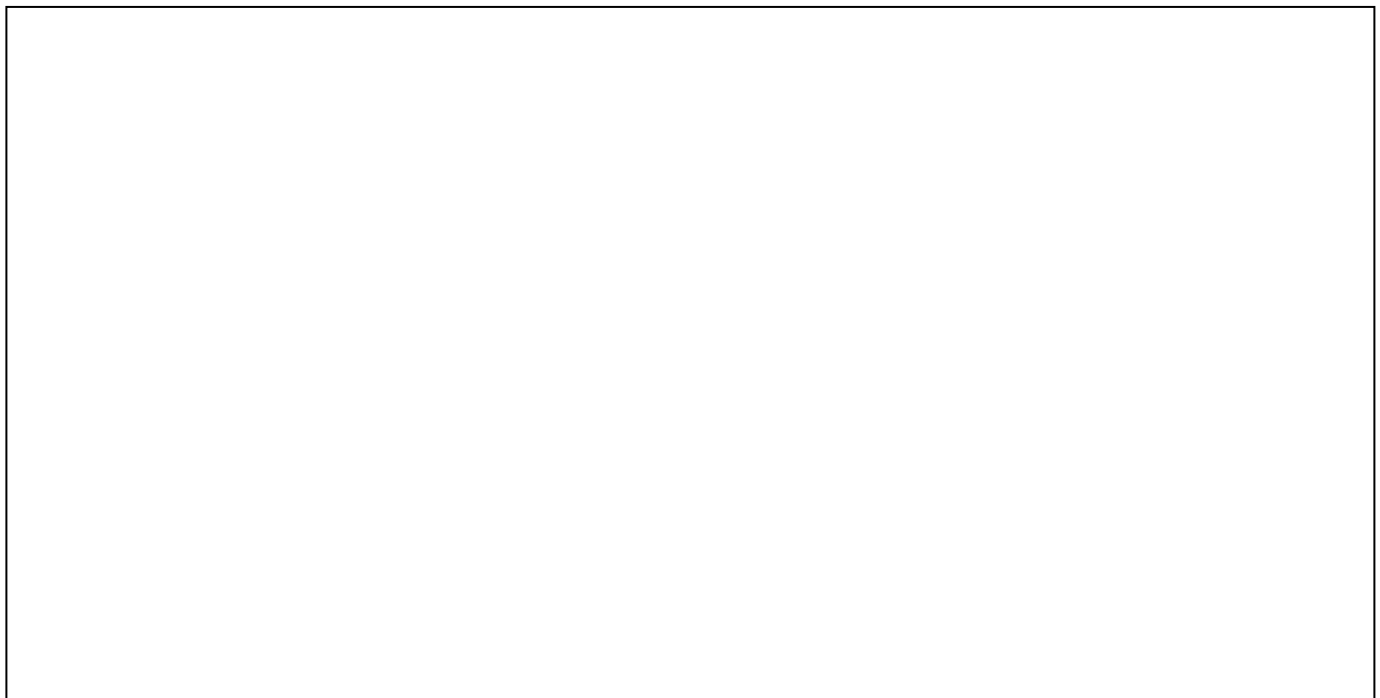
6. What variables or measures will you use to answer or explore your research question?

In this next section you will be guided through several steps that will help you develop a plan for your project. You should do this independently and review it with your research mentor.

TASKS AND TIME LINES

7. Right now, as you conceive of it, draw a conceptual flow chart or diagram of exactly what you will do in your research. This helps form the methods section of your proposal. You may come back and change this later, but start now.

8. Right now, as much as you can, associate times such as days and weeks with the tasks needed to achieve your research goals. Start with when it will end and work backwards. Include, for example, where you will go, what you will get or see, what tools you will need to collect information and where you will get them, and what your subjects will do and when. Think about who else will be involved from the IRB to your major professor and any staff involved or touched by this project

WHAT/TASK	WHEN	WHO
Example: Power analysis to determine sample size using information from published research on effect size – amount of change or difference anticipated	When I have my question or hypothesis finished	A biostatistician or my mentor

OUTCOMES – MEASUREMENTS

9. **What are the outcome measures of interest?** These are your variables.
Specifically what do you expect to change or to happen as a result of the intervention?

