The value of Bayesian statistics for assessing comparability

Timothy Mutsvari (Arlenda)
on behalf of EFSPI Working Group

Agenda

- Bayesian Methods: General Principles
  - Direct Probability Statements
  - Posterior Predictive Distribution
- Biosimilarity Model formulation
- Sample Size Justification
- Multiplicity
  - Multiple CQAs
- Assurance (not Power)

Bayesian Methods:

General Principles

Two different ways to make a decision based on

A

$$\Pr(\text{observed data} \mid \text{not biosimilar})$$

- Better known as the **p-value** concept
- Used in the **null hypothesis** test (or decision)
- This is the likelihood of the data assuming an hypothetical explanation (e.g. the “null hypothesis”)
- **Classical statistics** perspective (Frequentist)

B

$$\Pr(\text{biosimilar} \mid \text{observed data})$$

- **Bayesian** perspective
- It is the probability of similarity given the data

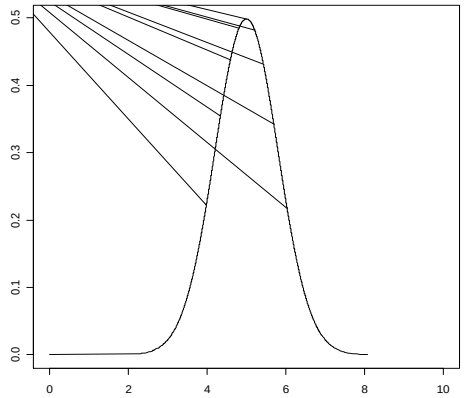
The Bayesian perspective allows to directly address the question of interest.

EFSPI

Bayesian Principle

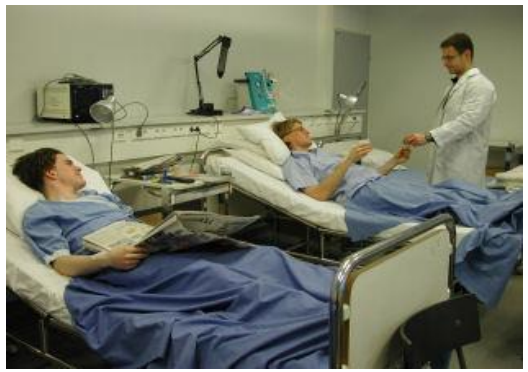
- After having observed the data of the study, the prior distribution of the treatment effect is updated to obtain the posterior distribution

PRIOR distribution



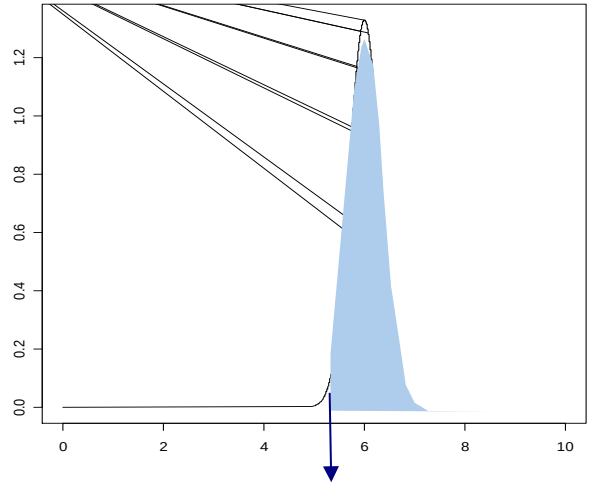
+

STUDY data



$\propto$

POSTERIOR distribution



P(treatment effect > 5.5)= P(success)

- Instead of having a point estimate (+/- standard deviation), we have a complete distribution for any parameter of interest

EFSPI

Posterior Predictive Distribution

- Given the model and the posterior distribution of its parameters, what are the plausible values for a future observation $\tilde{y}$?

- This can be answered by computing the plausibility of the possible values of $\tilde{y}$ conditionally on the available information:

$$p(\tilde{y}|data) = \int p(\tilde{y}|\theta)p(\theta|data)d\theta$$

- The factors in the integrant are
  - $p(\tilde{y}|\theta)$: it is given by the model for given values of the parameters
  - $p(\theta|data)$: it is the posterior distribution of the model parameter

EFSPI

Posterior Predictive Distribution - Illustration

1st , draw a mean and a variance from:

