Alternative to Resampling Methods in Maximum Likelihood Estimation for NLMEMs by Borrowing from Bayesian Methodology

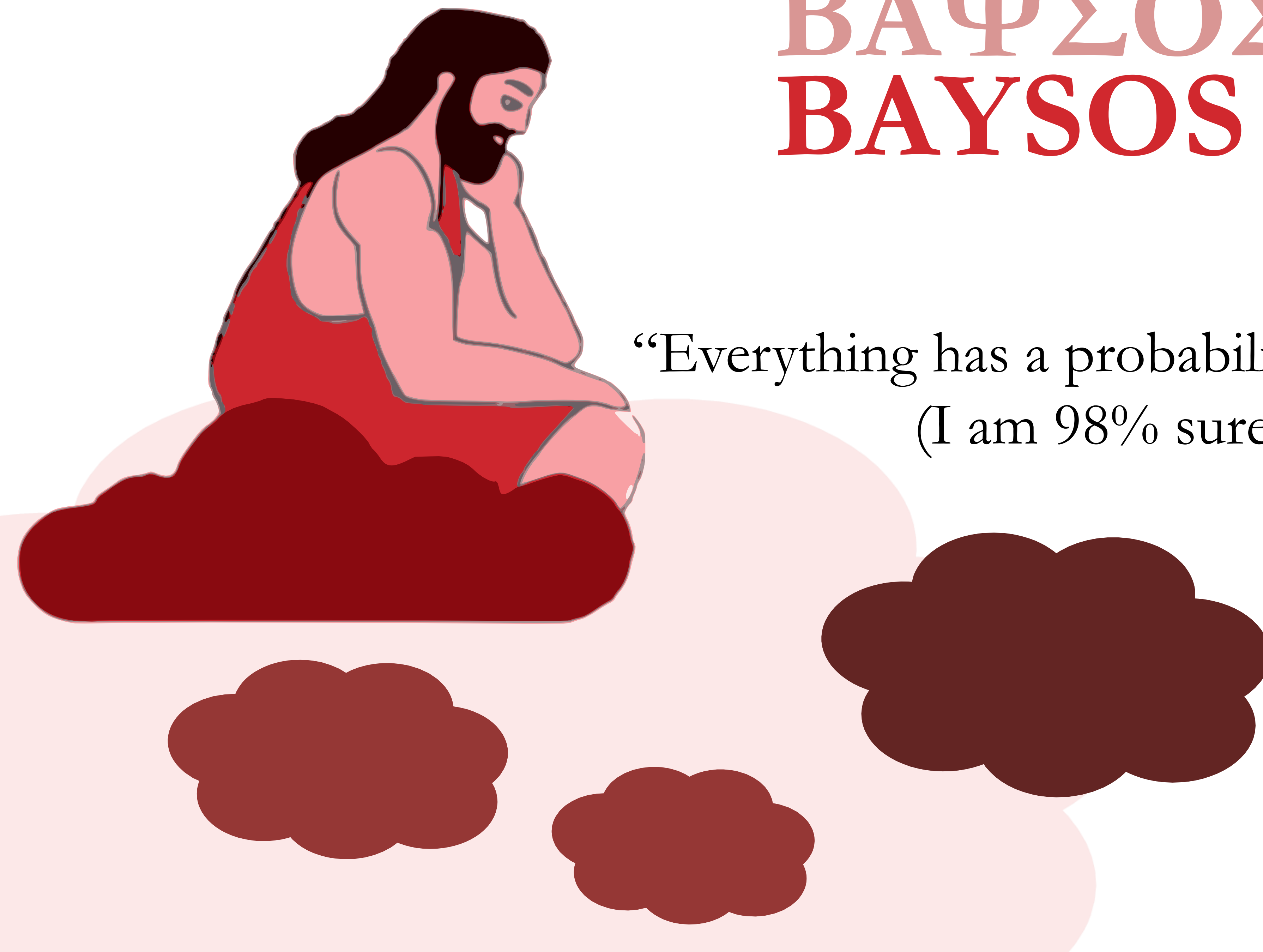
Sebastian Ueckert (1,2), Marie-Karelle Riviere (1), France Mentré (1)

(1) IAME, UMR 1137, INSERM and University Paris Diderot, Paris, France
(2) Pharmacometrics Research Group, Uppsala University, Uppsala, Sweden

ΦΡΕΘΥΕΝΤΙΑ
FREQUENTIA

"It's all about repetition"

ΒΑΨΣΟΣ
BAYSOS



"Everything has a probability
(I am 98% sure)"

ΛΙΚΕΛΙΗΟΟΔΟΣ
LIKELIHOODOS

“Likelihood, what else?”

