NIH StrokeNet Professional Development Seminar – Dec 7, 2017

When and How to Consult with a Statistician… (and a bit more)

Yuko Y. Palesch, PhD
Department of Public Health Sciences
Medical University of South Carolina

Conflict of Interest / Disclaimer

- I have been an applied biostatistician for almost 30 years, with majority of time spent in clinical trials for stroke with funding from the NIH.
- This presentation contains my personal biases and opinions.
- I am a Co-PI of the StrokeNet National Data Management Center (NDMC).

StrokeNet NDMC in Charleston, SC

Medical University of South Carolina (MUSC)

↓

College of Medicine (COM)

↓

Department of Public Health Sciences (DPHS)

↓

Data Coordination Unit (DCU)*

* Whence, the database software name, WebDCU™.

DCU Biostatistics Team

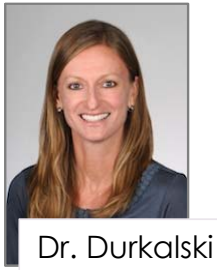

Dr. Yeatts


Dr. Palesch

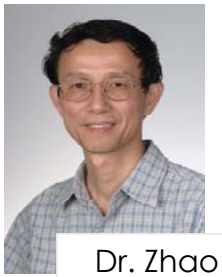

Dr. Martin


Dr. Meinzer

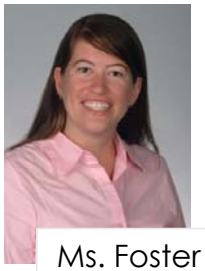

Dr. Elm

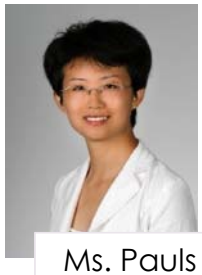

Dr. Durkalski


Dr. Ramakrishnan


Dr. Zhao


Ms. Arora


Ms. Foster


Ms. Pauls


Ms. Tillman


Ms. Gottfried


Ms. Teklehaimanot

DCU Biostatistics Team

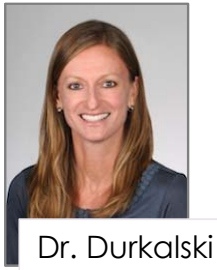

Dr. Yeatts


Dr. Palesch

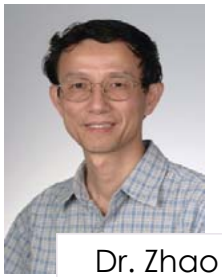

Dr. Martin


Dr. Meinzer


Dr. Elm

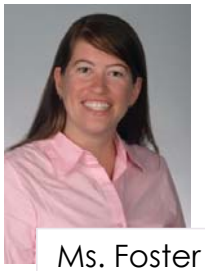

Dr. Durkalski

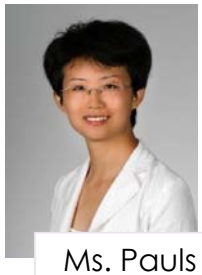

Dr. Ramakrishnan


Dr. Zhao


Dr. ??


