Name: KEY

HIV Vaccine study

Read the abstract for the Study "Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 Infection" (at end of assignment, and handed out in class). This study investigated whether a vaccine could prevent the acquisition of HIV. There are 4 treatment groups

1) Men who received vaccine
2) Men who received placebo
3) Women who received vaccine,
4) Women who received placebo.

All individuals were HIV negative when the study began. They carried out 2 analyses: one for men, one for women. The main response variable is "infection" rate, which can be thought of for each individual as "infected or not infected". For each person they also measured viral load, which is the number of viruses per red blood cell.

1) What type of statistical model would you use to analyze these data

Paired t-test 2-sample t-test ANOVA regression chi-squared test

Commented [l1]: Each analysis would be done with a 2-sample t-test. The individuals are not paired up in anyway. (see question below for an example of pairing.) Since there are only 2 groups its not an ANOVA

2) What kind of data is the response variable "infected or not infected"?

Categorical / factor / binary (if this was on a test I'd give partial credit for "qualitative")

Commented [l2]: "qualitative" usually refers to subject; I tend not to use this term, though other authors/professors might. I usually contrast "quantitative" with "categorical", though some contrasts "quant" with "qualitative"

2) What kind of data is the response variable "viral load"?

Numeric / quantitative / "quantitative numeric" / continuous

3) What is the predictor?

Group that got the treatment, each vaccinated males vs. placebo males.

3) What kind of data is the predictor?

Categorical / factor / binary

4) What is the **statistical term** for the difference in viral load between **men who received the placebo** and men who received the vaccine.

Effect size; an effect size is the difference between 2 means, such as 2 treatments

5) The 95% confidence interval for the difference in viral load between **men who received the placebo** and men who received the vaccine contains zero. What does this indicate?

There is not a "significant" difference between the 2 groups; the p value will be > 0.05

6) Did the authors do a good job reporting their results? Could they have done better?

Overall yes; they reported mean, effect sizes, confidence intervals, sample sizes.

7) What term could apply to the "**Exploratory … analyses"** they conducted?

p-hacking

8) Why do you think they "**adjusted for multiple … comparisons"** when they carried out their "Exploratory analyses"?

They were carrying out many tests which were not "planned" and so might be led astray by low p-values due to random chance. That is, a risk of Type-I error (rejecting the null, evening though its actually true.)

9) What are two methods they could use to correct for multiple comparisons? (actually only one would work in this case, but those details aren't important)

Bonferroni, Tukey

10) This study had a large sample size (>5000 people). If the sample size was smaller (say 50), would the confidence intervals around the effect size get bigger or smaller (Assuming they got the same effect size as with the larger sample size).

Confidence interval would be larger

11) Imagine a different study where they gave anti-retroviral therapy to people who acquired HIV during this study. They compared each person's viral load at the start of the study to their viral load a year later. What kind of statistical test would they use?

Paired t-test 2-sample t-test ANOVA regression chi-squared test

Placebo-Controlled Phase 3 Trial of a Recombinant Glycoprotein 120 Vaccine to Prevent HIV-1 Infection
https://academic.oup.com/jid/article/191/5/654/1234535/Placebo-Controlled-Phase-3-Trial-of-a-Recombinant

The Journal of Infectious Diseases, Volume 191, Issue 5, 1 March 2005, Pages 654–665,

Abstract
Background
A **vaccine** is needed to prevent human immunodeficiency virus type 1 (**HIV**-1) infection

Methods
A double-blind, randomized trial of a recombinant HIV-1 envelope glycoprotein subunit (rgp120) **vaccine** was conducted among **men who have sex with men**[highest risk group] and **among women at high risk** for heterosexual transmission of HIV-1. Volunteers received 7 injections of either **vaccine or placebo** (ratio, 2:1) over 30 months. The primary end point was HIV-1 seroconversion over 36 months

Results
A total of 5403 volunteers (5095 men and 308 women) were evaluated. **The vaccine did not prevent HIV-1 acquisition**: infection rates were 6.7% in 3598 vaccinees and 7.0% in 1805 placebo recipients; vaccine efficacy (VE) was estimated as 6% (**95% confidence interval**, −17% to 24%). There were no significant differences in viral loads, rates of antiretroviral-therapy initiation, or the genetic characteristics of the infecting HIV-1 strains between treatment groups. **Exploratory … analyses** [extra analyses done using demographic data from the study, such as race, age, income] showed nonsignificant trends toward efficacy in preventing infection in the highest risk (VE, 43%; n=247) and nonwhite (VE, 47%; n=914) volunteers (P=.10, **adjusted for multiple … comparisons**)

Conclusions
There was no overall protective effect. The efficacy trends in subgroups may provide clues for the development of effective immunization approaches