Or what are you going to analyze as a measure of change or difference?
How will you define your outcomes of interest?
For example, if you are determining differences in diagnosis and self-efficacy in adolescents in the inpatient unit compared to a group in the juvenile detention facility, you will want to determine how you will measure "diagnosis" and "self-efficacy." If you are examining the relationship between length of stay and medications at admission to a crisis unit, how will you define length of stay? How will you classify the medications and their use such as verifying that they have been in fact taken as prescribed.
Use the table on the next page to list the things you want to measure as an outcome or as an observed phenomenon, define them, and identify where or how you will get the information.

OUTCOME VARIABLE	DEFINITION	SOURCE

10. How will you collect this information?

11. If you are using information from existing records, what form will you use to record them? Do you have to create a data collection form (DCF) for your study? Does a DCF exist?

DCF	SOURCE AND OTHER INFORMATION

12. Name the Independent (what will not vary or change) and Dependent Variables (your outcome variables of interest measured as change in a condition, a number of actions or events, or a status). This will help you determine how you will analyze the data.

13. Are you comparing outcomes for different groups, such as inpatient length of stay by length of time since first known onset of the psychiatric condition, or one group with schizophrenia taking a particular medication with a matched group on a different regimen? List any grouping variables that will be independent variables or factors such as age, gender, presence or absence of another condition or another characteristic.

14. List the variables you will test as predictor or criterion variables, such as a study that hypothesizes that gestational age, number of previous pregnancies, and chronological age of the patient may predict level of depressive symptoms.

INCLUSION/EXCLUSION CRITERIA

15. List the characteristics or criteria by which you will select your records or subject for the study (inclusion criteria).

List those by which you will exclude records or subjects (exclusion criteria). (e.g. age, health condition, lab value, time periods, services criteria, locations)

INCLUSION	EXCLUSION

ACQUISITION AND RANDOMIZATION

16. How will you **acquire the sample** of records or persons? Where will you get the information or subjects, how will you recruit or access them?

17. If this is an **experimental design**, what are the experimental groups? How will you make random assignment?

REVIEW AND REVISE

18. **REVIEW:** Look back at the question or hypothesis you wrote on page one of the worksheets.

If necessary, REWRITE one sentence for your main research question or hypothesis. It must withstand the test of being focused and answerable.

19. Ask your research mentor. or colleague, to read the question you wrote and have them ask you questions about it. Have a discussion to refine these key items.

After revising or re-writing your primary question or hypothesis, review the drawing of your research question, methods and outcomes and the time-table in items 7 and 8. REVISE THESE IN THE SPACE BELOW IF YOU NEED TO DO THAT NOW.

(7) Flow chart or diagram of exactly what you will do in your research.

(8) Time-table

WHAT/TASK	WHEN	WHO
Power analysis to determine sample size using information from published research on effect size – amount of change or difference anticipated	When I have my question or hypothesis finished	A biostatistician

20.   REVIEW AND REVISE THE DCF LIST AND USE IT TO ANSWER THESE QUESTIONS

20.1.   What will your data collection forms look like? Attach samples. If you develop your own you must test them and adjust them before you begin the data collection.

20.2.   Are the collection instruments in the public domain, or is cost associated? Is a license required? It is important to not accept someone's opinion that you can use an instrument. You would want credit given to you if you developed an instrument. Also, there are many good public domain or other-student-developed instruments available.)

20.3.   How will you document validity and/or reliability of the measurement instruments?

20.4.   Will you attempt to "generalize" to a larger population or other populations?

20.5.   Which computer software will you use to enter data?

20.6.   How and where will data be stored?

20.7.   How will you protect the data from errors or tampering?

20.8.   How will data be guaranteed as confidential or anonymous?

20.9.   Identify specific data analysis approach and appropriate statistical tests of the data.

20.10. Identify data characteristics that may influence the analysis methods.

GUIDE TO THE CRITICAL EVALUATION OF A RESEARCH ARTICLE

ACTION: Read the abstract.

1a. What is your overall impression of the article?

1b. Do authors provide a strong rationale for conducting the study (e.g. is it an important clinical or public health issue)?

1c. Is it a study in the U.S.? _____ Does it matter? _____ How does it matter, or why not?

1d. If you can determine from the abstract, what is the study design?

1e. What is the primary research question or hypothesis?