Ms. Arora

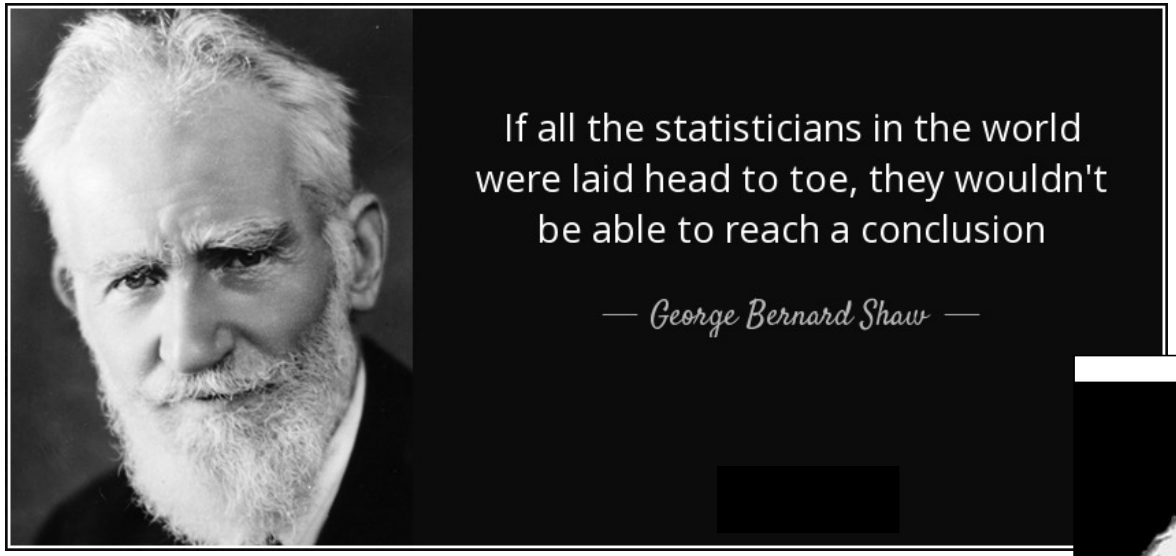

Ms. Foster


Ms. Pauls

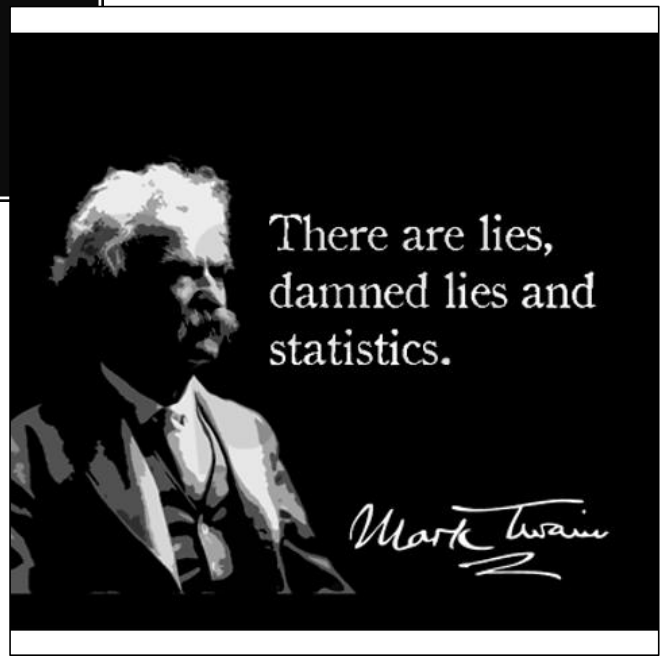

Ms. Tillman


Ms. Gottfried


Ms. Teklehaimanot


If all the statisticians in the world were laid head to toe, they wouldn't be able to reach a conclusion
— George Bernard Shaw —


"When I grow up, I wanna be a statistician!" said no one ever.
someecards
user card


There are lies, damned lies and statistics.
Mark Twain

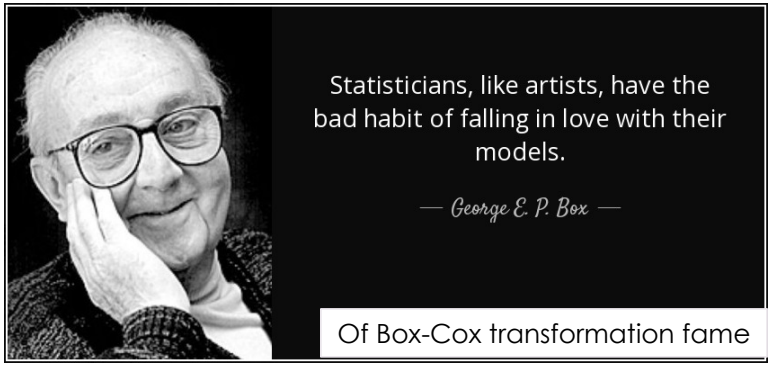
Myths about statisticians in biomedical research

