Predictions vs. Simulations

in early clinical development: a framework to evaluate the predictive probability of success based on NONMEM outputs

Bruno Boulanger, Arlenda S.A. & Université de Liège, Belgium
Astrid Jullion, UCB, Belgium
Pierre Lebrun, Arlenda S.A. & Université de Liège, Belgium

Introduction

Difference between Predictions and Simulations

How to perform predictions

How to perform predictions from NONMEM

An example of dose ranging study design

Major impact

Avoiding approximations

Conclusions

AGENDA

The objective

In Early Clinical development

- By means of modeling is to identify the range of doses, if any, that will guarantee:
  - Efficacy in future late phase trials
  - Safety in the future
- To minimize to risks of investing in low success but costly confirmatory trials
- To power/design adequately confirmatory trials to ensure success (if go decision is taken) ie efficacy and safety.

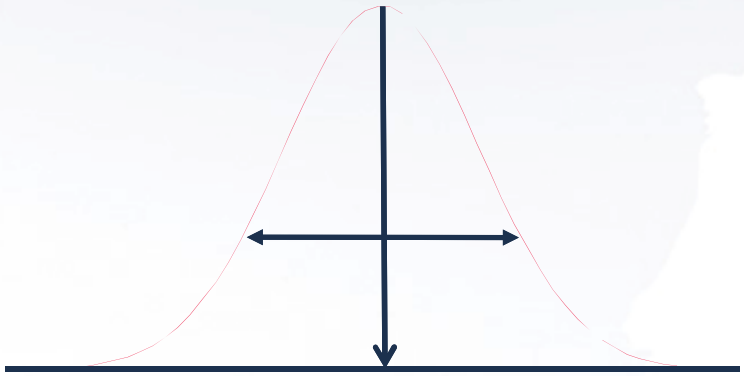
Facts

- The number of subjects or patients is usually limited
- The “PKPD” is estimated with uncertainty
- Pre-clinical, historical data or competition prior information is usually available
- Efficacy biomarkers are usually available

Difference Simulations/Predictions

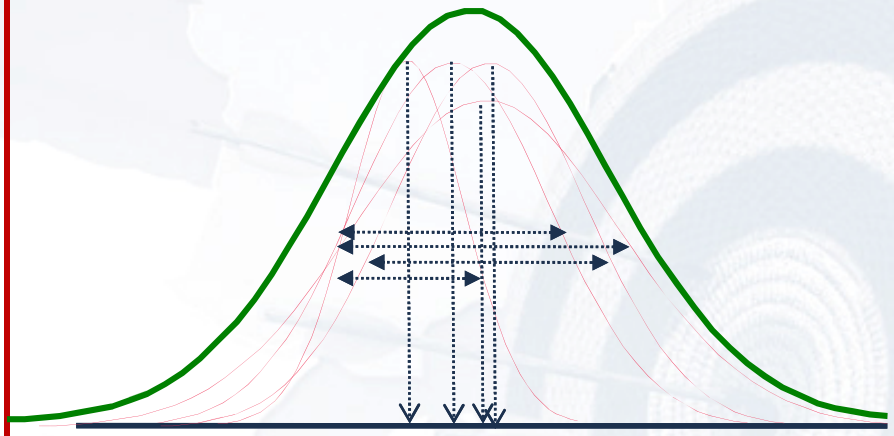
Simulations

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



Predictions

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



Bayesian Predictive Distribution

The Bayesian theory provides a definition of the Predictive Distribution of a new observation given past data.

$$p(\tilde{x}|data) = \iint_{\sigma^2\,\mu} p(\tilde{x}|\mu,\sigma^2,data) \times p(\mu,\sigma^2|data)\; d\mu d\sigma^2$$

Integrate over parameter distribution

Model

Joint posterior

$$= \iint_{\sigma^2\,\mu} p(\tilde{x}|\mu,\sigma^2) \times p(\mu|data) \times p(\sigma^2|\mu,data)\; d\mu d\sigma^2$$

$$= \iint_{\sigma^2\,\mu} p(\tilde{x}|\mu,\sigma^2) \times p(\sigma^2|data) \times p(\mu|\sigma^2,data)\; d\mu d\sigma^2$$

