A Bayesian Approach For Evaluating Equivalence Over Multiple Groups, and Comparison With Frequentist TOST

Jos Weusten, Ji Young Kim, Katherine Giacoletti, Jorge Vázquez, Plinio De los Santos

Center for Mathematical Sciences, MMD, Merck & Co., Inc., Rahway, NJ, USA

October 2022, BAYES 2022

14-Oct-2022

Authors

Jos Weusten

Ji Young Kim

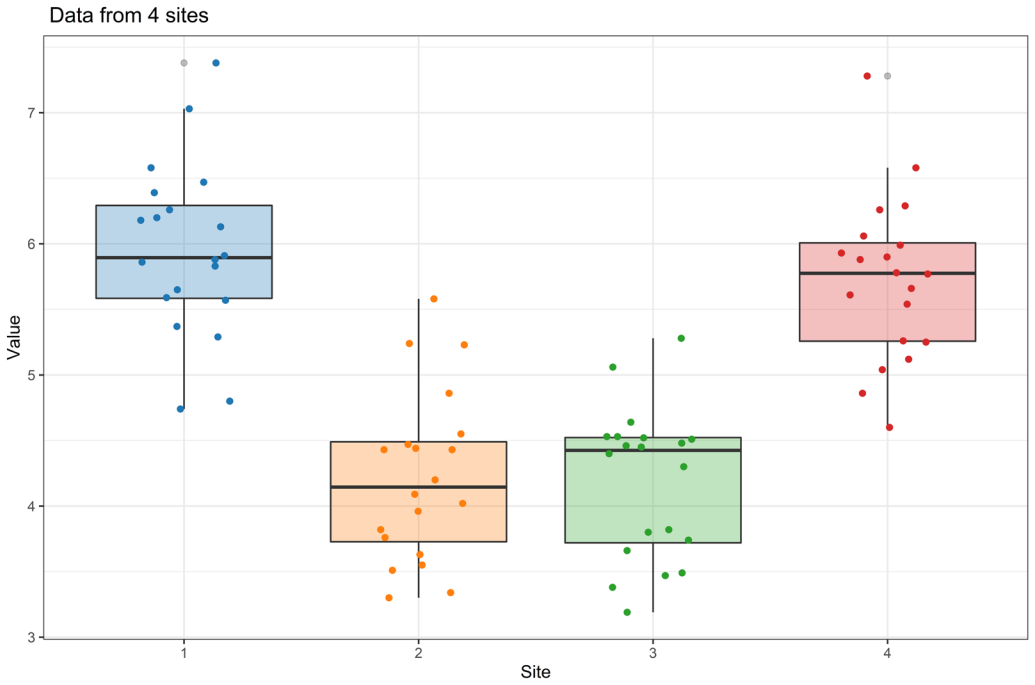
Katherine Giacoletti

Jorge Vázquez

Plinio De los Santos

Center for Mathematical Sciences (CMS) supports all aspects of MMD's business in both Commercialization and Supply: Pharmaceuticals, Vaccines and Biologics, Consumer Healthcare and Animal Health, through application of statistical and mathematical methods.

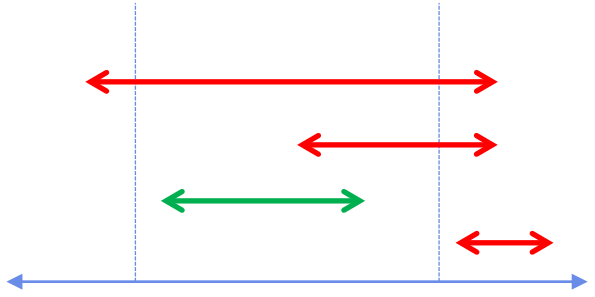
Equivalence across multiple groups



Manufacturing and testing of pharmaceutical products frequently occurs in **multiple facilities**.

It is of interest **to demonstrate equivalence** among **two or more** testing/manufacturing facilities.

2-group equivalence test is well established: **TOST** (Two One-Sided T-tests)



Equivalence Acceptance Criteria (EAC)

MERCK

Equivalence Test for means across many sites

Frequentist's Framework

H_0: $\max_{i,j}|\mu_i - \mu_j| \geq \Delta$ vs H_1: $\max_{i,j}|\mu_i - \mu_j| < \Delta$

Perform TOST on all pairs.

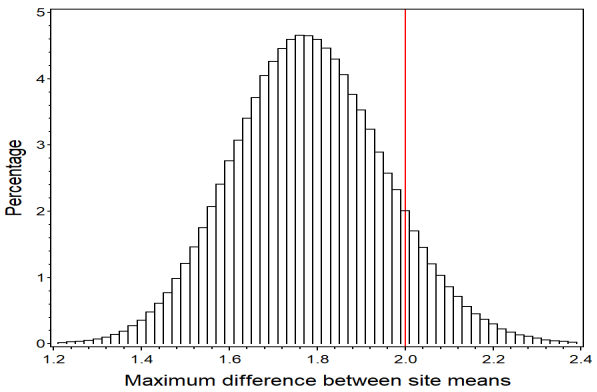
Multiple comparison adjustment?

