

Reanalyzing Pfizer's BNT162b2 Vaccine Efficacy Using Bayesian and Frequentist Approaches

Isaac Muhlestein, Jian Kang, Troy Russo

3/16/2024

Abstract

Write your abstract here.

Keywords

Maximum Likelihood Estimation, Bayesian Beta-Binomial Model, Parametric Bootstrap, Likelihood Ratio Test

Introduction / Background

Introduction

The COVID-19 pandemic, caused by the coronavirus SARS-CoV-2, has had a massive global impact since it first emerged in late 2019, affecting health, economies, and daily life worldwide. Developing safe and effective vaccines quickly became crucial to help curb the spread and reduce severe illness. One of the earliest breakthroughs was the mRNA-based BNT162b2 vaccine from Pfizer and BioNTech, which showed impressive efficacy in clinical trials (Polack et al., 2021). By December 2020, the U.S. Food and Drug Administration (FDA) granted Emergency Use Authorization (EUA) for BNT162b2 (FDA, 2020)—a major turning point in fighting the pandemic. This authorization was supported by data from a phase 2/3 randomized, placebo-controlled trial involving over 34,000 participants, which reported a 95% vaccine efficacy (VE) against symptomatic COVID-19 (Polack et al., 2021).

Vaccine efficacy trials rely on rigorous statistical frameworks to evaluate the reduction in disease risk between vaccinated and unvaccinated groups. Traditional frequentist methods, such as confidence intervals and p-values, have long dominated clinical trial analyses. However, Bayesian approaches are increasingly employed for their ability to incorporate prior evidence and provide probabilistic interpretations of efficacy (Berry et al., 2006). The original BNT162b2 trial used a Bayesian beta-binomial model with a Beta(0.700102, 1) prior to estimate VE, defined as

$$\psi = \frac{1 - 2\pi}{1 - \pi} \tag{1}$$

where π represents the probability of infection in the vaccine group relative to placebo (Polack, et al., 2021). This analysis yielded a posterior median VE of 95% (95% credible interval: 90.3–97.6%) and a $>99.99\%$ posterior probability that VE exceeded 30%. These results aligned with findings from other mRNA vaccines like Moderna’s mRNA-1273 (Baden et al., 2021).

Despite the strong evidence presented by the original Bayesian analysis, comparing Bayesian and frequentist methodologies offers unique insights. Frequentist methods, which avoid prior distributions and focus on long-run frequency, remain the regulatory standard for many trials (FDA, 2020).

A double analysis could strengthen our confidence in the results by examining similarities and differences between the two frameworks, particularly in estimating uncertainty intervals. For instance, while Bayesian credible intervals reflect posterior probability, frequentist confidence intervals rely on repeated sampling interpretations—a distinction with practical implications for interpretation (Wasserstein et al., 2019).

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

| Group | Cases | Sample Size |
|----------|-------|-------------|
| BNT162b2 | 8 | 17,411 |
| Placebo | 162 | 17,511 |
| Total | 170 | 34,922 |

| ## | Original Clinical Data | | | |
|----|------------------------|----------|--------------|-------|
| ## | Group | Infected | Not Infected | Total |
| ## | Placebo | 162 | 17349 | 17511 |
| ## | Vaccine | 8 | 17403 | 17411 |
| ## | Total | 170 | 34752 | 34922 |

The present study reanalyzes these data Table 1 using both Bayesian and frequentist approaches. The primary goals are to: (1) estimate VE and associated uncertainty under each framework, (2) compare the resulting efficacy point and interval estimates, and (3) evaluate the posterior probability (Bayesian) and p-values (frequentist) for the hypothesis that VE exceeds 30%. Our hypotheses are as follows:

Hypotheses:

- Both frameworks will confirm high VE ($>90\%$) for BNT162b2
- Uncertainty intervals will differ modestly between methods
- Both analyses will reject the null hypothesis ($VE \leq 30\%$) with overwhelming evidence

This comparative analysis not only reinforces confidence in BNT162b2’s efficacy but also highlights the complementary strengths of Bayesian and frequentist statistics in public health research.

