METHOD REVIEW OF THE BNT162B2 VACCINE PHASE III TRIAL

CHRIS CHEN, PETER LIU, YUCHEN SUN, AND YUEQI XU

Abstract. In this project, we analyze the vaccine efficacy of BNT162b2 via both Bayesian and Frequestist approaches. We propose multiple prior selection methodologies, and cross-compare them with popular Frequestist methods such as likelihood ratio test and Clopper-Pearson interval. We elaborate the strength and weakness of the proposed methods and critically analyze the prior choice in Polack et al. (2020). We further provide our own prior selection philosophy, which we consider more justifiable than that of Polack et al. (2020).

Keywords. COVID-19, Vaccine Efficacy, Bayesian, Frequentist, Prior Selection

1. Introduction

Coronavirus disease 2019 (Covid 2019), a contagious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), continues to rage in a worldwide pandemic. A return to normality has increasingly come to rely on the research and development of safe and effective vaccines.

In 2020, BioNTech and Pfizer conducted an multinational, placebo-controlled, observer-blinded, pivotal efficacy trail, which 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 BNT162b2[1] vaccine candidate. In the efficacy analysis at the first primary endpoint, among 36523 participants who had no evidence of existing or prior SARS-CoV-2 infection, 8 cases of laboratory-confirmed Covid-19 with onset at least 7 days after the second dose were observed among vaccine recipients and 162 among placebo recipients.

A Bayesian beta-binomial model with a minimally informative prior is used for primary efficacy endpoint. Polack et al. (2020) proposed a beta prior with shape parameters $(0.700102, 1)$ for $\theta = (1-\psi)/(2-\psi)$, where ψ is the vaccine efficacy (VE). The prior is centered at

[1]A lipid nanoparticle–formulated, nucleoside-modified RNA vaccine that encodes a prefusion stabilized, membrane-anchored SARS-CoV-2 fulllength spike protein.

$\theta = 0.4118$, corresponding to VE = 30%, which was claimed to be pessimistic. Their result shows BNT162b2 was 95% effective in preventing Covid-19, with a 95% Credible interval of $[90.3\%, 97.6\%]$.

In this project, we will attempt to answer the following question: "*What will be a proper statistical methodology to analysis vaccine efficacy?*" We will draw from the tools and knowledge we learned throughout the STAT 34x sequence to investigate issues such as prior choice and to compare with Bayesian and Frequentist methods.

2. Methodology

We will now outline the general problem we consider. Denote X as the number of cases of Covid-19 infection among vaccine recipients and Y among placebo recipients. Heuristically, suppose all trials are independent, and the probability of being infected with Covid-19 are consistent across all subjects within each group. A reasonable binomial model for X, Y is

$$X \sim Binom(n_1, \pi_1), \quad Y \sim Binom(n_2, \pi_2), \quad X \perp Y, \tag{1}$$

where π_1, π_2 are the infection probability across vaccine and placebo group with sizes $n_1 = 17,411$ and $n_2 = 17,511$ respectively. The vaccine efficacy (VE), which is defined in Polack et al. (2020) as

$$\psi \equiv 1 - \frac{\pi_1}{\pi_2}, \tag{2}$$

is our parameter of interest. Note that, the maximum of ψ is attained at 1 (when $x = 0$), whereas the minimum is attained at minus infinity (when $y = 0$, $x > 0$), so that $\psi \in (-\infty, 1]$. Given this nature, directly doing inference on ψ can be intractable.

Polack et al. (2020) shied away from analyzing ψ directly and considered a transformation of ψ as an alternative. Now, Denote $W = X|X + Y = n$, where n is the observed infected cases from all subjects. Given large sample sizes and low event rate, the binomial distribution of X, Y can be approximated by a Poisson distribution with $\lambda_1 = n_1\pi_1$ and $\lambda_2 = n_2\pi_2$. Under

such an approximation, the distribution of W will follow as

$$W \sim Binom\left(n, \theta = \frac{n_1\pi_1}{n_1\pi_1 + n_2\pi_2}\right). \tag{3}$$

A simple derivation of the distribution of W will be given in Appendix 5.1. In the vaccine trial, due to 1:1 randomization, approximately we have $n_1 \approx n_2$. Thus, a heuristic expression of θ is

$$\theta = \frac{n_1\pi_1}{n_1\pi_1 + n_2\pi_2} \approx \frac{\pi_1}{\pi_1 + \pi_2} = \frac{1-\psi}{2-\psi}. \tag{4}$$

Indeed, given a fixed number of observed cases, θ can be considered as the probability of a randomly chosen case will be in the vaccine group. After transformation, θ takes the form of a binomial proportion, which immediately enables prevalent Frequestist and Bayesian methods, which we will elaborate as follows.

2.1. **Bayesian Workflow.** The binomial model enjoys a conjugate prior with the form of a Beta distribution with α, β being the first and second shape parameters. The corresponding posterior follows a Beta-binomial distribution

$$W|X, n, \theta, \alpha, \beta \sim Beta(\alpha + X, \beta + n - X). \tag{5}$$

Selecting a proper prior is the cornerstone of any Bayesian workflow. We will now introduce four philosophies for prior selection.

2.1.1. *Uninformative Priors.* We will first propose two uninformative priors that does not incorporate subjective belief from practitioners, namely the Flat prior and the Jeffreys prior. The Flat prior considers the probability of each parameter being selected is uniformly distributed across the parameter space, which is $\theta \sim Beta(1, 1) \equiv Unif(0, 1)$.

The Jeffreys prior is another prevalent choice of uninformative prior in Bayesian analysis. In the case of Beta-binomial model, the Jeffreys prior takes the form of $\theta \sim Beta(1/2, 1/2)$.

Note that, for binomial model, the data has the least effect on the posterior when $\theta_{\text{true}} = 1/2$, and has greatest effect near the extremes, that is, when $\theta_{\text{true}} = 0$ or 1, which is very

likely in the vaccine efficacy trial. The Jeffreys prior compensates for this by placing more mass on the extremes and less mass in the middle. See Figure 3 for a visualization of the Jeffreys prior for binomial model.