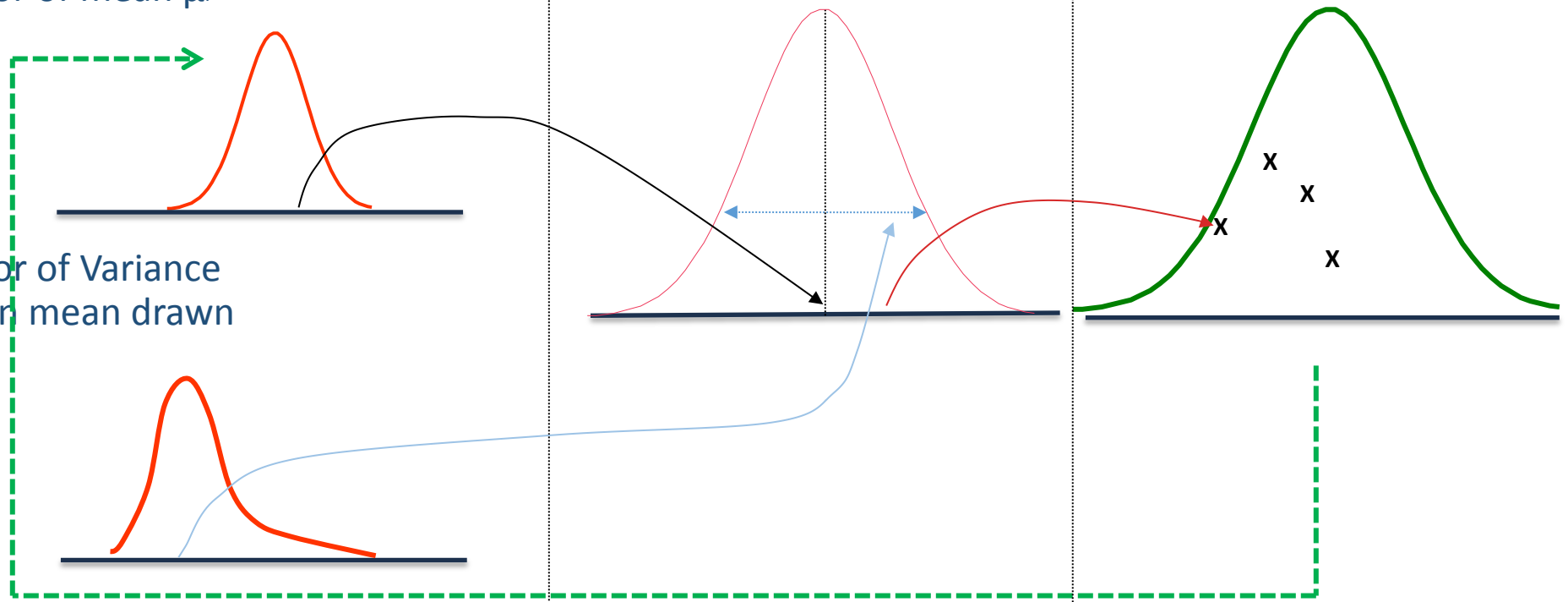
Posterior of mean μ_i

Posterior of Variance σ^2_i given mean drawn

2nd , draw an observation from the resulting distribution

$Y \sim \text{Normal}(\mu_i, \sigma^2_i)$

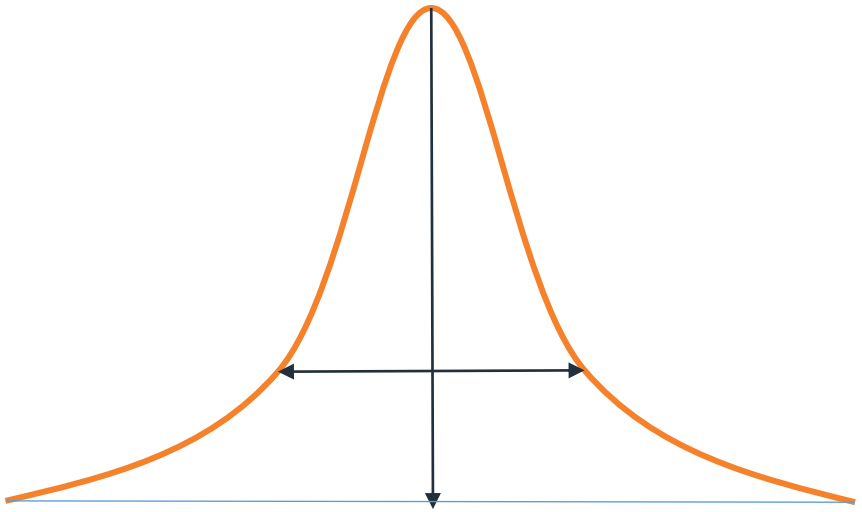
3rd , repeat this operation a large number of time to obtain the predictive distribution



Difference: Simulations vs Predictions

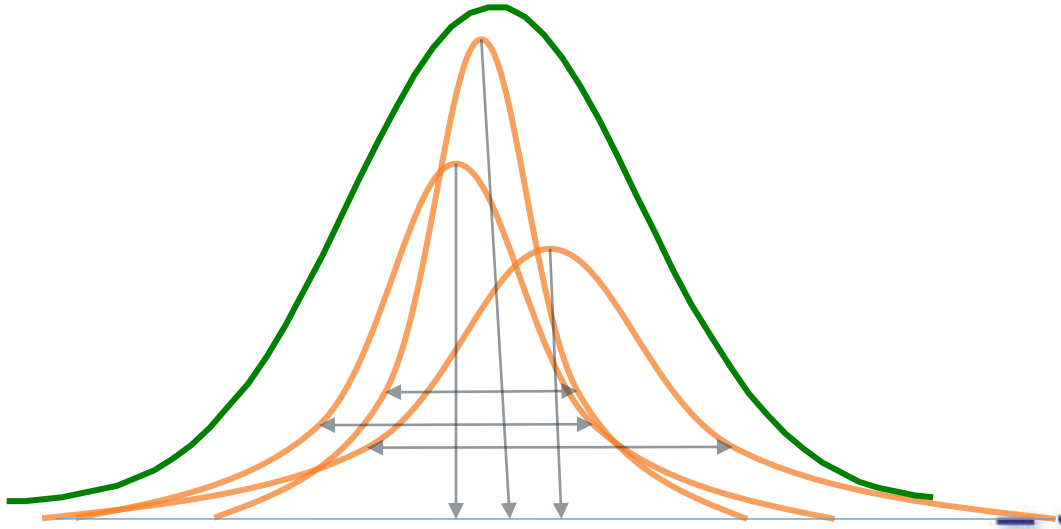
Monte Carlo Simulations

the “new observations” are drawn from distribution “centered” on estimated location and dispersion parameters (treated as “true values”).