Model

Marginal

Conditional

Sampling from Predictive Distribution

Simple case

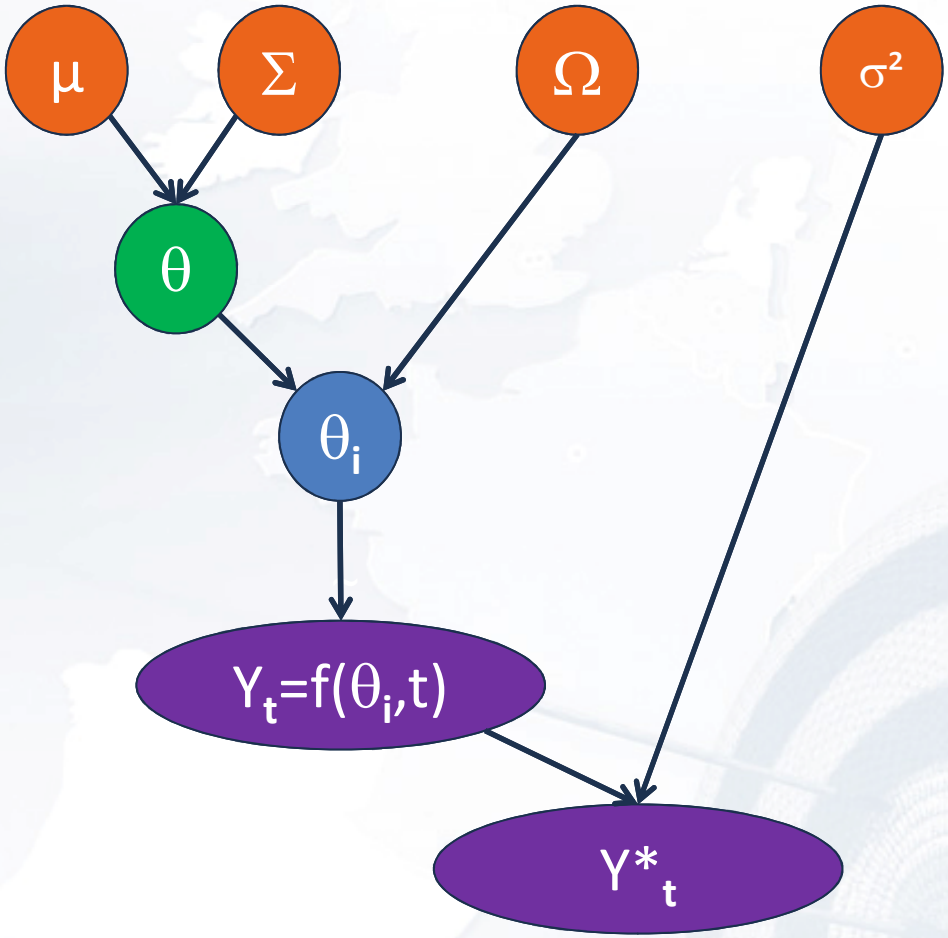
$$p(\tilde{x}|data) = \int_{\sigma^2}\int_{\mu} \underbrace{p(\tilde{x}|\mu,\sigma^2)}_{\text{Model}} \times \underbrace{p(\sigma^2|data)}_{\text{Marginal}} \times \underbrace{p(\mu|\sigma^2,data)}_{\text{Conditional}}\ d\mu d\sigma^2$$

This suggests the following algorithm for sampling from the predictive distribution for a simple model:

1. **Sample σ^2_i from $p(\sigma^2|data)$, ie a InvGamma**
2. **Sample μ_i from $p(\mu|\sigma^2_i,data)$, ie a Normal**
3. **Sample x_i from $p(x|\mu_i, \sigma^2_i)$, ie a $N(\mu_i, \sigma^2_i)$**

Hierarchical (mixed) PKPD Model

- PKPD Structural Model
  - $dA/dt = f_A(\theta,t)$
  - $dB/dt = f_B(\theta,t)$
  - $dC/dt = f_C(\theta,t)$
  - $dD/dt = f_D(\theta,t)$
- Hierarchical model
  - $\Theta_i \sim N(\Theta, \Omega)$
  - $\Theta \sim N(\mu,\Sigma)$
  - $\varepsilon_i \sim N(0,\sigma^2)$



Sampling for hierarchical/mixed model

When (a linear) hierarchical model is envisaged with several parameters, the following algorithm for sampling from the predictive distribution is to be envisaged:

1. Population level sampling
    1. τ_j from $p(\tau|data)$, ie a Wishart(ρ,υ) (note $\rho=(\upsilon\Sigma_\theta)^{-1}$)
    2. θ_j from $p(\mu|\tau_j,data)$, ie a MultNormal(μ, τ_j)
2. Individual level sampling
    3. φ_i from $p(\varphi|data)$, ie a Wishart(ψ,ν) (note $\psi=(\nu\,\Omega)^{-1}$)
    4. θ_{ji} from $p(\theta|\ \theta_j, \varphi_i,data)$, ie a MultNormal$(\theta_j, \varphi_i)$

    ➔ Predicted PKPD profile for individual i : $y_{i,t}=f(\theta_{ji},t)$
3. Residual error level sampling
    5. χ_j from $p(\chi|data)$, ie a InvChisq

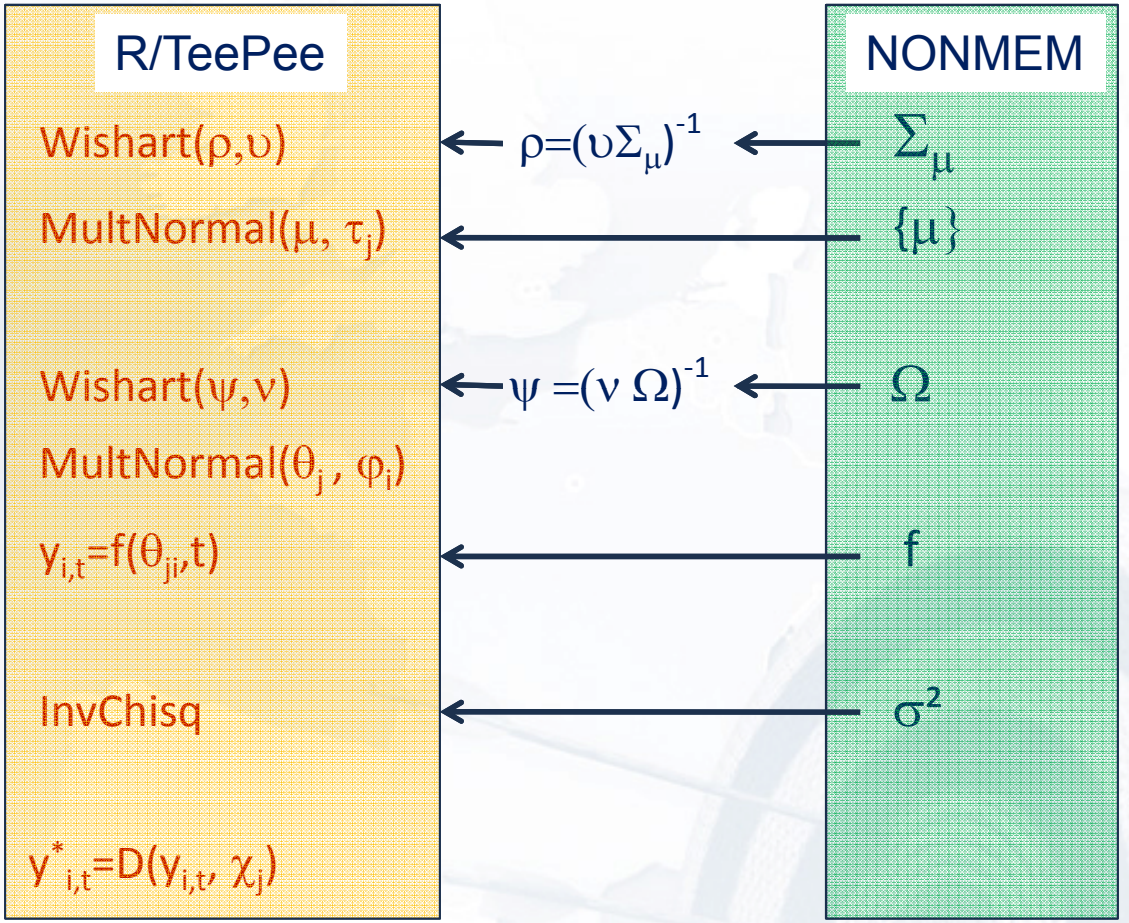
    ➔ Predicted PKPD observation for individual i : $y^*_{i,t}=D(y_{i,t}, \chi_j)$

Note: D distribution depend on error model (Normal, LogNormal,….)

How to proceed with NONMEM?

