# The posterior probability of in vitro dissolution equivalence

Stan Altan, Janssen R&D
Hans Coppenolle, Janssen R&D
*David.LeBlond@sbcglobal.net
Areti Manola, Janssen R&D
Linas Mockus, Purdue University
Steve Novick, MedImmune
John Peterson, GSK
Yan Shen, Janssen R&D
Jyh-Ming Shoung, Janssen R&D
Harry Yang, MedImmune

# Acknowledgements

This presentation is drawn from the following joint efforts:

1. Shen Y, LeBlond D, J Peterson, S Altan, H Coppenolle, A Manola, J-M Shoung (2011) A Bayesian Approach to Equivalence Testing in a Non-linear Mixed Model Context, Non-Clinical Biostatistics Conference Boston, MA Oct 19, 2011

2. LeBlond D, J Peterson, and S Altan (2011) The posterior probability of dissolution equivalence, Midwest Biopharmaceutical Statistical Workshop, Muncie IN, May 25, 2011

3. LeBlond D, S Altan, S Novick, J Peterson, Y Shen, H Yang (2015) In vitro Dissolution Curve Comparisons: A Critique of Current Practice, Dissolution Technologies, dx.doi.org/10.14227/DT230116P14

4. Novick S, Shen Y, Yang H, Peterson J, LeBlond D, Altan S (2015) Dissolution curve comparisons through the F2 parameter, a Bayesian extension of the f2 statistic. J Biopharm Stat. 2015;25(2):351-71. doi: 10.1080/10543406.2014.971175.

5. Altan S, D LeBlond, J Peterson, Y Shen, H Yang, and S Novick (2017) In Vitro Dissolution Curve Comparisons: A Critique of Current Practice and a Proposed Bayesian Test Statistic, poster presented at the Non-clinical statistics meeting, Boston

6. Mockus L and D LeBlond (expected, 2019) Bayesian methods for in vitro dissolution drug testing and similarity comparisons, in Bayesian Methods in Pharmaceutical Research (eds: Lesaffre E, Baio G, and Boulanger B), CRC Press

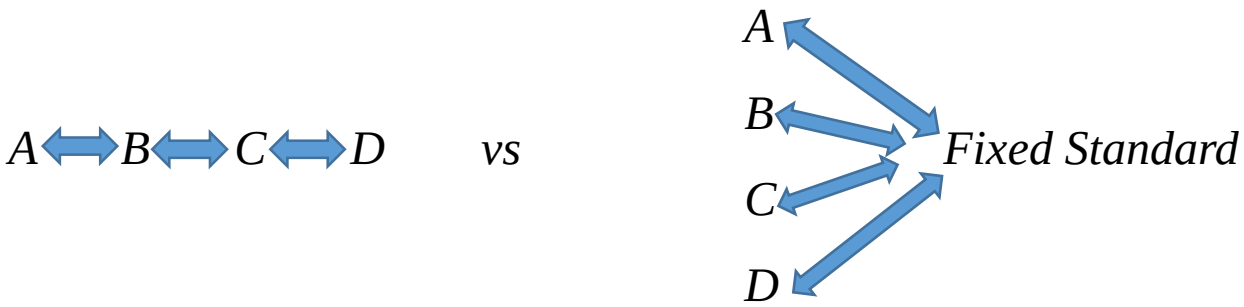
We are grateful to Gianluca Baio University College London, Bruno Boulanger Arlenda, and Emmanuel Lesaffre, KU Leuven for helpful suggestions.

# Outline

- Frame the dissolution equivalency question
- Normative requirements for a Bayesian approach
- Univariate vs multivariate criterion
- Model dependent approach (Weibull)
- Concept of minimum posterior probability of equivalence
- Importance of inference space
- Example
  - Compare 2 processes
  - Multiple lots/process & multiple profiles/lot
  - Hierarchical nonlinear model
  - Priors
  - MCMC inference
  - Univariate & Multivariate criteria
- Discussion points

