

# Comparing a t-Distribution-Based Bayesian Approach with the Two One-Sided t-Test (TOST) for Bioequivalence Studies in Real-World Datasets

Assigned  
Poster  
Number  
M1430-10-65

Jing Wang<sup>1</sup>, Gregory Campbell<sup>2,3</sup>, Meng Hu<sup>1</sup>, Kairui Feng<sup>1</sup>, Somesh Chattopadhyay<sup>4</sup>, Liang Zhao<sup>1</sup>, Carl Peck<sup>2,5</sup>

1. Division of Quantitative Methods and Modeling, Office of Research and Standards, Office of Generic Drugs, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD, USA

2. NDA Partners, a ProPharma Group Company, Washington, St Paul, MN, USA

3. GCStat Consulting LLC, Silver Spring, MD, USA

4. Division of Biometrics VIII, Office of Biostatistics, Office of Translational Sciences, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD, USA

5. Department of Bioengineering and Therapeutic Sciences, University of California at San Francisco, San Francisco, CA, USA

CONTACT INFORMATION: Carl Peck, ccpeck@icloud.com



## PURPOSE

- The two one-sided t-test (TOST)<sup>1</sup> has been used as a regulatory standard to evaluate average bioequivalence (BE).
- Peck, et al<sup>2</sup> compared the performance of TOST with a Bayesian procedure (BEST package<sup>3</sup>, which assumes the t-distribution, referred to as Bayes<sub>T</sub>) using simulated datasets and found that TOST and the Bayes<sub>T</sub> procedures perform comparably on Test/Reference (T/R) ratio data following the lognormal distribution, while Bayes<sub>T</sub> is more accurate, requiring fewer subjects when data follow the normal distribution or contain extreme values<sup>2</sup>.
- In this study, real-world BE study datasets were used to compare the performance of Bayes<sub>T</sub> with TOST in different BE scenarios, e.g., in term of normality of log(T/R) data and presence of extreme values.

## METHOD(S)

- 2,341 Real-world BE pharmacokinetic datasets (811 CMAX, 754 AUCT, and 776 AUCI) were assembled from two-treatment, two-period (2x2), crossover BE studies in abbreviated new drug applications (ANDAs) submitted to the US. Food and Drug Administration between 2020 and 2021.
- For TOST, the equivalence test was conducted by following the FDA's guidance<sup>4</sup>.
- Bayes<sub>T</sub> employs noninformative priors for t-distribution parameters and the Markov Chain Monte Carlo (MCMC) procedure to compute posterior distributions for these parameters.
- BE analysis by TOST and Bayes<sub>T</sub> was compared on the collected datasets, where the BE acceptance range of 0.8000-1.2500 was used for both TOST and Bayes<sub>T</sub> assessments.
- Shapiro-Wilk normality tests and extreme value assessments (3 interquartile range as criteria) were also applied to log(T/R) of the datasets to assess the characteristics of their distributions.

## RESULT(S)

- All datasets were split into four categories based on BE conclusions from TOST and Bayes<sub>T</sub> as follows: 1) TOST BE, Bayes<sub>T</sub> BE; 2) TOST not-BE(NBE), Bayes<sub>T</sub> NBE; 3) TOST BE, Bayes<sub>T</sub> NBE; and 4) TOST NBE, Bayes<sub>T</sub> BE.
- As shown in Table 1, BE conclusions from TOST and Bayes<sub>T</sub> agreed in 2,314 (98.85%, i.e., Categories 1 and 2) of the datasets and disagreed on BE in 27 datasets. For 26.8% of the total datasets, their log(T/R) rejected normality. Extreme values were identified in 17.5% of the total datasets.
- Table 2 shows details of the normality and extreme values results in each category.

Category 1 (Table 2a) - 27% of the datasets with log(T/R) rejected normality (605 datasets), and only 60% of them (369 datasets) included identified extreme values.

Category 2 (Table 2b) - for most of the 41 datasets, the log(T/R) failed to reject normality and no extreme values were identified.

Category 3 (Table 2c) - all log(T/R) failed to reject normality and no extreme values were identified. The estimated confidence/credible bounds for both approaches were close to either of BE boundaries (0.8000 or 1.2500).

Category 4 (Table 2d) - all log(T/R) rejected normality and 17 of 20 datasets included identified extreme values.

- In addition, we conducted analysis to compare the width of the Bayes<sub>T</sub> credible intervals (CrI) to the confidence intervals (CI) of TOST for all datasets. The results (Figure 1) show that for Category 4, the CrIs widths of Bayes<sub>T</sub> were much smaller than the CI widths of TOST. The width differences of the other three categories were all very close to 0.