1. Population level
   1. τ_j from $p(\tau|data)$
   2. θ_j from $p(\mu|\tau_j,data)$
2. Individual level
   3. φ_i from $p(\varphi|data)$
   4. θ_{ji} from $p(\theta|\ \theta_j, \varphi_i,data)$
   → (the model)
3. Residual error level
   5. χ_j from $p(\chi|data)$

→ Predicted observation :



A typical study

- A study conducted on patients (eg Phase IIa)
- A dose ranging study, using several doses
  - Placebo=0, Low=0.05, Medium=0.1, High=1
- Study to be sized to ensure 80% power Dose 0.1 > Placebo
- Simulations and predictions are performed
  - using the model and parameter estimates from previous Ph I studies
  - NONMEM population model was used
  - PK and PD parts estimated separately

A PK-PD model available

PK-PD Model

dA/dt =-KA*A+Ri

dB/dt = KA*A-CL/V1*B-Q/V1*B+Q/V2*C

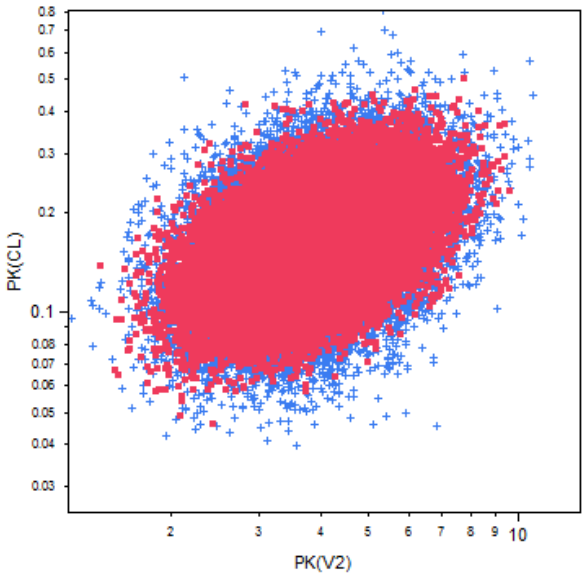
dC/dt = Q/V1*B-Q/V2*C

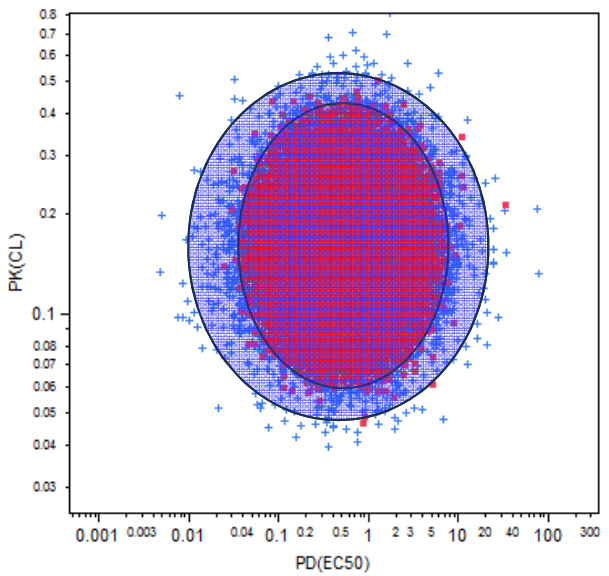
dD/dt = KE0*((B/V1)-D)

