Biostatistics in Biomedical and Clinical Research

Sandra Taylor, Ph.D.
Clinical and Translational Science Center
University of California, Davis
30 April 2013

Do I really need a statistician?

- "My study is really simple."
- "I don't have funding to support a statistician."
- "I don't know any statisticians."
- "A statistician will make my study more complicated or turn it into a statistics research project."

Goals and objectives

- **Understand the role of statisticians in biomedical research and clinical trials**
  - What we can do for you
  - When we should be involved
  - Why it's important
- **Provide pointers for working with statisticians**
- **Understand what resources are available and how to access those resources**

Where are statisticians involved in research projects/clinical trials?

Where are statisticians involved in research projects/clinical trials?

EVERYWHERE!

Why and when to work with a statistician?

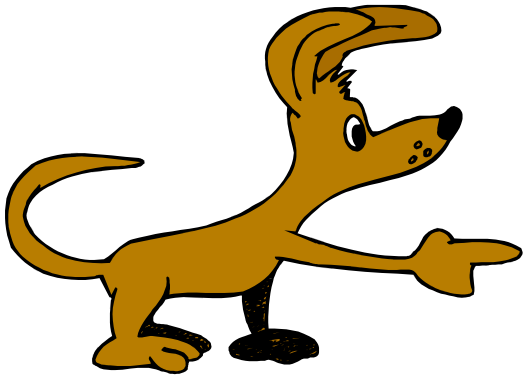
- **Planning the study**
  - Study design, randomization, sample size
  - Proposal preparation
- **Conducting the study**
  - Interim analyses, DSMB
- **Evaluating the results**
  - Conducting statistical analyses
- **Reporting the results**
  - Interpreting the results
  - Manuscript preparation

The Research Process

"Gee, I wonder if..."

Planning Stage Involvement

Planning Stage Involvement



Statistics can't fix a poorly designed study!

Planning Stage Activities
Not just sample size calculations

- Develop specific aims
- Identify endpoints and formulate testable hypotheses
- Identify confounding factors, biases
- Develop study design, randomization scheme, matching protocol
- Conduct sample size calculations and prepare statistical analysis plan

Formulating specific aims

- **Identify parameters of interest**
- **Specify testable hypotheses**
  - Determines statistical methods
- **Specific aims determine**
  - Study design
  - Sample size calculations
- **Identify fatal flaws**

Specific Aim 1: Determine if new treatment is better than standard care.

- **What constitutes *better?***
- **What measureable parameter reflects *better*?**
  - Survival, number of events, mean value?
- **What testable hypothesis addresses the specific aim?**
  - H_0: Mean cholesterol under new treatment does not differ from standard care
  - H_a: Mean cholesterol under new treatment differs from standard care