Bayesian Predictions

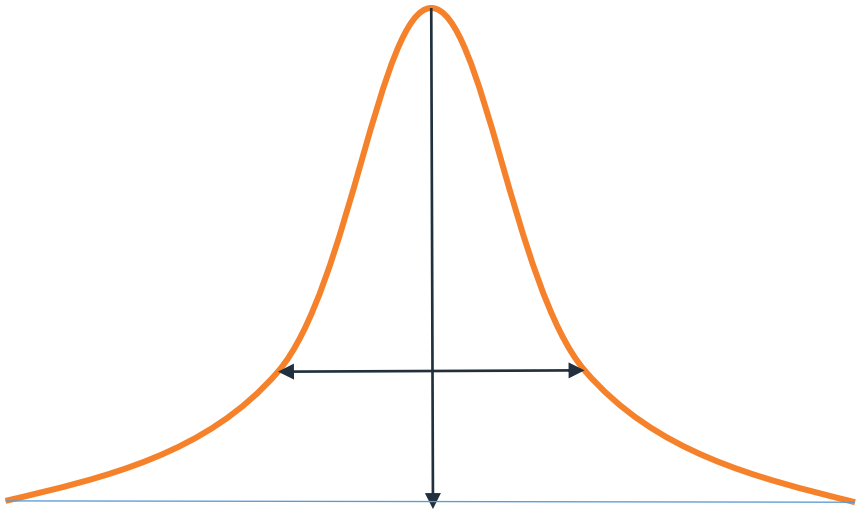
the uncertainty of parameter estimates (location and dispersion) is taken into account before drawing “new observations” from relevant distribution



Difference: Simulations vs Predictions

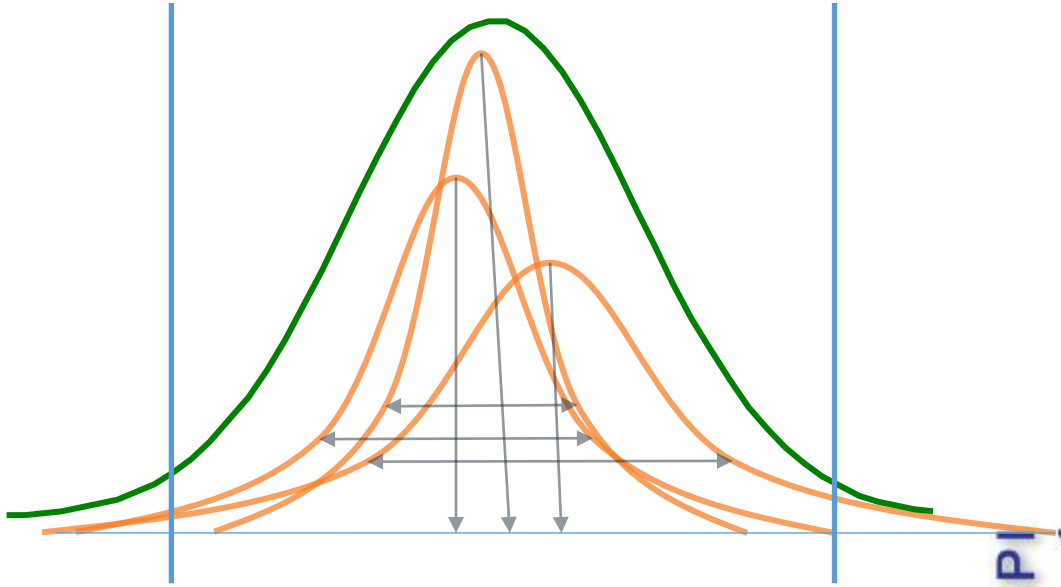
Monte Carlo Simulations

the “new observations” are drawn from distribution “centered” on estimated location and dispersion parameters (treated as “true values”).



Bayesian Predictions

the uncertainty of parameter estimates (location and dispersion) is taken into account before drawing “new observations” from relevant distribution



Why Bayesian for Biosimilarity?

- What is the question:
  - what is the probability of being biosimilar given available data?
  - what is the probability of having future lots within the limits given available data?

$$\Pr(\textbf{observed data} \mid \textbf{not biosimilar}) \quad \text{vs} \quad \Pr(\textbf{Biosimilar} \mid \textbf{observed data})$$

$$\Pr(\textbf{Future lots in limits} \mid \textbf{observed data})$$

- The question becomes naturally Bayesian
- Many decisions can be deduced from the posterior and predictive distributions
- In addition
  - leverage historical data (e.g. on assay variability)
  - Bayesian approach can easily handle multivariate problems