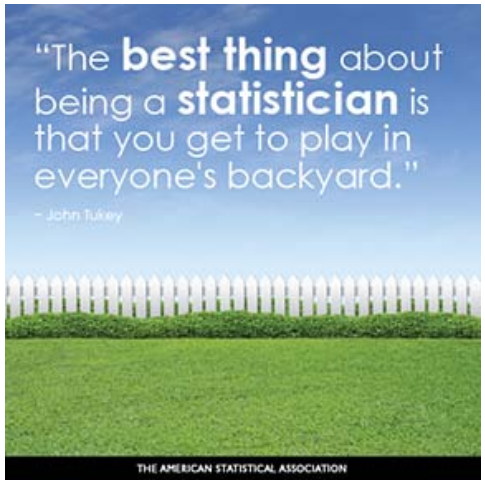
- Anyone with some statistics courses will do.
- Only need a statistician at the beginning (to give you the necessary sample size) and at the end (to do the analyses) of a study.
- A statistician is a service provider.

- You don't need to include them as authors, especially if you pay them.

Truths about (most) biostatisticians

- Most PhD statisticians train, on average, 4~6 years post-baccalaureate.
- Some get post-doc training.
- Love seeing our skills and knowledge put to practical use.





- Don't necessarily know everything and anything about statistics (e.g., not all of us are Bayesians or econometricians) – but very adaptable/flexible in application of the statistical skills and knowledge.
- Do more than just give you the required N and calculate *p*-values for the studies.
- Are your peers / colleagues.

When and what to look for in a statistician for your clinical trial?

- "Time is Brain" mantra applies to timing of when to solicit statistical help – the sooner, the better.
- Preferably, find a statistician who is familiar with (or at least with interest to learn about) your clinical area –
- Definitely, find a statistician who has clinical trials experiences – not just design and/or analysis but in the actual implementation.*
- Neurologists (some who are closet statisticians) and Statisticians (some who are doctor-wanna-be's), who have **mutual respect** for each other's expertise, make an awesome study team.

*Analogous to finding an architect who has actually "built" a structure.

Where to find a clinical trial statistician?

- Ask your mentors and colleagues at your institution.
- Inquire with Biostatistics department or group (e.g., CTSA) at your institution.
- Browse through published papers of clinical trials designs and/or results.
- Contact someone who has taught you or colleagues a clinical trials course, like instructors at the NINDS-sponsored Clinical Trials Methodology Course.
- Ask NDMC or other DCCs.
- Not always easy to find one…

How to work with a clinical trial statistician?

- In person meeting is the best.
- Sending a written synopsis of the project, and other relevant references prior to the first meeting would be helpful.
- Agree early on about expectations – role in the grant (e.g., co-PI or co-I), order of authorship in the paper, funding/financial issues, timeline, etc.
- Keep the ball moving… You ask for input, you get it, and then, not get back in touch for months is problematic (yes, it's a two-way street).
- Communicate regularly!
  - Ask questions until you understand the design/methods.
  - Keep him/her in the loop on all aspects of the project.
  - Remember, he/she is on your team as a collaborator.

Some random statistical issues in a nutshell

- Adaptive designs
- Sample size calculations
- P-values
- Interpretation pitfalls
- Big data - quality vs quantity
- Grant writing and budgeting

Adaptive Designs

"This really is an innovative approach, but I'm afraid we can't consider it. It's never been done before."

- Still innovative?
- Often useful for phase II trials when there're still many uncertainties about the intervention.
- Adaptive Designs ≠ smaller sample size, nor is it necessarily efficient.
- Frequent looks at the data may be vulnerable to unblinding, biases, etc.
- Implementation can be a real  !!

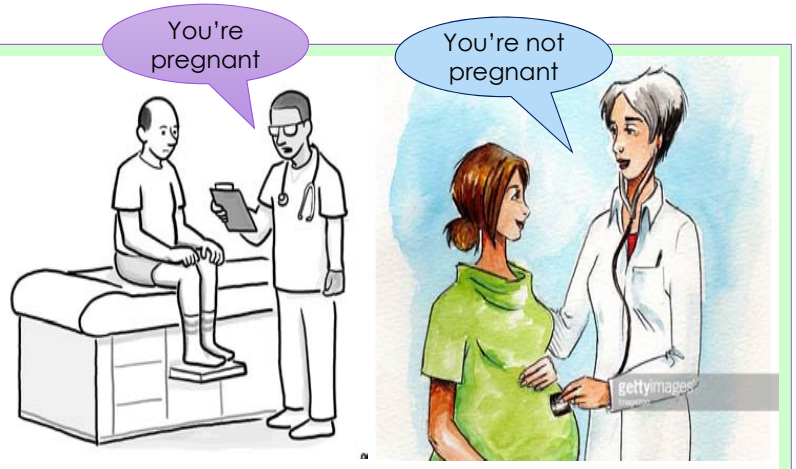
- Consider using it gingerly for phase III trials – don't make it so complicated such that it makes the study results difficult to interpret.
- Try to remember the KISS principle.