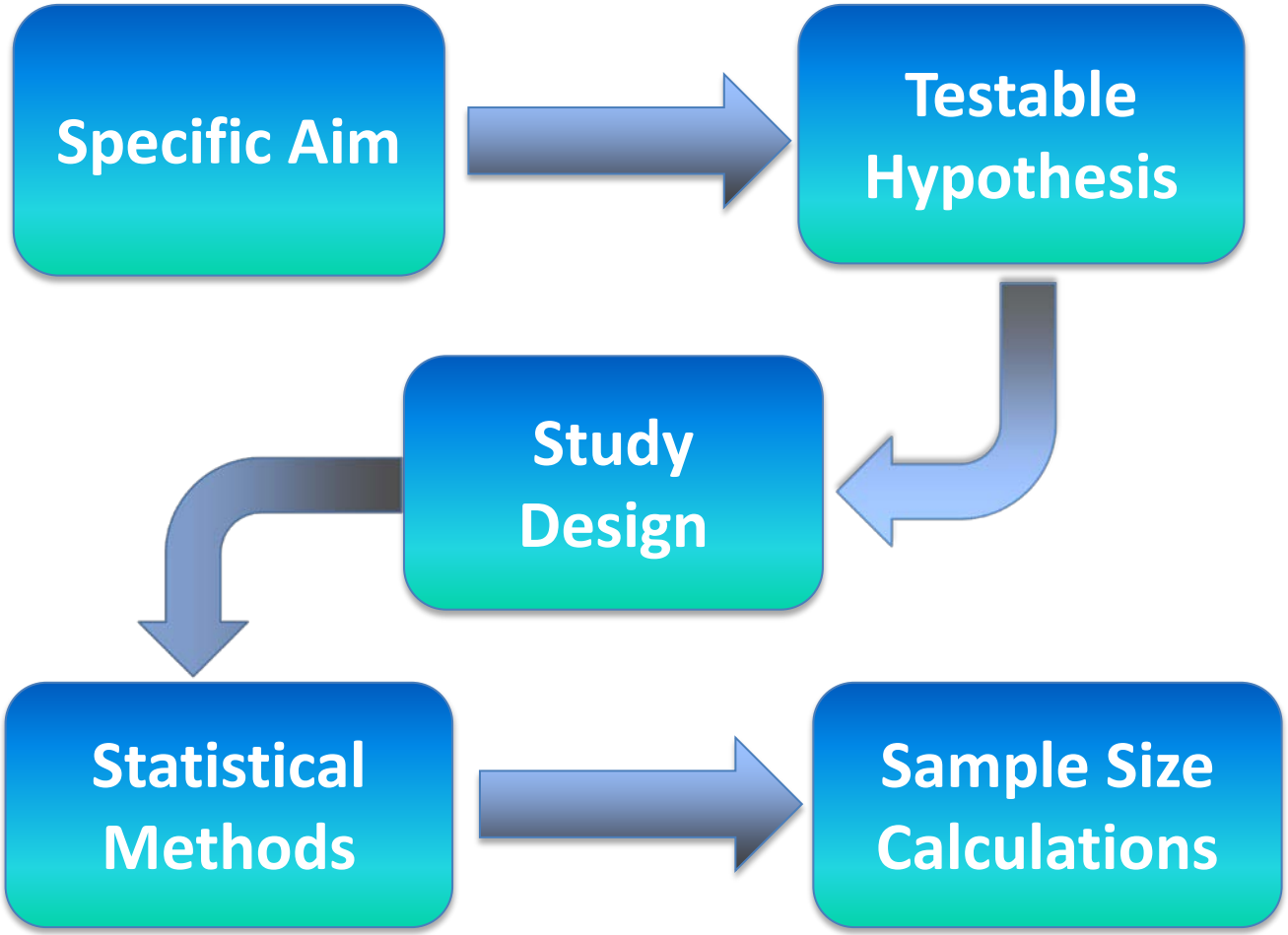
Study type and design considerations

- **Prospective Study**
  - Randomization
  - Stratification
  - Matching
  - Interim analyses, adaptive or sequential designs
- **Retrospective Study**
  - Confounding variables
  - Site effects and biases in data often overlooked

How many subjects do I need?

- Too few – insufficient power to detect differences
- Too many – unnecessary costs
- Statisticians need input from researchers to determine sample size requirements.

Sample size determination comes at the end of a series of steps.



How is sample size determined?

Depends on:

- **Specific aim – primary hypothesis of the study**
- **Study design**
  - These two influence the statistical test.
- **Effect size to be detected**
- **Variability of the response variable**
  - Researchers need to provide this information

Example: New medication study

- **Test: H_o: $\mu_{new} = \mu_{old}$ vs. H_a: $\mu_{new} \neq \mu_{old}$**
- **Design: Randomized into each arm**
- **Statistical method: *t*-test**

$$n = \frac{\left(Z_{\alpha/2} + Z_{\beta}\right)^2 \sigma^2}{\Delta^2}$$

- **σ^2 = variance; Δ = effect size to detect**

Researchers need to provide this information.

