Placebo-Controlled Efficacy Analysis of the COVID-19 Vaccine

Spring 2024

Oliver Brown, Josie Czeskleba, Luke VanHouten

Abstract

Keywords

Efficacy, Likelihood, COVID-19, Bootstrap, Estimators

Introduction

The COVID-19 pandemic has impacted millions globally, since the World Health Organization declared it a global pandemic on March 11th, 2020. Sensitive groups are at high risk of COVID-19 and its complications. Moreover, recent studies have shown that long-term COVID-19 or post-COVID-19 symptoms make individuals more susceptible to additional health conditions like diabetes, heart conditions, and neurological conditions (CDC, n.d.). Therefore a safe and effective vaccine is imperative to global health.

In December 2020, Pfizer and BioNTech received FDA Emergency Use Authorization for their two-dose BNT162b2 COVID-19 mRNA vaccine. The approval was based on a placebo-controlled, observer-blinded study involving participants aged 16 and older who were randomly assigned to receive either the placebo or the vaccine. The data provided leverage to prove that the BNT162b2 vaccine efficacy exceeds the FDA-required efficacy of 30%.(Polack et al. 2020).

Given the FDA's minimum efficacy requirement of 30%, this paper focuses on the statistical analysis of the vaccine's efficacy. We will utilize hypothesis testing, bootstrapping, and maximum likelihood estimation to evaluate the data. The thoroughness of this analysis is essential for ensuring reliable and valid findings, which will guide public health decisions based on robust evidence.

Statistical Methods

The data for positive and negative tests between the placebo and vaccine groups is taken from the previous research on the vaccine efficacy (Polack et al. 2020):

We denote the random variable T as the number of vaccinated individuals from the 170 COVID cases.

$$T \sim Binom(n = 170, \pi)$$

Test	Positive	Negative	Total
Vaccine	8	17403	17411
Placebo	162	17349	17511
Total	170	34752	34922

We can define $\pi = \text{P(Vaccine|COVID)} = \frac{\pi_1}{\pi_1+\pi_2}$, given that the sample sizes for the vaccine and placebo groups are approximately equal. Here, π_1 is the proportion of vaccinated individuals who got COVID and π_2 is the proportion of unvaccinated individuals who got COVID. Moreover, we define the vaccine efficacy as $\psi = \frac{1-2\pi}{1-\pi}$ (Senn 2021). We can then formulate the following hypothesis test:

$$H_0 : \psi_0 = 0.3$$
$$H_1 : \psi_0 \geq 0.3$$

We will use a maximum likelihood estimator to conduct a likelihood ratio test as well as bootstrapped confidence intervals to test our hypothesis.

Maximum Likelihood Estimator

The first method we used to analyze the BNT162b2 vaccine efficacy is maximum likelihood estimation. We found the maximum likelihood estimate for our efficacy parameter ψ, which we can call $\hat{\psi}_0^{mle}$. To calculate this estimator, we can first write the likelihood function of π based on the PDF of T:

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

Then we write π in the form $\pi = g(\psi)$, given that $\psi = \frac{1-2\pi}{1-\pi}$. We thus have that $\psi - \psi\pi = 1 - 2\pi$, which becomes $2\pi - \psi\pi = 1 - \psi$, which becomes:

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

Because ψ is the parameter we want to make our estimator,we can rewrite our likelihood function in terms of ψ, still based on the PDF of T for $t_{obs} = 8$:

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

Mathematically, it is difficult to do much with just this function, so we can log-transform it to convert it from a product to a sum. We can calculate the log-likelihood function for ψ as shown here:

$$\ell(\psi) = \ln(L(\psi) = \ln\left(\binom{n}{t}\right) + t\ln(1-\psi) - t\ln(2-\psi) - (n-t)\ln(2-\psi) = \ln\left(\binom{n}{t}\right) + t\ln(1-\psi) - n\ln(2-\psi)$$