EFSPI

Biosimilar Model formulation

Biosimilarity Model - Univariate Case

$CQA_{Test} \sim N(\mu_{Test},\ \sigma^2_{Test})$ → Model for Biosimilar

$CQA_{Ref} \sim N(\mu_{Ref},\ \sigma^2_{ref})$ → Model for Ref

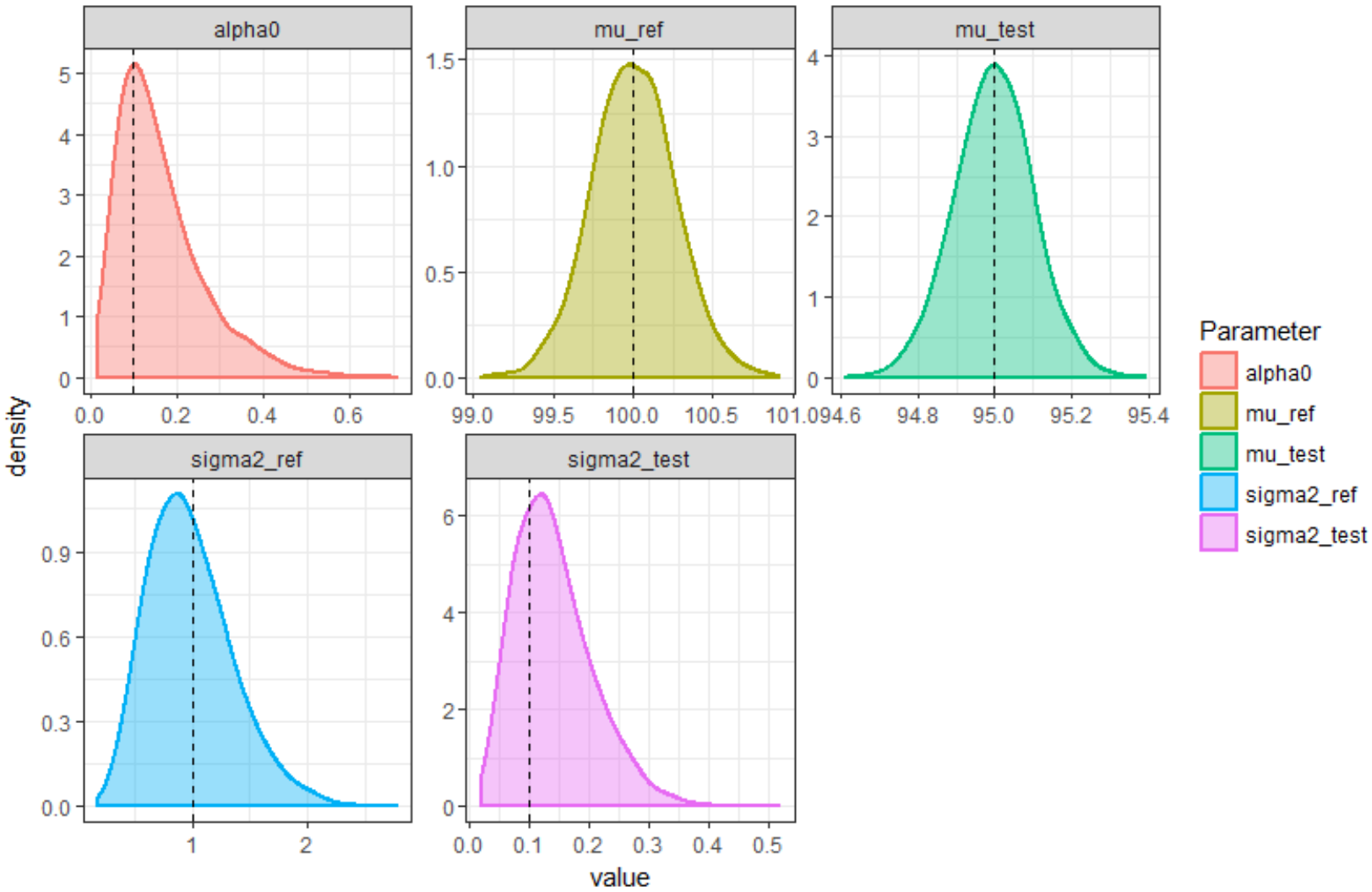
$\sigma^2_{Test} = \alpha_0 * \sigma^2_{ref}$ → Test will not extremely different from Ref

$\alpha_0 \sim$ Uniform (a, b), for well chosen a & b, e.g. 1/10 to 10

- From this model:
  - directly derive the PI/TI from predictive distributions
  - easily extendable to multivariate model
  - power computations are straight forward from predictive distributions

EFSPI

Model performance (compare to true pars)

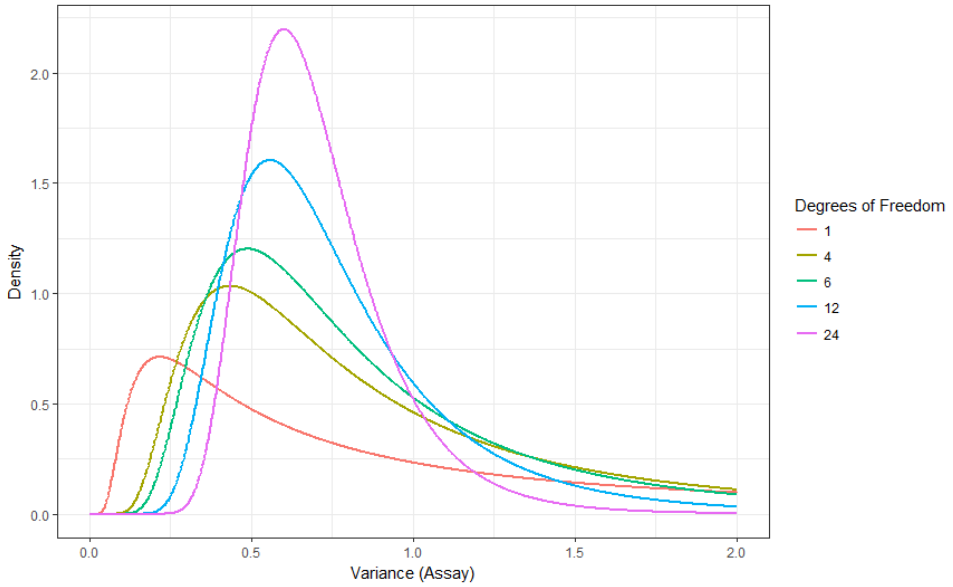


Biosimilarity Model - Univariate Case

- Variability can be decomposed to:

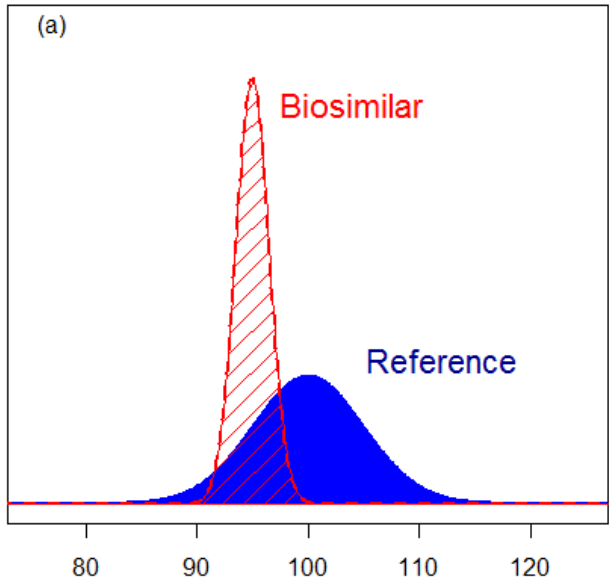
$$\sigma^2_{Test} + \sigma^2_{assay}$$
$$\sigma^2_{Ref} + \sigma^2_{assay}$$