- Published results
- Pilot data
- Clinically meaningful change

Sample size calculations may not be straight-forward

- **More complex designs require more complex calculations**
- **Examples:**
  - Longitudinal studies
  - Cross-over studies
  - Correlation of outcomes
- **Sometimes simulations are required**

Sample size calculations may not be straight-forward

- **More complex designs require more complex calculations**
- **Examples:**
  - Longitudinal studies
  - Cross-over studies
  - Correlation of outcomes
- **Sometimes simulations are required**

THERE IS NO MAGIC NUMBER.

Proposal Content

Proposal sections involving statistics

- **Sample size justification**
- **Statistical analysis plan**
  - Statistical methods for each aim

For clinical trials,

- **Interim analyses/Early stopping rules**

Engage a statistician to write or at least review these sections.

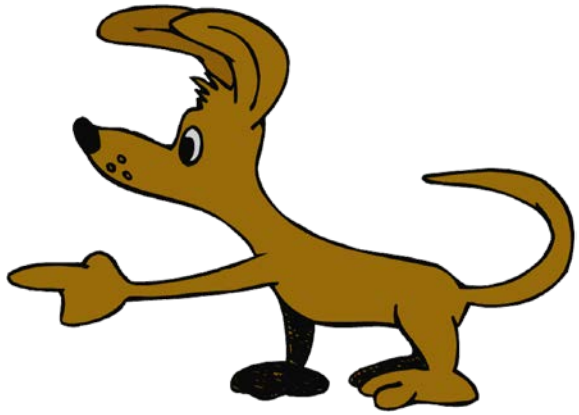
Common statistical problems in proposals

- Sample size justification absent or insufficiently justified
- Lack of statistical analysis plan for all aims, including secondary aims
- Inappropriate statistical analysis methods

Common statistical problems in proposals

- Sample size justification absent or insufficiently justified
- Lack of statistical analysis plan for all aims, including secondary aims
- Inappropriate statistical analysis methods

These issues can doom a proposal.



Data Collection and Compilation

www.ucdmc.ucdavis.edu/ctsc/redcap/

Getting Started | Latest Headlines | Subscribe to Deborah ... | Access 2010 And Share...

UC DAVIS HEALTH SYSTEM

UC Davis Health System | News | Jobs | Giving | UC Davis

search

Clinical and Translational Science Center

Health System > Clinical and Translational Science Center > REDCap

E-mail | Tweet | Facebook | Print | Share

- Clinical and Translational Science Center
- Application for Resource Use (AFRU)
- Investigators
- Biomedical Informatics
- Clinical Trials Resource Group
- Community Engagement and Research
- CTSC Clinical Research Center (CCRC)
- Pilot Translational and Clinical Studies
- Research Design, Biostatistics, and Clinical Research Ethics
- Research Education, Training, and Career Development
- Translational Technologies, Methodologies, and Resources
- Facilities, Cores, and Resources
- Evaluation
- Grant Facilitation
- Newsroom
- Research Networking
- Seminars and Events
- Contact Us

Application for Resource Use

REDCap (Research Electronic Data Capture)

View dates for upcoming REDCap training and drop-in worshops at the CTSC Event Calendar

What is REDCap?

REDCap (Research Electronic Data Capture) is a secure web application for building and managing online databases for research.

Launch REDCap

REDCap™

Click here to go to REDCap Survey

Using REDCap's stream-lined process for rapidly developing projects, you may create and design projects by constructing a 'data dictionary' template file in Microsoft Excel, which can be later uploaded into REDCap. REDCap provides audit trails for tracking data manipulation and user activity, as well as automated export procedures for seamless data downloads to Excel, PDF, and common statistical packages (SPSS, SAS, Stata, R). Also included are a built-in project calendar, a scheduling module, ad hoc reporting tools, and advanced features, such as branching logic, file uploading, and calculated fields.

How do I get access to REDCap?

You will only have to request access once. Access to individual projects is determined by the User Permissions Matrix submitted for each project.

1. Submit a Lotus Notes Online Access Request, instructions are on the How to request access page
2. Send (fax, inter-office mail, or scan and email) the signed REDCap End User Agreement to the CTSC informatics staff responsible for your project

How do I begin using REDCap?