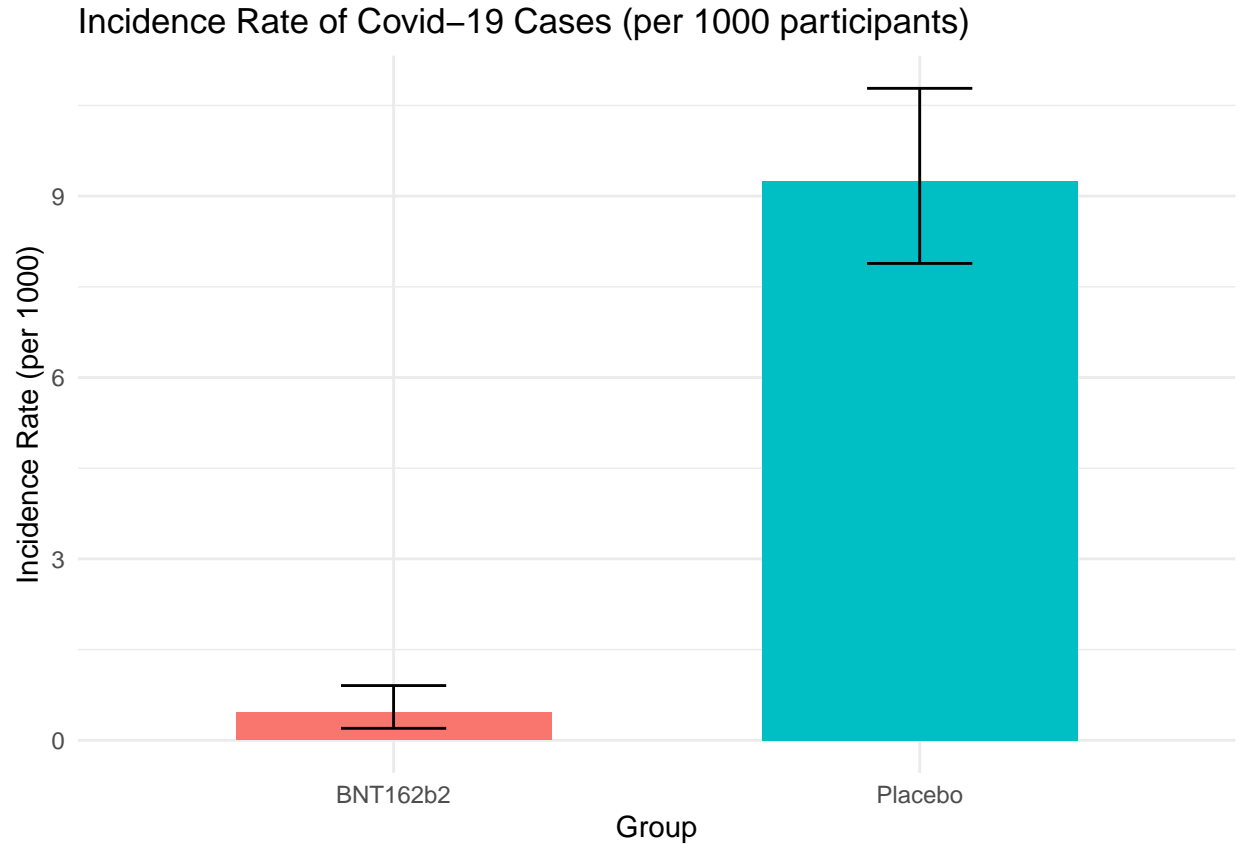


Table 2: Estimated Vaccine Efficacy (VE) based on observed attack rates.

| Group | Attack Rate (%) |
|--------------------|-----------------|
| BNT162b2 (Vaccine) | 0.046 |
| Placebo | 0.925 |
| Vaccine Efficacy | 95.03% |

Statistical Methods

Model

The experiment uses a binomial model targeting 170 infections. π , the probability measure of this model, estimates the proportion of infected individuals that had received the vaccination. Because

we are using a Binomial method, we are making assumptions that should be checked. We target 170 infections, so the constant trials assumption is met. Either the infected person received the vaccine or the placebo, so the assumption of only two outcomes is met. The parameter we are estimating, π , is assumed to have some unknown true value, an assumption we cannot strictly verify, but is reasonable to believe. Lastly, we assume that our observations are independent, which we once again cannot strictly verify, but given our large sample size and the number of individuals involved in the survey (*add number here*), this is a reasonable assumption. Thus, taking X to be the number of infections out of 170 where the patient is vaccinated, our experimental model is:

$$X \sim \text{Bin}(170, \pi)$$

However, the parameter that we are more interested in making inference on is ψ , which represents the efficacy of the vaccine. This parameter is calculated as

$$\psi = \frac{1 - 2\pi}{1 - \pi}$$

Cite Result from Introduction It is on this parameter that we are interested in making inference and drawing conclusions. Notably, the FDA (Food and Drug Administration) requires that in order for a drug to be released for wholesale production and use, it must be shown to have an efficacy greater than 0.3. Thus, our hypothesis tests will test how statistically significant the evidence is that the vaccine has an efficacy greater than 0.3.

$$H_0 : \psi = 0.3; \quad H_1 : \psi \neq 0.3$$

We use two statistical methods to make this analysis: the Frequentist maximum likelihood method, and the Bayesian inference method.

