BAYES BAYESIAN BIOSTATISTICS LYON 2019

TUESDAY 21ST FRIDAY 24TH MAY 2019

MARRIOTT HOTEL CITÉ INTERNATIONALE, LYON, FRANCE

SANOFI BIOMÉRIEUX PHARMALEX CONFIDENCE BEYOND COMPLIANCE

From Power and Assurance to Bayesian Power:

Application to probability of success

Arnaud Monseur| Manager , Pharmalex - Belgium

Karine Imberdis | Sanofi- Montpellier France

Didier Poirault| BioMérieux – Marcy L'Etoile France

BAYES 2019 - Lyon

May 2019

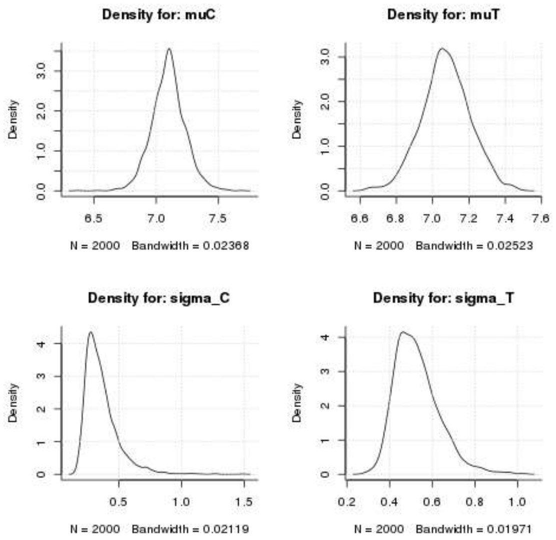
Introduction/ Context

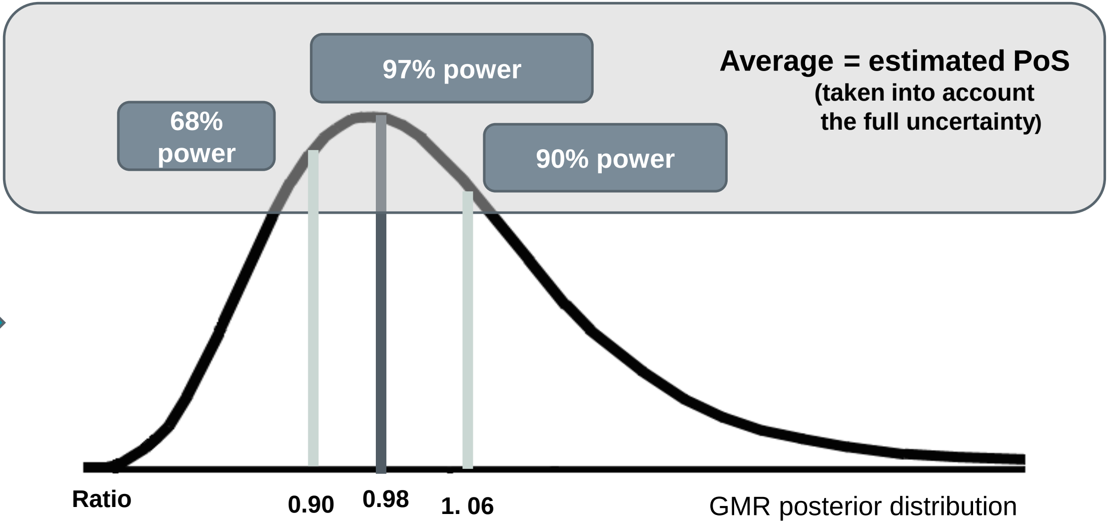
- **Bayesian network in France and Belgium (ARM of BAYES congress people!)**

- Study case presented:
  - In clinical early phase, estimation of the probability of success (PoS) of a second study taken into account the data of a first study:
    - Previous : Pilot study observed 24 patients (16 G1 and 8 G2))
    - Future: Similarity study with N=188 (94 patients per group G1 and G2)
    - Success rule criteria (frequentist) : 90% CI of the geometric mean ratio (GMR) is within [0.80; 1.25]
  - What is the PoS (probability of the success) for the future study based on the previous study data?