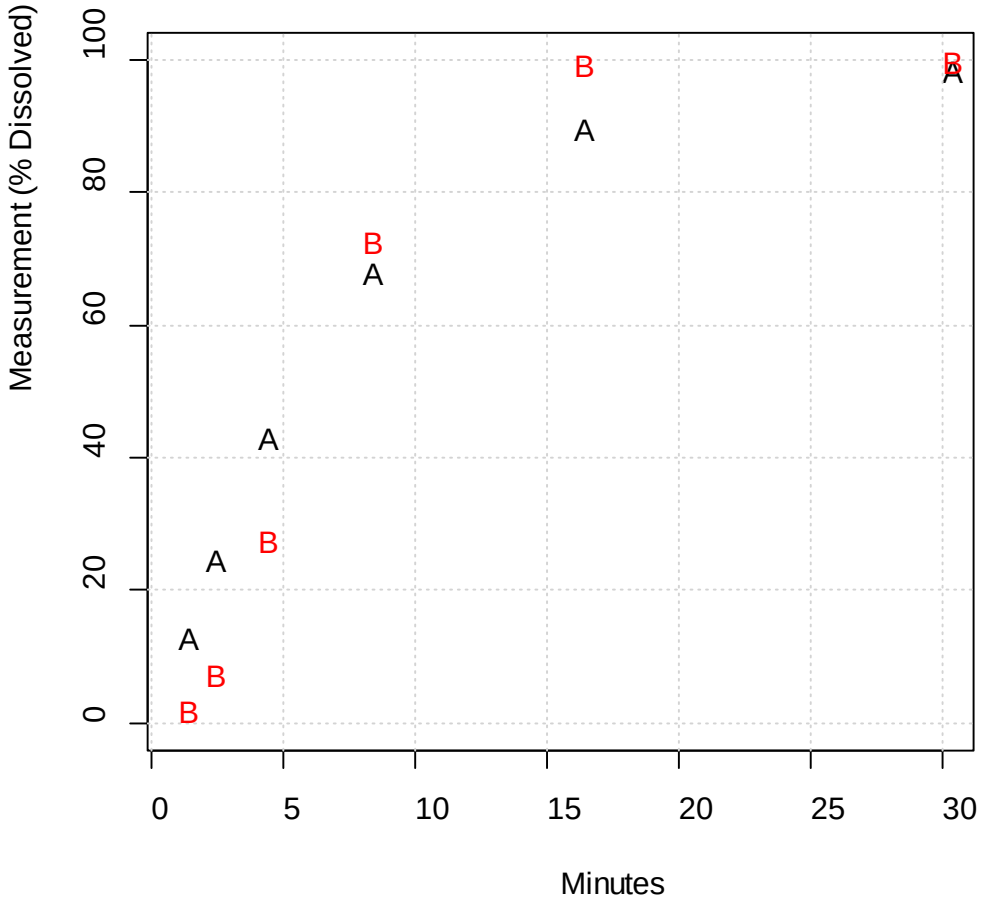
# Not considered here…

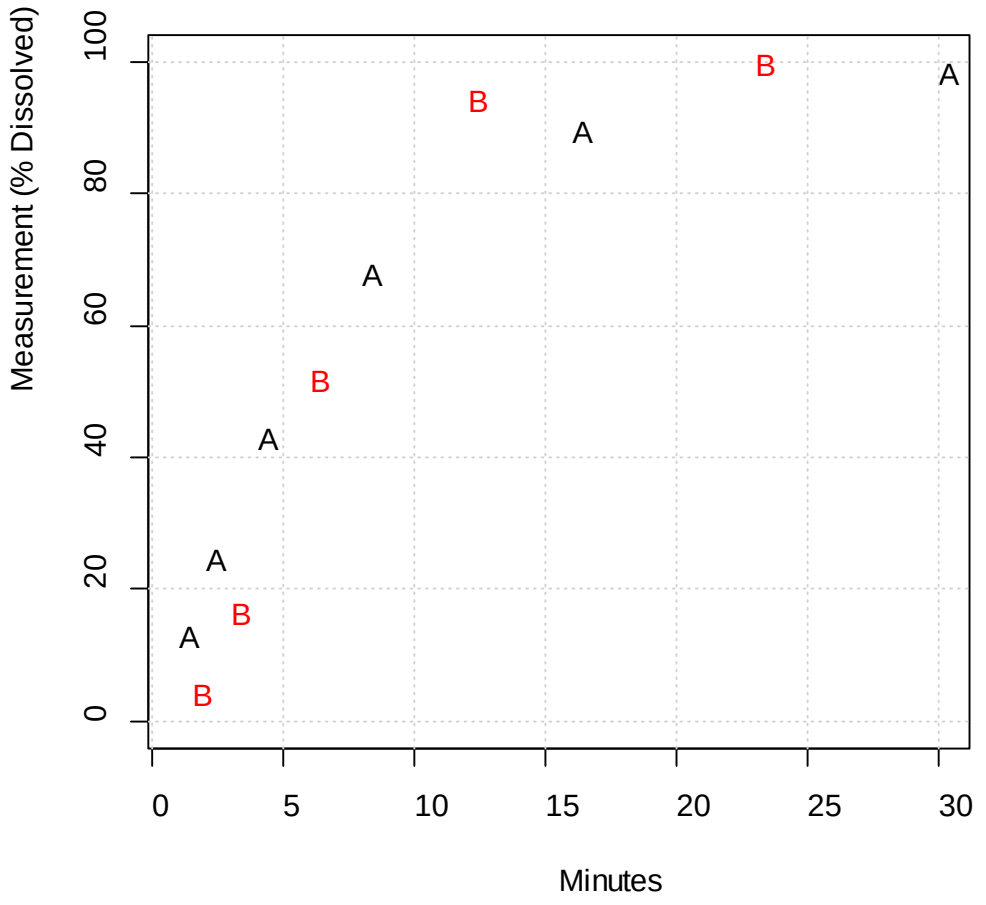
- Is IV dissolution even predictive of efficacy, safety, or quality?
- Is pairwise equivalence testing the best way?




- What is the “best” criterion of equivalence?
- Frequentist vs Bayesian

# Is B *equivalent* to A?



# Normative Requirements
## (Bayesian version[1])

1. A specified function, $\phi$, of the population parameters is used to define a fixed region of equivalence $\rho$.

2. The test of equivalence…
   - Makes an inference to the specific process under study
   - Is conditional on the justified population model (including prior knowledge), and observed data
   - Is based ***solely***[2] on the posterior probability, $\hat{P}$, that $\phi \subset \rho$

3. Equivalence is accepted iff $\hat{P} \geq P_{MIN}$
   - Requires a choice of $P_{MIN}$

[1]For a frequentist perspective, see Eaton ML, Muirhead RJ, and Steeno GS (2003) Aspects of the dissolution profile testing problem, Biopharmaceutical Report 11(2) 2-7.

[2]Calibration of frequentist performance of the test is useful, but is ***not*** the basis for an equivalence decision about the process under study.

# Normative Requirements
## (Bayesian version)

1. A specified function, $\phi$, of the population parameters is used to define a fixed region of equivalence $\rho$.

Possible candidates for $\phi$ and $\rho$ *( p time points )*:

$$f2 = 50 \cdot \log_{10}\left(\frac{100}{\sqrt{1+\frac{1}{p}\sum_{n=1}^{p}\left(\bar{y}_{B,n}-\bar{y}_{A,n}\right)^2}}\right) > 50 \ ?$$

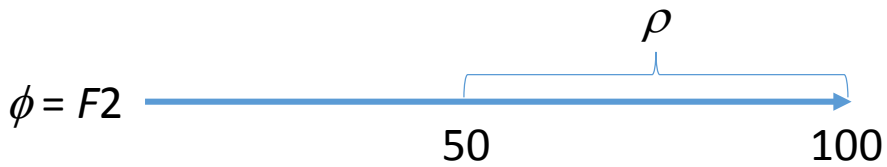
1

**No**. $f2$ is a function of data, not population parameters

$$F2 = 50 \cdot \log_{10}\left(\frac{100}{\sqrt{1+\frac{1}{p}\sum_{n=1}^{p}\left(w_{B,n}-w_{A,n}\right)^2}}\right) > 50 \ ?$$