2.1.2. *Mean & Variance Based Beta Priors.* Polack et al. (2020) proposed a minimally informative prior with a prior belief on an average vaccine efficacy of 30%. Minimally informativeness can be consider in regard to the variance of prior distribution. More specifically, the weaker the prior belief, the larger the variance, the less informative the prior is. Prior belief on vaccine efficacy is incorporated by Polack et al. (2020) as follows: by transformation in (4), a VE of 30% corresponds to a $\theta = 0.4118$. Thus correspondingly, the mean of θ is set as $\mu_\theta = 0.4118$. Then, by straight-forwardly specifying $\beta = 1$[2], a set of desired parameters of the beta prior is obtained by solving $\mu_\theta = \alpha/(\alpha + \beta)$, which yields

$$\theta \sim Beta(0.700102, 1). \tag{6}$$

Such a prior choice, with a variance of 0.0901, allows considerable uncertainty. The 95% interval for θ is $[0.005, 0.964]$, with the corresponding 95% interval for VE to be $[-2620\%, 99.5\%]$. In this case, naively choosing $\beta = 1$ results in a desired minimally informative prior with large variance. However, this may not be true in general and can be problematic in practice.

We will now propose a generalized prior selecting procedure based on Polack et al. (2020) with a user-defined VE average μ_ψ and variance σ^2. We will first transform μ_ψ into μ_θ accordingly by (4). Now, a desired beta prior on θ will follow as

$$\theta \sim Beta\left(\frac{\mu_\theta^2(1-\mu_\theta) - \mu_\theta\sigma^2}{\sigma^2}, \frac{(1-\mu_\theta)(\mu_\theta^2(1-\mu_\theta) - \mu_\theta\sigma^2)}{\mu_\theta\sigma^2}\right). \tag{7}$$

Such a prior will enable practitioners to control the mean and informativeness of the beta prior. A visualization of the relationship between informativeness and variance can be found at Appendix Figure 2.

[2]For $\mathbb{E}[\psi] < \infty$, it is required that $\beta > 1$. However, prior choices with $\beta \leq 1$ can also be considered as valid.

2.1.3. *Single Quantile & Variance Based Beta Priors.* An alternative way to incorporate prior belief is via specifying the quantile of prior distribution. Echo to mean based methods in Section 2.1.2, we will propose a prior selecting procedure with a user-defined VE quantile q_ψ and variance σ^2.

Similarly, we will first transform q_ψ into q_θ. Note that, unlike the previous case, given q_θ and σ^2, there are no closed form solutions for (α, β). Thus, a grid approximation is used to find a desired set of the shape parameters. The algorithm will scan through a fine grid of a subset of $\mathbb{R}^2$, and will stop if the differences in quantile and variance are less than a desired threshold. An `R` code implementation can be found in Appendix 5.3.

2.1.4. *Double Quantile Based Beta Priors.* Apparently, the prior belief can be interpreted via more than one quantile. Suppose for now that the practitioners express their prior beliefs over two quantile of vaccine efficacy, namely $q_{\psi,1}$ and $q_{\psi,2}$. Our first step will continue to transform $(q_{\psi,1}, q_{\psi,2})$ to $(q_{\theta,1}, q_{\theta,2})$. A desired choice of beta prior is then computed using the `LearnBayes::beta.select` function in `R`. More specifically, in this vaccine trial, we will provide a prior choice based on our own belief, which will be speculated in the following manner:

(1) The median of vaccine efficacy is 0. Thus θ has a median of 0.5;
(2) The $95th$ percentile of vaccine efficacy is 0.3. Thus the $5th$ percentile of θ is 0.4118.

Such a prior can be considered more pessimistic than the prior proposed in Polack et al. (2020). Moreover, it reflects a strong prior belief, with a variance of 0.0029.

In later result and discussions, we will elaborate the strength and weakness of these priors, and cross compare their estimation performances.

2.2. **Frequestist Workflow.** We will further introduce two Frequestist approaches reciprocal to the Bayesian analysis.

2.2.1. *Confidence Interval for Binomial Proportion.* Polack et al. (2020) considered the Clopper-Pearson method to construct confidence interval (CI) for the binomial proportion

θ. The CI is directly converted upon vaccine efficacy via transformation (4). The Clopper-Pearson CI guarantees the coverage rate is at least the nominal confidence interval by inverting the acceptance regions based on the exact p-value corresponding to the binomial distribution. This is a prevalent choice to analyze binomial proportion and is the default of `mosaic::binom.test` in `R`.

Clopper-Pearson method in Polack et al. (2020) yields an average of 95.0% for VE and a 95% CI of $[90.0\%, 97.9\%]$. Given the sample size is relatively large, we will further propose three alternative CI methods to compare with the proposed Clopper-Pearson: the Wald method, the Wilson's Score method, and the Plus-4 method. These methods are all built-in functions in `R` and can be applied efficiently. Closed formed formulas and implementation details can be found in Appendix 5.1.

2.2.2. *Likelihood Ratio Test.* Likelihood Ratio Test (LRT) is a popular likelihood-based method that is widely used to perform hypothesis testing and to construct confidence interval. In our case, the hypothesis[3] concerning vaccine efficacy ψ is

$$H_0 : \psi = 0.3, \quad H_1 : \psi \neq 0.3 \, . \tag{8}$$

By transformation via Equation (4), the pair of hypothesis can be re-written as

$$H_0 : \theta = 0.4118, \quad H_1 : \theta \neq 0.4118 \, . \tag{9}$$

Given the binomial model, the MLE of θ has the closed form solution $\widehat{\theta}_{MLE} = W/n$, thus the likelihood ratio is

$$\lambda = \frac{L(\theta_0)}{L(\widehat{\theta}_{MLE})} = \frac{(\theta_0)^w (1-\theta_0)^{n-w}}{(\widehat{\theta}_{MLE})^w (1-\widehat{\theta}_{MLE})^{n-w}}. \tag{10}$$

The binomial model for efficacy data enjoys a smooth likelihood function, a relatively large sample size, and a support that is invariant of choices of θ. Thus, by Theorem 2.1 of likelihood

[3]Intuitively, the hypothesis should be $H_0 : \psi \leq 0.3, H_1 : \psi > 0.3$. However, this will not satisfy the dimension constraints of the common LRT. See Casella and Berger (2001) Example 8.2.6 for a LRT framework that may be applied on this hypothesis. In this project, we will take the hypothesis in (8) for LRT heuristically.