ACTION: Scan or skim the entire article to gain an appreciation for the whole: Highlight main points in each section. Make note of new or unfamiliar terms. Search and define these terms to guide your reading.

2a. List new terms with a brief definition.

2b. Note your impression of the findings: (e.g. important for rare cases; new findings not reported before; preliminary and not definitive)

3. How the sample was drawn and how the study is powered?

Power analysis:
Sample size:
Groups?
Limitations?

Rate: Clear _____, Unclear _____

4. Does the study design suit the question/s? How? How would you improve the design?

5. Can you restate or draw a flow chart of each step used to conduct the research? Summarize it here:

6. What are the primary and secondary (if any) outcome measures or variables of interest? List all variables and their types.

_____
_____
_____
_____
_____

7. What statistical methods are used and how do the statistical analysis tools suit the type of data or not?

_____
_____
_____
_____

8. How did the researchers minimize variance, maximize validity and control for Type I or II errors?

_____
_____
_____
_____
_____

9. Is there a placebo role in this research? If so, does it contribute to the study validity?

_____
_____
_____
_____

10. What are the major findings of the study and what statistics were used to report the results?

_____
_____
_____
_____
_____

11. What are the strengths and limitations of the study?

_____
_____
_____
_____
_____

12. What did you need/use to evaluate the article (e.g. previously published article, medical/scientific terms, a questionnaire)?

_____
_____
_____
_____

13. Briefly state your interpretation of the major findings, and an assessment of their validity.

_____
_____
_____
_____
_____

14. What are the clinical implications of this study?

_____
_____
_____
_____
_____

TIERED RESEARCH COMPETENCY TABLE

TIER I BASIC RESEARCH COMPETENCY TO ACHIEVE BY THE END OF EACH YEAR AND SEMESTER From MSY 1 Semester 1, through completion of residency training:	MSY 1/2	MSY 2/1	MSY 2/2	MSY 3/2	MSY 4/2	PG 1/1	PG 1/2	PG 2/1	PG 2/2	PG 3	PG 4
Biostatistics and Study Design											
Recognize different types of the most common research designs	X					X					
Define the concept of validity of a measurement	X					X					
Define what is meant by reliability of a measurement	X					X					
Define what is meant by "sensitivity" and "specificity" of a testing method or instrument	X					X					
Identify and define basic, most commonly used types of variables	X					X					
Recognize why a power analysis is important for research	X					X					
Recognize the qualities of a well constructed research question		X					X				
Define the concept of validity of a research study		X					X				
Define the concept of reliability of a research study		X					X				
Define what is basically meant by a Type I error and a Type II error		X					X				
Use a decision tree to determine if a given research report used an appropriate statistical analysis for the question and the type of data		X					X				
Define the concept of the "power" of a study		X					X				
Identify what is needed to calculate/estimate sample size for a given study		X					X				
Answer questions to facilitate an evaluation of a simple to minimally complex research article		X					X				
Define the purpose of some of the most common basic statistical concepts and methods (e.g. normal distribution, skewness, kurtosis, variability, prevalence, incidence, Chi-Squared statistic, confidence intervals, Odds Ratios, Students' T, and Pearson correlation coefficient).		X					X				
Define the concept of t-scores and z-scores		X						X			
Distinguish between parametric and non-parametric statistics		X						X			
Identify some of the considerations for researchers in determining power and sample size			X					X			
Provide an oral interpretation of basic descriptive statistics for given research findings/observations			X					X			

TIER I BASIC RESEARCH COMPETENCY TO ACHIEVE BY THE END OF EACH YEAR AND SEMESTER From MSY 1 Semester 1, through completion of residency training:	MSY 1/2	MSY 2/1	MSY 2/2	MSY 3/2	MSY 4/2	PG 1/1	PG 1/2	PG 2/1	PG 2/2	PG 3	PG 4
Recognize a description of what is meant by "statistical assumptions" for selection of a statistical analysis tool			X					X			
Identify and provide examples of threats to the internal validity of a given study			X					X			
Generally understand the strengths and limitations of common research designs			X					X			
Recognize the use of moderately sophisticated statistical concepts (not the computation) (e.g. Linear Regression Analysis, Analysis of Variance, Cronbach's Alpha, and Factor Analysis)			X							X	
Identify the use and purpose of t-scores and z-scores			X							X	
Conducting Research											
Efficiently perform a focused literature search.			X				X	X			
May participate in the functions of a research team/experience in a lab or clinical research study			X	X			X				
Scientific writing, Presentation and teaching											
Identify essential components of a review of the literature for a given research question/topic	X					X					
Identify qualities of a well written abstract of a given research study	X					X					
Prepare and present an intramural review of the literature on a given selected research question/topic with a team/small group			X				X	X			
Incorporate research findings into a clinical case presentation		X						X			
Present an intramural evaluation of a simple to minimally complex research article			X					X			
May organize and present a research poster			X	X				X		X	
See TIER I Continued on next page											