Likelihood Inference

We can begin with the Likelihood function. Recall that the design of our experiment is binomial; we observe 170 infections and try to make inference on π , the proportion of infections from those

who have the vaccine. Assuming each observation is independent (which we have justified above), we can write our likelihood function in terms of π :

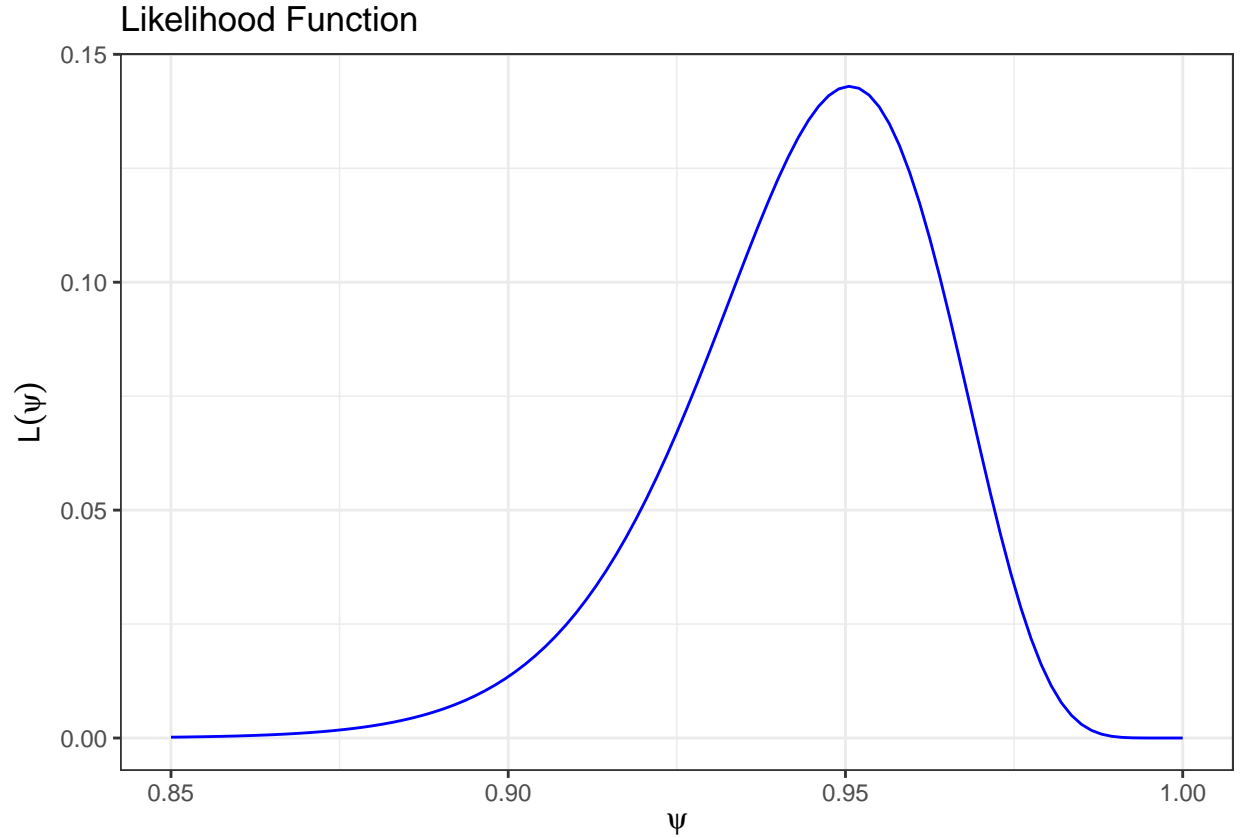
$$L(\pi) = \binom{170}{X} (\pi)^X (1 - \pi)^{170-X}$$

Where x is the observed number of infections where the patient had the vaccine.

We can rewrite this likelihood function in terms of ψ :

$$L^*(\psi) = \binom{n}{X} \left(\frac{\psi - 1}{\psi - 2} \right)^X \left(\frac{-1}{\psi - 2} \right)^{n-X}$$

We know that $0 \leq \psi \leq 1$, so we can plot $L^*(\psi)$ on this interval, using our observed values $n =$



170, $x = 8$:

To find the estimate of ψ that maximizes this likelihood function, we found the log likelihood and took the first derivative, a derivation that can be found in the Appendix. our final formula is given by:

Put Final Formula Here

Given our estimate, we now want to perform a likelihood ratio test on which to test our hypotheses.

We can initially find the likelihood ratio, Λ . It has formula given by:

$$\Lambda = \frac{L(\hat{\psi}_0^{mle})}{L(\psi_0^{null})}$$

#likelihood function:

#Confidence Interval for Psi

Bootstrap Confidence Interval for $\hat{\psi}_{MLE}$

In addition to the large-sample interval, we use a computational bootstrap approach to estimate a 95% confidence interval for $\hat{\psi}_{MLE}$. Suppose we observe

$$\hat{\pi} = \frac{T}{n}$$

from the sample, where T is the observed number of infections out of n participants. Then:

1. **Generate Resamples.** We repeatedly sample T^* from a Binomial($n, \hat{\pi}$) distribution. For each bootstrap draw:

$$\hat{\pi}^* = \frac{T^*}{n}, \quad \hat{\psi}^* = \frac{1 - 2\hat{\pi}^*}{1 - \hat{\pi}^*}.$$

2. **Form the Confidence Interval.** After repeating this process B times, let $\hat{\psi}_1^*, \hat{\psi}_2^*, \dots, \hat{\psi}_B^*$ be the resulting bootstrap estimates of $\hat{\psi}$. We then take the 2.5th and 97.5th percentiles of the $\hat{\psi}^*$ values to form the approximate 95% confidence interval:

$$[\hat{\psi}_{(0.025)}^*, \hat{\psi}_{(0.975)}^*].$$

This bootstrap procedure avoids strong distributional assumptions and can provide more reliable interval estimates for moderate sample sizes.