Our estimator is defined as where the log-likelihood function is maximized by the parameter ψ, and we can find this by computing $\arg\max_{\psi} \ell(\psi|t)$, or in other words, by taking the derivative of this function and setting it equal to 0. So we can calculate $\ell'(\psi) = 0$ here:

$$\frac{d}{d\psi}\ell(\psi) = \frac{d}{d\psi}\ln\left(\binom{n}{t}\right) + \frac{d}{d\psi}t\ln(1-\psi) - \frac{d}{d\psi}n\ln(2-\psi) = \frac{n}{2-\psi} - \frac{t}{1-\psi} = 0$$

We can then solve $\frac{n}{2-\psi} = \frac{t}{1-\psi}$. We get that $n - n\psi = 2t - t\psi$, which becomes $t\psi - n\psi = 2t - n$. Assuming that this critical point is a maximum, we have that our estimator estimator is $\widehat{\psi}_0^{mle} = \frac{2t-n}{t-n}$. We cannot use the estimator on its own to make much inference about our hypothesis. However, e can use our MLE to perform a likelihood ratio test, defined as $W = 2\left(\ell\left(\hat{\psi}_0^{mle}\right) - \ell\left(\psi_0^{null}\right)\right)$. A small W corresponds with statistical significance for rejecting the null hypothesis, and vice versa. For a sufficiently sized n, which we have as $n = 170$, the distribution $W \sim \chi_1^2$ holds, letting us compute a P-value for our hypothesis. We can compute this value of W here as:

$$W = 2\left(\left(\ln\left(\binom{n}{t}\right) + t\ln\left(1-\widehat{\psi}_0^{mle}\right) - n\ln\left(2-\widehat{\psi}_0^{mle}\right)\right) - \left(\ln\left(\binom{n}{t}\right) + t\ln\left(1-\psi_0^{null}\right) - n\ln\left(2-\psi_0^{null}\right)\right)\right)$$

This becomes:

$$W = 2t\ln\left(1-\widehat{\psi}_0^{mle}\right) - 2n\ln\left(2-\widehat{\psi}_0^{mle}\right) - 2t\ln\left(1-\psi_0^{null}\right) + 2n\ln\left(2-\psi_0^{null}\right)$$

Finally, we can set up an equation to find a 95% confidence interval for our parameter ψ:

$$\widehat{\psi}_0^{\text{mle}} \pm 1.96\sqrt{\frac{-1}{\ell''(\widehat{\psi}_0^{\text{mle}})}} = \widehat{\psi}_0^{\text{mle}} \pm 1.96\sqrt{\frac{-1}{\frac{n}{(2-\widehat{\psi}_0^{\text{mle}})^2} - \frac{t}{(1-\widehat{\psi}_0^{\text{mle}})^2}}}$$

Bootstrap

Results

For our MLE, we can plug in $t_{obs} = 8$ and $n = 170$ to $\widehat{\psi}_0^{mle} = \frac{2t-n}{t-n}$, we get $\widehat{\psi}_0^{mle} = \frac{16-170}{8-170} = \frac{77}{81} = 0.9506$. We can also use the Newton Raphson method to estimate ψ to get the same value, shown in the appendix.

For the likelihood ratio test, we can plug in our values $\widehat{\psi}_0^{mle} = 0.9506$, $\psi_0^{null} = 0.3$, $t = 8$, and $n = 170$ to our W to get:

$$W = 2(8)\ln(1-0.9506) - 2(170)\ln(2-0.9506) - 2(8)\ln(1-0.3) + 2(170)\ln(2-0.3) = 121.6012$$

Given our sample size is large we know $W \sim \chi^2_1$. Therefore, we can calculate P-value for our hypothesis as $P(W \geq 121.6012) = 0$. This P-value is zero, so we can reject the null hypothesis under our likelihood ratio test that the COVID-19 vaccine efficacy is 30%.