Bayesian Framework

Hypothesis: $\max_{i,j}|\mu_i - \mu_j| < \Delta$

Compute % of the posterior draws of the max absolute difference within $-\Delta$ to $+\Delta$.

	mu_1	mu_2	mu_3	mu_4	sigma2
1	5.965464	4.162494	4.224249	5.783101	0.4232861
2	5.976479	4.018961	4.128656	6.133597	0.4766082
3	5.692333	4.275053	4.134850	5.836720	0.4293945
4	5.841607	3.994508	4.358272	5.641781	0.3735788
5	5.930221	3.877991	4.182934	5.645526	0.3455996
6	5.840453	4.447853	4.462293	5.738037	0.5647765
7	5.974761	4.135773	4.222009	5.821567	0.3451791
8	5.487635	4.372068	4.394754	5.712490	0.4828538
9	5.933843	4.265793	4.251791	5.350941	0.4707030
10	5.826172	4.312465	4.511344	5.783250	0.3944685
11	5.722077	4.209600	4.176960	5.878313	0.5499560



We will illustrate and compare both methods for the scenario that the variables from all four sites are normally distributed and with a common variance.

MERCK

Example data: Frequentist analysis (EAC for the mean difference = 2)

Analyses on differences between sites			
Comparison site pairs	Difference between means	90% confidence intervals	
		No correction	Tukey correction
1 vs 2	1.66	[1.32 ; 2.00]	[1.18 ; 2.14]
1 vs 3	1.57	[1.23 ; 1.91]	[1.09 ; 2.05]
1 vs 4	0.01	[-0.33 ; 0.35]	[-0.47 ; 0.49]
2 vs 3	-0.09	[-0.43 ; 0.25]	[-0.57 ; 0.39]
2 vs 4	-1.65	[-1.99 ; -1.31]	[-2.13 ; -1.17]
3 vs 4	-1.56	[-1.90 ; -1.22]	[-2.04 ; -1.08]

MERCK

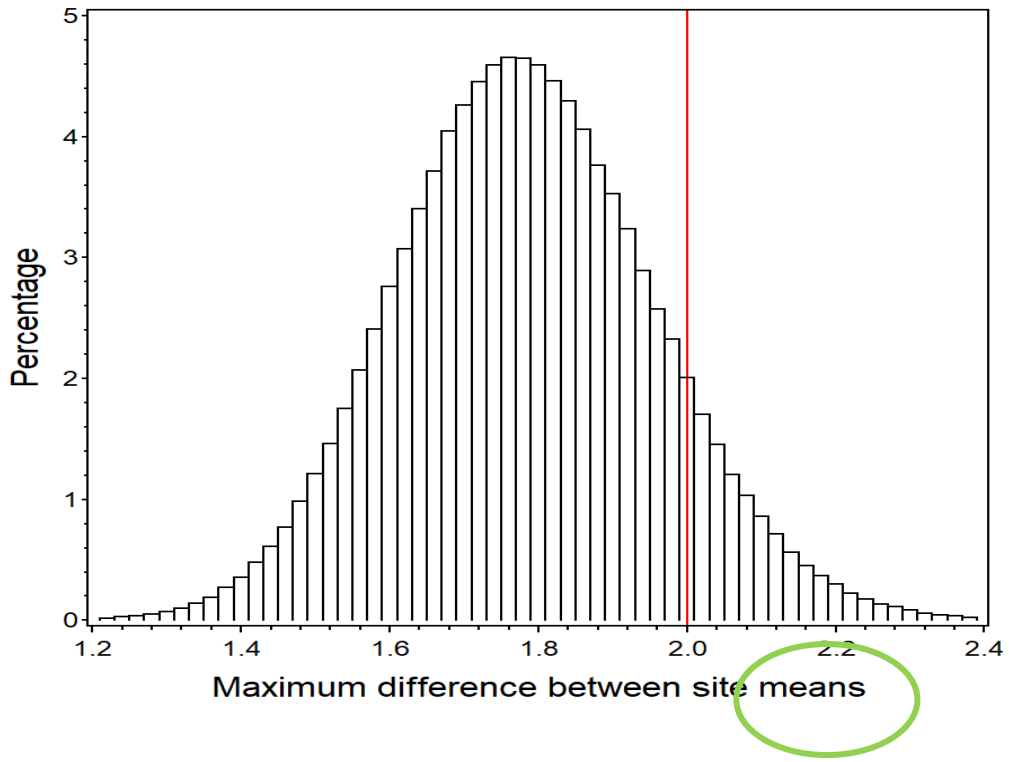
Example data: Bayesian analysis (EAC for the mean difference = 2)
(using non-informative prior for site means and a common variance)