ratio lecture note, under the null hypothesis we have

$$-2\log(\Lambda) \xrightarrow{D} Chisq(df = 1). \tag{11}$$

The validity of such an approximation is provided in Appendix Figure 7. A 95% confidence interval for θ can be calculated via

$$-\log(\lambda) < qchisq(p = 0.95, df = 1). \tag{12}$$

3. Results & data analysis

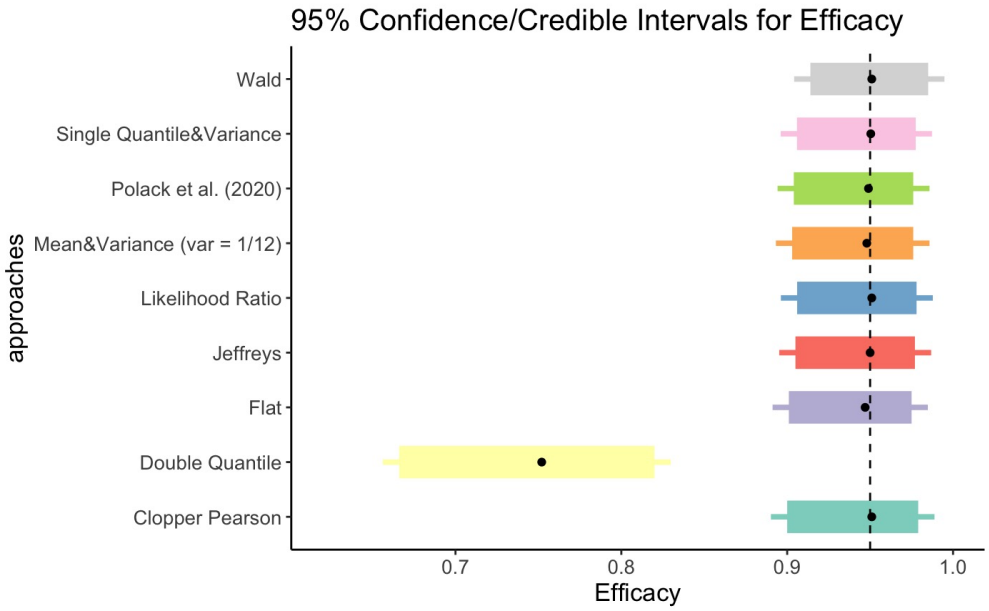

FIGURE 1. 95% Bayesian credible intervals and Frequentist CIs of vaccine efficacy.

3.1. **Results for Bayesian Approaches.**

3.1.1. *Uninformative Priors.* For the uninformative priors, we will not incorporate any prior information or beliefs on vaccine efficacy. With data dominating the posterior, the medians for ψ should be close to Frequestist estimates. The median of Flat prior and Jeffreys prior are 94.7% and 95.0% respectively, which are quite close to the expected 95.1% (with $\theta = 8/170$). Similarly, the 95% credible intervals of ψ are very close, with a credible interval of $[90.1\%, 97.5\%]$ from flat prior and $[90.5\%, 97.7\%]$ from Jeffreys prior.

3.1.2. *Mean & Variance Based Beta Priors.* We will consider two sets of parameters for this methodology. First, we considered the proposed prior in Polack et al. (2020) with $\alpha = 0.700102, \beta = 1$ (See Appendix Figure 4). The 95% credible interval for ψ is $[90.4\%, 97.6\%]$ and the posterior median is 94.9%.

Second, we considered an alternative prior proposed in Senn (2021), satisfying $\theta = 0.4118$ and a variance $\sigma^2 = 1/12$, which is the same as the variance of a Flat prior. Via calculation in Equation (7), we have $\alpha = 0.7852, \beta = 1.1215$. The corresponding 95% credible interval is $[90.3\%, 97.6\%]$, which coincides with the 95% credible interval of ψ in Polack et al. (2020), with a posterior median of 94.8%. This is expected, since both priors are weakly informative with the same prior mean.

3.1.3. *Single quantile & Variance Based Beta Priors.* We will consider a prior with a single quantile based prior belief as the median of VE is 50%. We will not specify a variance for this prior, and straightforwardly take $\beta = 1$, echoing the parameter choice in Polack et al. (2020). This corresponds to a beta prior with parameters $\alpha = 0.6293, \beta = 1$. (See Appendix Figure 5). The 95% credible interval for ψ is $[90.4\%, 97.7\%]$ and the posterior median is 94.9%.

3.1.4. *Double Quantile Based Beta Priors.* We will consider a strongly informative, pessimistic prior, which is elaborated in Section 2.1.4. This corresponds to the minimum VE requirement established by FDA and what we want to show how the VE is "at least better than". We can observe the prior distribution, posterior distribution, and likelihood function have three unique peaks with the prior on the rightmost position and likelihood at the leftmost position (See Appendix Figure 6).

The 95% credible interval of ψ is $[66.6\%, 82.0\%]$, which is significantly distinct from other intervals and does not include 95.1%. Intuitively this makes sense, since our prior beliefs are very inconsistent with the observed data, which indicates our prior beliefs were being too pessimistic. However, note that the credible interval is still a lot greater than 30%, which implies VE is greater than 30% with a probability at least of 95%.

Approach	Prior	Credible Interval	Median
Jeffreys	$Beta(0.5, 0.5)$	$[90.5\%, 97.7\%]$	95.0%
Flat	$Beta(1, 1)$	$[90.1\%, 97.5\%]$	94.7%
Double-Quantile	$Beta(43.06, 43.06)$	$[66.6\%, 82.0\%]$	75.2%
Polack et al. (2020)	$Beta(0.700102, 1)$	$[90.4\%, 97.6\%]$	94.9%
Mean & Variance ($\sigma^2 = 1/12$)	$Beta(0.7852, 1.1215)$	$[90.3\%, 97.6\%]$	94.8%
Single Quantile & Variance	$Beta(0.6293, 1)$	$[90.4\%, 97.7\%]$	95.0%

TABLE 1. 95% credible intervals and medians of ψ of different prior choices.

3.2. **Frequentist Approaches.**

3.2.1. *Confidence Interval for Binomial Proportion.* We observed the Wald interval behaves worse than the Clopper-Pearson's because of the relatively small sample size ($n = 170$), which results in less asymptomaticity. The 95% confidence intervals of ψ with different interval methods are provided in the table below.

