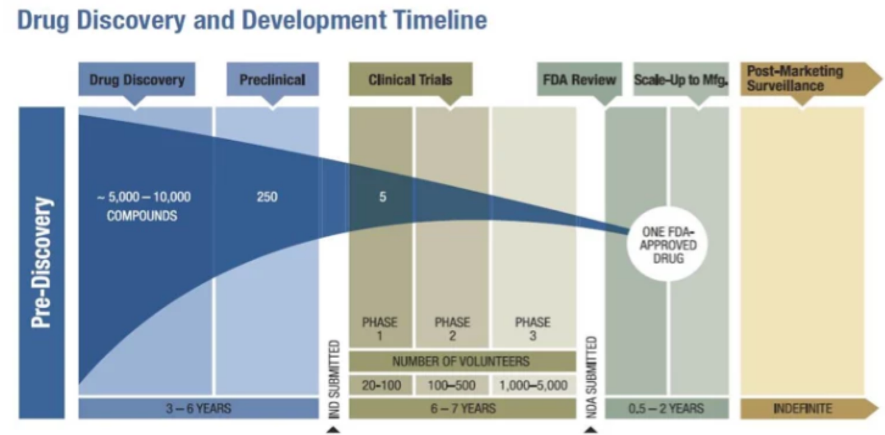
A Bayes Factor Approach to Noninferiority Trials

Amy Ly

Oregon State University

June 12, 2023

The Argument for Non-Inferiority Trials



- Ethical concerns?
- Worth the cost?

Motivation: Biocreep is Bad

Biocreep: An erosion in the level of improvement seen in new drugs after a series of NI trials because a worser therapy is incorrectly declared efficacious

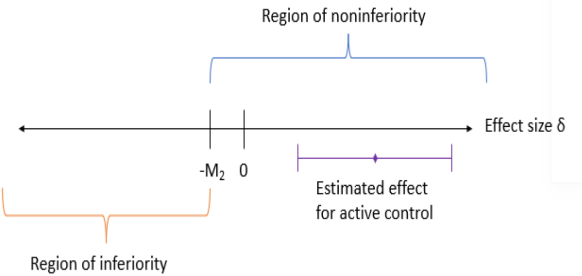
Factors influencing biocreep[1]:

- availability of historical data
- selection of active control
- improvement in standard care
- patient population characteristics

[1] Everson-Stewart & Emerson, 2010

What is the NI Margin?

NI margin: the amount by which the true effect of the new therapy is allowed to be worse than that of the active control



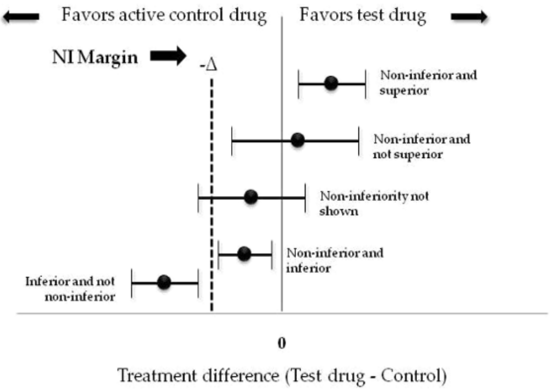
Fixed-Margin Approach: 95-95% Method

- 1st 95% refers to CI of estimated effect of control based on historical studies
- 2nd 95% refers to CI used to test H_0 in NI study

Frequentist Analysis Approach

Possible outcomes shown below[2].

- Perform a t-test.
- Calculate CI and assess where it is relative to the NI margin



[2] Schumi & Wittes (2011)

The Form of a Bayes Factor

- Posterior odds: how much we favor one hypothesis over another **after** observing the data
- Prior odds: how much we favor one hypothesis over another **before** we see the data
- BF: how much the data shifted the relative odds between two hypotheses

H_i refers to the set of assumptions used

$$\frac{p(H_1|y)}{p(H_0|y)} = \frac{p(y\,|H_1)}{p(y\,|\,H_0)} * \frac{p(H_1)}{p(H_0)}$$

Posterior odds

BF_{10} = likelihood ratio

Prior odds

$$BF_{10} = \frac{\int p(y\,|\theta_1, H_1) \; * p(\theta_1|H_1)d\theta_1}{\int p(y\,|\theta_0, H_0) \; * p\,(\theta_0|H_0)d\theta_0}$$