Finally, we can use our MLE estimator to calculate a 95% confidence interval for ψ:

$$0.9506 \pm 1.96\sqrt{\frac{-1}{-3124.317}} = 0.9506 \pm 0.0351$$

We are 95% confident that the true efficacy of the BNT162b2 vaccine lies in the interval [91.55%, 98.57%]. This supports our findings in the likelihood ratio test. There is convincing evidence that the true efficacy exceeds the FDA-required efficacy of 30%.

For our bootstrap, the analysis was performed with 10000 iterations to ensure a stable estimate of the vaccine efficacy. The observed efficacy, calculated from the original data, was consistent with the findings of Polack et al. (2020). The histogram of the bootstrap ψ values revealed a right-skewed distribution, indicating that most of the bootstrap samples support a high efficacy of the vaccine. The 95% confidence interval for ψ derived from the bootstrap distribution, ranged from 0.91 to 0.98, closely aligning with the Bayesian credible interval reported in the original study.

Conclusion

In conclusion, we can say with great confidence that the Pfizer-BioNTech BNT162b2 COVID-19 vaccine is effective at preventing positive COVID-19 tests, as the vaccine efficacy far exceeds the FDA minimum efficacy requirement of 30%. With about a 95% efficacy (based on the maximum likelihood estimate and the bootstrapped estimate for efficacy), this COVID-19 vaccine should be recommended for general use for disease prevention. We conducted various statistical tests to back up this claim, such as the likelihood ratio test and bootstrapped efficacy parameters, which both netted extremely small P-values and tight confidence intervals around the estimate for the vaccine efficacy. Our results match the findings of other research on this vaccine's efficacy, where Bayesian estimates were used to show that the vaccine is sufficiently effective (Polack et al. 2020). The next steps for this research would be to utilize more estimators to reach this conclusion, such as the method of moments estimator, and to corroborate the previous study using Bayesian methods. Other next steps would be to use these methods to identify vaccine efficacy for different diseases.

References

CDC. n.d. *Centers for Disease Control and Prevention*. Centers for Disease Control; Prevention. https://cdc.gov/coronavirus/2019-ncov/long-term-effects/index.html#:~:text=As%20a%20result%20of%20these,have%20not%20had%20COVID-19.https://cdc.gov/coronavirus/2019-ncov/long-term-effects/index.html#:~:text=As%2https://cdc.gov/coronavirus/2019-ncov/long-term-effects/index.html#:~:text=As%20a%20result%20of%20these,have%20not%20had%20COVID-19.0a%20result%20of%20these,have%20not%20had%20COVID-19.

Lee, Jack C., and Darius J. Sabavala. 1987. "Bayesian Estimation and Prediction for the Beta-Binomial Model." *Journal of Business & Economic Statistics* 5 (3): 357–67. https://doi.org/10.2307/1391611.

Polack, Fernando P., Stephen J. Thomas, Nicholas Kitchin, Judith Absalon, Alejandra Gurtman, Stephen Lockhart, John L. Perez, et al. 2020. "Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine." *New England Journal of Medicine* 383 (27): 2603–15. https://doi.org/10.1056/NEJMoa2034577.

Senn, Stephen. 2011. "You May Believe You Are a Bayesian but You Are Probably Wrong." *RMM* 2: 48–66. http://www.rmm-journal.de/.