- Synthesize assay historical data into informative prior for variability (all other pars being non-informative)

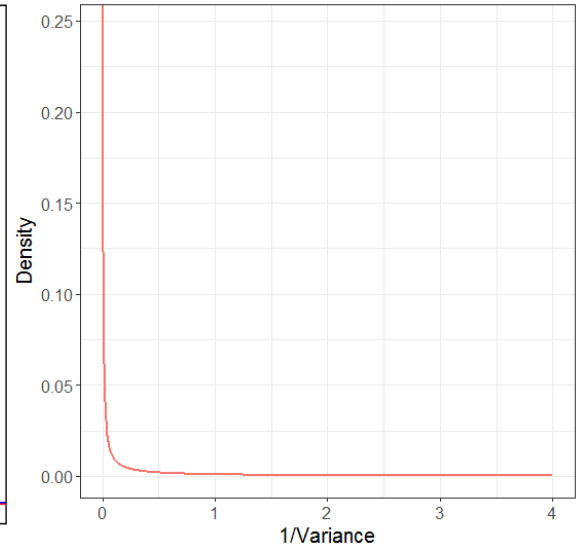


Bayesian PI/TI – Illustration (1)

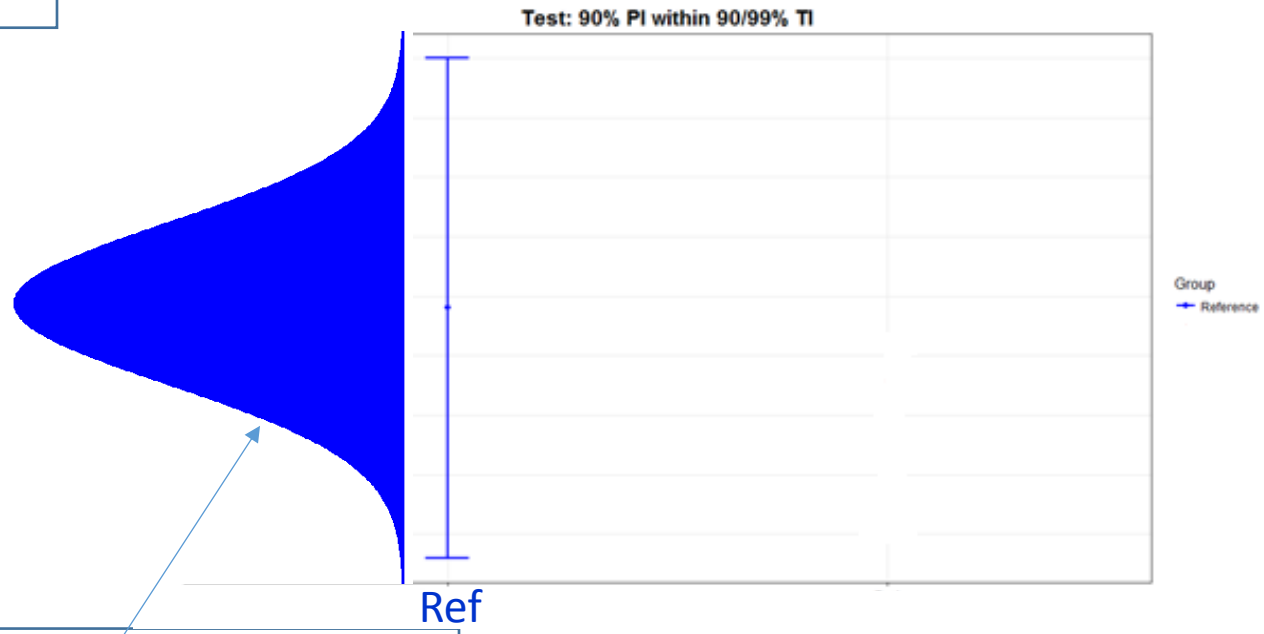
Likelihood



(non-informative **Prior** on all parameters)



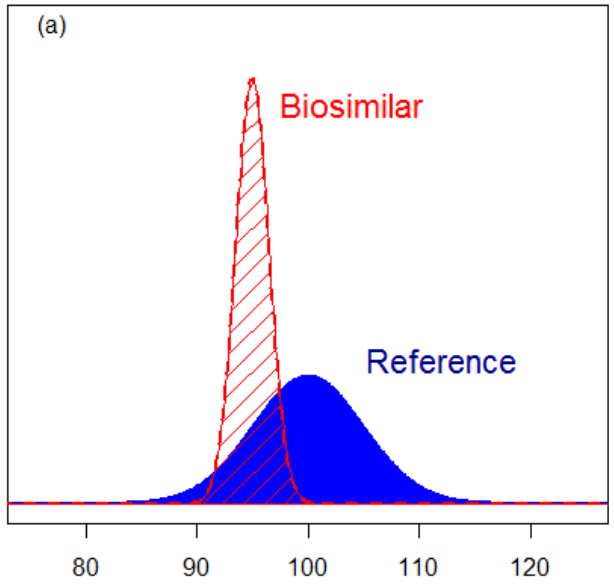
Predictive Distributions



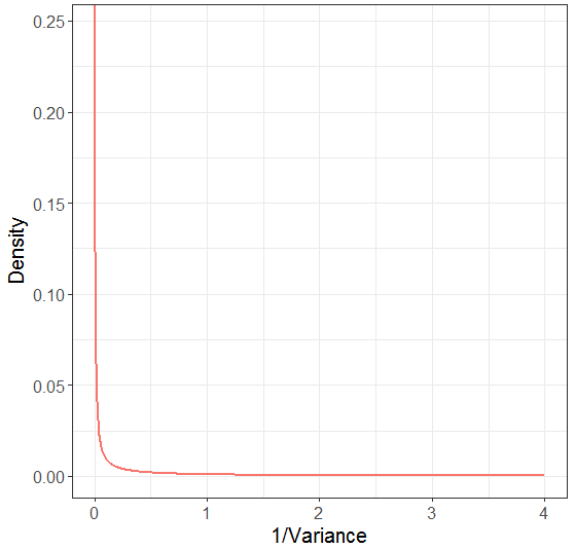
Predictive Distribution
➔ Tolerance Intervals (e.g. Wolfinger)