Sample Size Calculation - Need to know…

- Primary scientific hypothesis.
- Primary outcome measure and its statistical characteristics under the H_0 (e.g., distribution, mean, sd, etc).
- MCID - minimum clinical effect size you want to see that could lead to changing clinical practice. (This is implied in your H_A.)
- **Type I (α)** and **Type II (β)** error probabilities – know their interpretation under your hypothesis setting (e.g., superiority, non-inferiority, futility), and the consequences of committing these errors.
- Does α have to be 0.05? (NOTE: β can generally range from 0.1 to 0.2.)

Definition: $\alpha = \Pr(\text{reject } H_0 \mid H_0 \text{ true})$
$\beta = \Pr(\text{fail to reject } H_0 \mid H_A \text{ true})$

Superiority: $H_0: \mu_T = \mu_C$ vs $H_A: \mu_T > \mu_C$



False positive (**Type I** error)

False negative (**Type II** error)

P-value (a quick review)

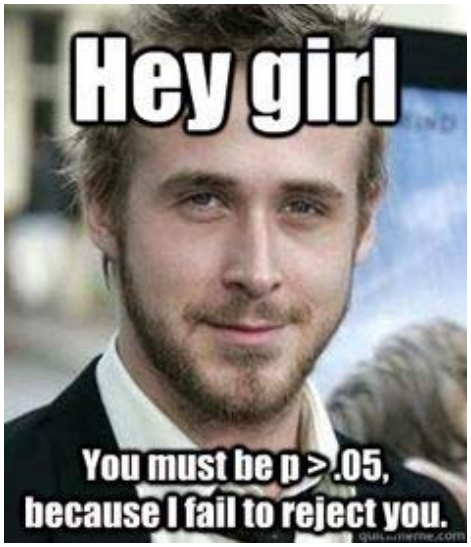
- Definition of ***p***-value: The probability of observing treatment effect (e.g., group difference in mean response) as extreme or more extreme (away from the H_0) if the H_0 is true. Hence, the smaller the ***p***- value, the more extreme or rare the observed data are, given the H_0 to be true. – i.e., **_p_-values are premised on the condition specified in the null hypothesis, as is the α value.**

- The ***p***-value obtained from the data is judged against the α. (NOTE: Remember that ***p***-values and α are not the same thing.)

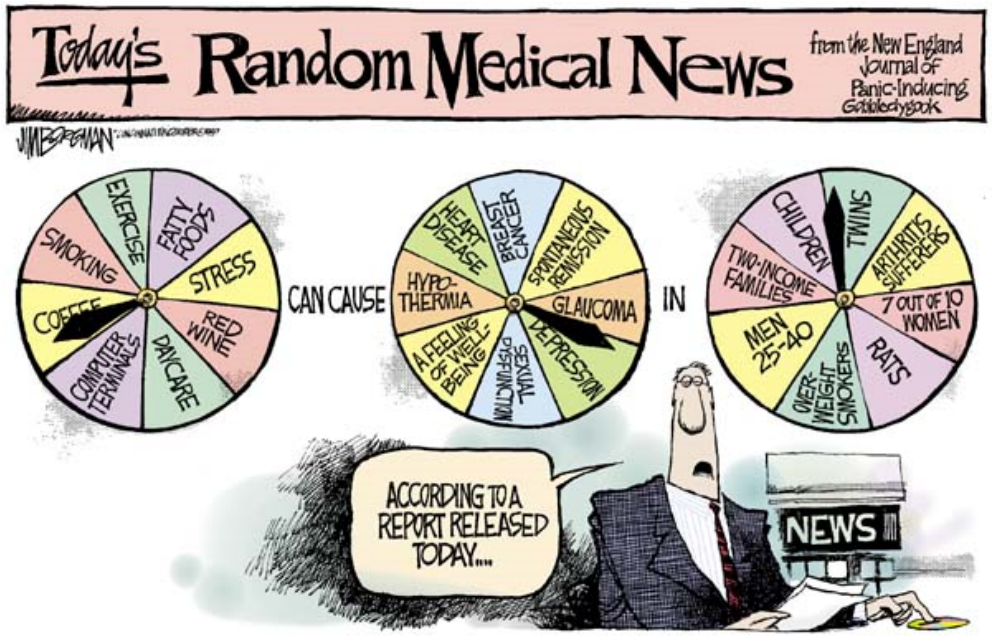
- Thus, if the ***p***-value < pre-specified α, then the data suggest that the study result is so rare under the H_0 that lead us to question the veracity of condition specified in the null hypothesis; hence, we reject the H_0.