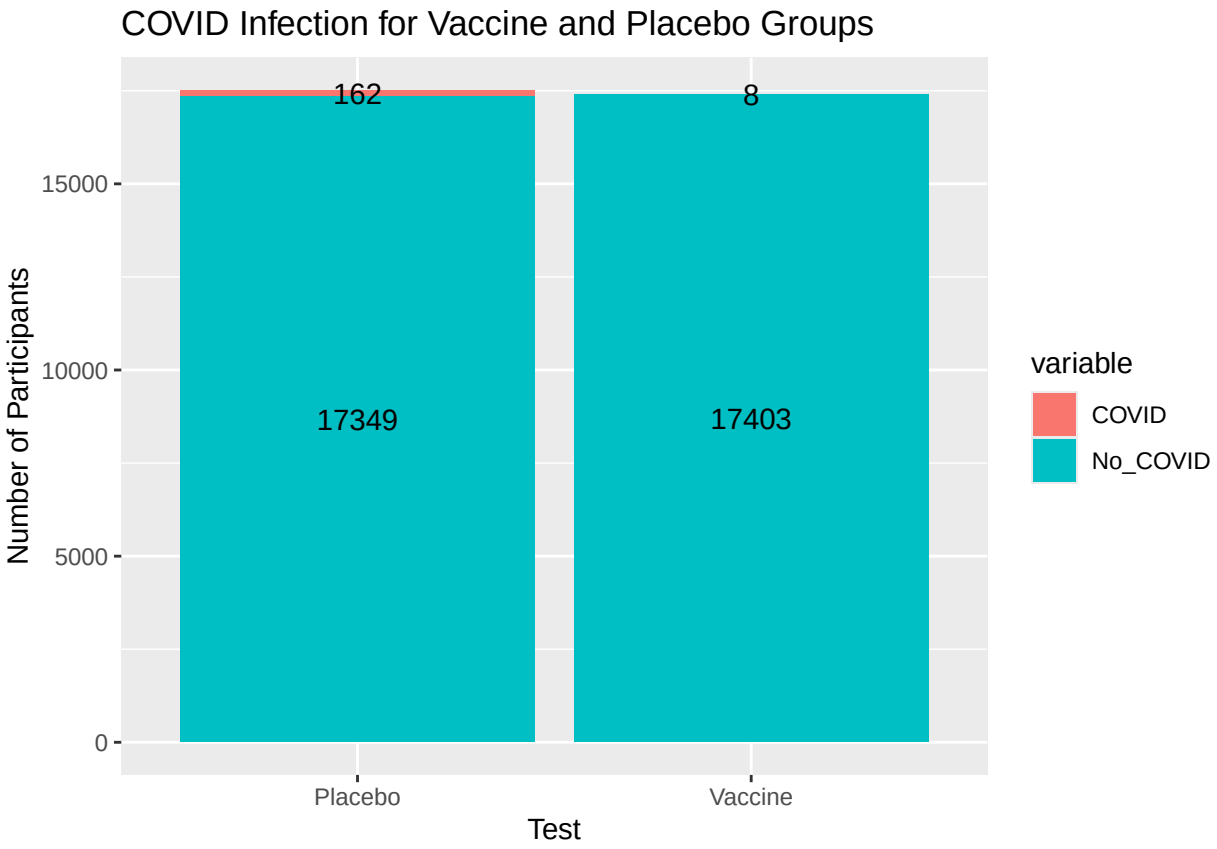
———. 2021. "S. Senn: 'Beta Testing': The Pfizer/BioNTech Statistical Analysis of Their Covid-19 Vaccine Trial (Guest Post)." *Error Statistics Philosophy*. https://errorstatistics.com/2021/01/17/s-senn-beta-testing-the-pfizer-biontech-statistical-analysis-of-their-covid-19-vaccine-trial-guest-post/.

Appendix

Visualizations

Stacked Barplot

```
ggplot(data_melted, aes(x = Test, y = value, fill = variable)) +
  geom_bar(stat = "identity") +
  geom_text(aes(label = value),
            position = position_stack(vjust = 0.5)) +
  labs(x = "Test", y = "Number of Participants",
       title = "COVID Infection for Vaccine and Placebo Groups")
```