TIER I BASIC RESEARCH COMPETENCY TO ACHIEVE BY THE END OF EACH YEAR AND SEMESTER From MSY 1 Semester 1, through completion of residency training:	MSY 1/2	MSY 2/1	MSY 2/2	MSY 3/2	MSY 4/2	PG 1/1	PG 1/2	PG 2/1	PG 2/2	PG 3	PG 4
Ethics											
Identify the basic principles and considerations for ethical research with human subjects	X					X					
Recognize the basic differences between the Statutes and Regulations of the Federal Drug Administration (FDA) and the Department of Health and Human Services (DHHS)	X					X					
Define the purpose of the Institutional Review Board			X				X				
Identify the essential elements governing human subjects research			X				X				
Discuss the role of a placebo treatment for a given study			X				X				
Identify and recognize strengths and limitations of selected epidemiologic research designs for clinical application		X	X				X				
Identify and discuss ethical issues in research pertaining to a clinical case scenario for a simple to minimally complex issue		X	X					X			
Recognize the essential components of a research study consent form			X					X			
Discuss and comment on ethical decisions in research aspects of clinical case scenarios			X					X			
Identify some of the differences between a patient and a research subject.			X	X				X		X	
Complete Human Subjects Ethics training (CITI or NIH)			X	X			X	X			
Complete Institutional Animal Care and Use Committee (IACUC) training for one model			X	X			X	X			
Research Leadership and Management											
No specific performance tasks in Basic Tier											

TIER II BASIC RESEARCH COMPETENCY TO ACHIEVE BY THE END OF EACH YEAR AND SEMESTER From MSY 1 Semester 1, through completion of residency training	MSY 1/2	MSY 2/1	MSY 2/2	MSY 3/2	MSY 4/2 & 5TH YR	PG 1/1	PG 1/2	PG 2/1	PG 2/2	PG 3	PG 4
Biostatistics and Study Design											
Describe the utility and characteristics of different research designs				X			X	X			
Formulate a research question								X	X		
Select a research design for practice and novel research questions				X				X			
Identify and discuss strengths and limitations of optional designs for a given research question							X				
Identify and provide an operational definition for primary outcome variables in a sample study								X			
Select outcome measures for selected sample research questions				X				X			
Select outcome measures for a novel research question								X	X		
Discuss strengths and limitations of different research designs according to the research question				X			X				
Discuss various potential threats to internal validity for a given study and determine ways to minimize these				X			X				
Discuss the importance of reliability of selected research studies and distinguish between theory building (inductive) studies, pilot studies, and definitive clinical studies				X			X				
Discuss sensitivity and specificity of various clinical tests/measures				X			X				
Interpret and discuss inferential statistics in simple to minimally complex research findings					X				X		
Determine whether the findings of a given research report are in keeping with the analysis of the primary research question				X				X			
Perform a simple practice power analysis for a given research hypothesis				X					X		
Identify primary outcome measures in a simple to minimally complex research study, and discuss their appropriateness and relevance to that selected research				X				X			
Discuss a power analysis and sample size calculation with a biostatistician for a sample study				X				X			
Evaluate the risk for a Type I and a Type II error in a given study				X	X			X	X		
Discuss how to control for error risks in a given study design					X				X		