Bayesian Inference

We can use a Bayesian approach to reanalyze the efficacy of the BNT162b2 vaccine. Let T be the number of infections observed among n vaccine recipients in the trial. We assume

$$T \mid \pi \sim \text{Binomial}(n, \pi)$$

where π is the probability of infection in the vaccine group. As in the original trial design, vaccine efficacy (VE) is defined by Equation (1):

$$\psi = \frac{1 - 2\pi}{1 - \pi}.$$

Under this model, the likelihood for T is proportional to $\pi^T(1 - \pi)^{n-T}$.

Prior Specification. We place a Beta(a, b) prior on π , thereby completing a Beta–Binomial model. As reported in the original trial, the Pfizer prior was Beta(0.700102, 1), while two alternative priors were elicited to achieve lower probabilities of $\psi \geq 0.3$. Specifically, these priors were chosen so that

$$P(\psi \geq 0.3) = 0.05 \quad \text{or} \quad 0.01,$$

respectively. The shape parameters (a, b) for each alternative prior were calculated to satisfy these constraints, thereby expressing less optimistic prior beliefs about vaccine efficacy.

Posterior Distribution. With a Beta prior and a Binomial likelihood, the posterior for π is

$$\pi \mid T = t \sim \text{Beta}(a + t, b + n - t).$$

Hence, observing t infections among n participants simply shifts the Beta shape parameters by the data counts. Posterior summaries (e.g., medians, credible intervals) for π follow from standard Beta distribution. Once obtained, π can be converted to ψ using Equation (1).

Credible Intervals and Hypothesis Testing. We report the 95% credible interval for π , defined

by the 2.5th and 97.5th percentiles of the posterior distribution. We similarly transform these percentiles into credible bounds for ψ . To assess whether $\pi < 0.5$ (equivalently, whether $\psi > 0$), we calculate

$$p\text{-value}_\pi = 1 - F_\pi(0.5),$$

where F_π is the posterior cumulative distribution function of π . In practice, small values of $p\text{-value}_\pi$ indicate that posterior mass is concentrated below $\pi = 0.5$, consistent with high efficacy. Parallel tests can be made on thresholds of ψ , such as verifying if $P(\psi > 0.3)$ exceeds a given target.

Results

Present your findings.

Bootstrap Interval for $\hat{\psi}$

Using a bootstrap procedure with $B = 1000$ resamples from a $\text{Binomial}(n, \hat{\pi})$ distribution where $n = 160$ and $\hat{\pi} = 8/170$ (from table 1), we computed $\hat{\psi}$ in each resample. The 2.5th and 97.5th percentiles of the resulting bootstrap distribution of $\hat{\psi}$ define the approximate 95% confidence interval. Table~3 shows the interval endpoints based on this approach.

Table 3: Bootstrap 95% Confidence Interval for $\hat{\psi}$

| | |
|---------------|--------|
| Lower (2.5%) | 0.9116 |
| Upper (97.5%) | 0.9809 |