3.2.2. *Likelihood Ratio Test.* The 95% confidence interval of the likelihood ratio test is [90.6%, 97.8%], with a mean of 95.1%. The validity of using a Chi-square approximation is verified using an empirical simulation with the p-value equals to 0.0002 (See Appendix 7).

Approach	Confidence Interval	Mean
Likelihood Ratio	$[90.6\%, 97.8\%]$	95.1%
Wald	$[91.4\%, 98.5\%]$	95.1%
Score	$[90.1\%, 97.5\%]$	95.1%
Plus-4	$[89.9\%, 97.7\%]$	95.1%
Clopper Pearson	$[90.0\%, 97.9\%]$	95.1%

TABLE 2. 95% confidence intervals and means of ψ

4. Discussion & Conclusion

We favor a Bayesian design over its Frequestist opponent attributed to the different statistical interpretations carried by the credible interval and confidence interval. A 95% Bayesian credible interval $[c_{\min}, c_{\max}]$ of ψ indicates that, given the observed data, the probability that the true ψ is bounded by $[c_{\min}, c_{\max}]$ is 0.95. In contrast, the boundaries of a 95% Frequestist

confidence interval does not carry any statistical meanings. The Frequestist CI can be interpreted in the following manner: "*were the vaccine trial repeated numerous times, the fraction of the calculated confidence intervals that encompass the true VE would tend towards 95%* (Cox and Hinkley, 1974)." Thus, the Bayesian credible interval is preferred as it directly evaluates the probability of the true vaccine efficacy without assuming the vaccine trial were conducted repeatedly.

We note that the transformation $T : \theta \to \psi$ takes the form of

$$T: \quad \psi = T(\theta) = \frac{1-2\theta}{1-\theta} = 2 - \frac{1}{1-\theta}, \quad 0 < \theta < 1, \tag{13}$$

which is a one-to-one, monotone decreasing, non-linear transformation, so that only quantile based information is invariant of transformation T.

In our analysis, inferences are made upon θ instead of directly on ψ. Thus, for statistical intervals, the interval boundaries are first calculated in terms of θ before being transformed into ψ. It arises the question that whether the statistical properties of the intervals will be preserved under such a transformation. Since T is monotone, the coverage rates for both Bayesian and Frequestist intervals will remain unchanged. However, the non-linear nature of T will potentially make the transformed Highest Density Interval (HDI) and Frequestist CI problematic. To see this, since T is non-linear, the HDI of ψ and θ doesn't necessarily share the same quantile as boundaries. Thus, the transformed HDI of θ may no longer be considered as the HDI of ψ. See Kruschke (2014) for a detail argument on HDI.

Similarly, Frequestist CI's are not preserved under non-linear transformation. For binomial proportion, only the center of CI, namely the MLE of θ, is invariant of T. In general, given the same data, the transformed CI of θ will not match the CI of ψ computed directly from the distribution of ψ.

Only the Equal Tailed Interval (ETI) is invariant under transformation T in terms of coverage rate and statistical interpretation, as it is solely constructed with the quantile of posterior distribution. Therefore, we suggest using ETI to construct Bayesian credible intervals based on these theoretical considerations.

Another potential problem is whether the prior beliefs will be preserved under transformation T. Since T is nonlinear, a Flat prior on θ will become informative after transformation, hence its prior beliefs on θ will not be transmitted to ψ. Similarly, a mean based prior on θ with $\mu_\theta = 0.4118$ will not correspond to a transformed prior on ψ with $\mu_\psi = 0.3$. To the best of our knowledge, only the quantile based, informative prior, along with the uninformative Jeffrey's prior, will preserve prior beliefs and remain invariant under monotone, nonlinear transformation.

We therefore consider the proposed prior in Polack et al. (2020) as inappropriate. As illustrated previously, prior beliefs of a mean based prior will not be preserved under the non-linear transformation. Instead, a quantile based prior, as proposed in Section 2.1.3, can be considered as a suitable alternative.

Furthermore, we don't reckon an average of VE at 30% represents any genuine vaccine belief of practitioners. Evidently, a vaccine candidate proceeding with a Phase III clinical trial should attain a vaccine efficacy of at least 50%. In other word, from the practitioners' prospective, they will *believe* the vaccine candidate is going to be at least 50% effective before any Phase III vaccine trial. We consider an appropriate prior should incorporate such an expertise belief.

Only heuristically, the prior belief in Polack et al. (2020) can be considered as a Bayesian version of null hypothesis significance testing (NHST) against the minimal FDA efficacy requirement so that $P(\psi > 30\% \mid \text{observed data}) > 99.99\%$. Such a NHST could have been done properly via Bayes factors. Unfortunately, this is not taken into account in Polack et al. (2020).

We will now outline our favored prior selection procedure. We will first specify the prior to follow a Beta distribution, so that it is conjugated with the proposed binomial model. For informativeness, since BNT162b2 is the one of the first mRNA based vaccine candidates that is studied in large-scale efficacy trial, the practitioners should not possess strong, evidence-based beliefs on it efficacy. Corresponding to such a belief, we will consider our prior as weakly informative, with relatively large variance.

We will now speculate the practitioner's belief on vaccine efficacy is at least 50%. We will further interpret such a belief in a quantile based manner, that is, we will suppose our prior satisfies a median of VE at 50%. For a valid expectation of ψ, we will set $\beta = 1$. A grid approximation will yield a set of desired parameter at $(0.6293, 1)$, with a variance of 0.0901. Given the efficacy data, such a prior corresponds to a 95% credible interval of ψ is given as $[90.4\%, 97.7\%]$, with a median of VE at 94.9%.

Further note that, the only reason for us choosing a conjugate beta prior is that we only intend to incorporate a single quantile based prior belief. For a second quantile belief, `beta.select` in Section 2.1.4 can be used to obtain a desired beta distribution. However, if two quantiles of a beta distribution is specified, then the shape parameters become deterministic (See Dorp (2003)). Therefore, if one wish to incorporate multiple quantile based prior beliefs with variance (informativeness) requirement, a beta prior might not be an optimal choice. In that case, other flexible priors, which can be manipulated to satisfy the quantile and variance requirements simultaneously, are preferred. With unconventional prior choices, the corresponding posteriors might not be of analytical form, but numerical algorithms such as MCMC can always be applied to sample from the posterior. It is also possible to directly consider priors with respect to ψ, and simply consider θ as an intermediate variable. This can be done efficiently via Bayesian analytical software such as `Stan`.