Ratio of prior-weighted averaged likelihoods with continuous θ_i

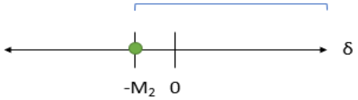
The BF in a NI Trial Setting

Treatment is **Noninferior**

H_0: $\theta_T - \theta_C = -M_2$

H_A: $\theta_T - \theta_C > -M_2$

$$BF_{0+}: \frac{P(y|\theta_T - \theta_C = -M_2)}{P(y|\theta_T - \theta_C > -M_2)}$$

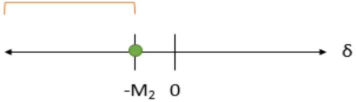


Treatment is **Inferior**

H_0: $\theta_T - \theta_C = -M_2$

H_A: $\theta_T - \theta_C < -M_2$

$$BF_{-0}: \frac{P(y|\theta_T - \theta_C < -M_2)}{P(y|\theta_T - \theta_C = -M_2)}$$

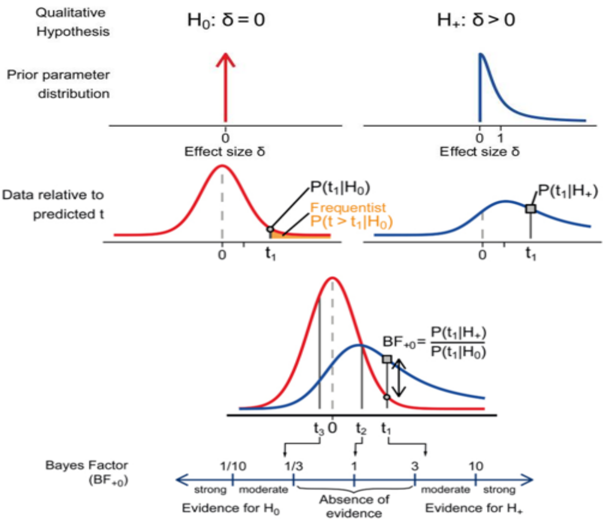


The larger the BF, the more evidence for supporting that the treatment is noninferior

$$BF_{+-} = \frac{BF_{0-}}{BF_{0+}} = \frac{\dfrac{P(y|\theta_T - \theta_C = -M_2)}{P(y|\theta_T - \theta_C < -M_2)}}{\dfrac{P(y|\theta_T - \theta_C = -M_2)}{P(y|\theta_T - \theta_C > -M_2)}} = \frac{P(y|\theta_T - \theta_C > -M_2)}{P(y|\theta_T - \theta_C < -M_2)}$$

The Informed T-Test

Source[3]



[3] Keysers et al, 2020

The Equation Used in the Code

For more details, check out *Informed Bayesian t-tests.* [4]

$$\mathrm{BF}_{10} = \frac{\int_0^\infty \left(1+\frac{n_y n_x}{n_y+n_x} g\right)^{-\frac{1}{2}} \exp\left\{-\frac{\mu_\delta^2}{2\left(\frac{n_y+n_x}{n_y n_x}+g\right)}\right\}\left[1+\frac{t^2}{(n_y+n_x-2)\left(1+g\frac{n_y n_x}{n_y+n_x}\right)}\right]^{-\frac{n_y+n_x-1}{2}} [A+B]\, p(g) \mathrm{d}g}{\Gamma\left(\frac{n_y+n_x-1}{2}\right)\left[1+\frac{t^2}{n_y+n_x-2}\right]^{-\frac{n_y+n_x-1}{2}}}$$

$p(g)$ corresponds to the density of an inverse-gamma distribution of the form

$$p(g) = \frac{\left(r^2\frac{\kappa}{2}\right)^{\frac{\kappa}{2}}}{\Gamma(\frac{\kappa}{2})} g^{-\frac{\kappa}{2}-1} \exp\left(-\frac{r^2\kappa}{2g}\right)$$

$$A = \Gamma\left(\frac{n-1}{2}\right) \; F_1\left(\frac{n-1}{2}; \frac{1}{2}; \frac{\mu_\delta^2 t^2}{2\left(\frac{n_y+n_x}{n_y n_x}+g\right)\left[(n_y+n_x-2)\left(1+\frac{n_y n_x}{n_y+n_x}g\right)+t^2\right]}\right)$$