Bayesian Results

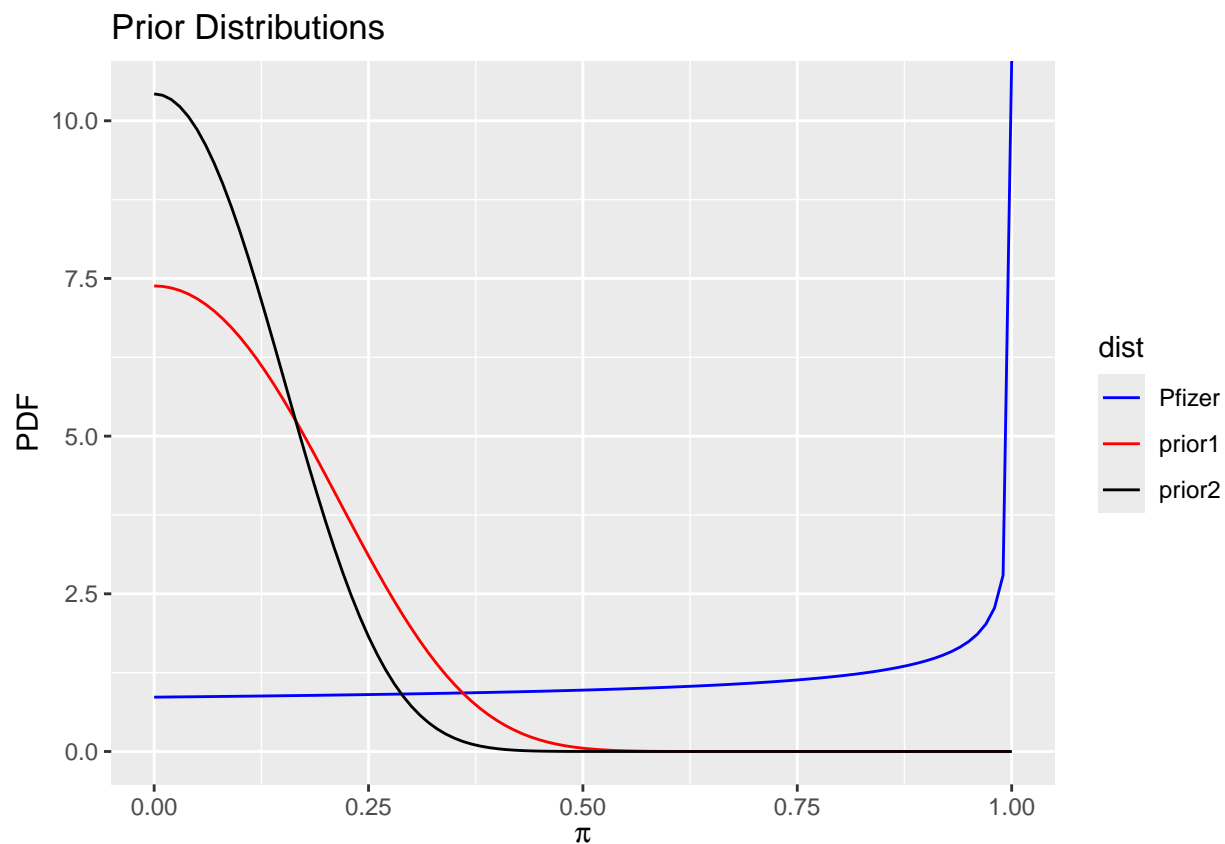


Figure 1: Comparison of the three priors (Pfizer, 0.05, 0.01)

Prior and Posterior Distributions. Figure 1 compares the initial priors for the Pfizer analysis ($\text{Beta}(0.700102, 1)$) and our two alternative priors ($\text{Beta}(43.04, 43.04)$ and $\text{Beta}(85.63, 85.63)$). As shown, the Pfizer prior places more mass toward large values of ψ , whereas the alternative priors are more tightly concentrated around $\psi = 0.25$. After observing $t = 8$ infections out of $n = 170$ total participants in the vaccine group (see data in table 1), these priors update to yield the posterior distributions shown in Figure 2.

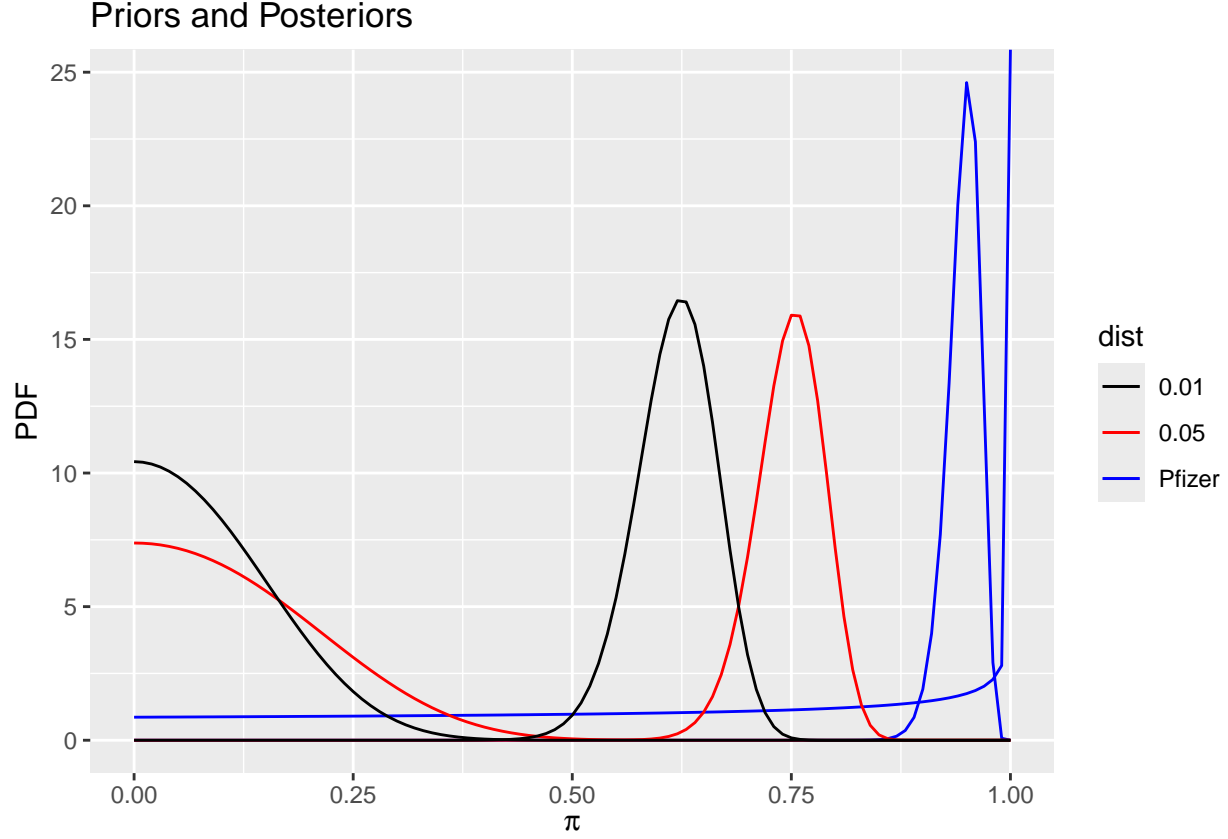


Figure 2: Comparison of the three priors (Pfizer, 0.05, 0.01) and their corresponding posteriors for ψ .