Introduction/ Context Assurance

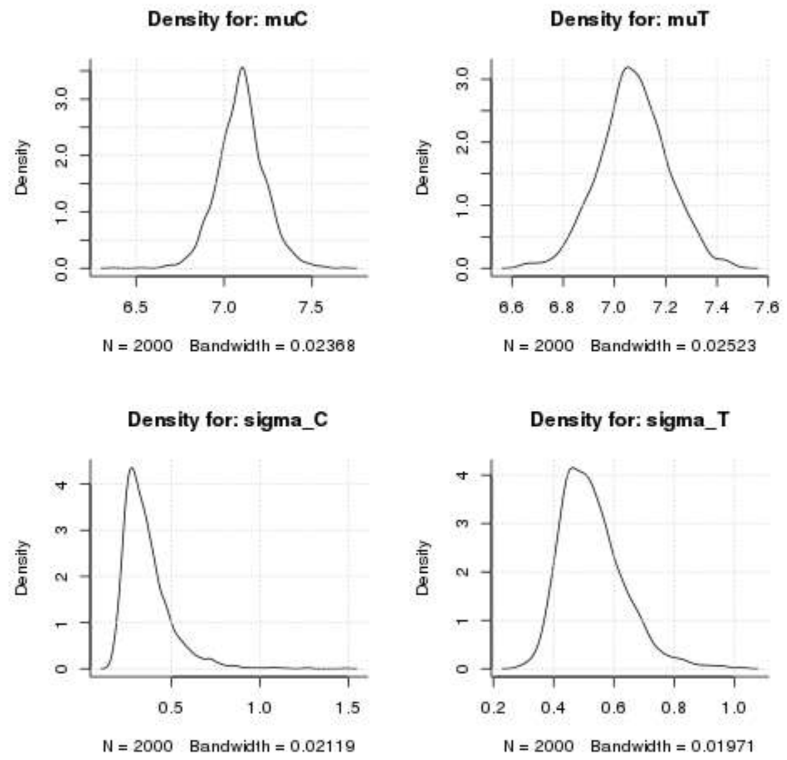
- **Easy to obtain the posterior distribution using a classical Bayesian analysis with non informative prior on the first pilot study data**





PoS = 58.5%

Introduction/ Context
Bayesian Power



- From each sample of the posterior densities:
  - Simulate a study
  - Assess whether it is a success yes or no
  - Repeat and take the mean overall all successes.

- ***PoS = 50%***

Different way to estimate the POS

- Explanation of difference?
- How to do concretely?
- What are the recommendations?

Power and Assurance: What is What ?

(Frequentist) Power

- Let R denote the rejection of the null hypothesis, the power is , assuming parameter values of $\boldsymbol{\theta} = \boldsymbol{\theta}^*$

$$\pi(\boldsymbol{\theta}^*, n) \coloneqq \Pr(R|\boldsymbol{\theta}^*, \boldsymbol{n})$$

- It is a conditional probability. It is conditional on the parameters of the model, e.g. the “true effect size” in a frequentist test and the sample size.

Assurance

- *“Assurance is the unconditional probability that a trial will lead to a specific outcome”*

$$\gamma(n) \coloneqq \int \pi(\theta, n) f(\theta) d\theta$$
$$\gamma(n) \coloneqq \Pr(R) = E_\theta[\pi(\boldsymbol{\theta})]$$

It is thus also a function of n (and eventually other nuisance parameters)

The assurance is the expected power over all possible values of theta (-> over its prior distribution…)

Assurance

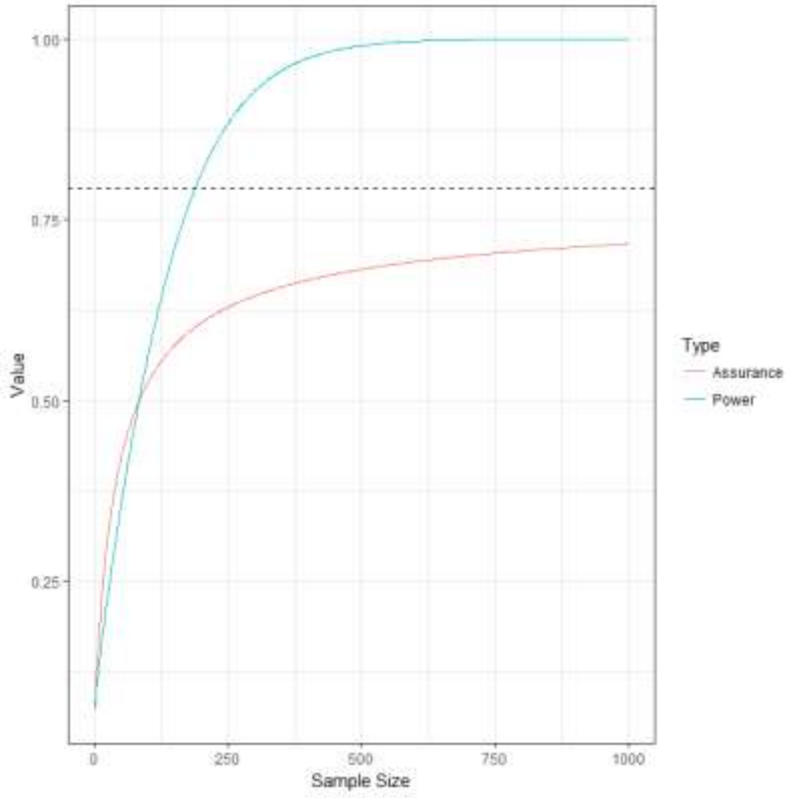
- Assurance does not necessarily converge to one when n increases