$$B = \frac{\mu_\delta t}{\sqrt{\frac{1}{2}\left(\frac{n_y+n_x}{n_y n_x}+g\right)\left[(n_y+n_x-2)\left(1+\frac{n_y n_x}{n_y+n_x}g\right)+t^2\right]}} \Gamma\left(\frac{n}{2}\right) \times F_1\left(\frac{n}{2}; \frac{3}{2}; \frac{\mu_\delta^2 t^2}{2\left(\frac{n_y+n_x}{n_y n_x}+g\right)\left[(n_y+n_x-2)\left(1+\frac{n_y n_x}{n_y+n_x}g\right)+t^2\right]}\right)$$

F_1 corresponds to the confluent hypergeometric function.

[4] Gronau et al, 2020

1. Background
2. NI Trial Analysis
3. **Case Study Reanalysis**
4. Robustness Check
5. Conclusion

Context

Test the noninferiority of the beta-lactam strategy to the beta-lactam-macrolide and fluoroquinolone strategies in treating clinically suspected community-acquired pneumonia (CAP), set in the Netherlands [5].

- Primary measure: 90-day mortality
- intention-to-treat analysis
- NI margin of 3%
- result based on 90% CI

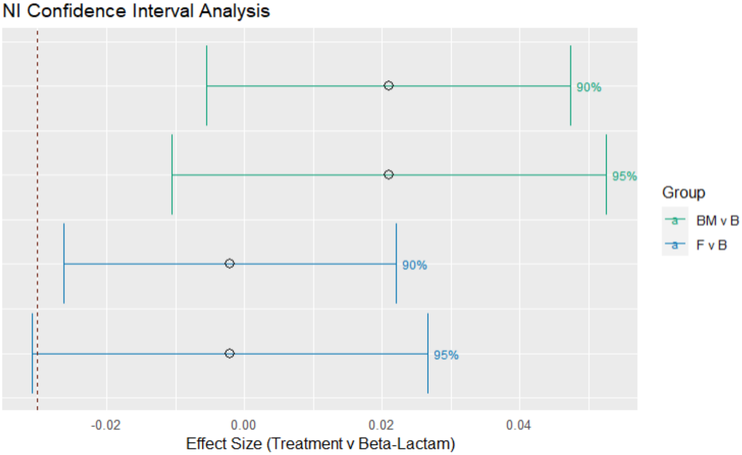
Treatment	Mortality Count	Sample Size	Mortality Rates (%)	Adherence rates (%)
Beta-lactam	59	656	9.0	93
Beta-lactam-macrolide	82	739	11.1	88
fluoroquinolone	78	888	8.8	92.7

Group comparisons are:

- beta-lactam-macrolides (BM) vs beta-lactam (B)
- fluoroquinolone (F) vs beta-lactam (B)

[5] Postma et al (2015)

Frequentist Analysis Result



90% CI's don't include NI margin => beta-lactam strategy is noninferior to the other alternative treatments.

Setting Up the BF Test

It should be more difficult to assess whether there is a statistically significant difference between beta-lactam and fluoroquinolone.

Beta-lactam is **Inferior**

$$H_0: p_F - p_B = -M_2$$

$$H_A: p_F - p_B < -M_2$$

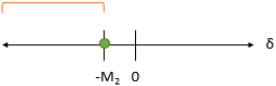
Beta-lactam is **Noninferior**

$$H_0: p_F - p_B = -M_2$$

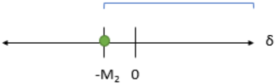
$$H_A: p_F - p_B > -M_2$$

If $\delta = p_F - p_B$ is negative, then the Beta-lactam treatment observed higher mortality counts and is inferior

$$BF_{-0}: \frac{P(y|\theta_F - \theta_B < -M_2)}{P(y|\theta_F - \theta_B = -M_2)}$$