2

**Yes.** $F2$ is a function of population parameters

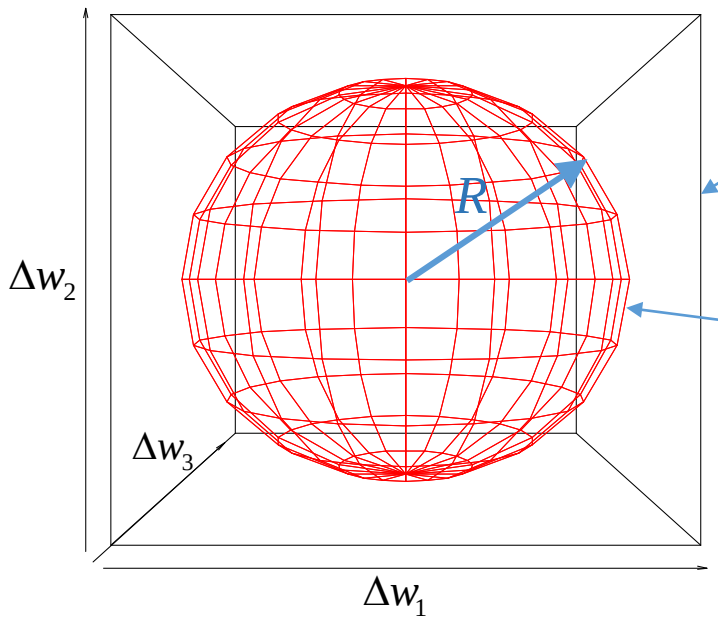



[1] Moore and Flanner (1996)
[2] Novick et al (2015)

# Normative Requirements
## (Bayesian version)

1. A specified function, $\phi$, of the population parameters is used to define a fixed region of equivalence $\rho$.

Possible candidates for $\phi$ and $\rho$ ( $p$=3 time points ):




$$\rho = \left( \begin{array}{r} -R < \Delta w_1 < +R \,\& \\ \phi = \mathbf{\Delta w} : -R < \Delta w_2 < +R \,\& \\ -R < \Delta w_3 < +R \end{array} \right) ?$$

$$\rho = \left( \phi : \phi = \sqrt{\mathbf{\Delta w}' \bullet \mathbf{\Delta w}} < R \right) ?$$

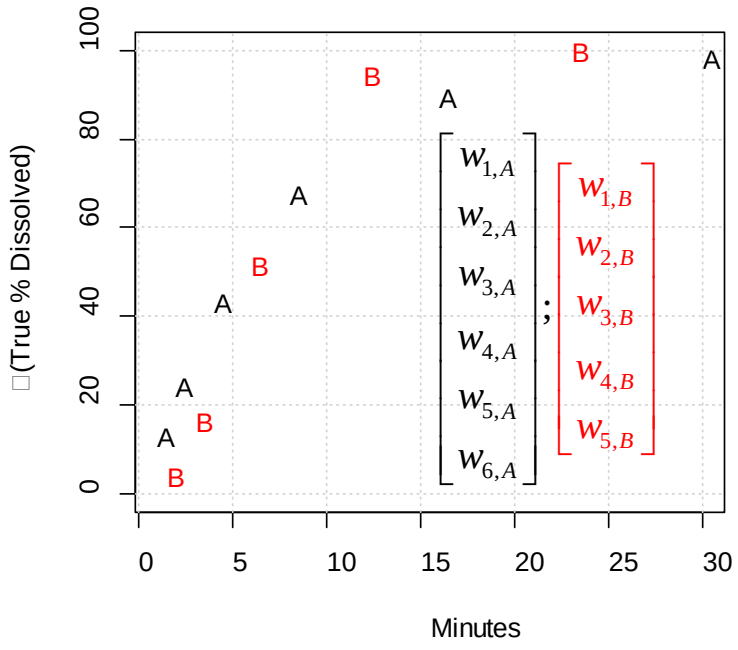
**Yes.** Both are functions of population parameters and constants, not dependent on measurement results

when $R = 99p$, $\phi \equiv F2$

# Normative Requirements
## (Bayesian version)

1. A specified function, $\phi$, of the population parameters is used to define a fixed region of equivalence $\rho$.

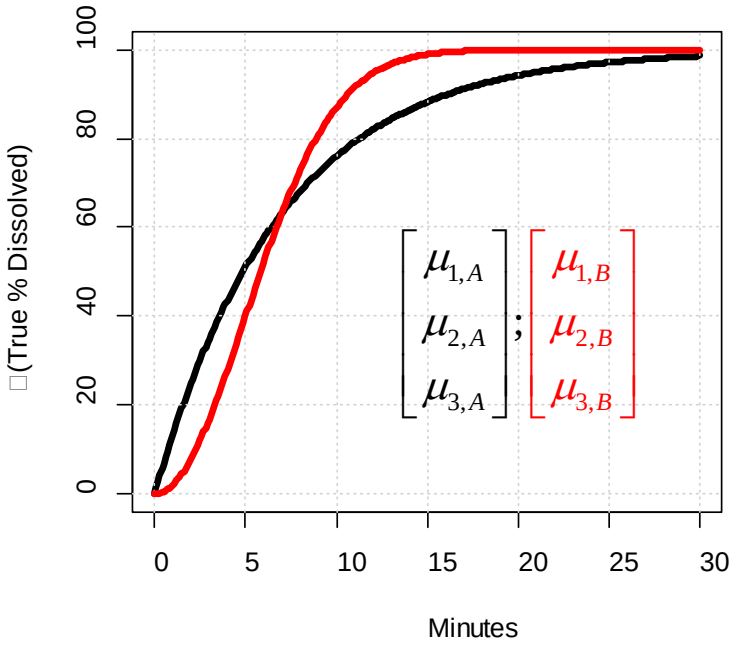
What about "model dependent" equivalence regions?