TIER II BASIC RESEARCH COMPETENCY TO ACHIEVE BY THE END OF EACH YEAR AND SEMESTER From MSY 1 Semester 1, through completion of residency training	MSY 1/2	MSY 2/1	MSY 2/2	MSY 3/2	MSY 4/2 & 5TH YR	PG 1/1	PG 1/2	PG 2/1	PG 2/2	PG 3	PG 4
Critique a research article from the perspective of statistical assumptions for the analysis performed						X				X	
Use software to compute simple descriptive to simple inferential statistical tests and discuss the results for that sample research				X				X	X		
Conducting Research											
Prepare an abstract for a given study									X		
Participate in journal club with specified competencies demonstrated					X			X			
Scientific writing, Presentation and teaching											
Apply rules of English usage, style, and composition				X						X	
Construct an abstract for a novel study in the manuscript stage										X	
Participate in writing research report					X					X	
Distinguish nuances between Clinical and basic science research issues				X	X					X	
Write thesis or project report					X					X	X
Present a mentored scientific poster for a novel research study					X						X
Present a critical review of a minimally to moderately complex research article in an intramural and in an extramural venue					X					X	
Ethics in Research											
Critique a consent form										X	
Attend an IRB meeting										X	
Develop a consent form with all essential elements									X	X	
Write and compile an Institutional Review Board proposal										X	
Discuss ethical issues in a human and animal research											
Identify major issues of placebos in clinical research						X				X	
Present a critique of a randomized, blinded, controlled clinical trial (RBCT)											X
Review and discuss a proposal for animal research for the institutional animal care and use committee (IACUC)											X
Locate and critique an epidemiologic study										X	
Identify and discuss ethical issues in research pertaining to a clinical case scenario for a minimally to moderately complex issue				X						X	

TIER II BASIC RESEARCH COMPETENCY TO ACHIEVE BY THE END OF EACH YEAR AND SEMESTER From MSY 1 Semester 1, through completion of residency training	MSY 1/2	MSY 2/1	MSY 2/2	MSY 3/2	MSY 4/2 & 5TH YR	PG 1/1	PG 1/2	PG 2/1	PG 2/2	PG 3	PG 4
Discuss issues of a transition for a patient from under medical treatment to a participant in research					X						X
Research Leadership and Management											
Demonstrate effectiveness as a research team member											X
Define the role of the PI and Research Coordinator					X				X		
Discuss and assist in developing a grant budget and an operating budget					X					X	X
Participate in research project management meetings					X					X	X

Tier III Advanced Competencies (CIPP "learning objectives" Bakken et al.)

Biostatistics and Study Design
Formulate a research question and operationalize variables
Categorize research designs and state purpose and limitations of each
Compare and contrast major study designs, e.g. case controlled, cross-sectional
Develop a research plan that demonstrates a grasp of threats to validity and reliability
Interpret inferential statistics
State the relationship between the chosen research design and the type of data collected
Collect and analyze data using appropriate sampling and biostatistical techniques
Conduct a "power analysis" and estimate a sample size
Take methods to protect against Type I and II errors

Conducting Research
Read and critique literature in your own scientific domain
Organize your own reference library using computer technology
Spend sufficient time developing and advancing one's own area of scientific knowledge and research
Demonstrate laboratory skills

Scientific writing, Presentation and teaching
Apply rules of English usage, style, and composition
Apply process strategies for organizing and drafting journal articles and grant proposals
Write journal articles, research proposals, and grant applications according to specified format guidelines
Accurately report research findings citing the strengths and limitations of studies
Lead a discussion of a critical review of a moderately complex to complex research article
Report research in an ethically responsible way; Write and use a consent form
Use word-processing software to prepare grants, journal publications and other scientific documents
Prepare and deliver poster presentations for professional conferences
Lead small and large group discussions
Deliver a focused and well-organized lecture; use computer technology in presentation
Plan a teaching activity with appropriate scope, sequence and focus

Ethics
Describe and apply the process of obtaining informed consent
Be sensitized to issues involving the integrity of research
Make a principled decision when faced with an ethical choice
Respect the diverse ethical challenges associated with minority populations
Discuss selected key issues of placebos in clinical research
State institutional and governmental policies regarding the ethical conduct of research
Identify institutional sources of support when faced with an ethical dilemma
Discuss professionalism and its implications in science

Leadership/Management
Establish a network of professional colleagues
Identify appropriate funding sources
Recruit, manage, and evaluate personnel involved with a study
Prepare a study budget, reports, and other administrative documents
Obtain protocol approval from the human subjects, research, and other review committees
Plan and adhere to a timeline for research projects
Demonstrate mentoring skills