Effective Methodological Presentation in Research Proposals

NIDUS Bootcamp 2019

Thomas G Travison

Hebrew SeniorLife Institute for Aging Research
Division of Gerontology, Beth Israel Deaconess Medical Center
Harvard Medical School
Boston, MA USA
tgt@hsl.harvard.edu

28 October 2019

Let's Talk Biostats and Experimental Design!

Objectives

- Explore general experimental design and quantitative concepts in developing research proposals
    - Focus on ‘statistical considerations’ section
- Spotlight specific tips and recommendations for use in delirium research

Caveats

- Focusing on subtle areas for improvement
- Two fields of knowledge covered in a few minutes
- The PI is the expert and ultimately responsible: you know best
- There will be misunderstanding
    - Best attitude: It's not them, it's you

Setting up the design: Some Considerations

- Phase of research
- Necessary innovation vs. excessive complexity
- Interplay between conceptual model and experimental design
- Appeal to literature / prior study
- Clinical significance

Phase of research: pilot / exploratory studies

Do

- Focus on feasibility endpoints
- Demonstrate that data capture will be sufficient to the study aims

Don't

- Devote excessive attention to classical statistical power
- Overstate the ability to perform a definitive analysis

Example: D Denny; Risk prediction for subsyndromal delirium in joint replacement

- Prospective cohort study, N = 70
- Delirium by CAM day 1, 2, 3 & 30 following surgery
- Aims
  - ... evaluate the use of a delirium risk classifier (DRC) to predict subsyndromal delirium ...
  - ... determine the relationships among preoperative DRC classification, subsyndromal delirium duration, and cognitive status ....
- Analysis: ANOVA, regression

Example: D Denny; Risk prediction for subsyndromal delirium in joint replacement

Analytic plan

The Delirium Risk Classifier will be applied to the sample and analyzed for predictive ability using decision curve analysis, a method for evaluating a prediction model that incorporates clinical consequences and does not require an external cohort. Decision curve analysis can be used to evaluate prognostic strategies that can be applied directly to a validation cohort (Vickers & Elkin, 2006)...

Power analysis

An a priori power analysis estimated 70 participants would be needed for statistical power of .80 with an alpha of .05 and the conventional effect size of 0.3 ($f2$ = .30; Cohen, Cohen, West, & Aiden, 2003) to detect significance while factoring in an estimated 7.5% loss due to attrition, and a 15% refusal rate based on previous work with similar samples (Denny & Lindseth, 2017; Denny & Such, 2018).

Sexy Innovation vs Problematic Complexity

In general, resist temptation to 'innovate' in analysis unless goals are explicitly the generation of new methodology.

Do

- Make use of methodology complex enough to address the question, population, etc.
- Demonstrate awarness of research trends
- Develop a backup plan in case of failure

Don't

- Fall in love with a design because it's so hot right now ("I simply MUST do a stepped wedge trial!").
- Neglect to motivate / support design choices

Example: M Kobayashi; CRT of multimodal comprehensive care for reduction of delirium incidence

- Multicenter controlled cluster RCT; N = 1370; Cluster number = ?
- Units of randomization: tertiary care hospitals

Design Features

A French multimodal comprehensive care methodology will be used as intervention.... nurse-specialists ... will provide lectures and bedside trainings of the methodology to all the nursing staff... In control arm, the nursing staff will provide conventional care...

Power computation

We estimated that a sample of 1370 patients would provide the study assuming a two-sided statistical significance of 0.05 and a power of 0.8 .... For cluster randomized controlled study, the key issue for sample size and power is the intra-class correlation coefficient which measures the similarity of the outcomes from individuals within the same cluster. It is currently under consideration about an intra-class correlation coefficient.

Conceptual model and experimental design

Do

- Have a conceptual model!
- Consider mechanistic / causal implications descending from said conceptual model
- Employ motivated, appropriate statistical control of potential confounders

Don't

- Dismiss mechanistic considerations b/c 'correlation is not causation' etc.
- Control for everything under the sun 'to determine the independent effect...'