$$w_{m,n} = Weibull\left(\boldsymbol{\mu}_m, t_n\right)$$

$$= \mu_{1,i}\left[1-\exp\left(-\left(\frac{t_n}{\exp\left(\mu_{2,n}\right)}\right)^{\exp\left(\mu_{3,n}\right)}\right)\right]; m = A, B$$

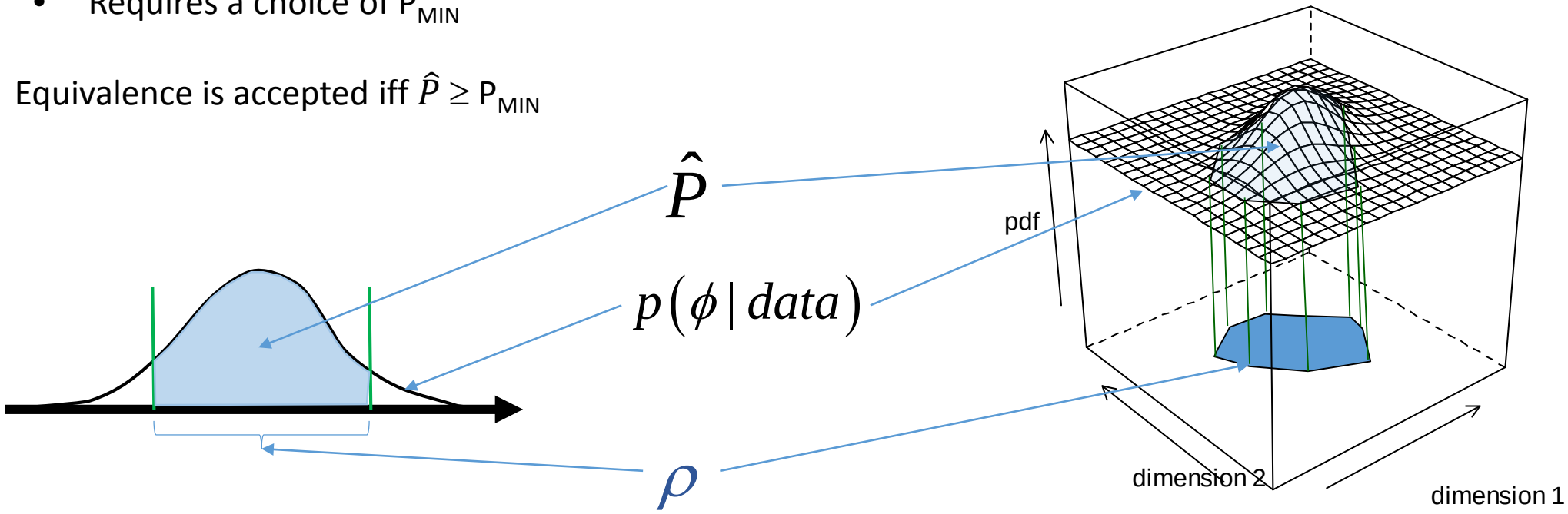
1. Dimensional reduction
2. $A$:$B$ mapping is 1:1
3. **Yes.** Define $\phi$ & $\rho$ in terms of $\boldsymbol{\mu}$ in place of $\mathbf{w}$
4. $\mu$'s have mechanistic interpretations



# Normative Requirements
## (Bayesian version)

2. The test of equivalence…
   - Makes an inference to the specific process under study
   - Is conditional on the justified population model (includin ; prior knowledge), and observed data
   - Is based *solely* on the posterior probability, $\hat{P}$, that $\phi \subset \rho$
   - Requires a choice of $P_{MIN}$

3. Equivalence is accepted iff $\hat{P} \geq P_{MIN}$



# Defining the inference space

| Study type[1] | Model of Lot(s) required? | Model of Process(es) required? | Inference space |
|---|---|---|---|
| Enumerative | Yes | No | Sampled Lot(s) only |
| Analytic | Yes | Yes | Lot(s) & Process(es) |

Should IV dissolution similarity studies be enumerative or analytic?

- Often a model of the process(es) is absent, yet…
- Inferences are applied to the process(es)

[1]Deming E (1975) On probability as a basis for action, The American Statistician 29(4), 146-152

# Example IV dissolution study

$$y_{m,l,u,n} \sim N\left(w_{m,l,u,n}, \sigma^2\right)$$

$$w_{m,l,u,n} = Weibull\left(\boldsymbol{\mu}_m + \boldsymbol{\varepsilon}_{m,l} + \boldsymbol{\delta}_{m,u}, t_n\right)$$

$$\boldsymbol{\varepsilon}_{m,l} \sim N_3\left(\mathbf{0}, \boldsymbol{\Sigma}_m^{Lot}\right)$$

$$\boldsymbol{\delta}_{m,u} \sim N_3\left(\mathbf{0}, \boldsymbol{\Sigma}_m^{Unit}\right)$$

