# Effect size, significance, modeling

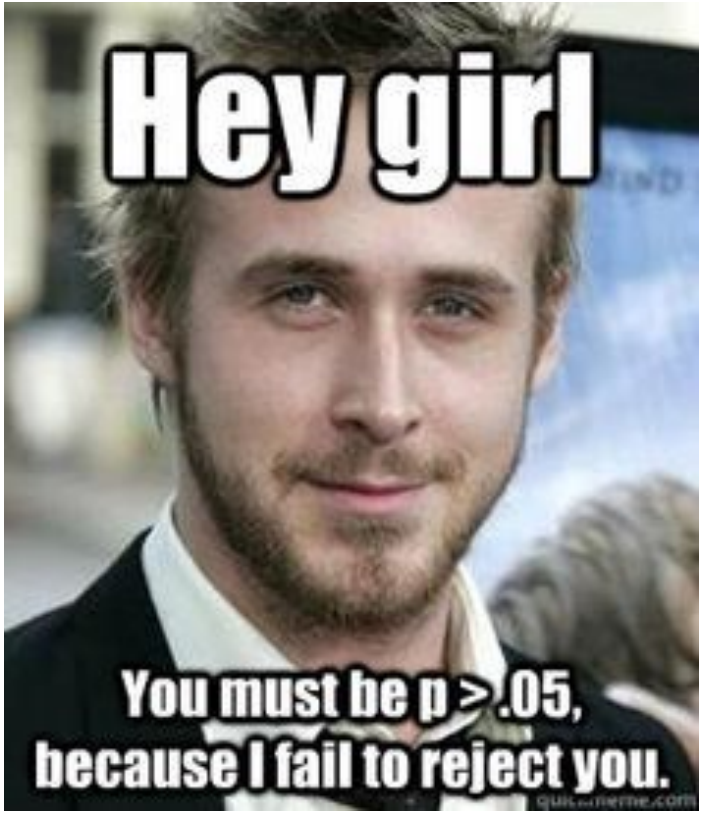
Brian Caffo, Jeff Leek, Roger Peng

@bcaffo

www.bcaffo.com

“Statistical significance is not practical significance”

- said by every statistics instructor ever





https://s-media-cache-ak0.pinimg.com/236x/1e/30/41/1e3041cbac01fb863fbd2edfc6ae9236.jpg
http://1.bp.blogspot.com/-r6q5-esz2Ws/U8fVhxXwwYI/AAAAAAAAE2E/H_pOlpYDrZE/s1600/1901960_10150397227781793_3141100807605620179_n.jpg

A very large epi study of nutrition finds a slightly significant result ($p = 0.049$) associating a hot dog consumption to colorectal cancer incidence

Does the large sample size bolster the evidence or hinder it?

A well done A/B test finds one ad campaign is not significantly better than another (p=0.056) for online purchases

Does the good study design bolster confidence in the lack of significance or not?

# Context is important

The size of the effect matters.

The context of the problem matters.

- What strength of evidence is required for the setting?
- What level of error is tolerable?
- What biases are likely present?
- Multiple comparisons

Hypothesis testing is a tool, but should rarely be used in isolation

There is no substitute for a critical review of results in context

Confidence intervals, study of effects should be used in addition to hypothesis tests

Three tools we'll discuss:

- Evaluate multiplicity concerns
- Compare effect sizes to other known effects
- Negative control (analyses where there's known to be no association)