Variable	Mean	Criterion	Fraction
$\mu_1 - \mu_2$	1.6601	$\lvert \mu_1 - \mu_2 \rvert < 2$	0.95
$\mu_1 - \mu_3$	1.5700	$\lvert \mu_1 - \mu_3 \rvert < 2$	0.98
$\mu_1 - \mu_4$	0.0099	$\lvert \mu_1 - \mu_4 \rvert < 2$	1.00
$\mu_2 - \mu_3$	-0.0900	$\lvert \mu_2 - \mu_3 \rvert < 2$	1.00
$\mu_2 - \mu_4$	-1.6501	$\lvert \mu_2 - \mu_4 \rvert < 2$	0.95
$\mu_3 - \mu_4$	-1.5601	$\lvert \mu_3 - \mu_4 \rvert < 2$	0.98
		All diffs < 2	0.89

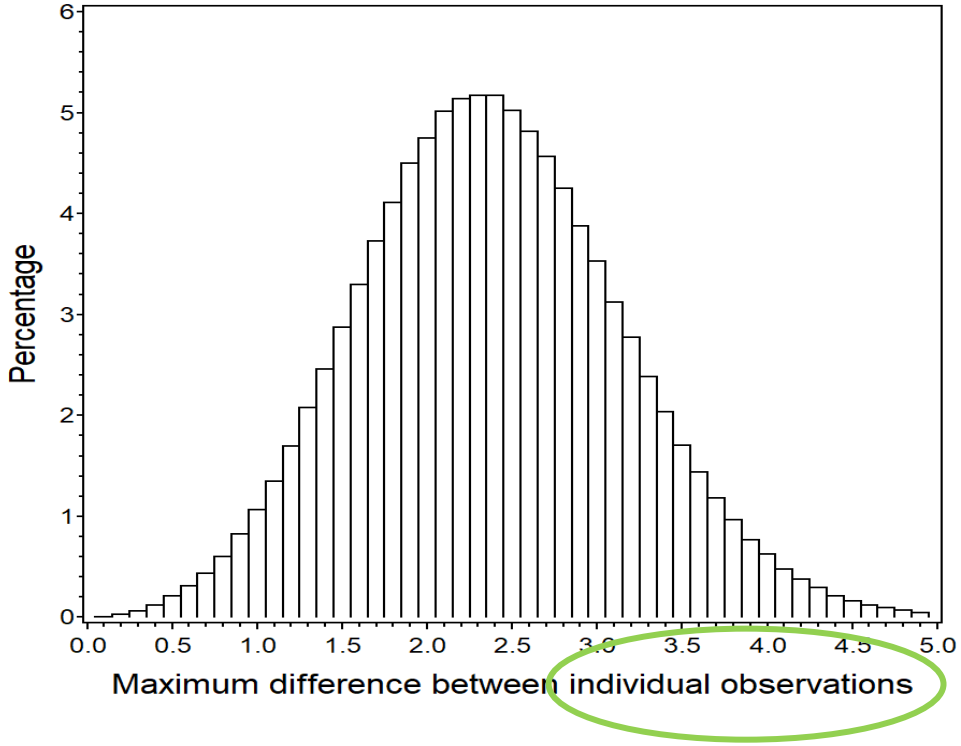
MERCK

Why Bayesian?

1. Rich information extracted from Bayesian analysis



Can visualize and quantify the **overall equivalence** among **means**



It is often more relevant to determine overall equivalence of **individual observations** across sites

MERCK

Why Bayesian?

2. Prior information can be used via Bayesian

Variable	Mean	Criterion	Fraction
$\mu_1 - \mu_2$	1.6600	$\lvert \mu_1 - \mu_2 \rvert < 2$	0.96
$\mu_1 - \mu_3$	1.5699	$\lvert \mu_1 - \mu_3 \rvert < 2$	0.99
$\mu_1 - \mu_4$	0.0101	$\lvert \mu_1 - \mu_4 \rvert < 2$	1.00
$\mu_2 - \mu_3$	-0.0902	$\lvert \mu_2 - \mu_3 \rvert < 2$	1.00
$\mu_2 - \mu_4$	-1.6499	$\lvert \mu_2 - \mu_4 \rvert < 2$	0.97
$\mu_3 - \mu_4$	-1.5597	$\lvert \mu_3 - \mu_4 \rvert < 2$	0.99
		All diffs < 2	0.92

Suppose that the analytical method used in the example had been validated at one of the sites, and that this validation yielded an estimate for the variance of 0.3 units, based on 80 degrees of freedom. This knowledge can be incorporated in the Bayesian analysis.