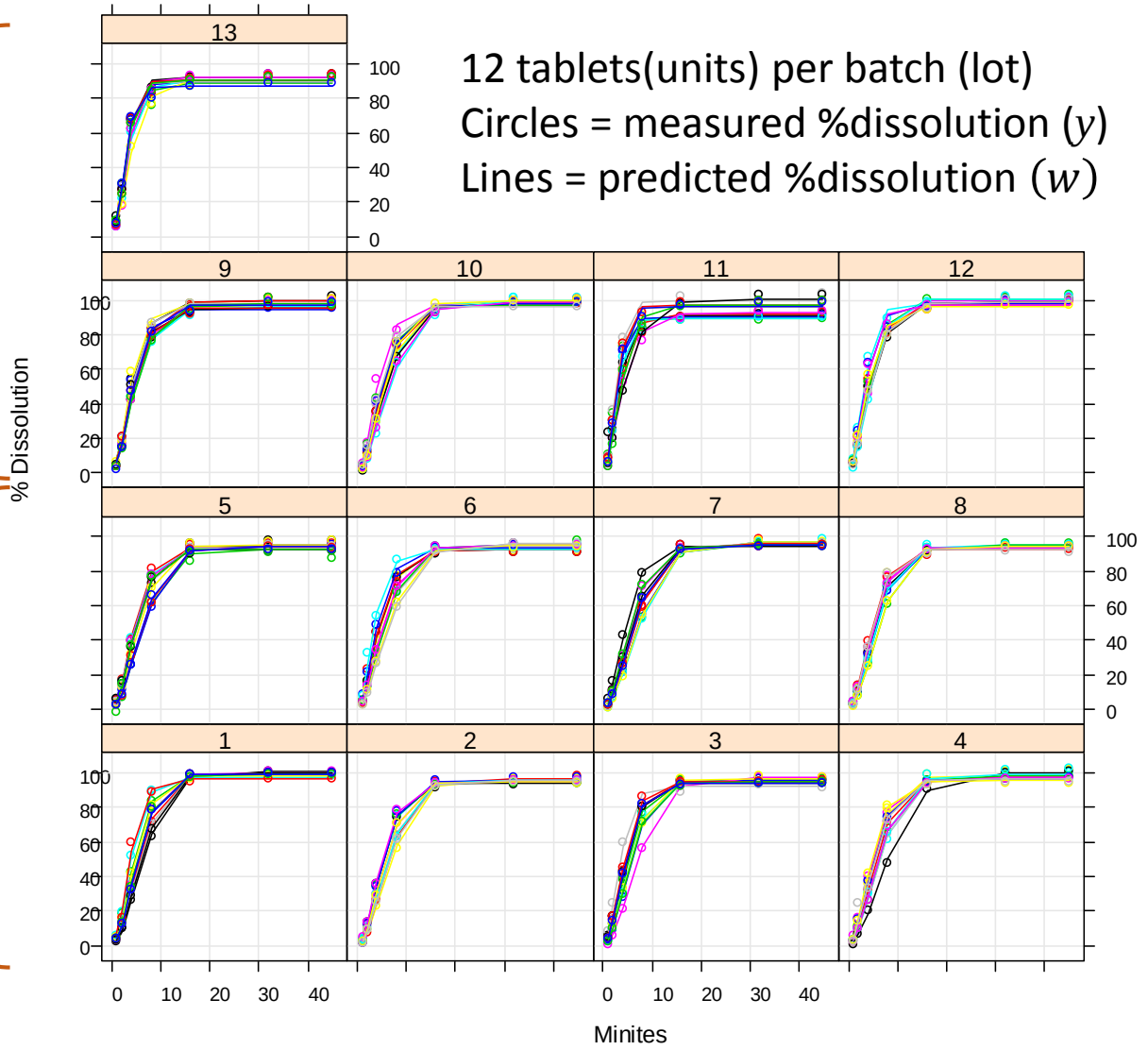
| Index | Variable | Levels |
|---|---|---|
| $m$ | Process | A,B |
| $l$ | Lot | 1,2,…,13 |
| $u$ | Unit | 1,2,…,156 |
| $n$ | $t$ (min) | 1,2,…,7 |

5 lots made using process B

8 lots made using process A

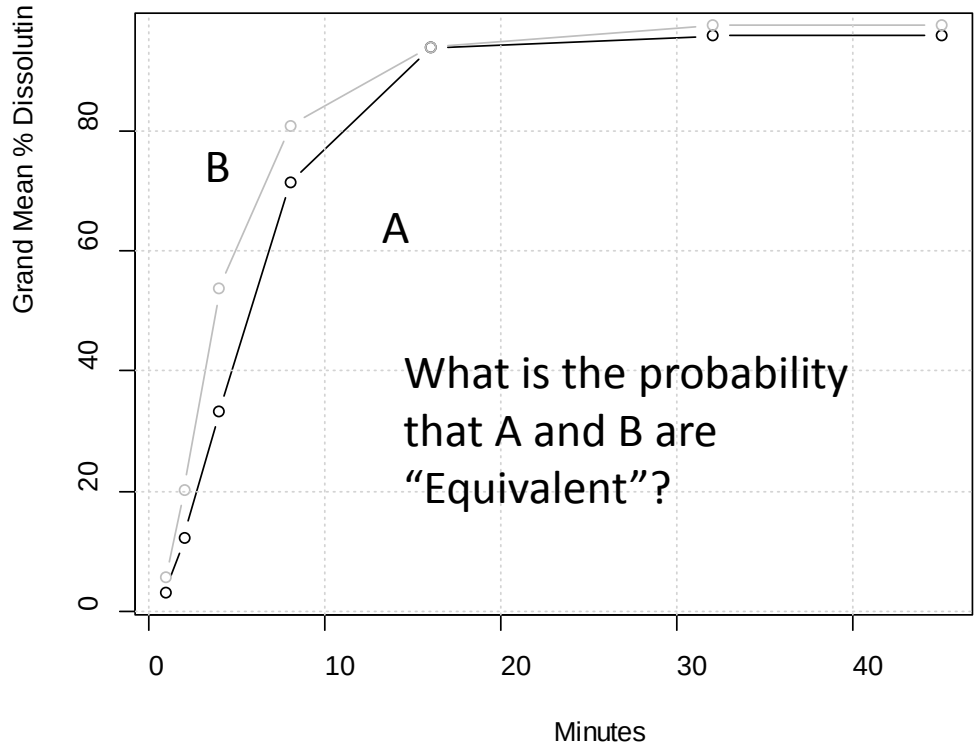
**Batches 1-8 are from Site 1 and Batches 9-1**

12 tablets(units) per batch (lot)
Circles = measured %dissolution ($y$)
Lines = predicted %dissolution ($w$)