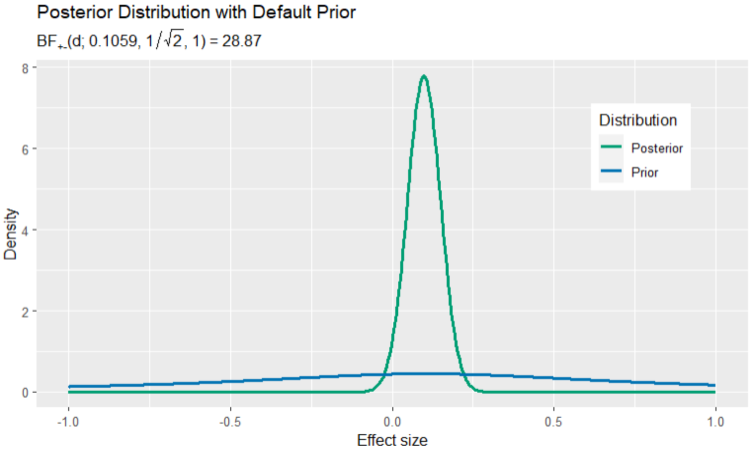
$$BF_{0+}: \frac{P(y|\theta_F - \theta_B = -M_2)}{P(y|\theta_F - \theta_B > -M_2)}$$



BF_{+-} is the ratio of marginal likelihood under the hypothesis that beta-lactam is noninferior to marginal likelihood under the hypothesis that beta-lactam is inferior.

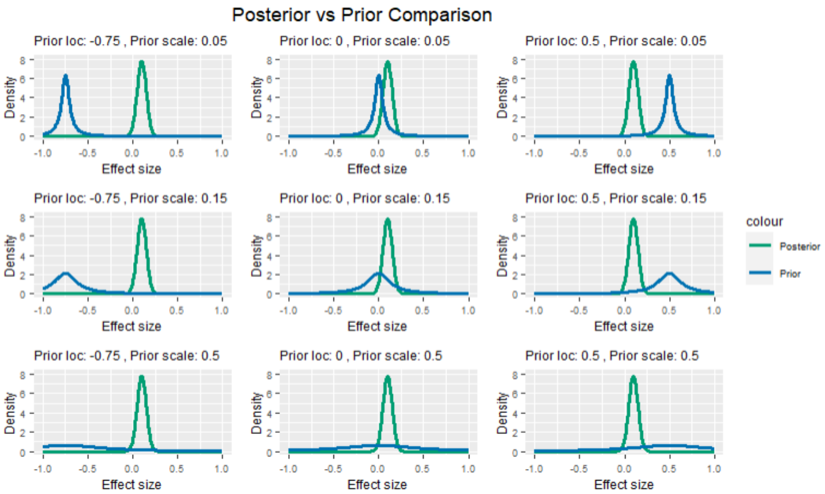
$$BF_{+-} = \frac{BF_{0-}}{BF_{0+}} = \frac{\dfrac{P(y|\theta_F - \theta_B = -M_2)}{P(y|\theta_F - \theta_B < -M_2)}}{\dfrac{P(y|\theta_F - \theta_B = -M_2)}{P(y|\theta_F - \theta_B > -M_2)}} = \frac{P(y|\theta_F - \theta_B > -M_2)}{P(y|\theta_F - \theta_B < -M_2)}$$

BF Analysis Result

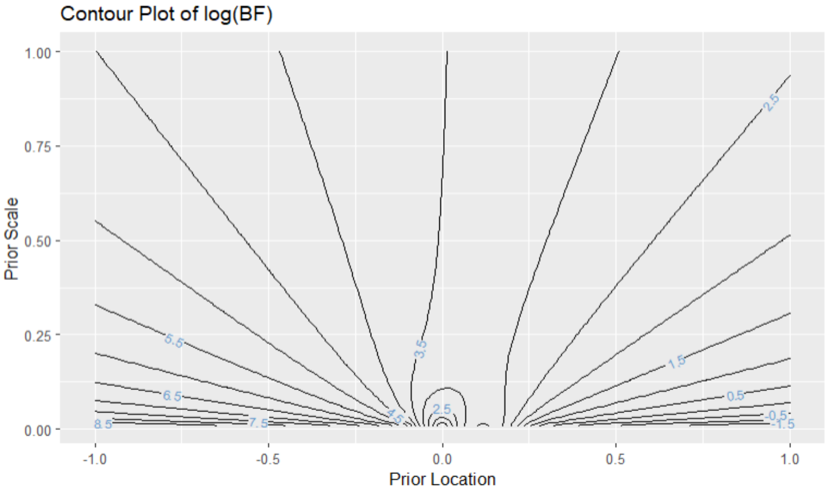


BF_{+-} indicates the data is about 28 times more likely under the noninferiority hypothesis.