- Intuitively this is due to the fact that there can exist non-zero probability that we do not reject the null hypothesis.

- More precisely if the parameter space associated with “not H1” is a non-negligible set then assurance will not converge to one.

An Example: A simple T- test

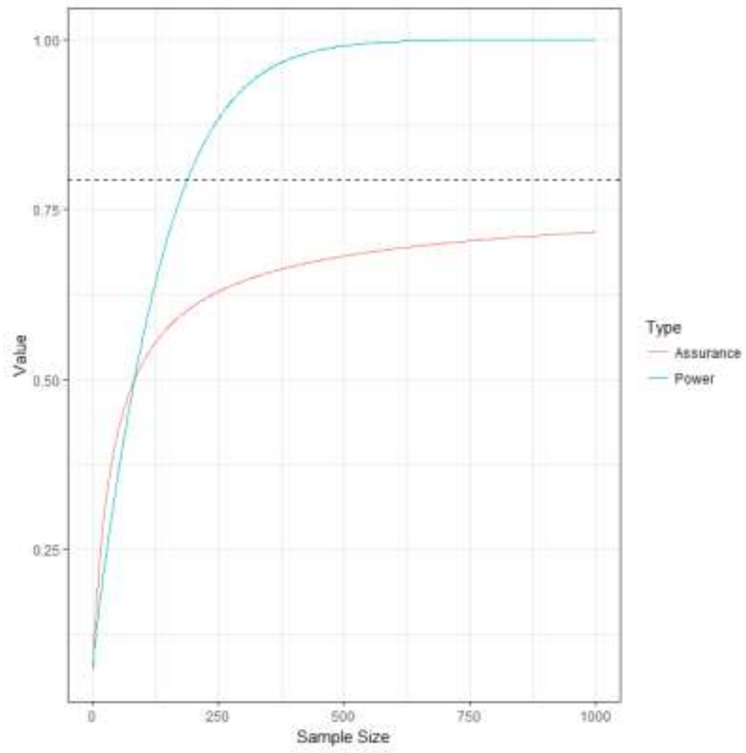
The assurance converges to 0.793, that is the prior probability that the new drug is indeed superior. (assumed in this example)



Link to Probabilities of Success

- This can also be expressed in terms of probability of success quite easily

- Imagine that in the previous example rejecting the null hypothesis is considered as a success. This implies:
  - Frequentist Power will always allow to have a perfect probability of success if the “true effect” is in the rejection region
  - Assurance will never give a higher probability of success than that which is assumed in the prior belief
  - /!\ At this stage assurance is still a frequentist concept ! /!\