**Batch numbers are shown above each panel**

# Grand average observed dissolution



# Priors used

$$y_{m,l,u,n} \sim N\left(w_{m,l,u,n}, \sigma^2\right)$$

$$w_{m,l,u,n} = Weibull\left(\boldsymbol{\mu}_m + \boldsymbol{\varepsilon}_{m,l} + \boldsymbol{\delta}_{m,u}, t_n\right)$$

$$\boldsymbol{\varepsilon}_{m,l} \sim N_3\left(\mathbf{0}, \boldsymbol{\Sigma}_m^{Lot}\right)$$

$$\begin{bmatrix} \left(\sigma_{1,m}^{lot}\right)^2 & 0 & 0 \\ 0 & \left(\sigma_{2,m}^{lot}\right)^2 & 0 \\ 0 & 0 & \left(\sigma_{3,m}^{lot}\right)^2 \end{bmatrix}$$

$$\boldsymbol{\delta}_{m,u} \sim N_3\left(\mathbf{0}, \boldsymbol{\Sigma}_m^{Unit}\right)$$

$$\begin{bmatrix} \sigma_{1,m}^{unit} & 0 & 0 \\ 0 & \sigma_{2,m}^{unit} & 0 \\ 0 & 0 & \sigma_{3,m}^{unit} \end{bmatrix} \begin{bmatrix} 1 & & \\ \rho_{12,m} & 1 & \\ \rho_{13,m} & \rho_{23,m} & 1 \end{bmatrix} \begin{bmatrix} \sigma_{1,m}^{unit} & 0 & 0 \\ 0 & \sigma_{2,m}^{unit} & 0 \\ 0 & 0 & \sigma_{3,m}^{unit} \end{bmatrix}$$

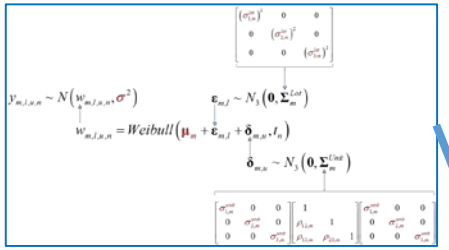
| Parameters | Minimally informative[1] Prior distribution |
|---|---|
| $\mu$'s (3/process) | Uniform(-∞,+∞) [2] |
| Common $\sigma$ | Uniform(0,+∞) [2] |
| $\sigma^{lot}$'s and $\sigma^{unit}$'s (6/process) | Uniform(0,+∞) [2] |
| Unit correlation matrix (3 $\rho$'s/process) | Uniform density over all 3x3 correlation matrices |

[1]If scientifically justified, informative priors would be preferable

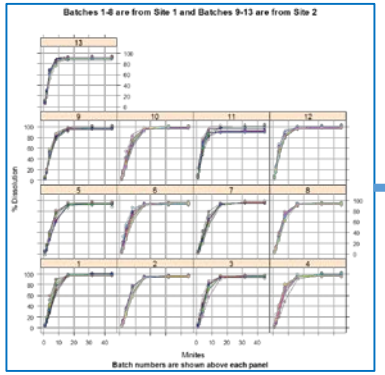
[2]Use of these improper priors led to posterior distributions indistinguishable from those obtained using wide proper prior distributions

# MCMC inference

likelihood



data




priors

| Parameters | Minimally informative[1] Prior distribution |
|---|---|
| $\mu$'s (3/process) | Uniform(-∞,+∞)[2] |
| Common $\sigma$ | Uniform(0,+∞)[2] |
| $\sigma^{Lot}$'s and $\sigma^{Unit}$'s (6/process) | Uniform(0,+∞)[2] |
| Unit correlation matrix (3 $\rho$'s/process) | Uniform density over all 3x3 correlation matrices |

Stan[1]

4 chains
- 4,000 draw warmup/chain
- 4,000 draw posterior sample/chain

Due diligence
- Convergence?
- Mixing?
- Prior impact?
- Effective sample size?
- MC std dev low?

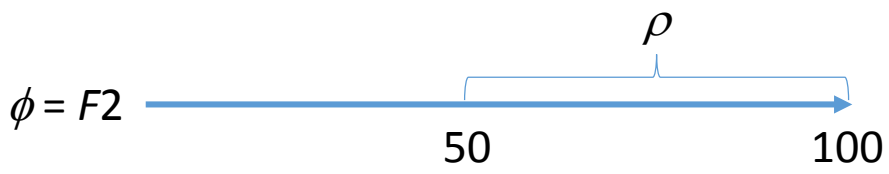
not OK

OK

16,000 draw sample from the joint posterior of all model parameters

Calculate quantities of interest for each draw → 16,000 draw sample from joint posterior of quantities of interest