ΠΗΑΡΜΑΞΟΜΕΤΡΙΞΥΣ
PHARMACOMETRICUS

“Can't I get some help for Baysos?”

Confidence intervals

- Quantify "confidence" in parameter estimate for parameter θ_k ($\theta = (\mu, \omega^2, \sigma^2)^T$)

μ… fixed effects ω^2… IIV variances σ^2… RUV variances

	Calculation	Assumption
Asymptotic (asymp)	`$COVARIANCE (NONMEM)`	$\hat{\theta}_k \sim \mathcal{N}\left(\theta_T, \mathcal{I}_k^{-1}\right)$
Log-likelihood profiling (llp)	`llp` (PsN)	$\frac{\mathcal{L}(\hat{\theta}_k; y)}{\mathcal{L}(\theta_1, \ldots, \hat{\theta}_k, \ldots \theta_p; y)} \sim \chi_1^2$
Non-parametric bootstrap (boot)	`bootstrap` (PsN)	$\hat{F}_n(y; x) \sim F(y; x)$ x … covariates

More approaches: Parametric bootstrap, multidimensional llp, sampling importance resampling, …

P-values

- Quantify significance level of a hypothesis test

$$H_0: \theta_k = \theta_0 \qquad H_1: \theta_k \neq \theta_0$$

Computational effort

Assumptions*

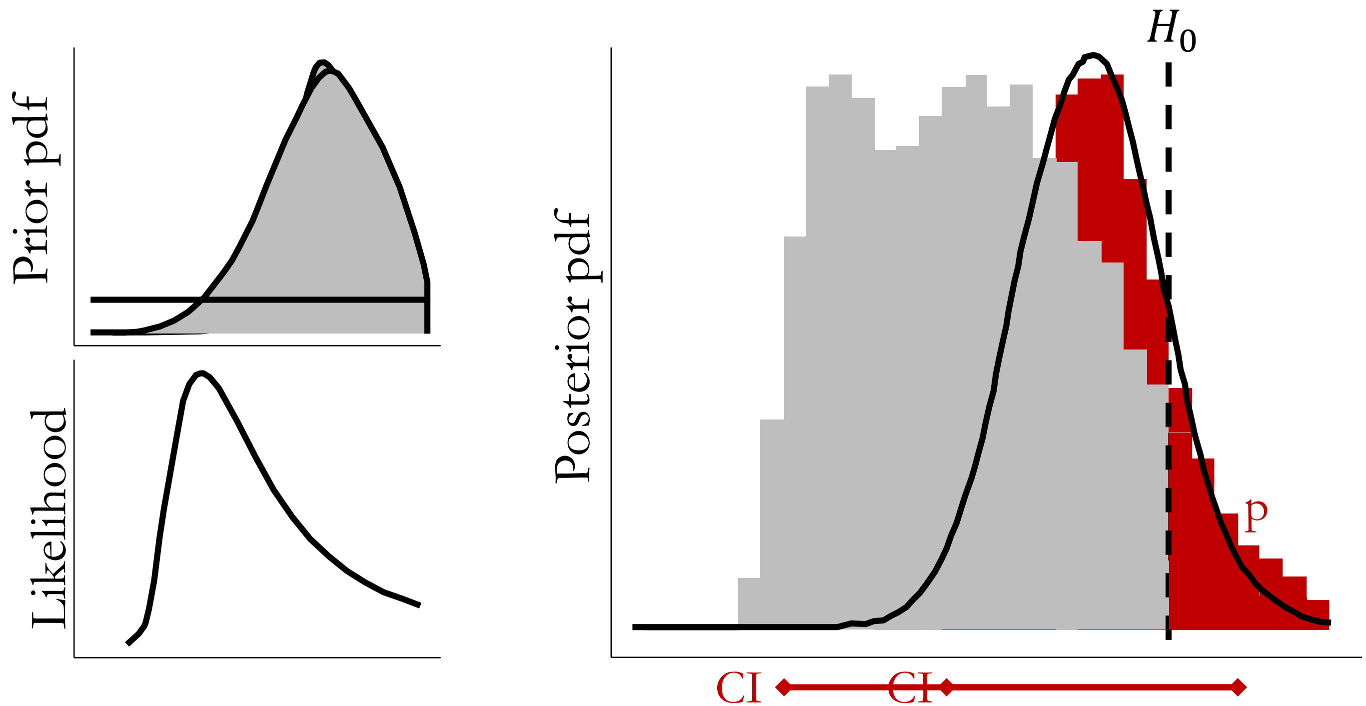
Calculation	Assumption under H_0	
`$COVARIANCE (NONMEM)`	$\hat{\theta}_k \sim \mathcal{N}\left(\theta_0, \mathcal{I}_k^{-1}\right)$	Wald test (wald)
Estimation of full & reduced model	$\frac{\mathcal{L}(\hat{\theta}_k; y)}{\mathcal{L}(\theta_1, \dots, \theta_0, \dots \theta_p; y)} \sim \chi_1^2$	Log-likelihood ratio test (lrt)
`randtest` (PsN)	$F(y, x) \sim F(y, x_\pi)$ x … covariates π … permutation	Permutation test (perm)