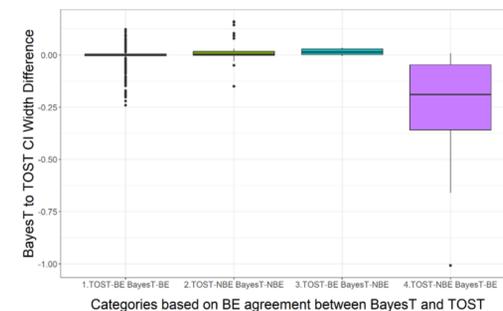


Figure 1. Width difference between the credible/confidence intervals (CrI/CI) of Bayes<sub>T</sub> and TOST for the categorized datasets based on the agreement of BE conclusions of the two methods.

## CONCLUSION(S)

- In 2,341 ANDA datasets, Bayes<sub>T</sub> and TOST agree on BE in 98.85% cases.
- When TOST rejected BE and Bayes<sub>T</sub> concluded BE, datasets are more likely to reject normality and/or include extreme values in log(T/R), which violates the underlying assumptions of TOST and may compromise the performance, while Bayes<sub>T</sub> seems to be more robust in this scenario. The results are consistent with Peck, et al<sup>2</sup>, indicating that the t-distribution based Bayesian approach can be a useful complementary tool for datasets rejecting normality and/or include extreme values in log-transformed T/R ratios.
- The cases with inconsistent BE conclusions from TOST and Bayes<sub>T</sub> warrant further analysis.

Table 1. Overview summary of the datasets used in this study (N=2341), in terms of normality, extreme values, and BE agreement between TOST and Bayes<sub>T</sub>

Tests	Results Categories	Number of Datasets
Normality	Log(T/R) rejected normality	628 (26.8%)
	Log(T/R) failed to reject normality	1713 (73.2%)
Extreme Values	No identified extreme values	1932 (82.5%)
	Extreme values identified	409 (17.5%)
Agreement of BE Conclusions between TOST and Bayes <sub>T</sub>	Category 1 - TOST BE vs. Bayes <sub>T</sub> BE	2273 (97.1%)
	Category 2 - TOST NBE vs. Bayes <sub>T</sub> NBE	41 (1.8%)
	Category 3 - TOST BE vs. Bayes <sub>T</sub> NBE	7 (0.3%)
	Category 4 - TOST NBE vs. Bayes <sub>T</sub> BE	20 (0.9%)

Table 2. Details of the normality and extreme value results in datasets categorized based on the agreement of BE conclusions between TOST and Bayes<sub>T</sub>: (a) Category 1 TOST BE vs. Bayes<sub>T</sub> BE; (b) Category 2 TOST NBE vs. Bayes<sub>T</sub> NBE; (c) Category 3 TOST BE vs. Bayes<sub>T</sub> NBE; (d) Category 4 TOST NBE vs. Bayes<sub>T</sub> BE.

(a)				(b)			
Category 1	No identified extreme values	Extreme values identified	Total	Category 2	No identified extreme values	Extreme values identified	Total
Log(T/R) failed to reject normality	1648	20	1668	Log(T/R) failed to reject normality	37	1	38
Log(T/R) rejected normality	236	369	605	Log(T/R) rejected normality	1	2	3
Total	1884	389	2273	Total	38	3	41

(c)				(d)			
Category 3	No identified extreme values	Extreme values identified	Total	Category 4	No identified extreme values	Extreme values identified	Total
Log(T/R) failed to reject normality	7	0	7	Log(T/R) failed to reject normality	0	0	0
Log(T/R) rejected normality	0	0	0	Log(T/R) rejected normality	3	17	20
Total	7	0	7	Total	3	17	20

## REFERENCES

- Schuurmann DJ. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J Pharmacokinet Biopharm. 15:657-680.
- Peck, C., Campbell, G., Yoo, I., Feng, K., Hu, M., & Zhao, L. (2022). Comparing a Bayesian Approach (BEST) with the Two One-Sided t-Tests for Bioequivalence Studies. The AAPS Journal, 24(5), 97.
- Kruschke JK. (2013). Bayesian estimation supersedes the t test. J Exp Psychol Gen. 142:573-603.
- <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/statistical-approaches-establishing-bioequivalence-0>

## ACKNOWLEDGEMENTS

We want to give our special thanks and memoriam to Dr Gregory Campbell, who contributed substantially to this project. Dr. Campbell passed away on May 25, 2023, and he will be deeply missed.