The CTSC Biomedical Informatics Program has developed a six-step process to create a database. We encourage researchers to begin working with biostatisticians early.

To get started, please complete the CTSC Application for Resource Use.

Six-step process to create a REDCap Database (hover over each box for details)

1. Introductory Meeting
2. Create Data Dictionary
3. Data Review Meeting
4. Create Database
5. Final Review
6. Begin Entering Data!

- REDCap & REDCap Survey
  - REDCap
  - REDCap Survey
- Getting Started
  - CTSC Application for Resource Use
  - REDCap Project Level Agreement (PDF)
  - REDCap Text for IRB and Grant Submissions
  - Citing REDCap and the CTSC in Publications
  - Frequently Asked Questions
- Data Dictionary
  - How to Create the Data Dictionary (Video) (14 min)
  - Sample Data Dictionary (CSV)
  - User Permissions Matrix (Excel)
- Accessing REDCap
  - How to request access
  - REDCap End User Agreement (PDF)
  - UCDHS Confidentiality Agreement (PDF)
  - Create New UC Davis Kerberos Account
  - UC Davis Temporary Affiliate Form

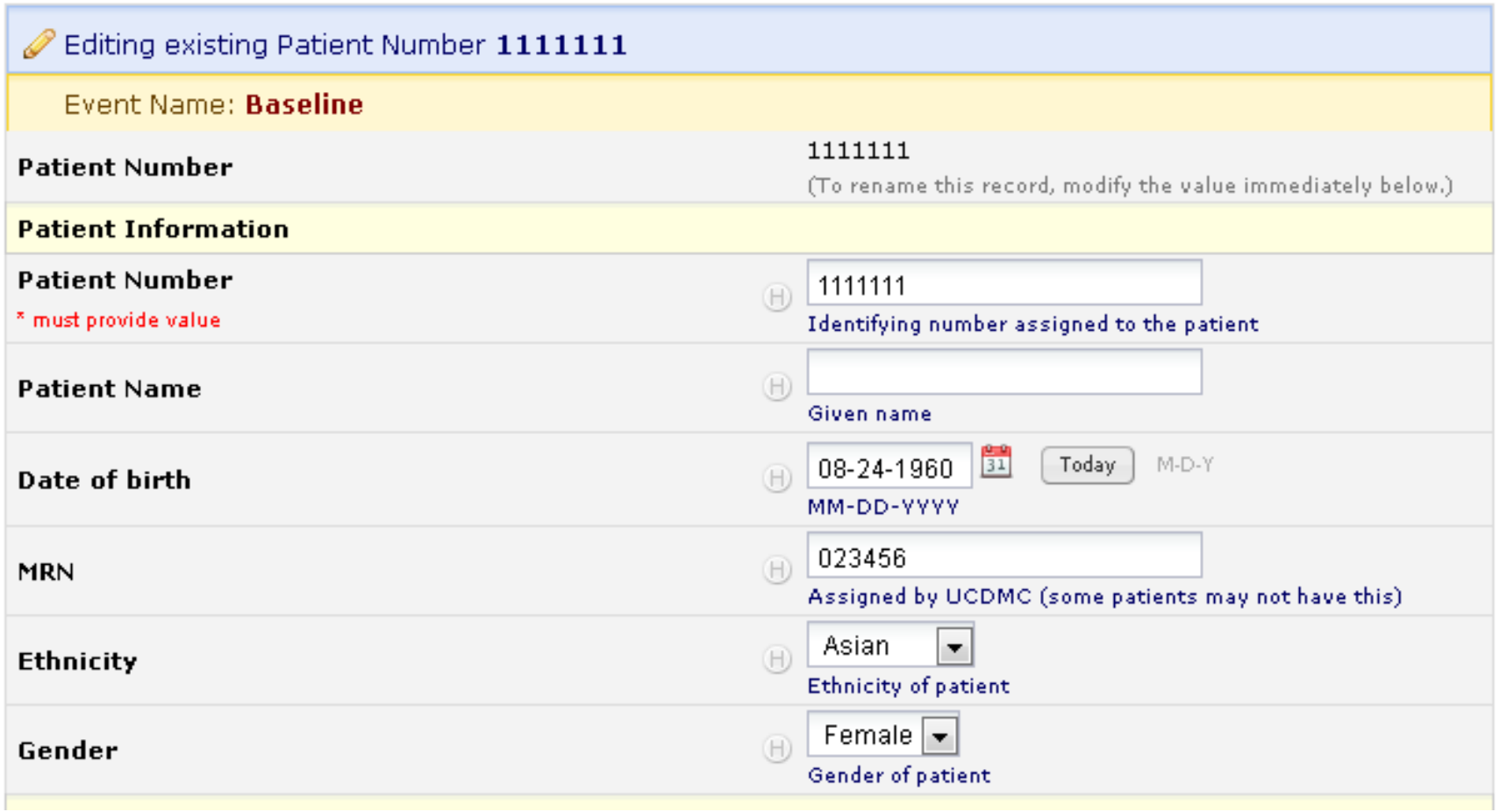
Data Collection and Compilation

- **Valid results require**
  - Collection of accurate data
  - Clear and accurate data compilation
- **Create workable and documented data sets**
- **QA/QC procedures**
  - Validation during data entry
  - Periodically audit the data