[1] http://mc-stan.org/users/documentation/

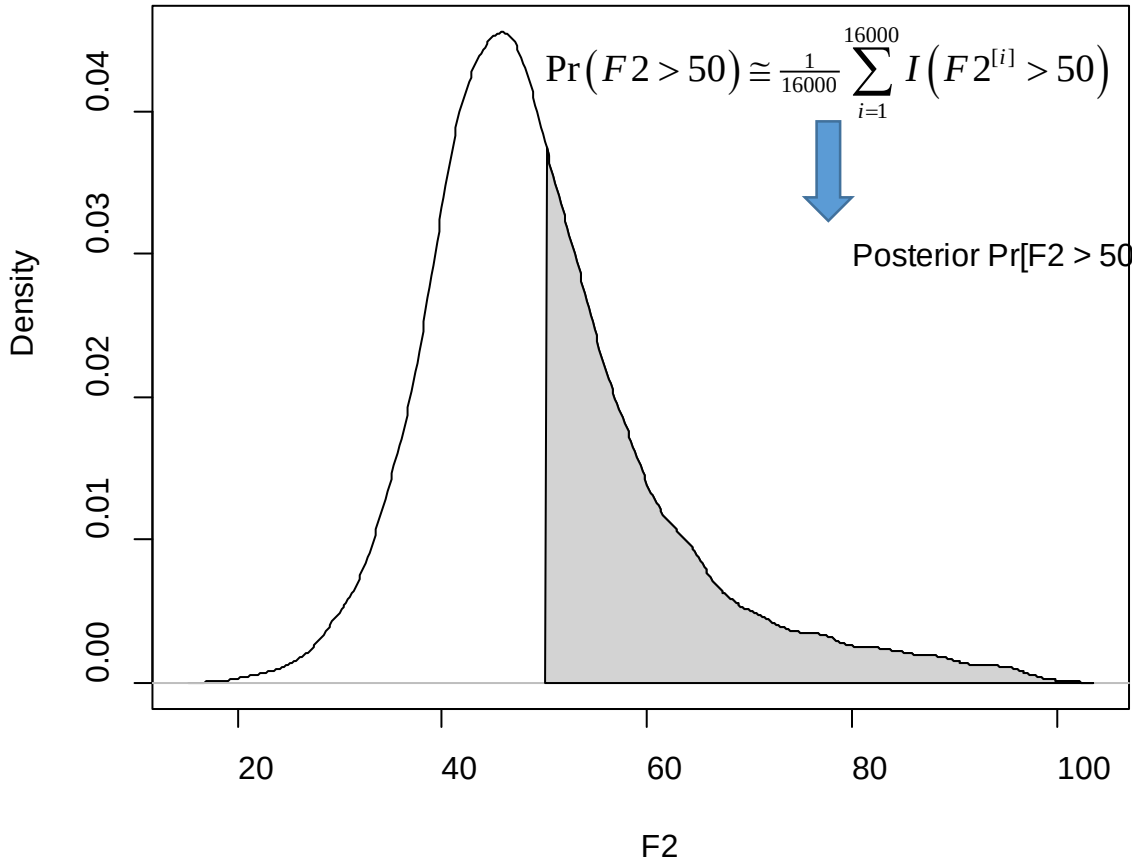



$$F2^{[i]} = 50 \cdot \log_{10}\left(\frac{100}{1+\frac{1}{N_{times}}\sqrt{\sum_{n=1}^{N_{times}}\left(w_{B,n}^{[i]} - w_{A,n}^{[i]}\right)^2}}\right); i = 1,\ldots,16000$$

$$w_{m,n}^{[i]} = Weibull\left(\boldsymbol{\mu}_m^{[i]}, t_n\right); m = A, B; i = 1,\ldots,16000$$

Posterior sample: $\boldsymbol{\mu}_m^{[i]}; i = 1,\ldots,16000$

**Kernal Density estimate of F2 p**



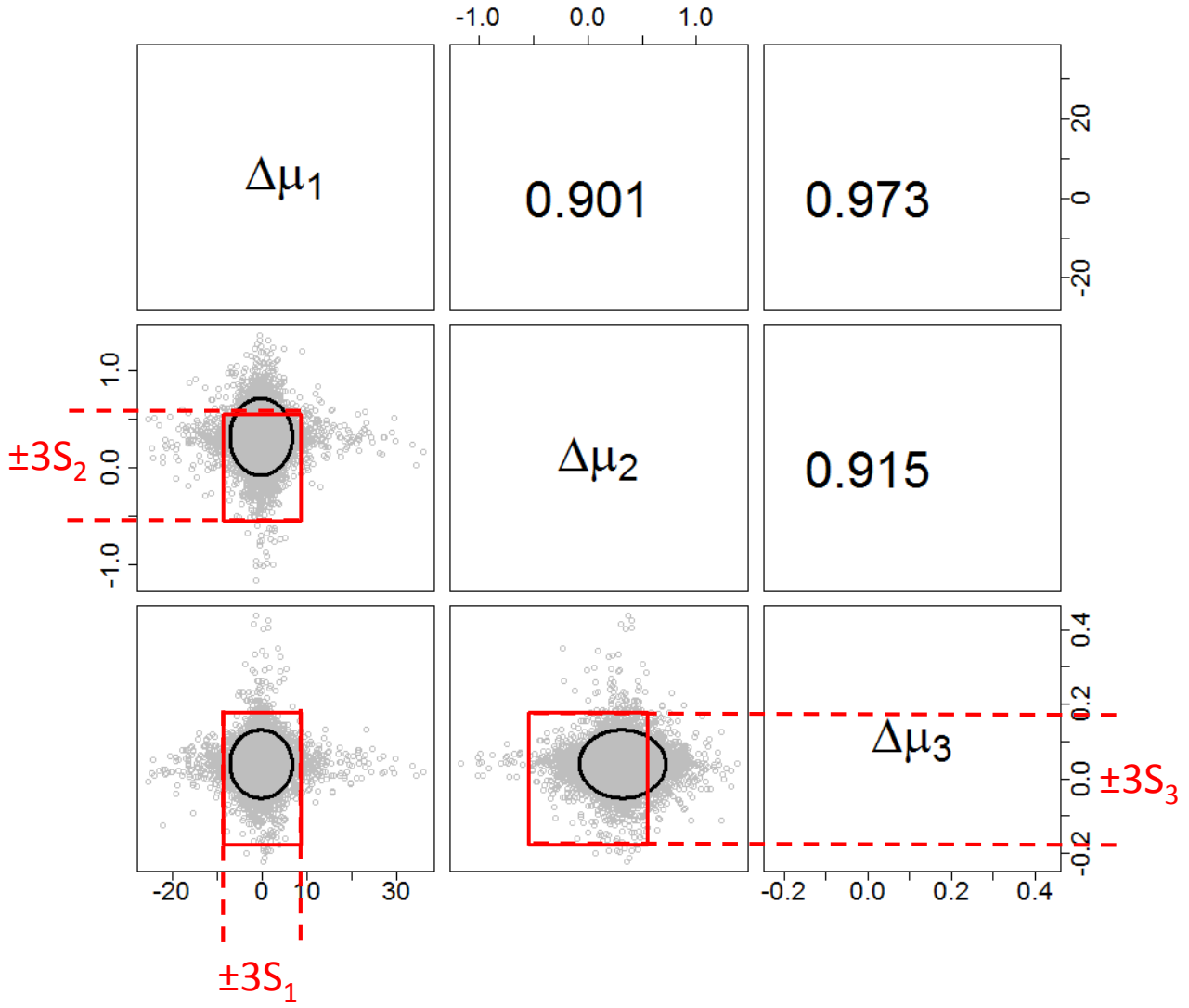
$$\rho = \left( \begin{array}{r} -3S_1 < \Delta\mu_1 < +3S_1 \ \& \\ \phi = \boldsymbol{\Delta\mu} : -3S_2 < \Delta\mu_2 < +3S_2 \ \& \\ -3S_3 < \Delta\mu_3 < +3S_3 \end{array} \right)^{1}$$

$$\Pr\left[\boldsymbol{\Delta\mu} \in \rho\right] \cong \frac{1}{16000} \sum_{i=1}^{16000} \prod_{p=1}^{3} I\left(-3S_p \le \Delta\mu_p^{[i]} \le +3S_p\right)$$

$$= 0.895$$

$$\boldsymbol{\Delta\mu}_m^{[i]} = \boldsymbol{\mu}_B^{[i]} - \boldsymbol{\mu}_A^{[i]}; i = 1,\ldots,16000$$