Let’s Make things Bayesian

Going Further to Bayesian Methods

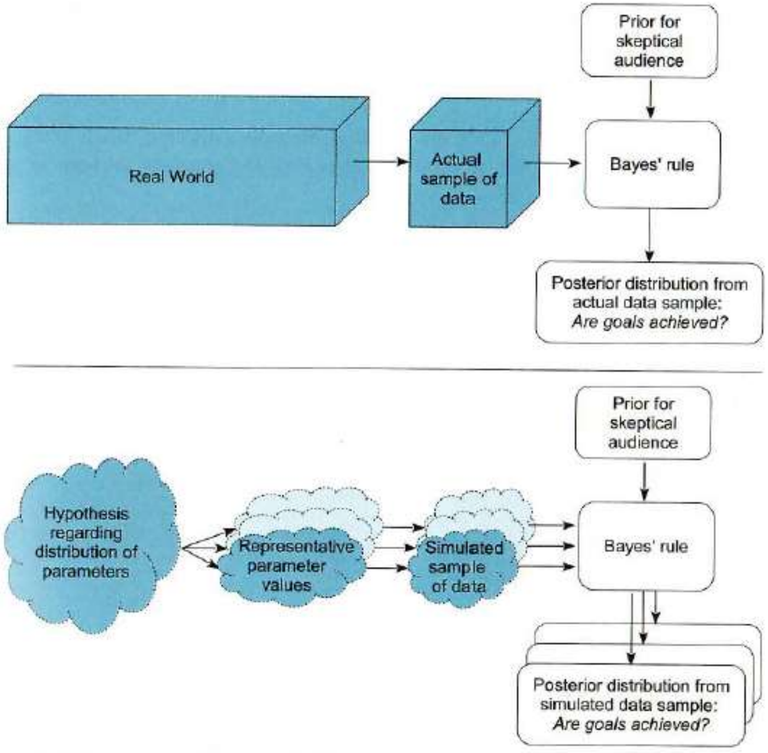
- Assurance is a relatively old concept
  - It predates the golden age of Bayesianism that we now live in

- Power calculation in the Bayesian methodology (Chapter 13: Goals, Power, and sample size p359-398 in Doing Bayesian Data Analysis A tutorial with R,JAGS and Stan, John K.Kruschke, Edition 2, 2015 ,Elsevier)
  - Differentiates power in the NHST world (Null Hypothesis Significance Testing) which is the rejection of the null from Bayesian power which can accommodate other goals and sampling plans and hypotheses