How the Prior Affects the Posterior Distribution

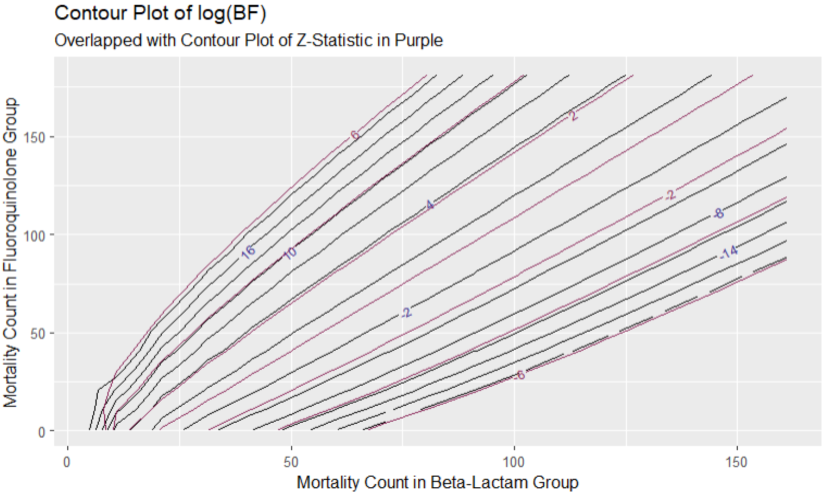


Does Choice of Prior Matter?

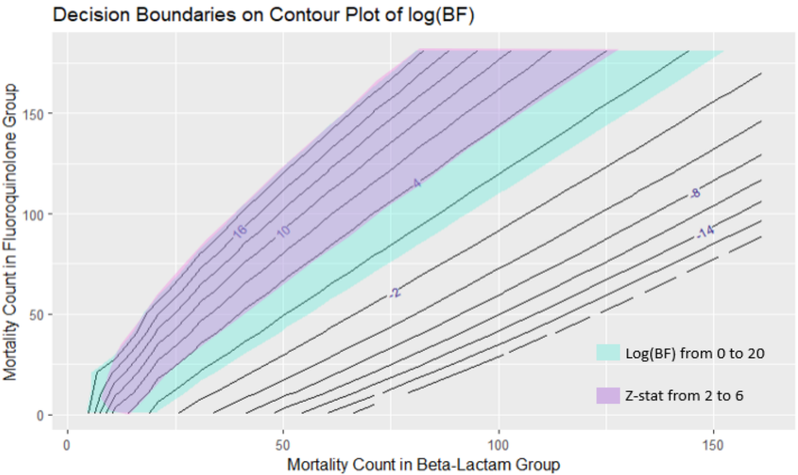


True Effect Sizes Matters

Overlap with the Frequentist Decision



Decision Boundaries for Noninferiority



Lessons Learned

To summarize:

- Qualitative conclusions between the Frequentist and BF testing methods were similar for this case study.
- Could be useful for exploratory studies.

Follow up:

- How would the BF method impact the rate of biocreep?

Thank you!

- Sarah, for guiding me
- Family and friends, for supporting me
- Stats department, for teaching me
- Colleagues, for their flexibility

References

1. Everson-Stewart, S., & Emerson, S. (2010). Bio-creep in non-inferiority clinical trials. Stat Med, doi: 10.1002/sim.4053. PMID: 20809482.
2. Schumi, J., & Wittes, J. T. (2011). Through the looking glass: understanding non-inferiority. Trials, https://doi.org/10.1186/1745-6215-12-106.
3. Keysers C, Gazzola V, Wagenmakers EJ. Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. Nat Neurosci. 2020 Jul;23(7):788-799. doi: 10.1038/s41593-020-0660-4. Epub 2020 Jun 29. Erratum in: Nat Neurosci. 2020 Nov;23(11):1453. PMID: 32601411; PMCID: PMC7610527.
4. Gronau, Q. F., Ly, A., & Wagenmakers, E.-J. (2020). Informed Bayesian t-Tests. The American Statistician, DOI:10.1080/00031305.2018.1562983.
5. Postma, D., van Werkhoven, C., van Elden, L., Thijsen, S., Hoepelman, A., Kluytmans, J., . . . Bonten, M. (2015). CAP-START Study Group. Antibiotic treatment strategies for community-acquired pneumonia in adults. N Engl J Med, doi: 10.1056/NEJMoa1406330. PMID: 25830421.