dE/dt = kout*Base*(1-EMAX*C1/(EC50+C1))-kout*E

Parameters are sampled

- for Simulations
- for Predictions





Remarks

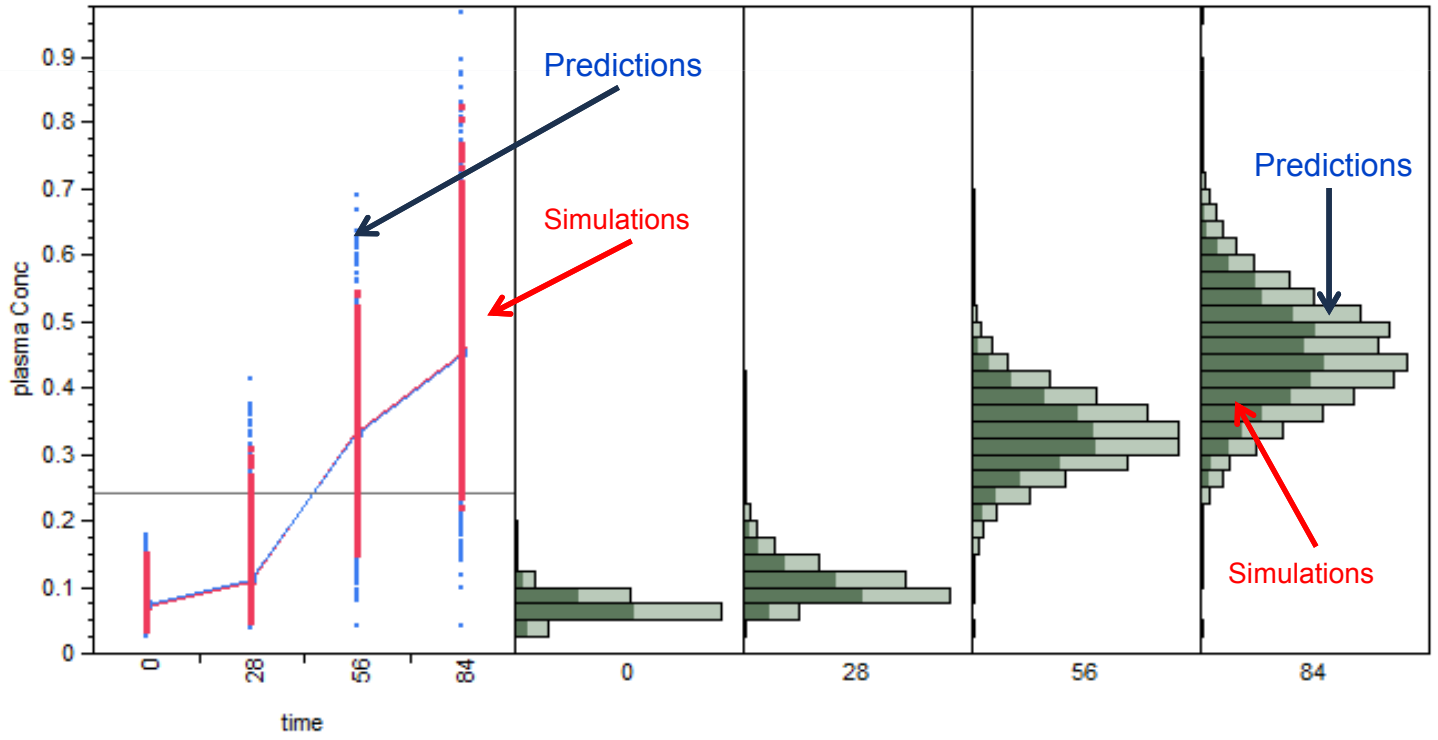
- Given models are mostly nonlinear, these samplings distributions are approximations.
  - No alternative when working from frequentist estimates
  - Approximations are reasonable (Aitchison, 1975)

- Working with full Bayesian methods (Bugs) provides directly joint “correct” posterior distributions.
  - Bayesian models provide “naturally” prediction, the very objective in early phases

Average and spread

Average are the same but spread is larger for predictions than for simulations

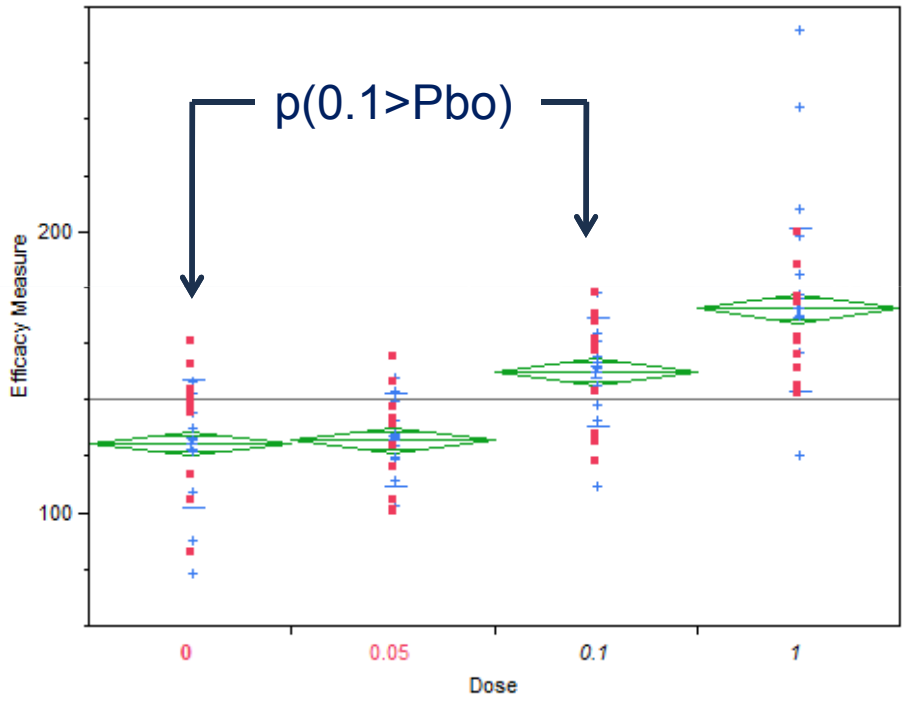
➔ Power of confirmatory trial !



