

# Final Project Instructions

In December 2020, Pfizer and BioNTech successfully obtained a US FDA Emergency Use Authorization (EUA) to begin distributing their two-dose vaccine (BNT162b2) for SARS-CoV-2 (“COVID-19”). In this paper, you will analyze their data, as reported by Polack et al. [here](#).

In brief, this placebo-controlled, observer-blinded, efficacy study randomly assigned persons 16 years of age or older in a 1:1 ratio to receive two doses, 21 days apart, of either placebo or the candidate, BNT162b2 vaccine, which was based the mRNA approach, as Moderna used, but not Janssen (Johnson and Johnson) or AstraZeneca. The primary efficacy endpoint was the time of laboratory-confirmed COVID-19. A total of 170 infections were targeted.

The relevant data from Table 2 of the paper referenced above is summarized below:

Table 1: Vaccine Efficacy against Covid 19 at least 7 days after second dose in patients without evidence of infection

Group	Cases	No. of subjects
BNT162b2	8	17,411
Placebo	162	17,511
Total	170	34,922

In the paper, the authors mention that they used a beta binomial model with the transformation

$$\pi = \frac{\pi_v}{\pi_v + \pi_p}$$

where  $\pi_v$  and  $\pi_p$  are the probabilities of an infection in the vaccine and placebo groups respectively.

The parameter  $\pi$  was modeled with a *Beta*(0.700102, 1) prior. (Note the six decimal places!!)

The vaccine efficacy was estimated to be 95% with the 95% Bayesian credible interval reported as [90.3%, 97.6%] and the posterior probability that vaccine efficacy exceeded 30% was larger than 0.9999.

Your task is to analyze these same data, but using both Bayesian and Frequentist methods.

1. Each group must turn in ONE copy of their final paper (typeset) by 11:59 PM on Monday June 5, 2023 in Gradescope.
2. The paper (excluding the appendix) should be no more than 13 pages, 11 pt font and double spaced. The complete report (including the appendix) should be no longer than 22 pages. You may use the YAML in the instruction document as your header.
3. Use an appropriate scientific writing style. Don't be overly formal, but not overly casual either. Find that happy medium.
4. Your paper should contain the following sections.
  - a. **Title** (come up with something catchy) and authors (group members listed in alphabetic order or by contribution)

*Assessment: Does the title give an accurate preview of what the paper is about? Is it informative, specific, precise?*

- b. **Abstract:** The abstract provides a brief summary of the entire paper (background, methods, results and conclusions). The suggested length is no more than 150 words. This allows you approximately 1 sentence (and likely no more than two sentences) summarizing each of the following sections. Typically, abstracts are the last thing you write.

*Assessment: Are the main points of the paper clearly and succinctly described?*

- c. **Keywords:** four-five summary words that are relevant to the topic and your methods.
  - d. **Introduction:** In this section you are providing the background of the research area and arguing why it is interesting and significant. This section relies heavily on literature review (prior research done in this area and facts that argue why the research is important). This whole section should provide the necessary background leading up to a presentation (in the last few sentences of this section) of the research hypotheses that you will be testing in your study. Well-accepted facts and/or referenced statements should serve as the majority of content of this section. Typically, the background and significance section starts very broad and moves

towards the specific area/hypotheses you are testing.

*Assessment:*

- *Does the background and significance have a logical organization?*
- *Does it move from the general to the specific?*
- *Has sufficient background been provided to understand the paper?*
- *How does this work relate to other work in the scientific literature?*
- *Has a reasonable explanation been given for why the research was done? Why is the work important? Why is it relevant?*
- *Does this section end with statements about the hypothesis/goals of the paper?*

e. **Statistical Methods:** begin with the statistical model, hypothesis of interest and then describe at least three different approaches you will use to analyze the data: Bayesian, likelihood and ???. Each subsection should contain formulas and details of calculations related to the method. For example, the Bayesian section might discuss prior elicitation/specification, and give the posterior calculations. You might also examine sensitivity to the choice of prior distribution and/or consider other non informative priors.

*Assessment:*

- *Could the analyses be repeated based on the information given here?*
- *Is the material organized into logical categories.*

f. **Results:** Typically, results sections start with descriptive statistics, e.g. what percent of the sample is male/female, what is the mean GPA overall, in the different groups, etc. Figures can be nice to illustrate these differences! However, information presented must be relevant in helping to answer the research question(s) of interest. Typically, inferential (i.e., confidence intervals, significance tests) statistics come next. Do not give computer output here! This should look like a peer-reviewed journal article results section. Tables and figures should be labeled, embedded in the text, and referenced appropriately. The results section typically makes for fairly dry reading. It does not explain the impact of findings, it merely highlights

and reports statistical information.

#### *Assessment*

- *Is the content appropriate for a results section? Is there a clear description of the results?*
- *Are the results/data analyzed well? Given the data in each figure/table is the interpretation accurate and logical? Is the analysis of the data thorough (anything ignored?)*
- *Are the figures/tables appropriate for the data being discussed? Are the figure legends and titles clear and concise?*

g. **Discussion/Conclusion:** discuss your findings in practical terms, compare and contrast the results with each other and with the Pfizer analysis, identify strengths and weaknesses of the methods used. The burning question is whether the prior used by Pfizer was non-informative or was it overly optimistic?

#### *Assessment:*

- *Do the authors clearly state whether the results answer the question (support or disprove the hypothesis)?*
- *Were specific data cited from the results to support each interpretation? Do the authors clearly articulate the basis for supporting or rejecting each hypothesis?*

h. **References:** List at least three references for your paper. These can be other articles on the subject/data, or mathematical references for the statistical techniques that you used. Each item you list here should be cited/referenced at least once in the actual text. For example, [here](#) is a blog by Stephen Senn that really helped me understand the problem. He's a great writer and you could read his other blogs on the vaccine trials.

#### *Assessment*

- *Are the references appropriate and of adequate quality?*
- *Are the references cited properly (both in the text and at the end of the paper)?*

h. **Appendix:** Include important R code, R output and additional plots/figures. Any plots you include in your report should be cited in the body of the report. There should be a reason for

the plot – something you want the reader to see. Note: tables and figures which are pertinent to your analysis should be included in the body of your report, not in the appendix. The reader (Harshil and I) should not have to look at your appendix for main results/figures.

## **5. Rubric**

- Each section mentioned above will be assigned points.
- Some general criteria we will use may include evidence of synthesis of ideas from the entire sequence, accuracy of data analysis, conclusions and discussion, overall clarity and presentation, originality, creativity, writing quality and organization
- Peer review: Each group member will complete a peer evaluation form providing feedback on the contribution of group members and scoring them on various criteria.

## **6. Expectation of team work**

Assigning roles to each team member is a vital part of ensuring your team works smoothly together. For example, in a three person group, you could designate to each member one of the following roles:

- Project manager: someone who has vision and can help with breaking things into tasks and identifying key due dates,
- Editor: someone who takes charge of the final document, proofreads sections and ensures that there are no inaccuracies.
- Meeting facilitator: someone who is good at guiding conversation can help ensure that the meetings stay productive and also democratic

The drafting of the main sections of the written report must be assigned to different members of the group. For example, in a 3-person group, one person might draft the first outline of the Introduction, another the Statistical Methods, the third the Results section and Discussion. Then you may all collaborate on editing and improving each section. Everyone should read all sections and suggest changes/edits, but one person will most likely be responsible for making all the editorial changes. The assignment of writing responsibilities should be noted on the peer review form.