Inverse Gamma(θ_1 = 80, θ_2 = 24)

The probabilities changed slightly when using an informative prior for the variance

MERCK

Why Bayesian?

3. Flexibility in the modeling

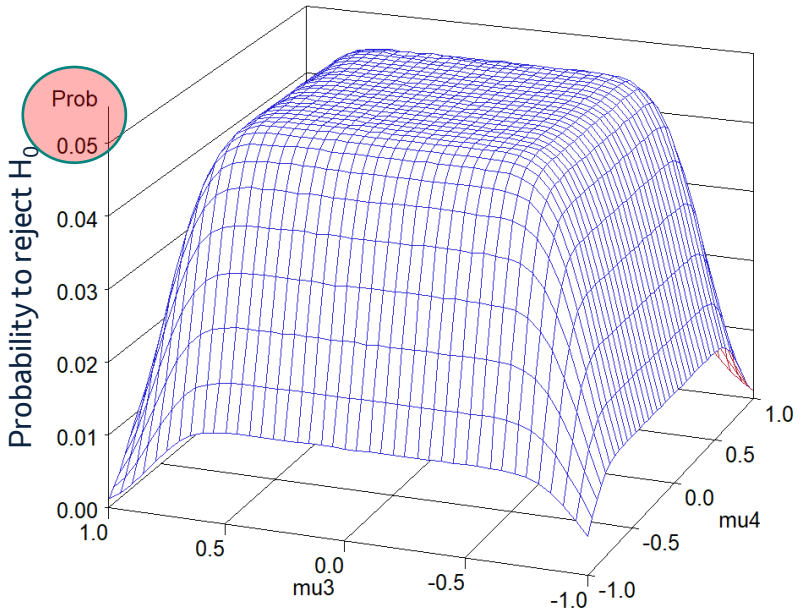
1. There is no need to mathematically derive the distribution of the estimator of the quantity being tested.

2. Can be flexibly applied to

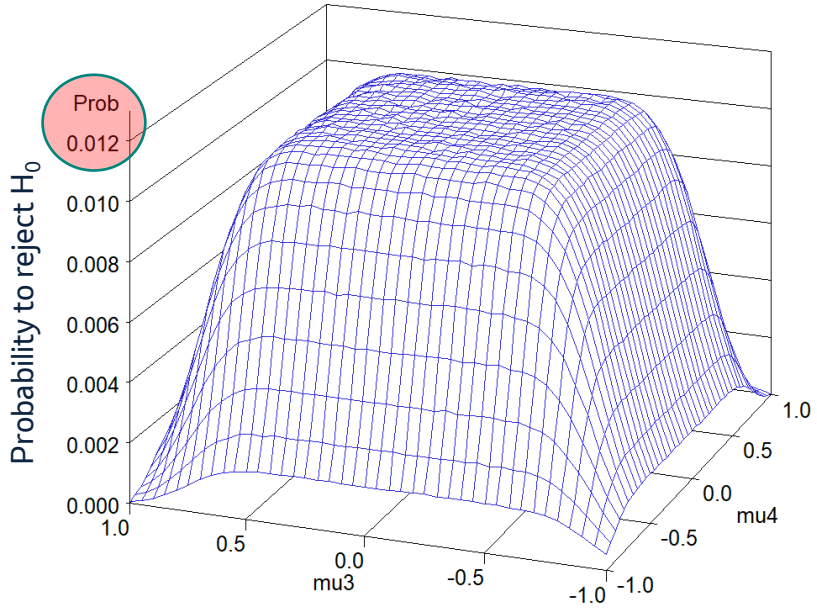
   - ✓ Any function of parameter(s)
   - ✓ Nonlinear models
   - ✓ Multivariate responses
   - ✓ Non-normal responses

Frequentist' equivalence test: No need for multiplicity correction

4-site example: $\mu_1 = -1$, $\mu_2 = 1$, and both μ_3 and μ_4 somewhere in-between.
$\sigma = 0.5$, n=20, EAC=2



No correction:
Type I Error with no correction is still maintained as ≤ 5%



Tukey correction:
Type I Error with Tukey correction is less than 1.1%
→ **Unnecessarily conservative**

Note: We believe this statement would hold in a wider scope, but so far the simulations have been performed only for the scenario of normal distributions with a common variance.

MERCK

Summary

- Bayesian strategy is very useful to deal with the problem of multi-group equivalence, providing more knowledge and flexibility than a hypothesis test as used in the Frequentist approach:

- We also showed that even in the Frequentist's multi-site equivalence test, there is no need to make multiplicity corrections.

Thank you

Merck Co., Inc.

E-mail: ji.young.kim1@merck.com