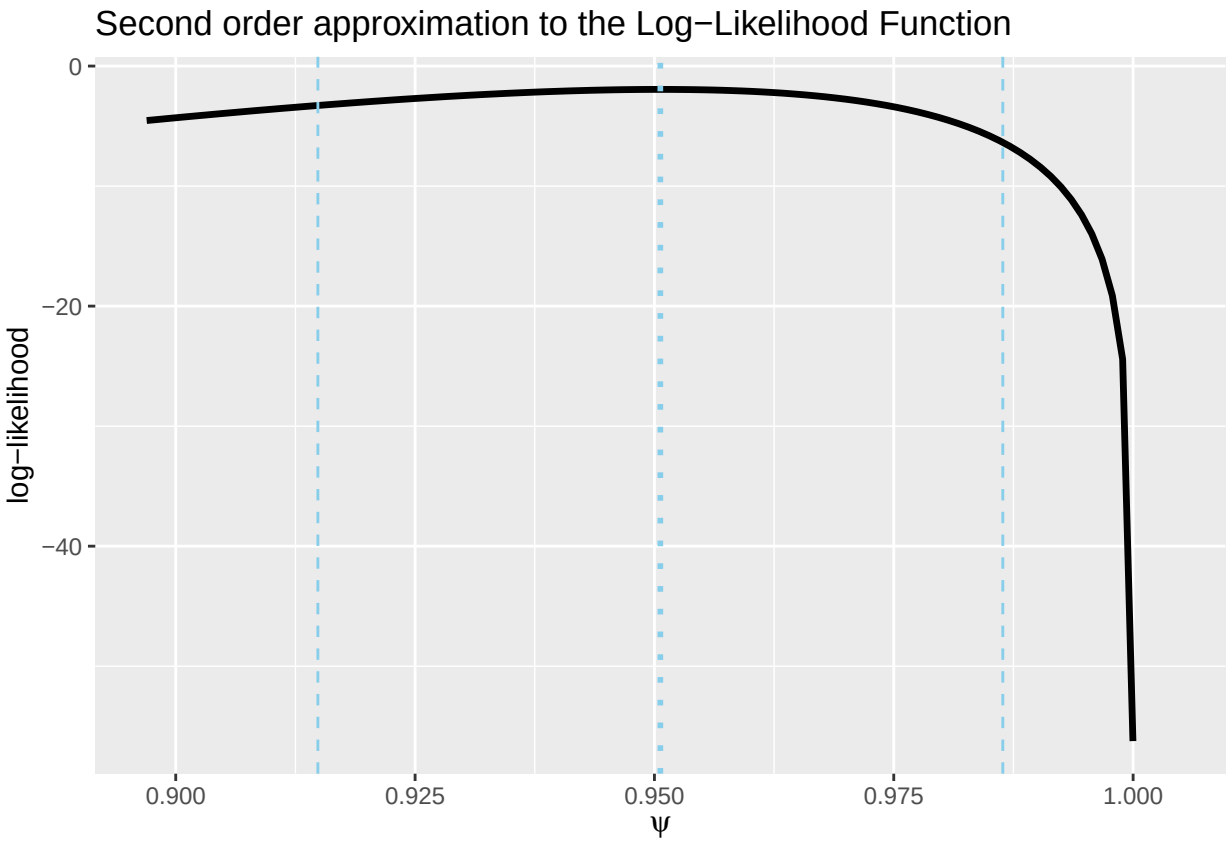


Newton Rhapson MLE Approximation

```
loglik <- function(psi, T, n){
  if (psi > 1 | psi < 0)
    return(NA)
  else
    return(log(choose(n, T)) + (T * log(1 - psi)) - (n * log(2 - psi)))
}

estimation <- maxLik2(loglik = loglik, start = 0.55, method = "NR", tol = 1e-4,
                      T = 8, n = 170)

print(estimation)
```

```
## Maximum Likelihood estimation
## Newton-Raphson maximisation, 6 iterations
## Return code 2: successive function values within tolerance limit (tol)
## Log-Likelihood: -1.944994 (1 free parameter(s))
## Estimate(s): 0.9506174
```

```r
plot(estimation) +
    labs(title = "Second order approximation to the Log-Likelihood Function",
         x = expression(psi))
```



Likelihood Ratio Test Calculation

```r
W = (2 * 8 * log(1 - (77 / 81))) - (2 * 170 * log(2 - (77 / 81))) -
    (2 * 8 * log(1 - 0.3)) + (2 * 170 * log(2 - 0.3))

p_value <- pchisq(q = W, df = 1, lower.tail=F)
```

Here, W is 121.6012 and the corresponding P-value is $2.8222944 \times 10^{-28}$.