References

Casella, G. and Berger, R. (2001). *Statistical Inference*. Duxbury Resource Center.

Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics*. Chapman & Hall, London, England.

Dorp, J., M. T. (2003). Parameter specification of the beta distribution and its dirichlet extensions utilizing quantiles.

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan, second edition.*

Polack, F., Thomas, S., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J., Marc, G., Moreira Jr, E., Zerbini, C., Bailey, R., Swanson, K., Roychoudhury, S., Koury, K., Li, P., Kalina, W., Cooper, D., Frenck, R., Hammitt, L., and Gruber, W. (2020). Safety and efficacy of the bnt162b2 mrna covid-19 vaccine. *New England Journal of Medicine*, 383.

Senn, S. (2021). S. senn: "beta testing": The pfizer/biontech statistical analysis of their covid-19 vaccine trial (guest post).

5. Appendix

5.1. **Formulas.**

(1) Find distribution of W: Since the sample sizes in each group are large and the event rates are small, a Poisson approximation to a binomial can really make sense here. In other words, if we have

$$X \approx Poisson(n_1\pi_1); \quad Y \approx Poisson(n_2\pi_2),$$

then

$$\begin{aligned}
P(W=w) &= P(X=w|X+Y=n) \\
&= \frac{P(X=w \cap X+Y=n)}{P(X+Y=n)} \\
&= \frac{P(X=w \cap Y=n-w)}{P(X+Y=n)} \\
&= \frac{P(X=w)P(Y=n-w)}{P(X+Y=n)} \quad X, Y \text{ independent} \\
&= \frac{(e^{-n_1\pi_1}(n_1\pi_1)^w)/w! \times (e^{-n_2\pi_2}(n_2\pi_2)^{n-w})/(n-w)!}{(e^{-(n_1\pi_1+n_2\pi_2))}(n_1\pi_1+n_2\pi_2)^n/n!} \\
&= \frac{n!}{w!(n-w)!}\frac{(n_1\pi_1)^w(n_2\pi_2)^{n-w}}{(n_1\pi_1+n_2\pi_2)^n} \\
&= \binom{n}{w}\left(\frac{n_1\pi_1}{n_1\pi_1+n_2\pi_2}\right)^w\left(\frac{n_2\pi_2}{n_1\pi_1+n_2\pi_2}\right)^{n-w} \\
&= \binom{n}{w}\theta^w(1-\theta)^{n-w}, \quad w=0,1,\ldots,n, \text{ where } \theta=\frac{n_1\pi_1}{n_1\pi_1+n_2\pi_2}.
\end{aligned}$$

Therefore,

$$W \sim Binom\left(n, \theta = \frac{n_1\pi_1}{n_1\pi_1 + n_2\pi_2}\right). \tag{14}$$

(2) The $100 - (1-\alpha)\%$ confidence intervals for θ using Wald, Wilson Score, and Plus-4 methods are shown as follows:

Wald:

$$\widehat{\theta} \pm z_{\alpha/2}\sqrt{\frac{\widehat{\theta}(1-\widehat{\theta})}{n}}, \qquad \widehat{\theta} = \frac{W}{n}. \tag{15}$$

Wilson's Score:

$$\frac{\widehat{\theta} + z^2_{\alpha/2}/2n \pm z_{\alpha/2}\sqrt{\widehat{\theta}(1-\widehat{\theta})/n + z^2_{\alpha/2}/4n^2}}{1 + w^2_{\alpha/2}/n}. \tag{16}$$

Plus-4:

$$\tilde{\theta} \pm z_{\alpha/2}\sqrt{\frac{\tilde{\theta}(1-\tilde{\theta})}{n+4}}, \qquad \tilde{\theta} = \frac{w+2}{n+4}. \tag{17}$$

(3) Some details:

(a) Prior beta distribution has parameters (α, β).

(b) Data (X) is observed from a binomial distribution with parameters (N, θ).

(c) Posterior is a beta distribution with parameters $(X + \alpha, N - X + \beta)$.

(d) $\hat{\theta}_{MLE}$ is $\frac{X}{N}$.

(e) Mean of the prior is $\frac{\alpha}{\alpha+\beta}$.

(f) Mean of the posterior is $\frac{X+\alpha}{N+\alpha+\beta}$.

$$\begin{aligned}
\mathbb{E}(\theta|X) &= \frac{X+\alpha}{N+\alpha+\beta} \\
&= \frac{X}{N+\alpha+\beta} + \frac{\alpha}{N+\alpha+\beta} \\
&= \frac{X}{N} \times \frac{N}{N+\alpha+\beta} + \frac{\alpha}{\alpha+\beta} \times \frac{\alpha+\beta}{N+\alpha+\beta} \\
&= \hat{\theta}_{MLE} \times W + \text{Prior Mean} \times (1-W)
\end{aligned}$$

The larger the alpha and beta choices, the more prior distribution's mean influences the posterior distribution. The smaller the choices, the more the MLE estimate of θ influences the posterior distribution.

5.2. **Figures.** Figures in the Appendix are listed below.