* Does not hold for bootstrap at small sample sizes[1,2]

(1) Thai et al. 2012

(2) Niebecker et al. 2013

Bayesian approach



- Bayesian "equivalents":
  - 95% CI: interval 2.5th - 97.5th percentile (95 % credibility interval)
  - P-value: probability parameter is larger or smaller than H_0 value
- Wide uniform or improper priors → posterior ∝ likelihood[1] → use for CIs & p-values

1. Bolstad WM (2013)

MCMC for CIs & p-values

Hamiltonian Monte-Carlo (HMC) in STAN[1] for sampling

- HMC very efficient (PAGE 24 (2015) Abstr 3677)
- Does not require conjugate priors for efficiency
- Supports improper priors (unlike WinBUGS, JAGS)

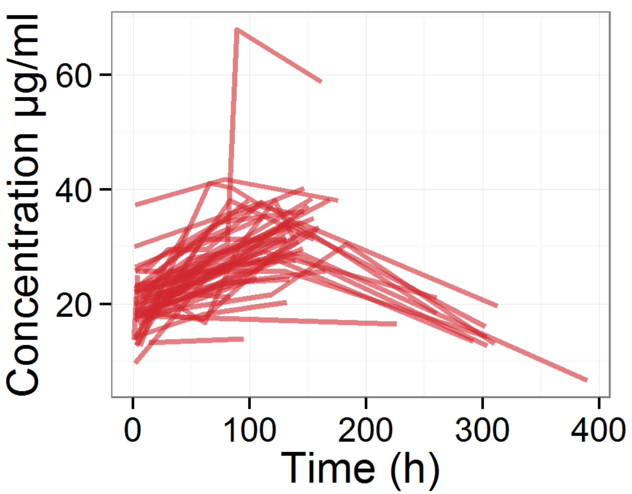
- **Approach:**

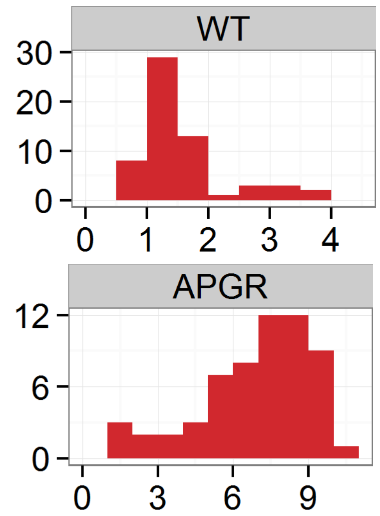
1. Estimate maximum likelihood (ML) model parameters (here NONMEM)
2. Implement model in STAN (uniform, improper priors)
3. Initialize MCMC chain at ML estimates
4. Obtain approx. 1000 samples for all parameters (effective sample size)

1. Stan Development Team (2015)

Case study

- Phenobarbital[1]
  - 59 individuals
  - 1-6 observations per individual
  - 1 compartment IV bolus
  - Log-normal IIV for V and Cl
  - Additive residual error
  - Covariates: WT & APGR
- Estimation method: FOCE