Posterior Summaries (for ψ). Table~4 presents the *approximate* posterior medians and 95% credible intervals for ψ under each prior (calculated by transforming the corresponding values for π).

Table 4: Approximate posterior medians and 95% credible intervals for ψ under each prior.

| Prior | Posterior Median | 95% Credible Interval (Approx) |
|--------|------------------|--------------------------------|
| Pfizer | 0.9486 | [0.9035, 0.9763] |
| 0.05 | 0.7523 | [0.6662, 0.8197] |
| 0.01 | 0.6227 | [0.5232, 0.704] |

Hypothesis Testing (for ψ). For each posterior, we computed the probability $P(\psi \leq 0.3)$, which serves as a Bayesian p-value under the hypothesis

$$H_0 : \psi \leq 0.3 \quad \text{vs.} \quad H_1 : \psi > 0.3.$$

(This threshold can be changed as needed, depending on your target efficacy.) Table~5 shows these values. We define

$$p\text{-value} = P(\psi \leq 0.3),$$

where F_ψ is the posterior cumulative distribution function of ψ . The estimated Bayesian p-values for each prior are:

Table 5: Bayesian p-values for testing $H_0 : \psi \leq 0.3$ under each posterior.

| Prior | Bayesian p-value |
|--------|------------------|
| Pfizer | 0 |
| 0.05 | 0, 0 |
| 0.01 | 0, 0 |

Discussion and Conclusion

Bootstrap

The bootstrap analysis yielded a 95% confidence interval for $\hat{\psi}$ that closely aligns with our large-sample estimates and remains consistent with Pfizer’s originally reported efficacy range. By resampling from a $\text{Binomial}(n, \hat{\pi})$ distribution, the bootstrap circumvents any strong assumptions about the distribution of $\hat{\psi}$ (beyond the binomial sampling model). This robustness adds credibility to our inference, especially in moderately sized samples where asymptotic approximations can be less reliable.

Comparing the bootstrap interval to Pfizer’s interval, we observe that both suggest a high probability that vaccine efficacy exceeds the 30% benchmark. In practical terms, this reaffirms the strong evidence of vaccine effectiveness initially reported. Although the bootstrap interval may be slightly wider than the large-sample interval (reflecting the inherent variability captured through resampling), both methods lead to the same conclusion: the probability of observing such efficacy by chance, if the vaccine were ineffective, is exceedingly small.

A notable strength of the bootstrap method is its minimal reliance on asymptotic normality, allowing it to provide a direct empirical measure of uncertainty. However, a potential drawback is its computational cost—particularly if many replicates are needed or if the underlying data are

extremely large. Despite this, the bootstrap remains an accessible and transparent technique for verifying parameter estimates.

Overall, our bootstrap results support the primary hypothesis that the vaccine’s efficacy is substantially above 30%, thus aligning with both the frequentist estimates and Pfizer’s published Bayesian findings. This convergence of evidence across multiple analytic strategies enhances confidence in the robustness of our conclusions regarding BNT162b2 efficacy.

Bayesian

Our Bayesian analysis provides consistent evidence that the vaccine efficacy ψ remains high. As shown in **Figure 2**, updating the Pfizer prior $\text{Beta}(0.700102, 1)$ and two alternative priors $\text{Beta}(43.04, 43.04)$ and $\text{Beta}(85.63, 85.63)$ with the observed data ($t = 8, n = 170$) yields posterior distributions for ψ that are heavily skewed toward large efficacy values.

Table 3 details the approximate numerical results for ψ . Even though the alternative priors initially place more mass around intermediate efficacy, the data strongly pull the posterior medians upward. Although the alternative priors shift the posterior to moderately lower efficacy, the observed infection count (only 8 out of 170) still imposes a high central estimate for ψ . This suggests that once sufficient data are incorporated, the final posterior distribution is relatively robust to prior differences.

Finally, **Table 4** presents Bayesian p-values for testing a clinically relevant hypothesis such as $H_0 : \psi \leq 0.3$. For all three priors, $P(\psi \leq 0.3)$ is effectively zero, implying near-certain rejection of H_0 . This finding aligns with the frequentist conclusion of strong efficacy, further bolstering overall confidence in the result.

Overall, the Bayesian framework yields direct probability statements about clinically meaningful thresholds (e.g., $\psi > 0.3$). Given sufficiently informative data, the resulting posterior estimates remain consistent across varying prior beliefs. These results corroborate Pfizer’s original analysis, highlighting that BNT162b2 efficacy remains high across a range of reasonable priors.

Bibliography

Brown, B. (2024). *Lecture Title*. Lecture slides, Course Name, University Name.