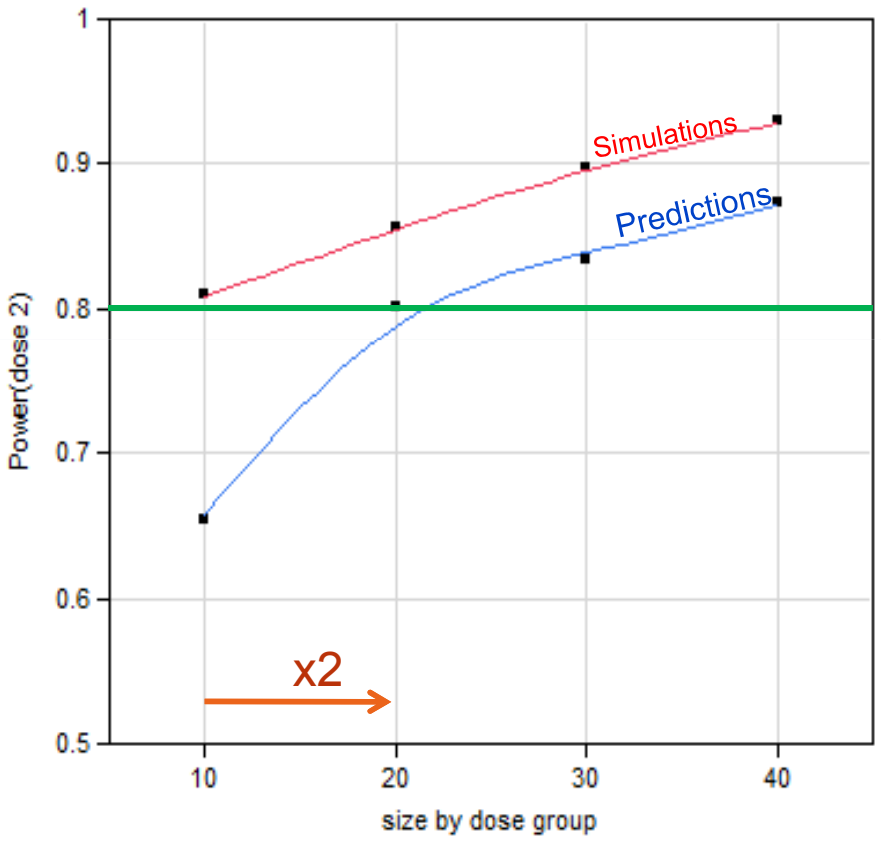
Sizing the study

The sample size is determined to ensure 80% power that dose 0.1 will be significantly greater than Placebo at 8 weeks.

Simulations and Predictions are used to power the study



Impact on Power



Ignoring parameter uncertainty (simulations)

- **10** pts/ group required

Taking into account parameter uncertainty (predictions)

-**20** pts/ group required

Predictions and Power

- Risk underpowering is high when simulations are based on limited sample size and using estimates as “true” values for simulations.
- There are ~50% chance parameters being worst than estimates, so ~50% chance of underpowering.
  - Note: average industry, 50% phase III fail mainly because of lack of demonstrated efficacy.
  - S.J. Wang, FDA, Shanghaï 2008, recommends taking uncertainty of parameters into account when powering.
- This is particularly important in early phases when decision should be made about future confirmatory studies.
- Think about safety! Ignoring uncertainty could lead to underestimate risks.

Arlenda
your partner in drug development

Conclusions

- Trial Prediction taking into account parameter uncertainty is recommended in early phases when “simulating” and powering late phase trials from limited number of patients.
- Assuming non-informative priors, a pretty good approximation is feasible in R using appropriate marginal and conditional distributions (Wishart, MultNormal, InvChisq,…)
- Using Bayesian methods in early phase is preferred because it provides directly predictions without approximations.
- Using Bayesian methods allows to integrate prior information from pre-clinical works and literature.

bruno@arlenda.com

THANK YOU