MLE Confidence interval calculation

```r
se <- sqrt(-1 / ((170 / (2 - 0.9506)^2) - (8 / (1 - 0.9506)^2)))

upper_lim <- 0.9506 + 1.96 * se
lower_lim <- 0.9506 - 1.96 * se
```

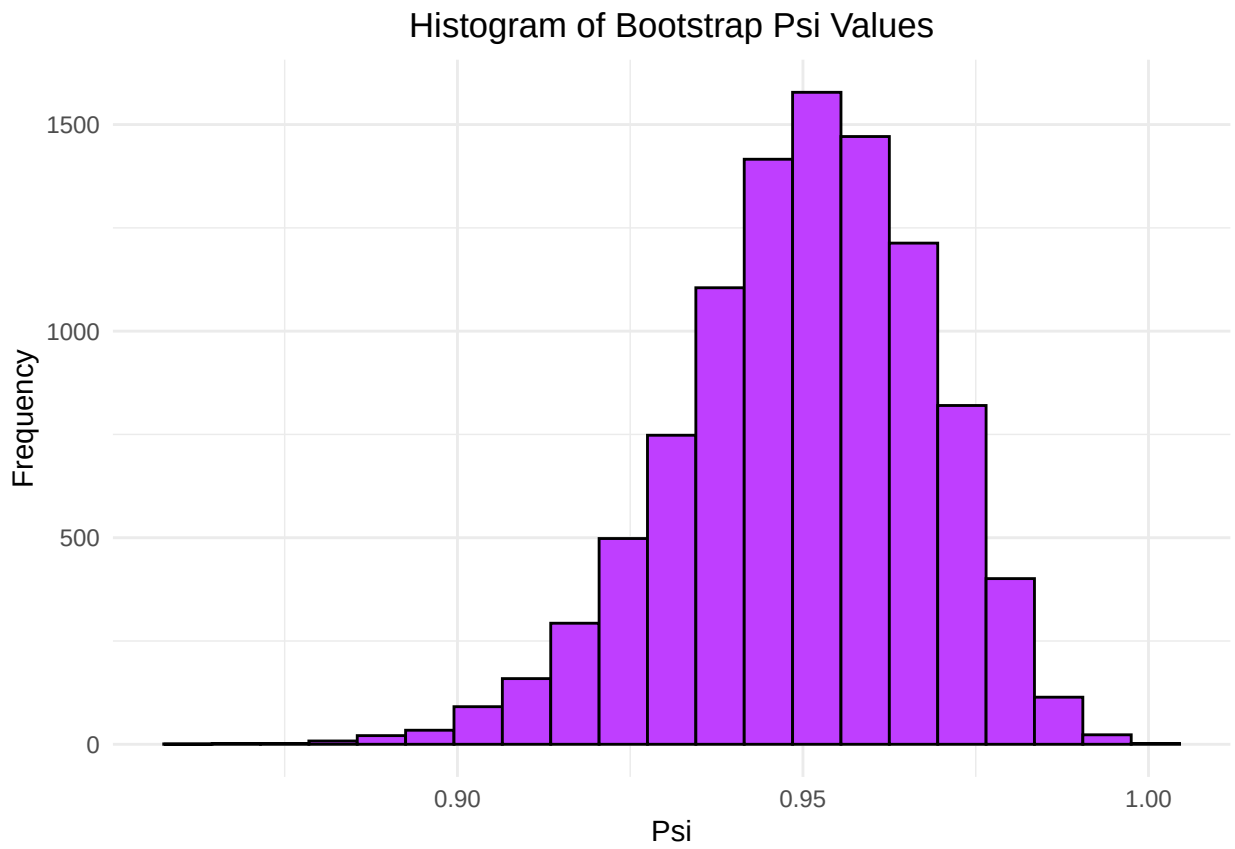
The calculated confidence interval is [0.9155, 0.9856] with a standard error of 0.0179.