Bayesian PI/TI – Illustration (2)

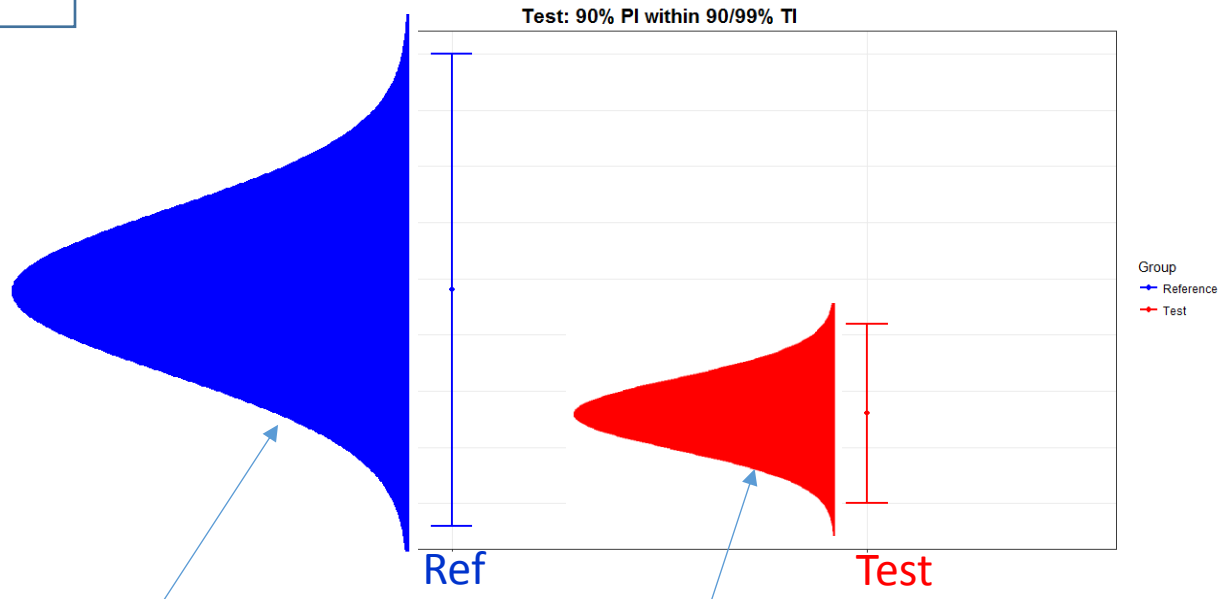
Likelihood



(non-informative **Prior** on all parameters)



Predictive Distributions

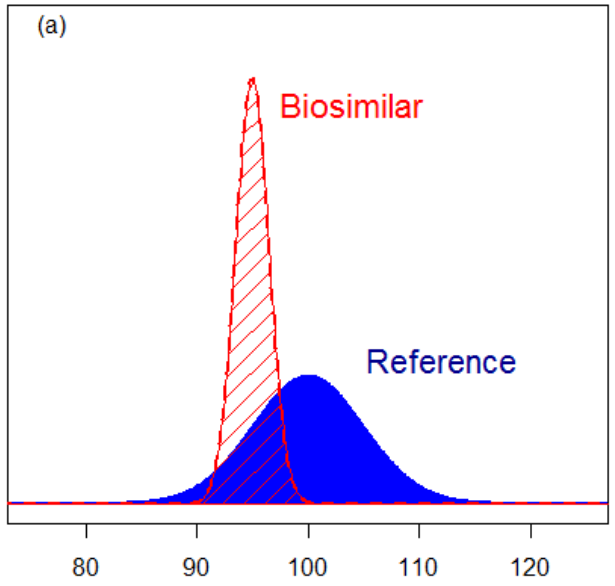


Predictive Distribution
➔ Tolerance Intervals (e.g. Wolfinger)

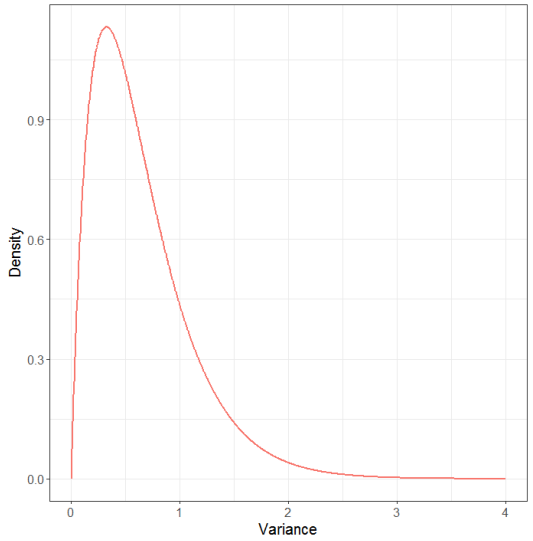
Predictive Distribution
➔ Prediction Interval

Bayesian PI/TI – Illustration (3)