Doe, J. (2020). Title of the Paper. *Journal Name*, 12(3), 45-67.

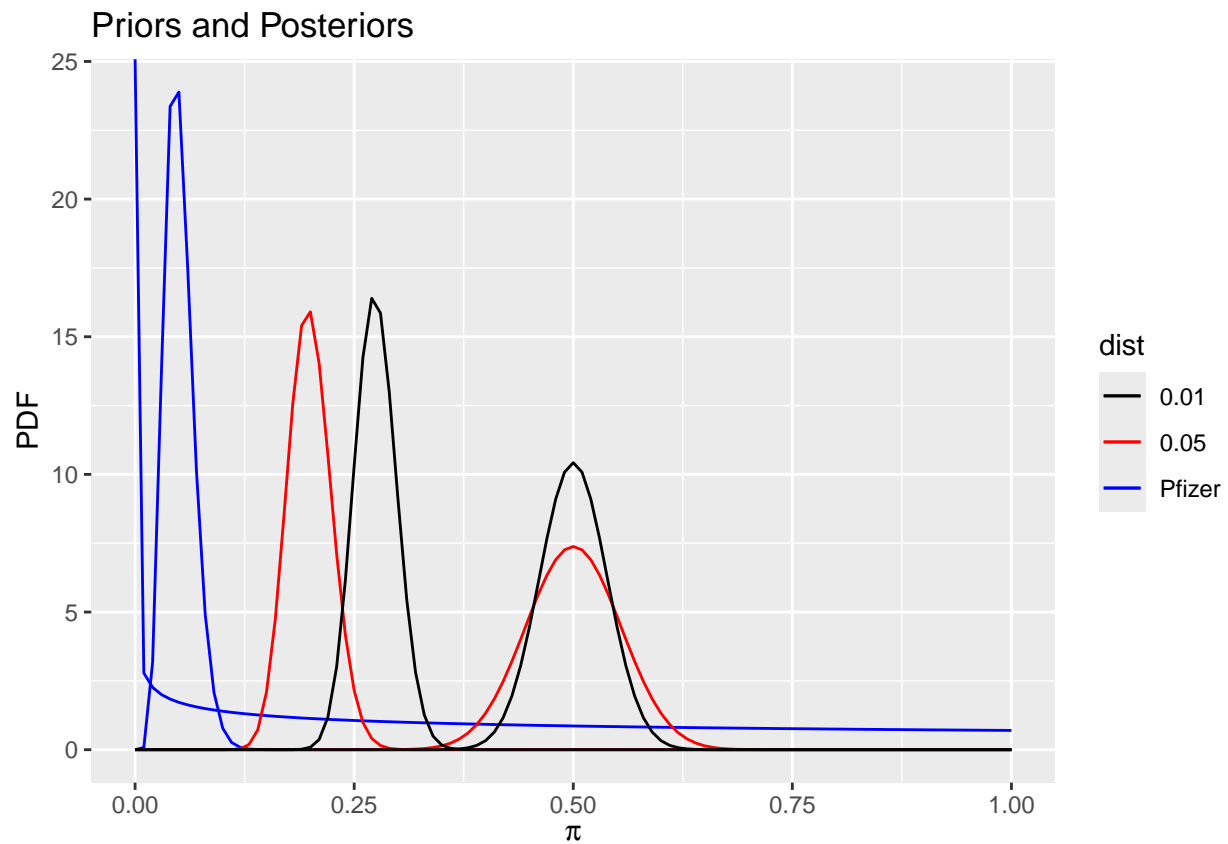
Last, F., & Last, F. (2025). *Book Title*. Publisher.

Smith, A., & Johnson, C. (2023). *Title of the Online Article*. Retrieved from <https://www.example.com>.

Appendix

Code

Plots



Proofs

If applicable, include detailed mathematical derivations or additional theoretical explanations.

Proof of the Maximum Likelihood Estimator for $\hat{\psi}_0^{mle}$:

We can begin with our Likelihood function

$$L(\pi) = \binom{n}{X} (\pi)^X (1 - \pi)^{170-X}$$

Where x is the observed number of infections where the patient had the vaccine. Notably, however, we are not hoping to draw inference on π , but rather on ψ , the efficacy of the vaccine. We can write this parameter in terms of π .

$$\psi = \frac{1 - 2\pi}{1 - \pi}$$

In order to write our Likelihood function in terms of ψ , we can rewrite ψ in terms of π :

$$\pi = \frac{\psi - 1}{\psi - 2}$$

Now we rewrite our likelihood function.

$$L^*(\psi) = \binom{n}{X} \left(\frac{\psi - 1}{\psi - 2} \right)^X \left(\frac{-1}{\psi - 2} \right)^{n-X}$$

Now that we have written our full likelihood function in terms of ψ , we can write the log likelihood function.

$$\ell^*(\psi) = \log \binom{n}{X} + X \log(\psi - 1) - X \log(\psi - 2) - (n - X) \log(\psi - 2) + (n - X) \log(-1)$$

$$\ell^*(\psi) = \log \left(\binom{n}{X} (-1)^{n-X} \right) + X \log(\psi - 1) - n \log(\psi - 2)$$