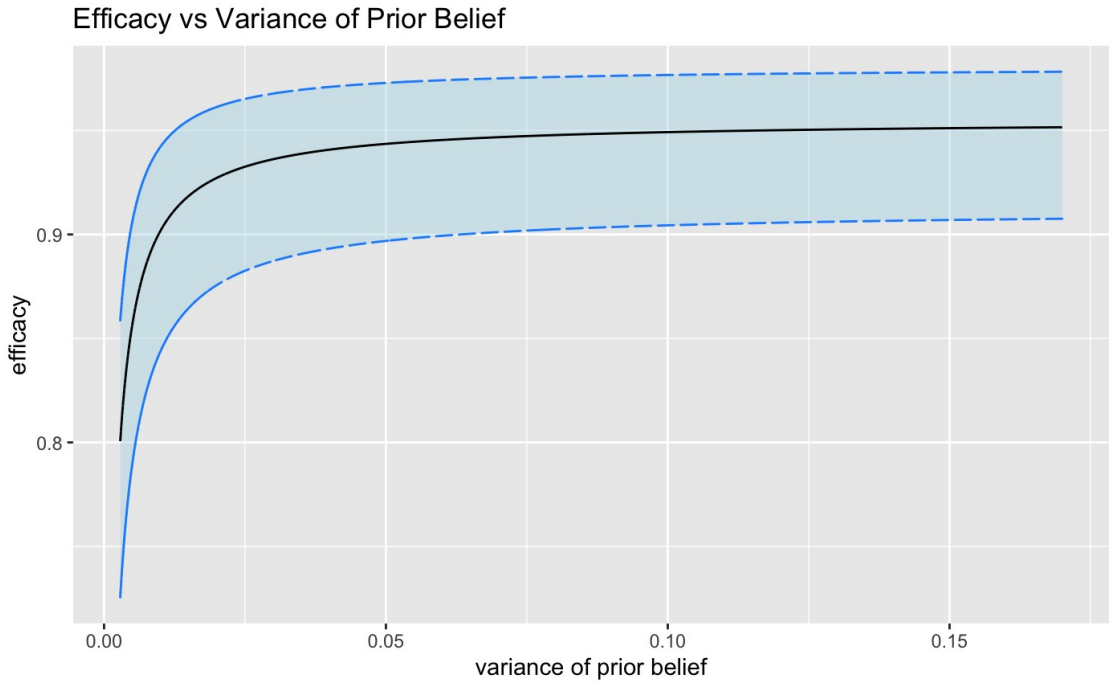

FIGURE 2. Relationship between 95% credible interval of vaccine efficacy and variance of the prior.

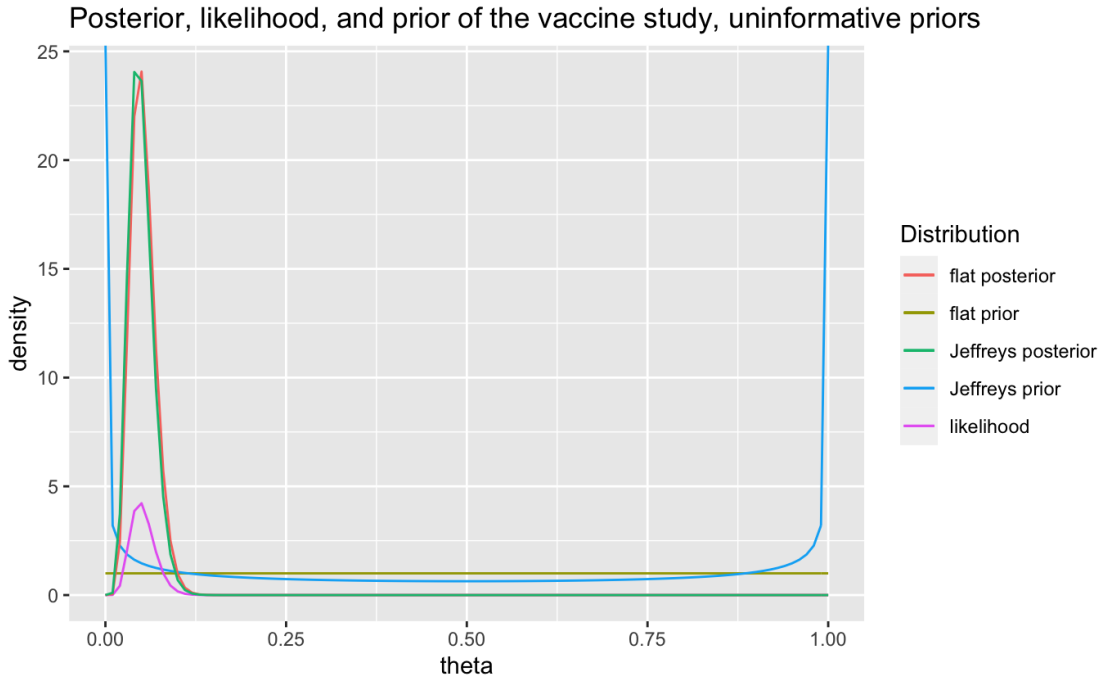

FIGURE 3. Posterior, likelihood, and prior of the vaccine study with uninformative priors.

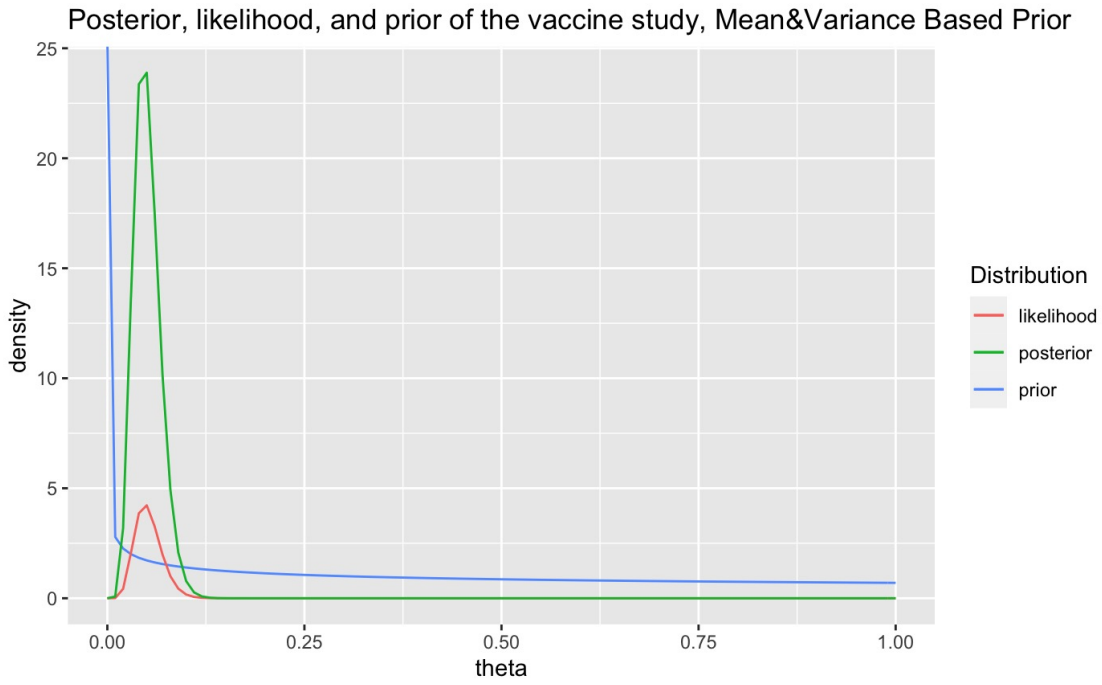

FIGURE 4. Posterior, likelihood, and prior of the vaccine study with mean & variance based prior.