  - Conduct internal validation of final data

Bad Scene!

	A	B	C	D	E	F	G	H	I	J	K	L
1	**Comparison of Drug A and Drug B**											
2	Drug A	Age of Patient	Patient	Height	Weight	24hrhct	blood pressure	tumor	Race	Date	complications	
3			Gender	(inches)	(pound)			stage		enrolled		
4												
5	1	25	Male	61"	>350	38%	120/80	2-3	Hipanic	1/15/99	no	
6	2	65+	female	5'8"	161	32	140/90	II	White	2/05/1999	yes	
7	3	?	Male	120cm		12	>160/110	IV	Black	Jan 98	yes, pneumonia	
8	4	31	m	5'6"	obse	40	40 sys 105 dias	?	African-Am	?		
9	5	42	f	>6 ft	normal	39	missing	=>2	W	Feb 99		
10	6	45	f	5.7	160	29	80/120	NA	B	last fall	n	
11	7	unknown	?	6	145	35	normal	1	W	2/30/99	n	
12	8	55	m	72	161.45	12/39	120/95	4	African-Am	6-15-00	y	
13	9	6 months	f	66	174	38	160/110	3	Asian	14/12/00	y	
14	10	21	f	5'								
15												
16	Drug B											
17	1	55	m	61	145	normal	120/80 120/90	IV	Native Am	6/20/	3	
18	2	45	f	4"11	166	?	135/95	2b	none	7/14/99	n	
19	3	32	male	5'13"	171	38	140/80	not staged	NA	8/30/99	n	
20	4	44	na	65	?	40	120/80	2	?	09/01/00	n	
21	5	66	fem	71	0	41	140/90	4	w	Sep 14th	y, sepsis	
22	6	71	unknown	172	199	38	>160/110	3	b	unknown	y, died	
23	7	45	m	?	204	32	40 sys 105 dias	1	b	12/25/00	n	
24	8	34	m	NA	145	36	130	3	w	July 97	n	
25	9	13	m	66	161	39	166/115	2a	w	06/06/99	n	
26	10	66	m	68	176	41	1120/80	3	w	01/21/58	n	
27												
28	Average	45		65	155	38						

RedCAP is user-friendly alternative.