Bootstrap

```r
vaccine <- data %>%
  filter(Test == "Vaccine")

placebo <- data %>%
  filter(Test == "Placebo")

n_vaccine <- vaccine$COVID + vaccine$No_COVID
n_placebo <- placebo$COVID + placebo$No_COVID

prop_vaccine <- vaccine$COVID[1] / n_vaccine
prop_placebo <- placebo$COVID[1] / n_placebo

observed_pi <- prop_vaccine/(prop_vaccine + prop_placebo)

observed_psi <- (1 - 2*observed_pi)/(1 - observed_pi)
```

```r
n_bootstrap <- 10000
bootstrap_psis <- numeric(n_bootstrap)
set.seed(123)

for (i in 1:n_bootstrap) {
  vaccine_sample <- sample(c(0, 1), size = n_vaccine, replace = TRUE,
                           prob = c(1 - prop_vaccine, prop_vaccine))
  placebo_sample <- sample(c(0, 1), size = n_placebo, replace = TRUE,
                           prob = c(1 - prop_placebo, prop_placebo))

  prop_vaccine_boot <- mean(vaccine_sample)
  prop_placebo_boot <- mean(placebo_sample)

  bootstrap_pi <- prop_vaccine_boot / (prop_vaccine_boot + prop_placebo_boot)

  bootstrap_psis[i] <- (1 - 2 * bootstrap_pi) / (1 - bootstrap_pi)
}
```

```r
bootstrap_df <- data.frame(psi = bootstrap_psis)

ggplot(bootstrap_df, aes(x = psi)) +
  geom_histogram(binwidth = 0.007, fill = "darkorchid1", color = "black") +
  labs(title = "Histogram of Bootstrap Psi Values", x = "Psi", y = "Frequency") +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5))
```




```r
overall_ci <- quantile(bootstrap_psis, c(0.025, 0.975))

ci_data <- data.frame(
  Iteration = 1:n_bootstrap,
  Lower = numeric(n_bootstrap),
  Upper = numeric(n_bootstrap)
)

for (i in 1:n_bootstrap) {
  sample_psis <- sample(bootstrap_psis, n_bootstrap, replace = TRUE)
  ci_data$Lower[i] <- quantile(sample_psis, 0.025)
  ci_data$Upper[i] <- quantile(sample_psis, 0.975)
}

print(overall_ci)
```

```
##      2.5%     97.5%
## 0.9110567 0.9814894
```

```r
plot_data <- ci_data[seq(1, n_bootstrap, by = 300), ]

ggplot(plot_data, aes(y = Iteration)) +
```

```
geom_vline(xintercept = observed_psi, linetype = "solid", color = "red") +
geom_text(aes(x = observed_psi-0.0125,
              y = max(Iteration) + 300, label = "Observed Psi"),
          color = "red", hjust = -0.05) +
geom_segment(aes(yend = Iteration, x = Lower, xend = Upper), color = "navy") +
geom_vline(xintercept = overall_ci[1], linetype = "dashed", color = "red") +
geom_vline(xintercept = overall_ci[2], linetype = "dashed", color = "red") +
geom_text(aes(x = overall_ci[1] + 0.0129,
              y = max(Iteration) + 300, label = "Lower Bound"),
          color = "red", hjust = 1.1) +
geom_text(aes(x = overall_ci[2] - 0.0129, y = max(Iteration) + 300,
              label = "Upper Bound"), color = "red", hjust = -0.1) +
labs(title = "Confidence Intervals of Bootstrap Samples",
     y = "Iteration",
     x = "Confidence Interval",
     subtitle = "Each line segment represents a 95% confidence interval for a bootstrap sampl
theme_minimal() +
theme(plot.title = element_text(hjust = 0.5),
      plot.subtitle = element_text(hjust = 0.5, size = 10),
      axis.text.y = element_text(hjust = 1))
```