P-value (continued)

- For a study with α =0.05 and p>0.05 (i.e., not significant), note that "failure to reject H_0" does not prove that the treatment groups are equal with respect to the outcome, i.e., you don't "accept H_0".
- Don't say, "There was no difference in the treatment groups…", unless your hypotheses were set up to prove this, like an equivalence design.
- Moral of the story: Put the research hypothesis that you want to prove in the alternative.



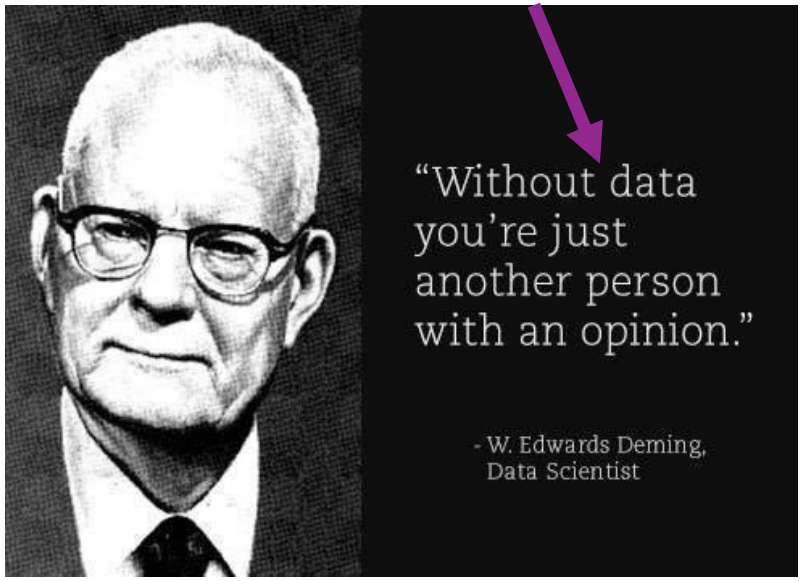
Interpretation Pitfalls



- Gives science (and statistics) a bad rap.
- Effect on subsequent randomized trials….?

Big Data – Quality vs Quantity

good quality



- Be careful about using EMR, survey and registry data without understanding how the data were collected.
- Be careful about "meta-analysis" using patient level data – make sure you are concatenating apples and apples – example of "baseline" NIHSS in IMS 3 vs MR CLEAN in the context of IV-tPA treatment timing.
- Also, you can show statistical significance if you have large enough N – be cautious of over-powered analysis that has no clinical value.

Grant Writing and Budgeting (related to stats)

- **DON'T procrastinate!**
- If you are relatively new to grant writing, strongly recommend having an experienced mentor. StrokeNet (NCC, NDMC, WGs) also can and will help.
- Get the draft of the near-final Specific Aims and Research Strategy sections ASAP to the statistician – tough for statistician to write his/her section in a vacuum.
- FYI - Items included in the NDMC budget for StrokeNet trials include: Personnel Effort (Statisticians, DMs, PMs, Programmers, Neuroimaging Managers); Travel; Supplies; and **On-Site Monitoring costs.**

- NDMC moving more towards remote monitoring to save on travel costs, and to central monitoring (by DMs and statisticians) to reduce on-site monitoring time.

Finally, the biggest myth about statisticians …

Boring

Dull

Nerdy

Humorless

And that's "normal" for him…

Thank you for your attention!