Statistical Considerations in Grant Writing

By: Aniko Szabo, PhD

The Medical College of Wisconsin is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

The Medical College of Wisconsin designates this Live activity for a maximum of 1.0 *AMA PRA Category 1 Credit(s) ™*. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

Hours of Participation for Allied Health Professionals

The Medical College of Wisconsin designates this activity for up to 1.0 hours of participation for continuing education for allied health professionals.

Financial Disclosure

- In accordance with the ACCME® standard for Commercial Support Number 6, all in control of content disclosed any relevant financial relationships. The following in control of content had no relevant financial relationships to disclose.

Name:	Role in Meeting:
Ruta Brazauskas, PhD	Planning Committee
Haley Montsma, BBA	Planning Committee
Aniko Szabo, PhD	Speaker

Learning Objectives

- Understand statistical issues relevant to grant applications
- Become aware of common mistakes and pitfalls
- Learn best practices and solutions for common problems

Evaluation Forms

Your opinion matters!

Help us plan future meetings, by completing and submitting your evaluation forms.

Thank you.

My grant-related background

- Statistical co-investigator on over a dozen grants
  - Basic science, human studies, clinical trials
- Wrote statistical sections for even more grants
  - and saw drafts, unfunded versions, etc
- Statistical reviewer for NIH study sections
  - Clinical and Integrative Cardiovascular Sciences [*CICS*]
  - Biomedical Methods and Research Design [*BMRD*]
  - NIEHS special panel on nanomaterials

Outline

- Specific aims and hypotheses
- Specifying the study population
- Selecting a study design
- Defining outcome measures
- Sample size calculations
- Data analysis plan

Specific aims and hypotheses

- **The hypotheses drive all the statistical aspects of the grant proposal**
- "Specific" aims can be relatively vague, but hypotheses have to be specific
  - Specific aim
    - Explore the effects of XYZ…
    - Characterize the properties
  - Hypothesis
    - XYZ will increase W, but will not affect U
- Always state the working hypothesis (ie what you would like to show), not the null hypothesis

Study population

- Conceptual target population

- Operationalization
  - Inclusion/exclusion criteria
    - How will their presence/absence validated
  - Source of recruitment
    - Patients presenting in the clinic
    - Female Wistar rats
  - Number of potential subjects available

Controls

- **Ideal: differ from the study group only in the study variable**
- Straightforward situation
  - Randomized prospective study
    - Many basic science experiments fall into this category
- Subject as his/her/its own control
  - Regression to the mean can cause spurious results
  - Cross-over designs can correct for this
    - Carry-over effects
  - Within-subject change can be used as an outcome with another control group

Controls in case-control studies

- Cases are subjects with a certain condition, Controls are subjects without that condition
  - What about associated characteristics?
  - Diagnostic process might be relevant
    - Eg diagnosis of cancer is highly dependent on screening practices
- Cases are subjects who received a treatment, Controls are those who did not
  - *Why* did the controls not receive the treatment?
    - Had more/less advanced disease
    - Need to be able to identify who was eligible

Study design

- Prospective vs retrospective
- Observational vs experimental
- Case-control studies
  - Matched vs unmatched
  - Matching can result in either gain or loss of power
  - Unmatched studies still recruit comparable controls, but there is no individual level matching
- Randomization and blinding
  - Rarely mentioned in basic science studies, though highly relevant

Study design

- Interim analysis
  - Check for statistical significance at preplanned time-points during the study
    - Significance levels have to be adjusted
  - Can be useful if there is substantial uncertainty about the expected effect
  - Rare in basic science studies, but could be useful
- Internal pilot
  - Quantities need for power calculation are estimated based on a small initial sample size
  - Needs to be planned in advance: a small adjustment of significance level might be needed

Defining the outcome measure

- Measureable and well defined
  - Measurement procedures
  - Timing
    - Too many repeated measurements just complicate the result
    - Two points at ends are sufficient to show change
- Separate primary versus secondary outcomes
  - Sample size calculation is guided by the primary outcome
  - Primary outcome should be
    - Most clinically/biologically relevant
    - Likely to show a difference

Sample size calculations

Type of Study	Sample Size	Detectable Difference	Power
Retrospective	Fixed known	Fixed Clinically / biologically important difference	Computed
Retrospective	Fixed known	Computed	Fixed (90, 80)%
Prospective	Computed	Fixed Clinically / biologically important difference	Fixed (90, 80) %
Observational	Computed	Fixed Desired precision of estimate	Not applicable

Sample size calculations

- Need to specify the details:
  - Planned analysis method
  - Significance level and power
  - Effect size
  - Additional assumptions (baseline rate, variance, correlation,…)
- Biological versus technical replicates
  - Spend money/effort on largest source of variability
    - Usually the biological variability
  - More than 3 technical replicates are rarely useful

Sample size calculations

- Needs “nuisance” parameters beyond expected treatment difference
  - Variance of measurement
  - Probability of event in control group
  - Correlation of repeated measures
- Source of estimates
  - Preliminary data!!!
  - Literature review
    - Help others! Report SD of key quantities in publications!
  - Educated guess

Sample size fallback options

- Sample size calculation will be done by XYZ (letter of support attached)
- Plan a small pilot study that will provide the data for the sample size calculation
  - Eg plan to update the calculations based on the results of Aim 1
- Plan an internal pilot
- Refer to Cohen's standard effect size scale