Example: S Williams Roberson; EEG entropy as a risk marker for ICU delirium and LTCI

- 'Prospective, observational case-control study'; N = 50 ('9 per group')
- Aim 1: Determine factors influencing EEG entropy during critical illness
- Aim 2: Determine long-term trajectory of entropy
- Aim 3: Evaluation relationship between entropy and LTCI

Conceptual Model / Experimental Design

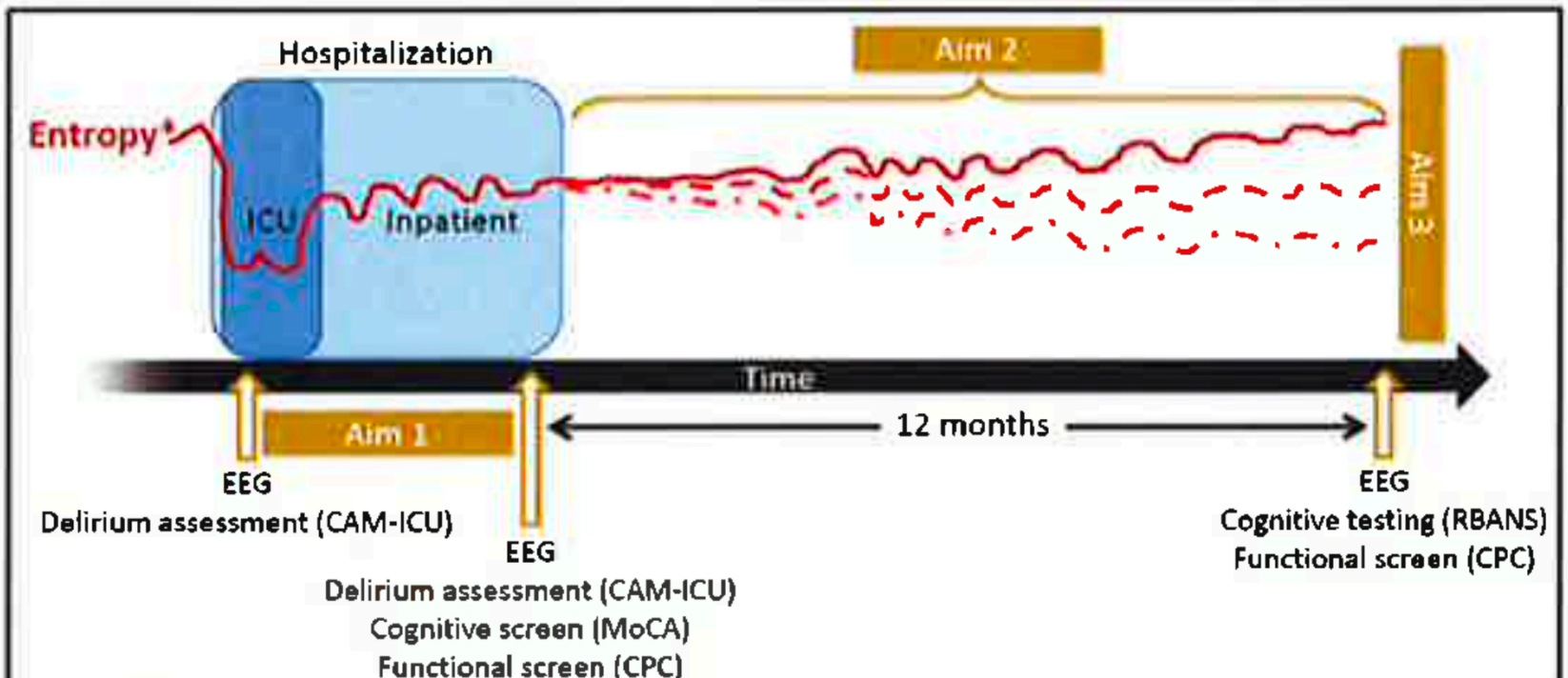

Figure 1: Characterizing EEG Entropy During and After Critical Illness.