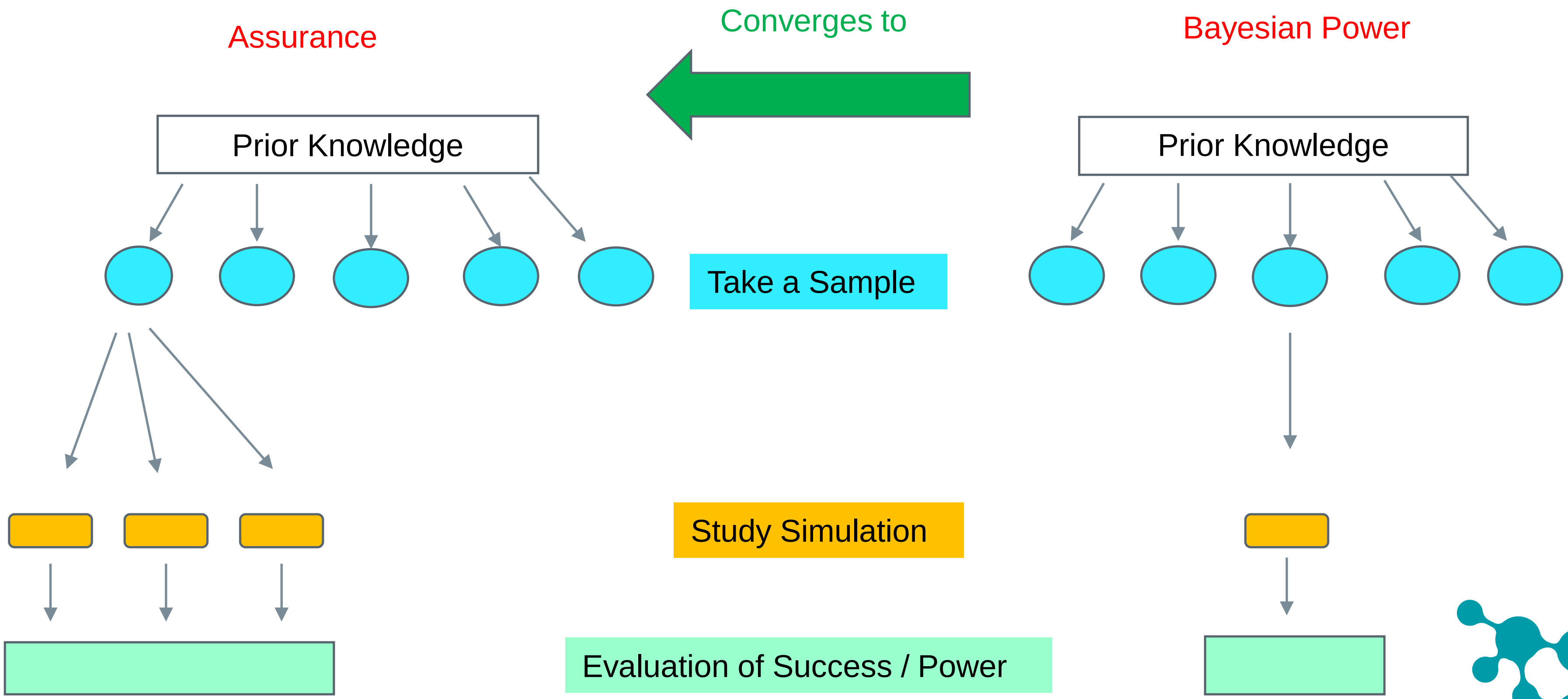
Going Further to Bayesian Methods

- From the hypothetical distribution of parameter values, generate representative values.
- Generate data from these values
- Compute the posterior on this sample using appropriate Bayesian analysis
- Tally , from the posterior if the goal is achieved
- Repeat

- THIS IS NEARLY ASSURANCE
  - It is assurance where power at each value of the prior is approximated by a unique sample (1-0)



Bayesian Power vs Assurance



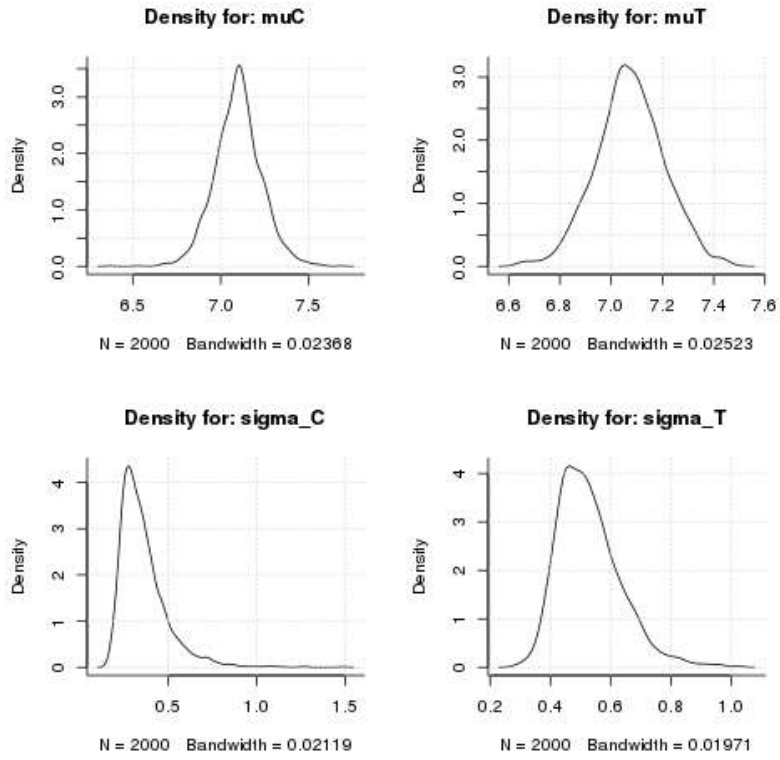
Back To the Real Case Study

Different Methods to Compute PoS

- Assurance
  - Directly from posterior distribution:
    - Calculation of power for each difference and precision of the posterior distribution
    - PoS=Mean of power
- Bayesian Power (as described by Kruschke)
  - From Posterior Distribution:
    - Calculation of the success or failure for difference and precision of the posterior distribution
    - PoS = Mean number of successes
  - It is an assurance methodology where the power is computed on an unique sample