Leverage informatics and biostatistics expertise

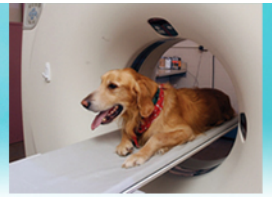
- **Medical informatics group can**
  - Create forms for data collection
  - Extract information from EMR
- **Use inter-disciplinary team to determine what information to collect and how**
  - Investigator, practitioners, biostatistician, informatics specialist
  - Ensures information is collected and compiled in a manner that facilitates analysis

Data Do's and Don'ts

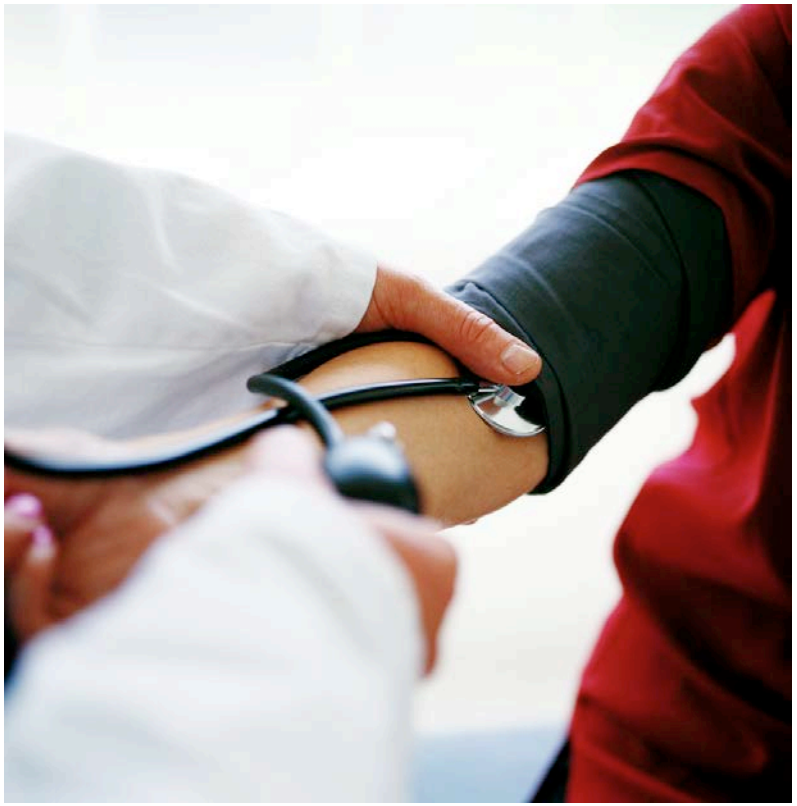
- **Use RedCAP where possible**
- **Assign unique ID to each subject and use consistently**
- **Remove all PHI prior submitting to statistician**
- **Unambiguously and consistently specify missing values**
  - Avoid using "0" or blanks for missing values
- **Avoid free text fields**







Go forth and collect data!



Statistician activities during the study

- **Sometimes limited involvement by statistician**
- **Involvement can include**
  - Conducting interim analyses
  - Serving on DSMB
- **Some study designs entail periodic reassessments and statistician will necessarily be involved during the study**
  - Two-stage, adaptive or sequential trials

Analysis and Reporting

Analyzing the data

- **Conduct statistical analyses**
  - Data validation
  - Run statistical tests
- **Interpret the results**
- **Prepare tables and figures to illustrate findings**

Working with a Statistician to Analyze your Data

Range of support provided

One-time
Advice Only

Select methods
Conduct all analyses

If statistician analyzes the data…

- **Remove PHI**
- **Provide “clean”, documented data set**
- **Allow sufficient time for analysis**
  - Rule of Thumb is 4 to 6 weeks
- **Provide references from similar studies if available**
- **Iterative and interactive process**

Report/Publication Preparation

- Craft statistical analysis section
- Contribute to results section
- Generate tables and figures

Biostatistics Resources at UC Davis

- **Clinical and Translational Science Center**
  - Biostatistics Workshop: 12-1 on Tuesdays

    sltaylor@ucdavis.edu
  - Biostatistics Core
    - Assist with study design, grant writing, data analysis and interpretation
    - Application for Resource Use (AFRU) http://www.ucdmc.ucdavis.edu/ctsc/
- **Division of Biostatistics**

Other CTSC Resources

- Biomedical Informatics
- Clinical Research Center
- Clinical Trials Resource Group
- Community Engagement
- Clinical Research Ethics
- Translational Technologies
- Pilot Studies Funding

www.ucdmc.ucdavis.edu/ctsc/