Entropy is a measure of **complexity (variability or spontaneity) in the EEG signal. It is an indicator of **healthy, wakeful brain function** and is decreased in delirium, dementia and under general anesthesia.*

Example: S Williams Roberson; EEG entropy as a risk marker for ICU delirium and LTCI

Statistical Analyses

Descriptive statistics will be used to analyze the clinical and demographic data and determine the distribution of ApEn in our population. We will evaluate the relationship among these data (Aim 1) using univariate and multivariable linear regression models. Independent variables will include age, sex, weight, BMI, BSA, IQCODE, delirium during EEG (present/absent/unable to assess due to coma), delirium type (hyperactive/hypoactive/normoactive), level of arousal at EEG (Richmond Agitation/Sedation Score), admission diagnosis, sedatives, 12-hour sedative doses, presence of acute renal or hepatic failure, respiratory failure, sepsis, acidosis and cardiogenic shock. Dependent variables will include mean global (across the entire head) and region-specific ApEn for the duration of the EEG recording and variance in ApEn over the duration of the recording. For Aim 2 we will map the median and distribution of global and regional ApEn values at enrollment, at discharge and at 12-month follow-up across participants as well as within groups by presence/absence of delirium during hospitalization. Here, the independent variables are time point and CAM-ICU. To evaluate the relationship between EEG entropy and LTCI (Aim 3) we will use univariate and multivariable linear regression, with ApEn at discharge and at 12-month follow-up as independent variables and indexed 12-month RBANS scores as dependent variables. To evaluate possible relationships with cognition-related quality of life indicators as dependent variables, ordinal logistic regression will be used. All statistical analyses will be performed using R [22].

Literature and Prior Study

Do

- Use existing literature to inform design
- Emphasize in detail one's own prior study and feasibility assessments

Don't

- Rely upon exact replication of a prior result
- Neglect to connect the dots between prior study, power assessment, and planned analyses

Example: A Silva; Validation study for ANT-I

Power computation

A feasibility study ... concluded for the estimation of the number of patients eligible for this kind of study to be 600/year ... Half of the eligible sample would fit the no-delirium group and 7% of the patients will have delirium without any neurodegenerative condition, fitting in the delirium group .... A sample size of 50 achieves 80 percent power to detect a significant difference between the groups ... Sensitivities and specificities will be calculated for each screening test and test and Confidence Interval (CI) testing significance was 95%. Cohen effect sizes will be used when interpreting correlation coefficients and correlations higher than .05 will be classified as "large". Linear regression analyses will be carried out with delirium as the predictor variable and the ANT scores as dependent variables. Correlation models between ANT-I scores and NFL levels will be also analysed.

Argument for clinical significance

Do

- Demonstrate that study is properly powered
  - Sample sufficient for high likelihood of success under anticipated conditions
- Show that anticipated effect of interest in clinically meaningful

Don't

- Neglect more subtle effects if they are also clinically meaningful
  - Sweet spot: power for **minimum clinically important difference**
- Forget to reference secondary endpoints to the degree that implications rely upon them

Example: S LaHue; Association study of anti-aging factors, oxidative stress, and post-op delirium

Sample Size

At UCSF, there were 2859 patients (fiscal year 2016-2017) and 3040 (fiscal year 2017-2018) admitted to the surgery service for at least 48 hours; a significant number of these patients were at least 65 years old. The compounds in this grant have not been studied in delirium ... estimated effect sizes and power calculations are based on their study in other neurological diseases. Differences in NAD+ levels were studied in multiple sclerosis where mean NAD+ was 17.9 ±3.2 µg/ml (controls) and 9.9 ±2.9 µg/ml (relapsing remitting multiple sclerosis).... effect size 8 µg/ml and standard deviation 3 µg/ml yields a total number of 14 participants, doubled to a goal enrollment of 28.

Theme and summary: The project is a coherent whole.

Consider closely and emphasize the *connective tissue* that binds the literature to the concentual model to the experimental design to the analytic plan.

- Phase of research
- Conceptual Model
- Prior study
- Feasibility
- Aims
- Population
- Design
- Clinical Meaningfulness

Let's discuss biostatistics and experimental design!