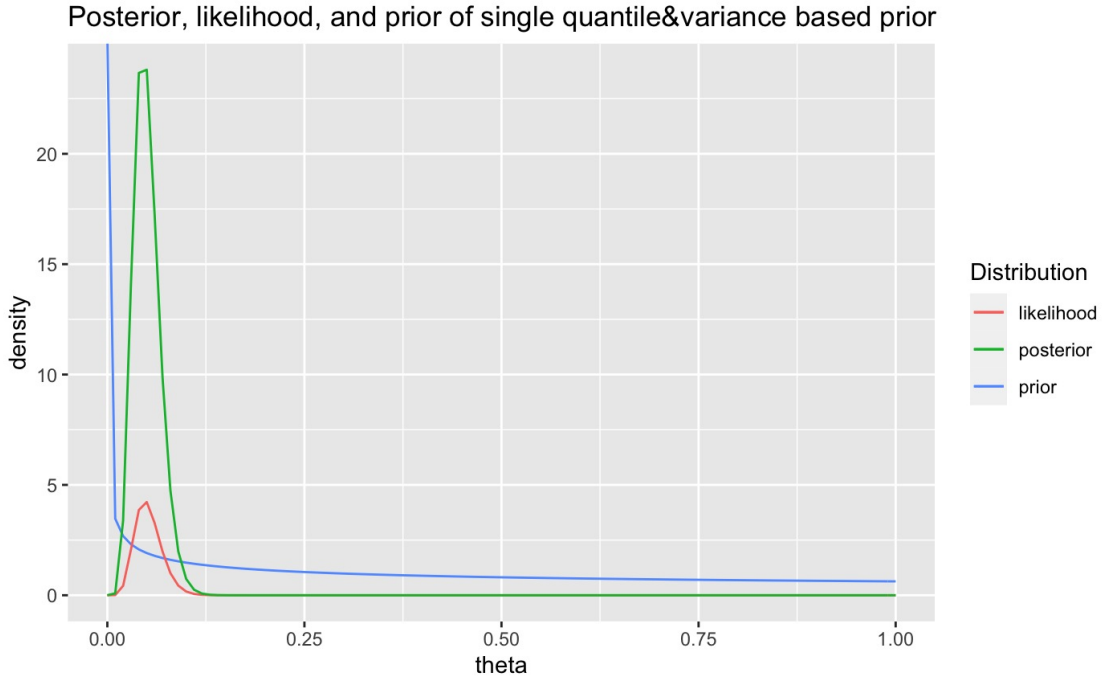

FIGURE 5. Posterior, likelihood, and prior of the vaccine study with single quantile & variance based prior.

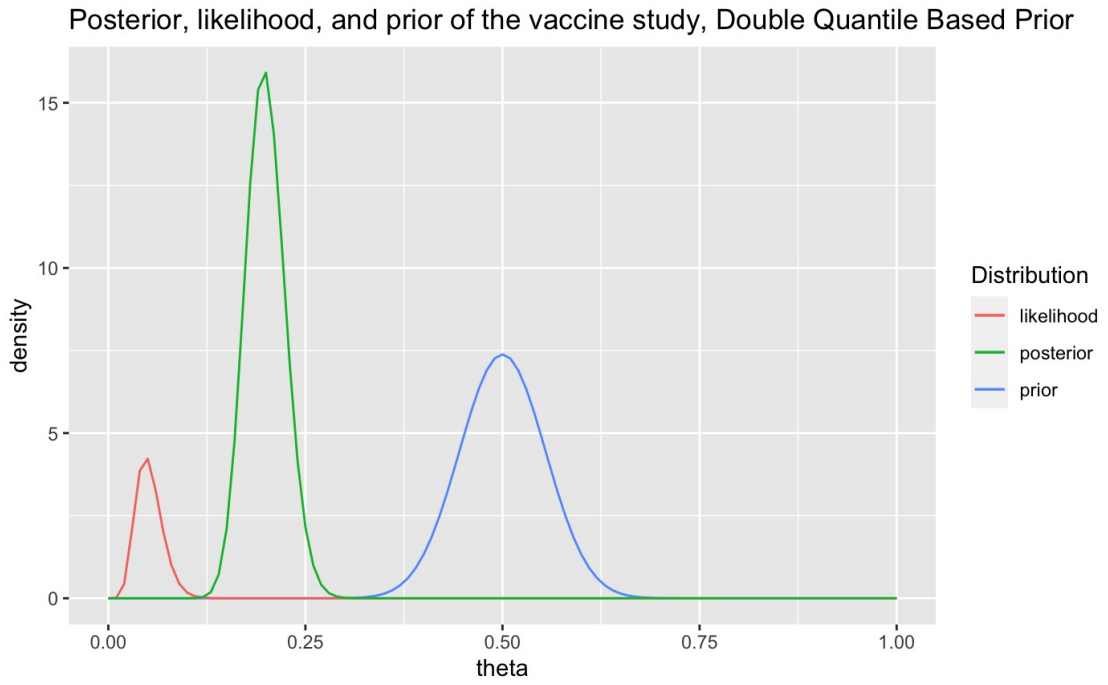

FIGURE 6. Posterior, likelihood, and prior of the vaccine study with double quantiles based prior.