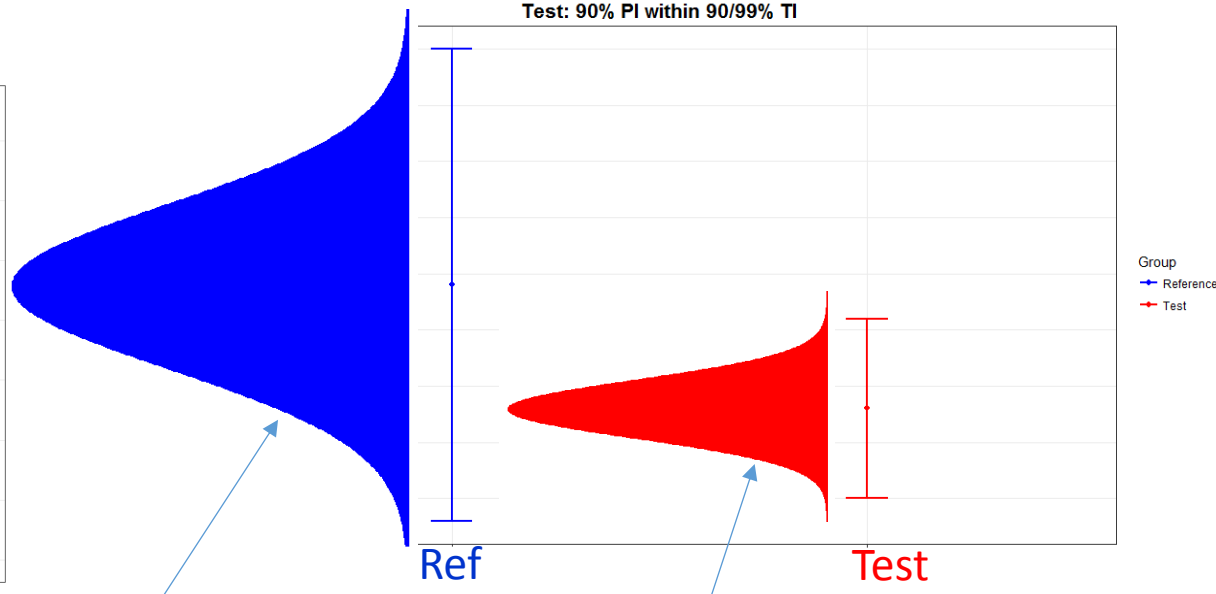
Likelihood



Prior
(informative on validated assay variance)



Predictive Distributions



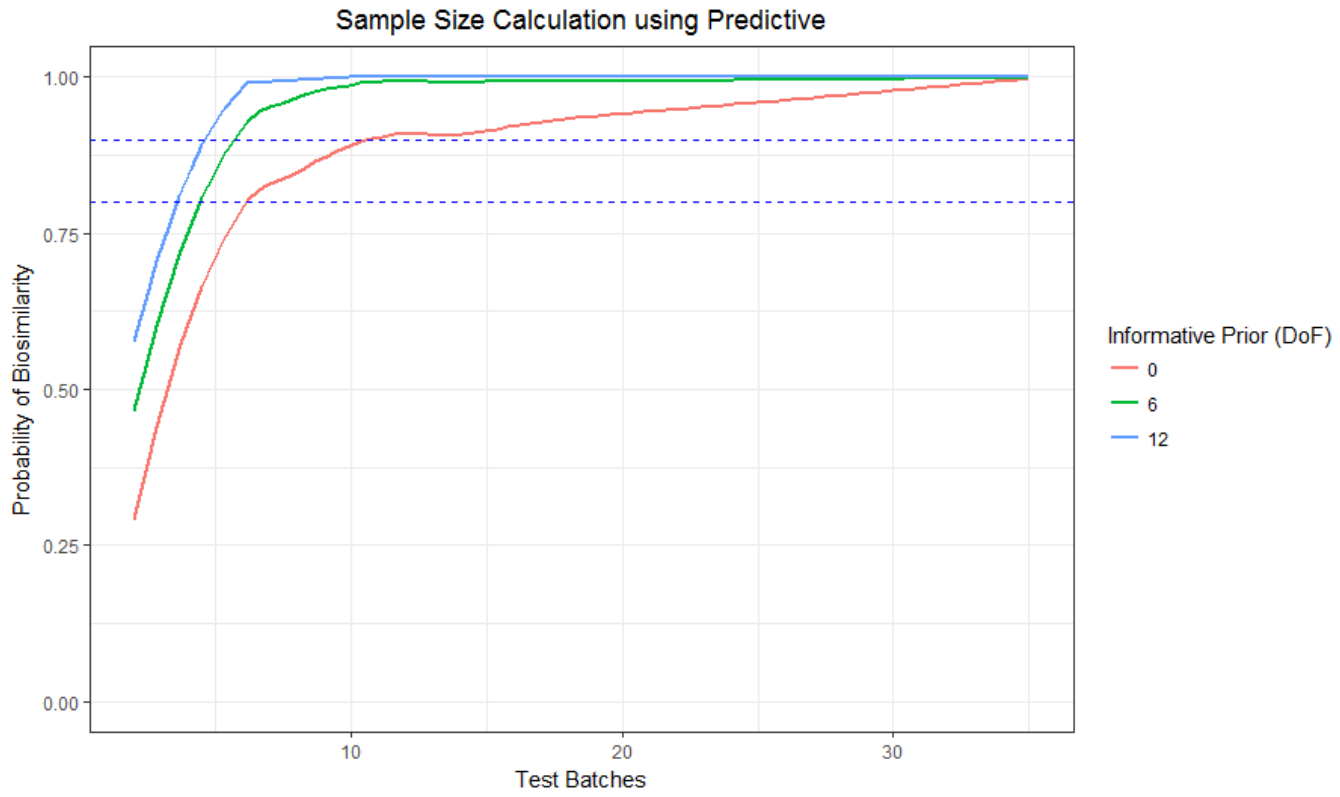
Predictive Distribution
➔ Tolerance Intervals (e.g. Wolfinger)

Predictive Distribution
➔ Prediction Interval

Sample Size Calculation

Sample size for Biosimilarity Evaluation

- Sample Test data from the predictive
  - How many new batches given past results to be within specs?