Posterior sample: $\boldsymbol{\mu}_m^{[i]}; i = 1,\ldots,16000; m = A, B$




[1]Sathe, Tsong & Shah (1996) suggest $S_p$ be obtained from reference process total standard deviation, ignoring estimation error

$$\rho = \left( \phi = \boldsymbol{\Delta\mu} : \boldsymbol{\Delta\mu}' \hat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{\Delta\mu} \le \chi^2_{3,0.99} \right)$$

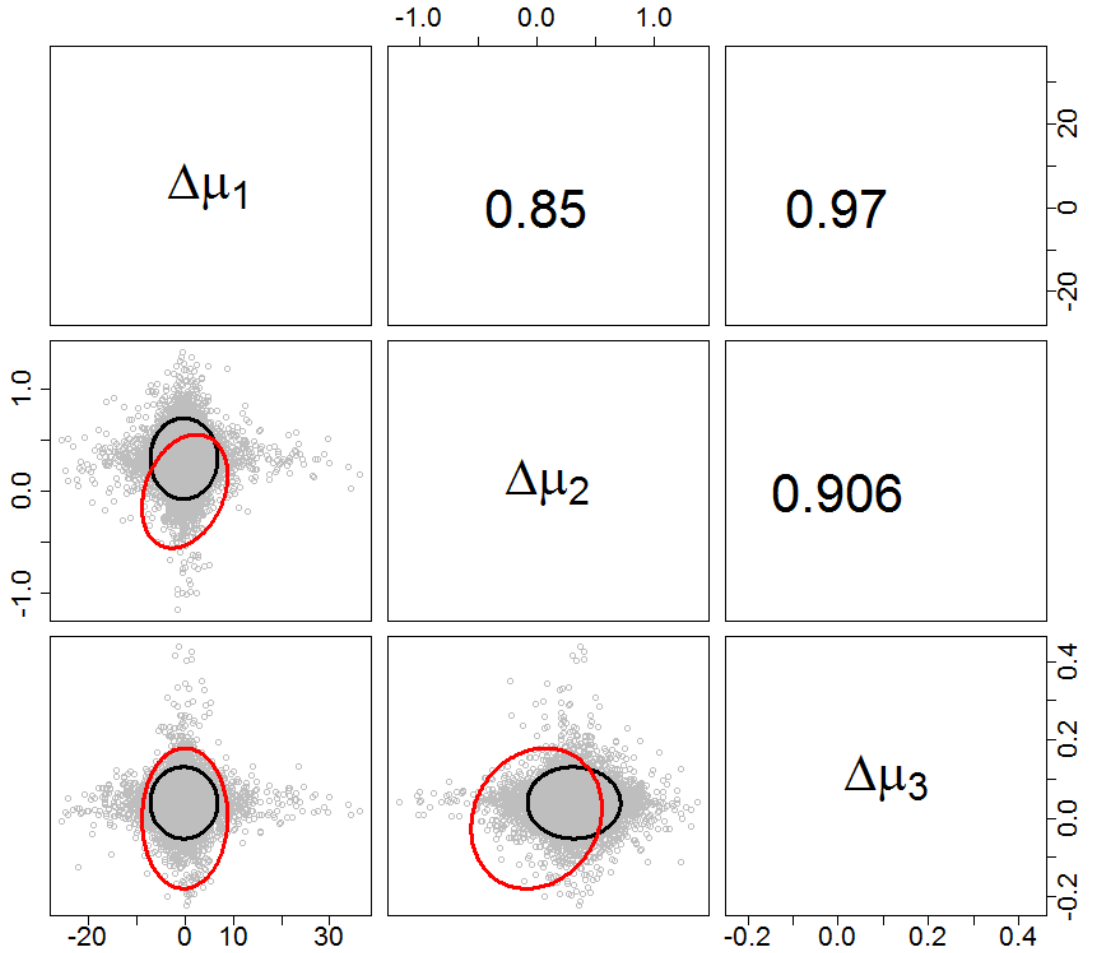
$$\underbrace{\frac{1}{16000} \sum_{i=1}^{16000} \left( \boldsymbol{\Sigma}^{[i]}_{Lot,A} + \boldsymbol{\Sigma}^{[i]}_{Unit,A} \right)}^{1}$$

$$\Pr\left[ \boldsymbol{\Delta\mu} \in \rho \right] \cong \frac{1}{16000} \sum_{i=1}^{16000} I\left( \boldsymbol{\Delta\mu}^{[i]\prime} \hat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{\Delta\mu}^{[i]} \le \chi^2_{3,0.99} \right)$$

$$= 0.880$$

$$\boldsymbol{\Delta\mu}^{[i]} = \boldsymbol{\mu}_B^{[i]} - \boldsymbol{\mu}_A^{[i]}; i = 1,\ldots,16000$$

Posterior sample: $\boldsymbol{\mu}_m^{[i]}; i = 1,\ldots,16000; m = A, B$




[1]An arbitrary choice for illustration. Should be fixed based on product specific requirements and scientific knowledge.

# For discussion…

- Does this Bayesian paradigm provide a useful framework for IV dissolution profile (and/or other) similarity comparisons?
- What would regulators think?
- What would sponsors think?
- How could a link to current practice (e.g., f2, MSD) be assured?
- How should the parameter function, $\phi$, and the equivalence region, $\rho$, be chosen?
- Should the equivalence region, $\rho$ , be fixed across studies or be estimated from individual study data?
- If scientifically justifiable prior knowledge is available, should it be used?