

Statistical consulting/collaboration

starting the conversation among our students

Erik Barry Erhardt

Assistant Professor
Department of Mathematics and Statistics
University of New Mexico

November 18, 2011

A few projects in my consulting career

Feb 2006, **UNM Biology Department.**

Department of Biology Course Fees Feasibility Study. "... this has allowed us to net an additional \$55,000 this semester in funding for student classes and labs." — Heather Paulsen, 17 Jan 2007

Feb – Aug 2007, **UniRac, Inc.**

For process development, improvement, and control of PVs Universal Flat Roof Solution. RapidRac G10 PV Mounting System, introduced Sep 2007.

Jul 2010, **Director of Quality Management, Lovelace.**

Analysis of provider risk of complications.



A few projects in my consulting career

Apr 2010 – Nov 2010, VeraLight, Inc.

Demographic subgroup analysis of pre-clinical trial data, analysis plan for FDA clinical trial.

4/26/2011 received Health Canada license approval.

7/28/2011 received CE mark approval

“As I went through the analysis plan and the justification documents, I was reminded how much you’ve helped us over the past several months: your contributions are evident in multiple places. Your work has been very valuable to us, and I’m grateful that you’ve continued to find time to work with us.”

— Edward Hull, PhD. 13 Jan 2011

Nov – Dec 2010 Aaftab Jain.

Modeling for improved inference of wind farm mortality.



Four basic questions guide investigations

- 1 What is the question?
- 2 Can it be measured?
- 3 Where, when, and how will you get the data?
- 4 What do you think the data are telling you?

Cross-sectional versus longitudinal consulting

Cross-sectional (advisory)

- brief, somewhat superficial conversation between consultant and investigator
- statistical advice for data already collected or for analyses already carried out

Longitudinal (interactive)

- long-term relationship between the statistician and investigator
- frequently collaborative and collegial and may lead to creative application of current statistical methods or the development of new methods

Consulting session

Structure: have a beginning, a middle, and an end.

- Begin with problem context and definition,
 - investigator introduces problem context and content
- move to resolution and solution,
 - consultant reviews the assumptions and (requested) methods
 - keep in mind the issue introduced may not be the one that ultimately should be addressed
- and conclude with a summary and actions.
 - discuss the next steps
 - allocate responsibilities
 - summarize, make next appointment

Consulting session

The hour meeting: prep and scope

- structuring a meeting to have a beginning/middle/end can be challenging in the common fixed-length (hour) consulting session
- **ideally**, for structure and preparation, ask the investigator to provide a short write-up of the topic, or to send papers beforehand
- try to judge what can and cannot be accomplished in the meeting time

Consulting session

The hour meeting: problem and goal

- exposition of the problem and context often takes longer than the investigator expects
- both people need to correctly understand the problem and the goal to answer the right question
- try to figure out what the question *should be*, not just what the question is.

Consulting session

The hour meeting: wrap up

- Complete the session by writing a brief, one-page summary of the session, including:
 - description of the subject area and problem
 - statistical issues
 - decisions and recommendations
 - action items

Asking questions

- good questions imply
 - you are understanding the content area,
 - you sense where the discussion is going,
 - and want to agree on the conversation's direction
- questions should be focused, answerable, and in the spirit of humility and desire for knowledge
- the investigator should verify the consultant understands the problem completely by asking the consultant to summarize the problem

Make distinctions

clarify these

Item	Choice 1	Choice 2
intent of study	exploratory	confirmatory
endpoint	measurable	not well defined
source of data	random	convenience
sampling unit	simple	multiple
randomization	single	multiple
model	fixed	random or mixed
type of study (1)	intrinsic	comparative
type of study (2)	cross-sectional	longitudinal
epidemiologic study	cohort	case-control
categories	nominal	ordered
numerator/denominator	rate	proportion, ratio

Make distinctions

Focus!

- if the investigator presents a large number of undifferentiated endpoints, clarify which are the key ones
- for example, which endpoints will be emphasized in an abstract?

Know yourself

Don't worry

- the consulting experience should be mutually beneficial, both learning and exchanging information
- an honest person knows what he knows and doesn't know
- investigators have a way of probing the limits of the consultant's knowledge
- in fact, they frequently come because they were not able to apply the standard examples found in many textbooks

Level of advice

- communication between people is hard, and between disciplines even harder
- a large part of an initial consulting session will be spent in learning terms and vocabulary
- initially this will largely be educating the consultant, but at some point it will be necessary to introduce new concepts to the investigator.
- for both, it is most helpful to avoid jargon, define terms carefully, and present information at an appropriate level
- use words, images, diagrams, equations, analogies — anything that helps bring understanding
- keep things as simple as possible, but not more simple

Agree on assignment of responsibilities

- a consultation generates expectations and responsibilities
- be clear who does what, and when
- data analysis: who will enter the data, verify data entry, decide which statistical package, what analysis, who will perform analysis, ... ?
- everyone is happier if rules and limits are clear
- be clear about authorship or coauthorship on the paper, or acknowledgment, and try to discuss up front
- if the statistician has made a substantial contribution, coauthorship may be warranted
- Warning: make sure you don't get credit if your advice about methods was not taken and the analysis was wrong



Any software will do

- use any available and easy-to-use statistical package or book
- some may be superior for one reason or another, but it's better that statistician and investigator can use and do the analysis than struggling with software