Multiplicity

Extension of the univariate case

Bayesian - Multivariate CQA Model

- Let
  - $\boldsymbol{X}$ be $n \times k$ matrix of observations for test.
  - $\boldsymbol{Y}$ be $m \times k$ matrix of observations for ref.

$$\begin{pmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \end{pmatrix} \sim \boldsymbol{MVN}\left(\begin{pmatrix} \boldsymbol{\mu_T} \\ \boldsymbol{\mu_R} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma_T} & \boldsymbol{\Sigma_{RT}} \\ \boldsymbol{\Sigma_{RT}} & \boldsymbol{\Sigma_R} \end{pmatrix}\right)$$

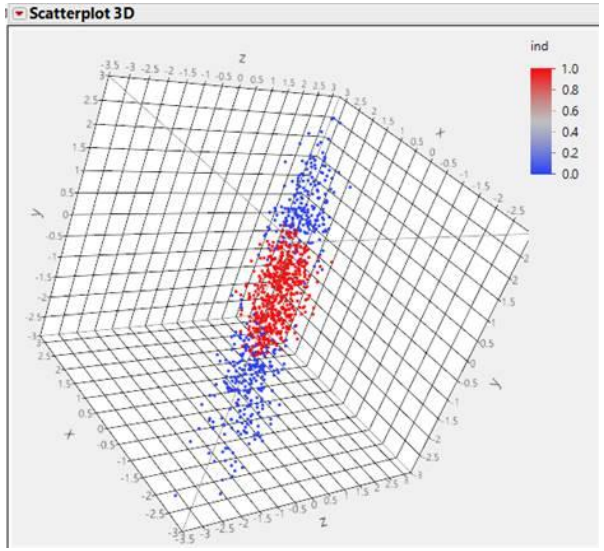
- Any test FDA Tier1, FDA Tier2 or PI/TI can be easily computed

- $\Pr([\mathbf{Test} - \mathbf{Ref}]|\mathbf{Data}) \sim \boldsymbol{MVN}([\boldsymbol{\mu_T} - \boldsymbol{\mu_R}], [\boldsymbol{\Sigma_T} + \boldsymbol{\Sigma_R} - \mathbf{2} * \boldsymbol{\Sigma_{RT}}])$

EFSPI

Multivariate CQA Model

- Use Ref. predictive to compute the limits of *k* CQAs
- Compare the Test data from *k* CQAs to the limits
- To get the joint test:
  - Calculate the joint acceptance probability



Assurance

(not Power)

Assurance (Bayesian Power)

- Unconditional probability of significance given prior - O'Hagan et al. (2005)

- Expectation of the power averaged over the prior distribution
  - 'True Probability of Success' of a trial

- In Frequentist power is based on a particular value of the effect
  - A very 'strong' prior

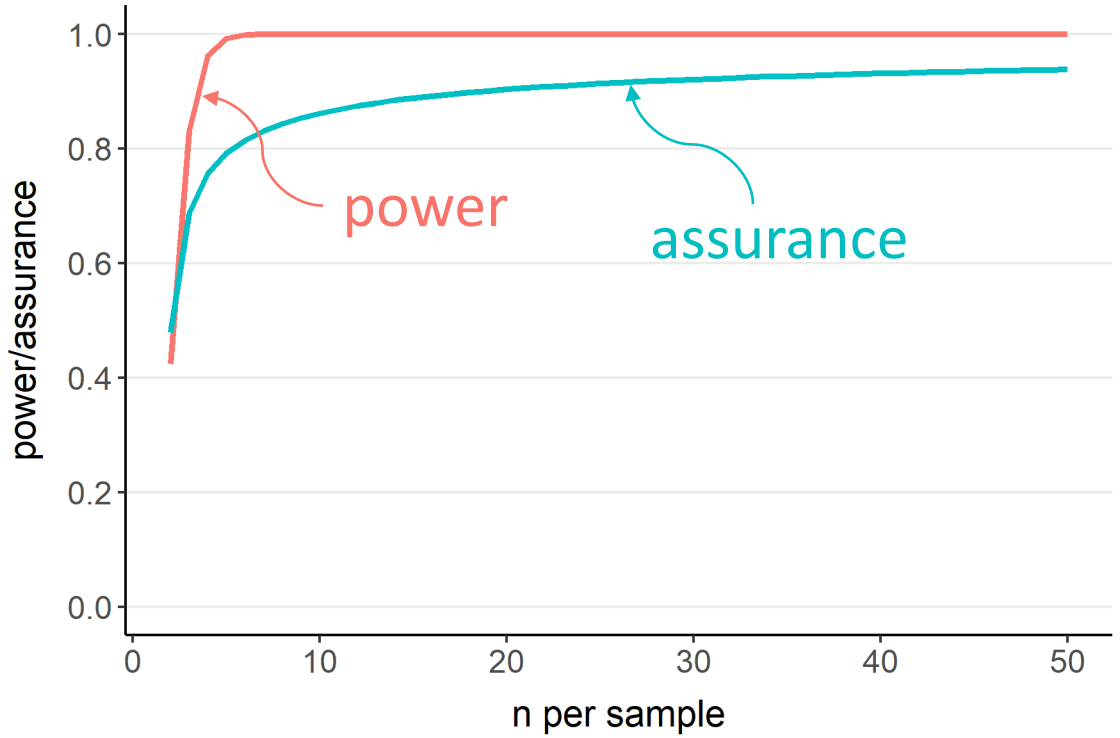
EFSPI

Power vs assurance

independent samples t-test ($H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$)

bayesian approach (assurance)

- In order to reflect the uncertainty, a large number of effect sizes, i.e. $(\mu_1 - \mu_2)/\sigma_{\text{pooled}}$, are generated using the prior distributions.
- A power curve is obtained for each effect size
- the expected (weighted by prior beliefs) power curve is calculated



Conclusions

- Using Bayesian approach:
  - I can directly derive probabilities of interest
  - Uncertainties are well propagated
- Bayesian predictive distribution answers the very objective
  - probability of biosimilar given data
  - future lots to remain within specs
  - leverage historical data → save costs
- Informative priors can be justified and recommended
- Correlated CQAs
  - Easily compute joint acceptance probability

EFSPI

When SIMILAR is not the SAME!