To maximize this equation, we take its derivative with respect to ψ :

$$\frac{d}{d\psi} \ell^*(\psi) = \frac{X}{\psi - 1} - \frac{n}{\psi - 2}$$

Setting this derivative finds the maximum:

$$0 = \frac{X}{\hat{\psi}_0^{mle} - 1} - \frac{n}{\hat{\psi}_0^{mle} - 2}$$

$$\frac{n}{\hat{\psi}_0^{mle} - 2} = \frac{X}{\hat{\psi}_0^{mle} - 1}$$

$$n\hat{\psi}_0^{mle} - n = X\hat{\psi}_0^{mle} - 2X$$

$$(n - X)\hat{\psi}_0^{mle} = n - 2X$$

$$\hat{\psi}_0^{mle} = \frac{n - 2X}{n - X}$$

Proof of Likelihood Ratio Test Value:

Recall that when performing a hypothesis test for a maximum likelihood estimator, we are interested in calculating W , which is given by

$$W = 2 \ln \left[\frac{L(\hat{\psi}_0^{mle})}{L(\psi_0^{null})} \right]$$

The log of the ratio of two likelihoods is equal to the difference of their respective log-likelihoods:

$$W = 2 \left(\ell(\hat{\psi}_0^{mle}) - \ell(\psi_0^{null}) \right)$$

$$\ell^*(\psi) = \log \left(\binom{n}{X} (-1)^{n-X} \right) + X \log(\psi - 1) - n \log(\psi - 2)$$

$$W = 2 \left(\left(\log \left(\binom{n}{X} (-1)^{n-X} \right) + X \log(\hat{\psi}_0^{mle} - 1) - n \log(\hat{\psi}_0^{mle} - 2) \right) - \left(\log \left(\binom{n}{X} (-1)^{n-X} \right) + X \log(\psi_0^{null} - 1) - n \log(\psi_0^{null} - 2) \right) \right)$$

$$W = 2 \left(\left(X \log(\hat{\psi}_0^{mle} - 1) - n \log(\hat{\psi}_0^{mle} - 2) \right) - \left(X \log(\psi_0^{null} - 1) - n \log(\psi_0^{null} - 2) \right) \right)$$

$$W = 2 \left(X \log \left(\frac{\hat{\psi}_0^{mle} - 1}{\psi_0^{null} - 1} \right) + n \log \left(\frac{\psi_0^{null} - 2}{\hat{\psi}_0^{mle} - 2} \right) \right)$$

$$W = 2X \log \left(\frac{\hat{\psi}_0^{mle} - 1}{\psi_0^{null} - 1} \right) + 2n \log \left(\frac{\psi_0^{null} - 2}{\hat{\psi}_0^{mle} - 2} \right)$$

Derivation of the Large Sample Confidence Interval:

As calculated above, our likelihood function is given by

$$\ell(\psi) = \log \binom{170}{8} + 8 \log(\psi - 1) - 8 \log(\psi - 2) - 162 \log(\psi - 2) + 162 \log(-1)$$

Taking the derivative with respect to ψ ,

$$\frac{d}{d\psi} \ell(\psi) = \frac{8}{\psi - 1} - \frac{170}{\psi - 2}$$

And the second derivative is

$$\frac{d^2}{d\psi^2} \ell(\psi) = -\frac{8}{(\psi - 1)^2} + \frac{170}{(\psi - 2)^2}$$

The Fisher information is the negative expectation of the second derivative of the log-likelihood.

So for $X \sim \text{Binomial}(n, \pi)$,

$$I(\psi) = \mathbb{E} \left[-\frac{d^2 \ell}{d\psi^2} \right] = \frac{\mathbb{E}[X]}{(1 - \psi)^2} - \frac{n}{(2 - \psi)^2}.$$

Substituting $\mathbb{E}[X] = n\pi = n \cdot \frac{1-\psi}{2-\psi}$,

$$I(\psi) = \frac{n(1 - \psi)}{(2 - \psi)(1 - \psi)^2} - \frac{n}{(2 - \psi)^2} = \frac{n}{(2 - \psi)(1 - \psi)} - \frac{n}{(2 - \psi)^2}.$$

Factoring out $\frac{n}{2-\psi}$,

$$I(\psi) = \frac{n}{2 - \psi} \left(\frac{1}{1 - \psi} - \frac{1}{2 - \psi} \right).$$

Combining terms with a common denominator,

$$I(\psi) = \frac{n}{2 - \psi} \cdot \frac{(2 - \psi) - (1 - \psi)}{(1 - \psi)(2 - \psi)} = \frac{n}{(1 - \psi)(2 - \psi)^2}.$$

So we get

$$I(\psi) = \frac{n}{(1-\psi)(2-\psi)^2}.$$

We know that our large sample confidence interval is given by

$$\hat{\psi}^{mle} \pm z_{\alpha/2} \cdot \frac{1}{\sqrt{nI(\hat{\psi})}}$$

We may now plug in our estimates and the given parameters such as sample size and significance level.