Ethics

- ethics precedes, guides, and follows consultation
- some ethical principles are straightforward: don't lie, don't fudge or cook data, don't deliberately slant the report of a study
- others are more challenging, such as research conflicting with the statistician's value system: such as research involving animals and pain
- if you don't want to be involved, communicate early and in a facilitative manner, and suggest a colleague who would be willing to help
- withdraw as gracefully and promptly as possible if you must, such as if the investigator does not want to follow your advice, or if study is not salvageable



Be proactive

- statisticians in many ways are the methodologists of the scientific method
- this creates obligations, responsibilities, and opportunities
- statistical consulting is best done at the start of research or production projects
- don't only respond to their requests, consider a bigger picture and make recommendations

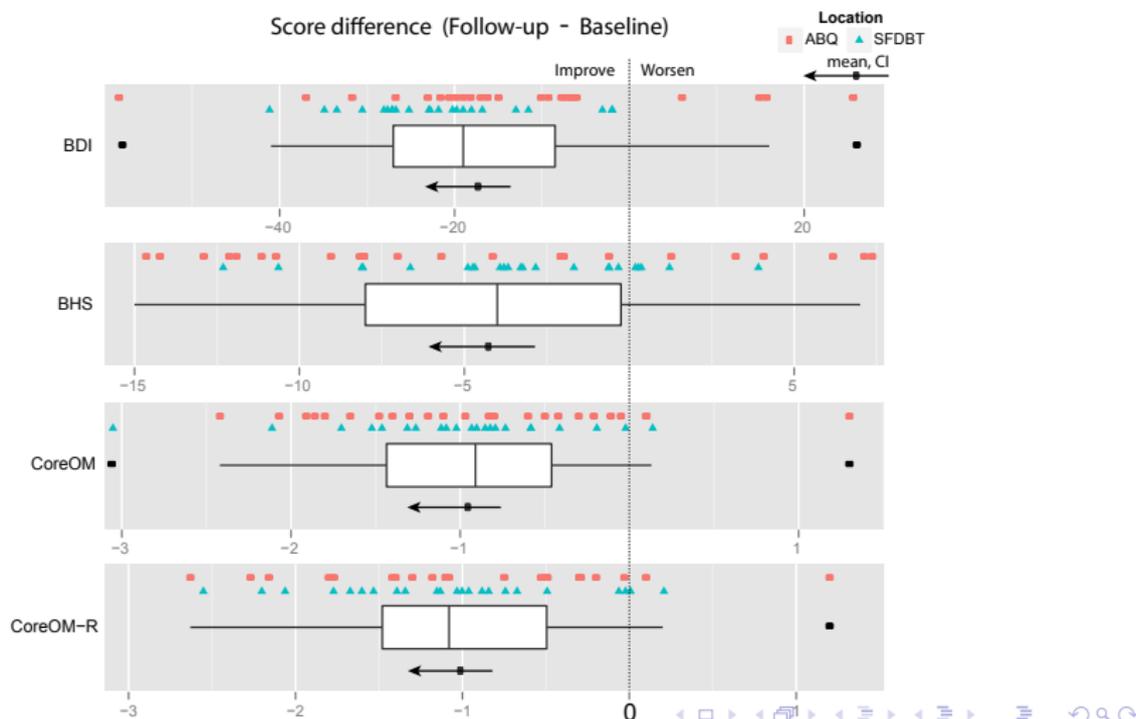
Listen to, and heed the advice of experts in the field

- Sir David Cox's consulting advice (selections from *Liaison*, of the Statistical Society of Canada)
- ① Collaborate rather than consult. Take interest in the subject matter. If the investigator seems misguided, retreat as soon as politeness and practicality permit. Discreetly determine how much understanding of statistical issues the investigator has.
- ② Aim to see some raw data, to understand the measurement processes involved, and to have some appreciation of the general quality of the data.

Cox's advice, continued

- 3 Enquire into aspects of the study design that might have bearing on the appropriate analysis.
- 4 Begin with very simple methods.
- 5 If possible, end with simple methods.
- 6 Since nice ideas for analysis often do not work the first time, be prepared to do some modification.
- 7 Do not be frightened to make strong assumptions. When a preliminary answer has been obtained, then consider which of the assumptions may be crucial.
- 8 Take considerable care over presentation and conclusion.

Advice, presentation example



Cox's advice, continued

- 9 Review acknowledged work before submission.
- 10 Occasionally, very rarely one hopes, be prepared to say that the data are incapable of throwing useful light on the issues involved.
- 11 If more than 10% of what you do ends up by being directly useful, you are doing well.
- 12 If the investigator begins by saying he has a trivial little problem which he is sure you will be able to sort out immediately, don't altogether believe him.

Practical details

- Negotiate your rate.
 - hourly, for each phase of project, or for entire project
 - I charge companies more than individuals
 - I charge more for projects that aren't fun
 - I may reduce charge for work leading to publications
- Add NM gross receipts tax, 7% in Albuquerque, 5.125% to 8.6875% elsewhere
- Tracking and billing your time
 - date/description/time invoice spreadsheet (example)
 - email as pdf each month
 - to avoid contested charges, establish trust through communication, detailed feedback, and specific notes
- Follow up and express gratitude

Primary reference

van Belle G. Statistical rules of thumb. New York, NY: John Wiley and Sons, Inc, 2002. Chapter 10.