Cohen's standard effect sizes

Definition of effect size

- Continuous data:

$$ES = \frac{\mu_1 - \mu_2}{\sigma}$$

- Binary data:

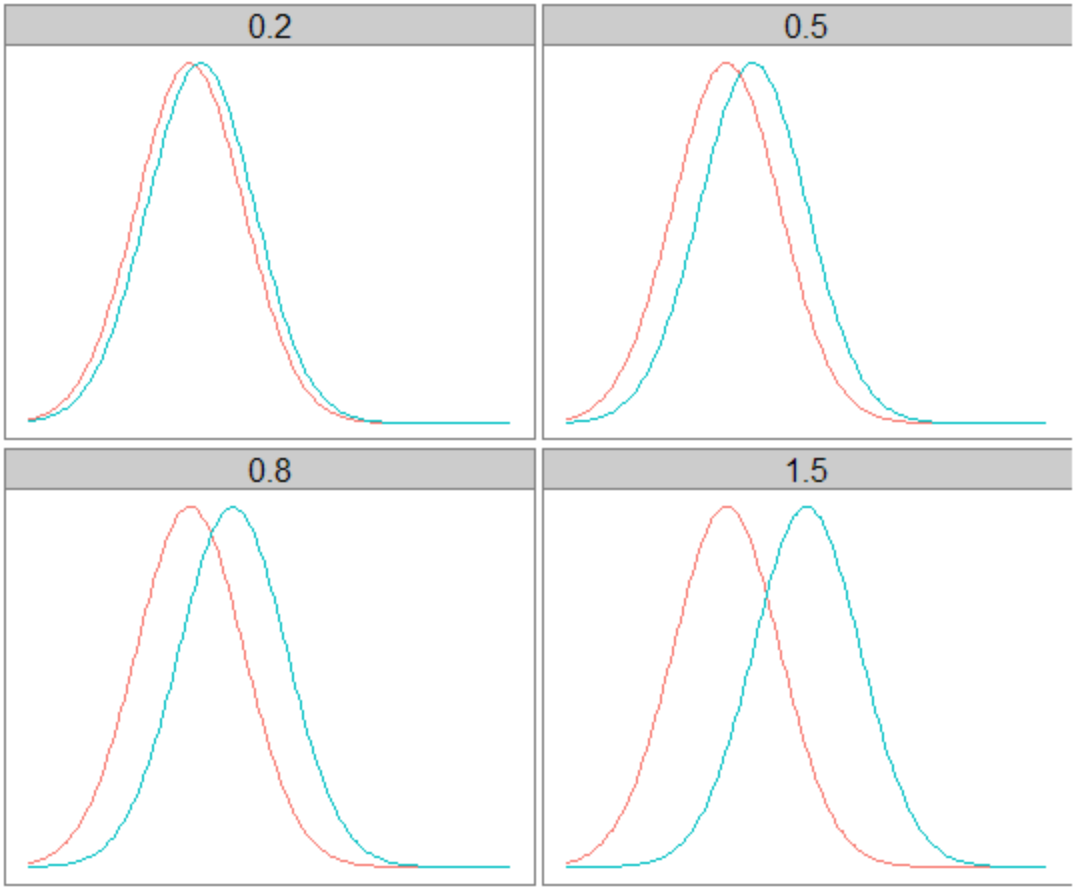
$$ES = 2\left(\sin^{-1}\sqrt{p_1} - \sin^{-1}\sqrt{p_2}\right)$$

Qualitative description	Standard effect size	Group size for a 2-sample test (α=5%,pwr=80%)	Examples	
			Normal data σ=1	Binary data
Small	0.2	394	μ = 3.0 vs 3.2	p = 10% vs 17%
Medium	0.5	64	μ = 3.0 vs 3.5	p = 10% vs 30%
Large	0.8	26	μ = 3.0 vs 3.8	p = 10% vs 44%
Very large*	1.5	8	μ = 3.0 vs 4.5	p = 10% vs 75%

Sample size for 2-sample comparison: $N_{\text{group}} \approx \left(\frac{4}{ES}\right)^2$

(α=5%, power=80%)

Cohen's standard effect sizes



If classification is the goal, small effect sizes are useless

Sample size software

- Online calculators
  - See upcoming talk
- G*Power 3
  - Free Windows/Mac program from the University of Düsseldorf
  - Needs some statistical sophistication
- Consult your friendly neighborhood statistician

Statistical analysis plan

- The goal is to convince the reviewers that you can analyze the data
  - Not all details are necessary
  - Showing awareness of statistical aspects and capability to address issues is important
  - If grant contains many experiments, consider having a separate section with overall analysis approach
- Some details can go outside of main proposal
  - “Protection of Human Subjects” and “Vertebrate Animals” sections have no page limits
  - “Facilities & Other Resources” can describe availability of statistical consulting
  - Data collection forms can be in the Appendix

Common mistakes

- Complicated data with no mention of statistics
- Mistakes/misinterpretations in the analysis of preliminary data
- Statistical methods do not match the study design
  - Statistical methods clearly copied from another grant
  - Matched design vs unmatched analysis
- Incorrect statistical plan
  - Plan to “sample until significance”
  - Superficial plan emphasizing minor details over substantative issues

Thank you!