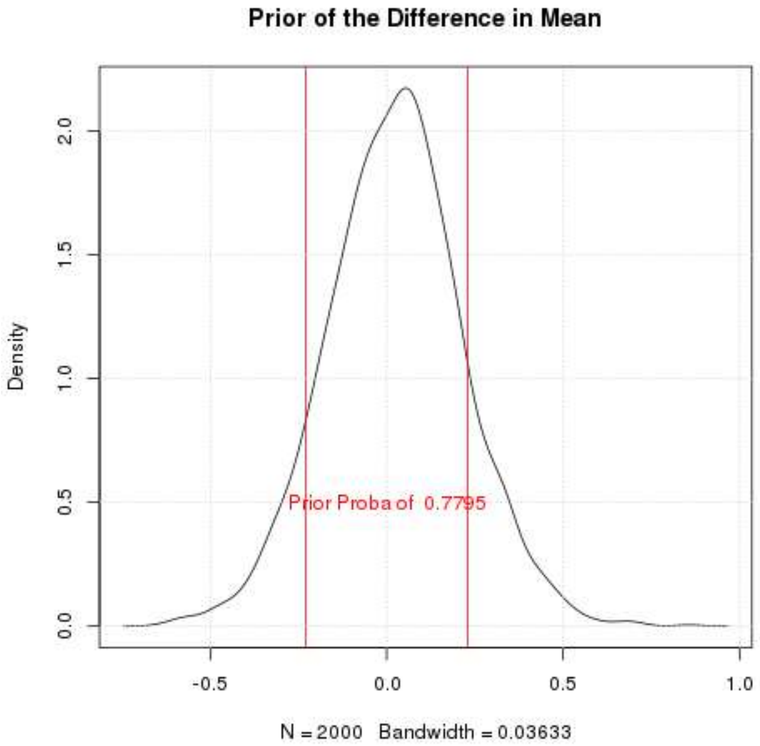
Posterior distributions from previous studies estimation

- From historical data we can obtain the following distributions.
- These are our prior probabilities for constructing our beliefs of future studies
- The aim is to assess if the 90% credible interval of the difference between the two means is contained within the interval [-0.23;0.23]



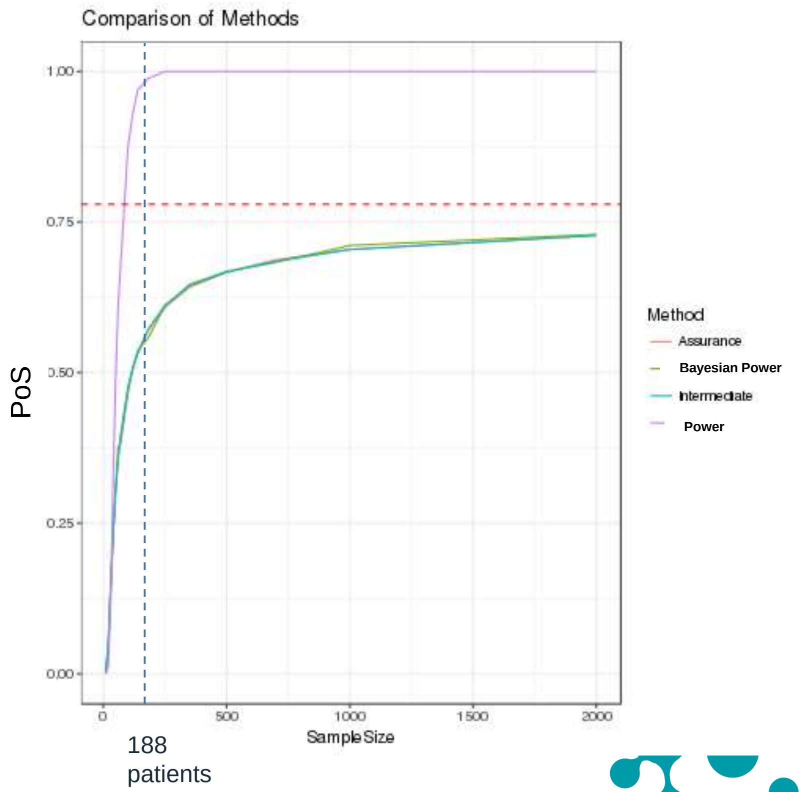
Comparison of Methods: A practical case

- The difference in mean has a prior probability of .7795 of being within the defined equivalence margins



Comparison of Methods: A practical case

- We see that the profiles for the power built on the predictive methodology converges to 1 very quickly.
- For very small sample sizes the three other methods actually yield a higher “power”
- The three other methods converge to +-0.78, that is the prior probability that the difference between the two groups is actually within the equivalence margins defined
- Dashed line represent the case with a total sample size of 188 patients



Conclusion

Conclusion

- Power in the frequentist view does not account for prior believes
- Power converges to 1 if the condition is in the acceptance region
- Bayesian Power and Assurance (asymptotically equivalent) do not converge to one because under the prior believes, there could exist a chance that we will never be able to show what we desire
- This reflects better the true world probabilities