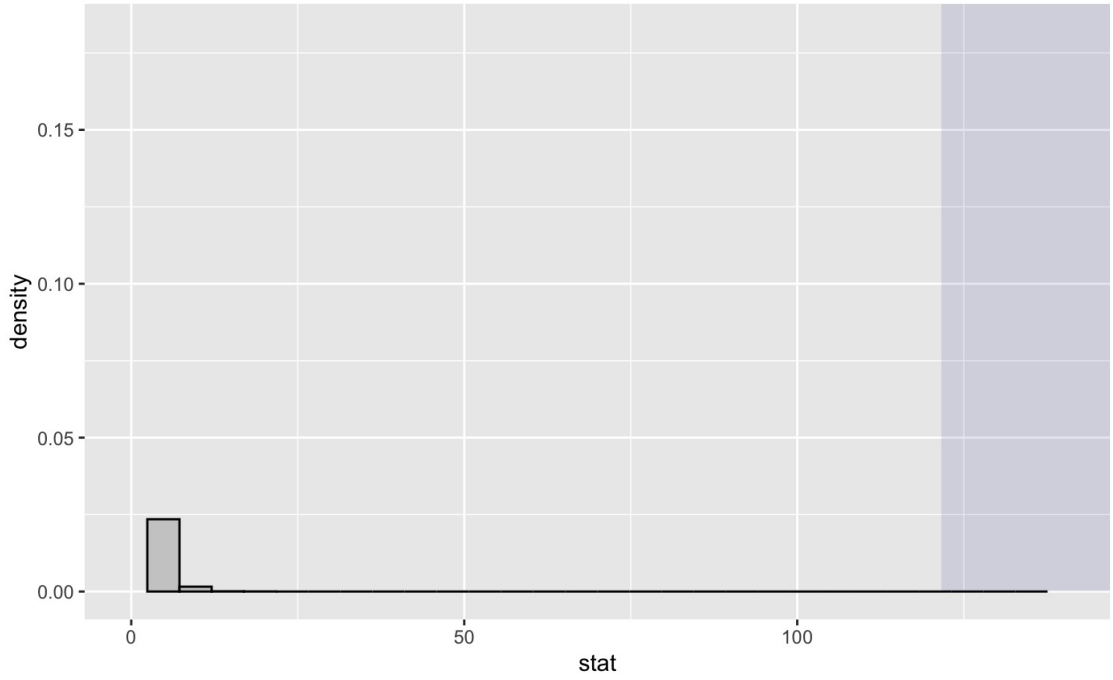

FIGURE 7. Empirical Chisq approximation.

5.3. **Codes.** Transform between ψ and θ:

```
# convert from theta to psi
topsi <- function(theta) {
  (1 - 2 * theta) / (1 - theta)
}
# convert from psi to theta
totheta <- function(psi) {
  (1 - psi) / (2 - psi)
}
```

Find α and β satisfies the given quantile and variance constraint:

```
# grid approximation
psi <- 0.5
theta <- totheta(psi)
beta_var <- function(a, b){
  a*b/((a+b)^2*(a+b+1))
}

# var = 1/8
var = c()
alpha <- seq(0.384545, 0.384546, 0.000001)
beta <- seq(0.55167, 0.5517, 0.000001)
for(i in alpha){
  for(j in beta) {
    med = qbeta(0.5, i, j)
    var = beta_var(i, j)
    if (abs(med-theta) < 0.000001 & abs(var-1/8) < 0.000001){
      print(paste(med, var, i, j))
    }
  }
}
```

Find credible interval for any given prior $Beta(\alpha, \beta)$:

```
a <- alpha
b <- beta
w <- 8
n <- 170
post_beta <- n - w + b
post_alpha <- w + a
lower <- topsi(qbeta(0.975, shape1 = post_alpha, shape2 = post_beta))
upper <- topsi(qbeta(0.025, shape1 = post_alpha, shape2 = post_beta))
CI <- c(lower, upper)
```

Find confidence interval for ψ using multiple methods:

```
library(mosaic)
library(tidyverse)

wald <- confint(binom.test(x = 8, n = 170, ci.method = "Wald"))
```

```
score <- confint(binom.test(x = 8, n = 170, ci.method = "score"))
plus4 <- confint(binom.test(x = 8, n = 170, ci.method = "Plus4"))
cp <- confint(binom.test(x = 8, n = 170))

transform.confint <- function(lower.theta, upper.theta){
     tibble(lower = (1-2*upper.theta)/(1-upper.theta), upper =
        (1-2*lower.theta)/(1-lower.theta))
}

wald.psi <- transform.confint(wald$lower, wald$upper)
score.psi <- transform.confint(score$lower, score$upper)
plus4.psi <- transform.confint(plus4$lower, plus4$upper)
cp.psi <- transform.confint(cp$lower, cp$upper)
psi.confint.matrix <- rbind(wald.psi, score.psi, plus4.psi, cp.psi)

psi.confint.matrix$method <- c("Wald", "Wilson Score", "Plus 4", "Clopper
    Pearson")
psi.confint.matrix %>% dplyr::select(method, lower, upper)
```

Code realization of Figure 1 via ggplot2:

```
approaches <- c('Double Quantile', 'Polack et al. (2020)', 'Mean&Variance
    (var = 1/12)', 'Flat', 'Jeffreys', 'Likelihood Ratio', 'Wald', 'Clopper
    Pearson', 'Single Quantile&Variance')
lowerci <- c(0.666, 0.904, 0.903, 0.901, 0.905, 0.906, 0.914, 0.900, 0.9059)
upperci <- c(0.820, 0.976, 0.976, 0.975, 0.977, 0.978, 0.985, 0.979, 0.9775)
median <- c(0.752, 0.949, 0.948, 0.947, 0.950, 0.951, 0.951, 0.951, 0.9504)
CIs <- data.frame(
  approaches <- approaches,
  lower <- lowerci,
  upper <- upperci,
  median <- median

 pd <- position_dodge(0.78)
ggplot(CIs, aes(y=approaches)) +
  geom_errorbar(data=CIs, aes(xmin=lower, xmax=upper, color=approaches),
      width=.1, position=pd, size = 8) +
  geom_vline(xintercept = 0.950, linetype="dashed") +
  geom_point(data=CIs, aes(x=median), position=pd) +
  xlim(0.62, 1) +
  xlab("Efficacy") +
  ggtitle("95% Confidence/Credible Intervals for Efficacy") +
  scale_color_brewer(palette="Set3") +
  theme_bw() +
  theme(panel.border = element_blank(), panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(), axis.line = element_line(colour =
           "black"),
        text = element_text(size = 13), legend.position = "none")
```