95% CIs

- Methods: asymp, llp, boot*, bayes

P-values

- WT & APGR on Cl
- Methods: wald, lrt, perm*, bayes

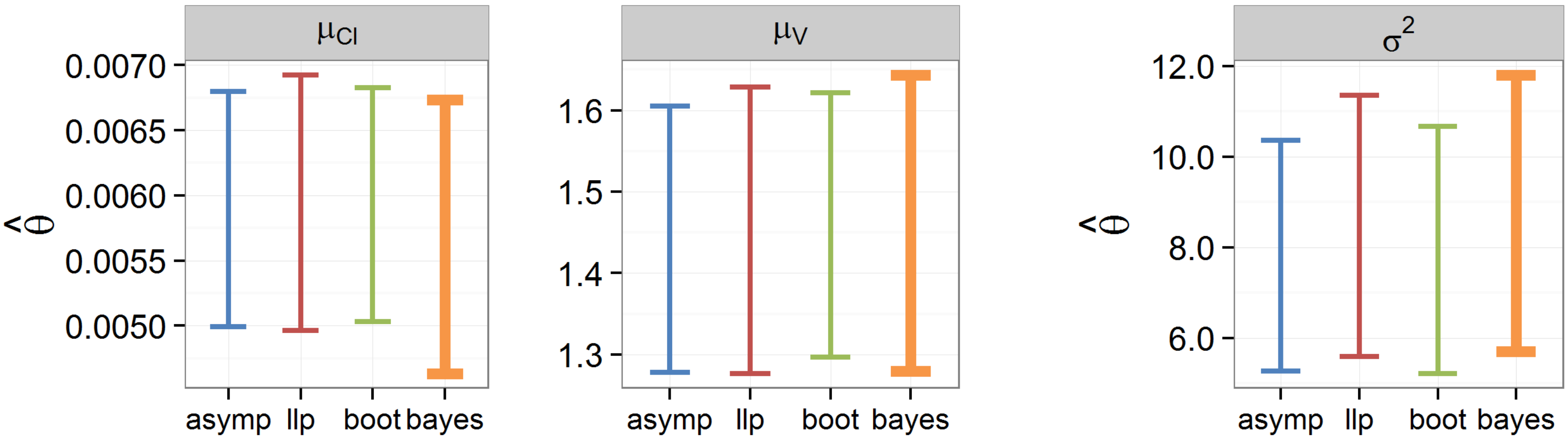
STAN: bayes

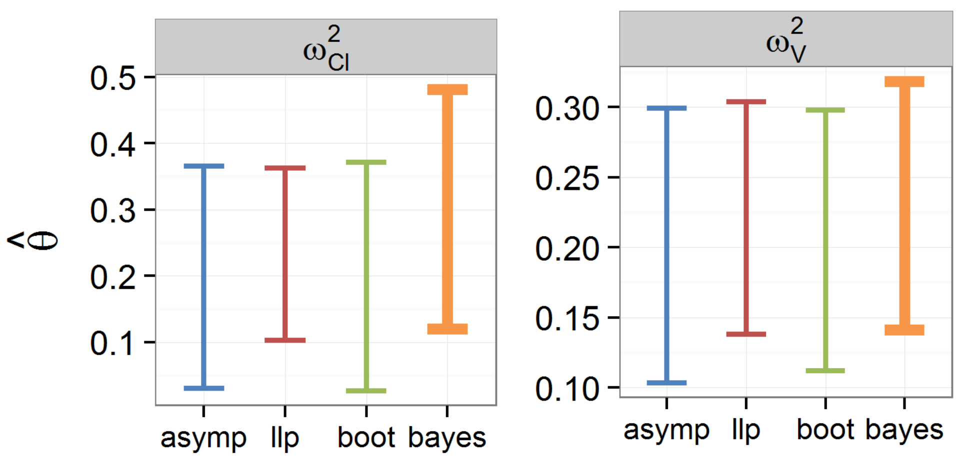
NONMEM/PSN: asymp/wald, llp/lrt, boot/perm

**1000 samples*

1. Grasela & Donn 1985

Phenobarbital – CIs & p-values

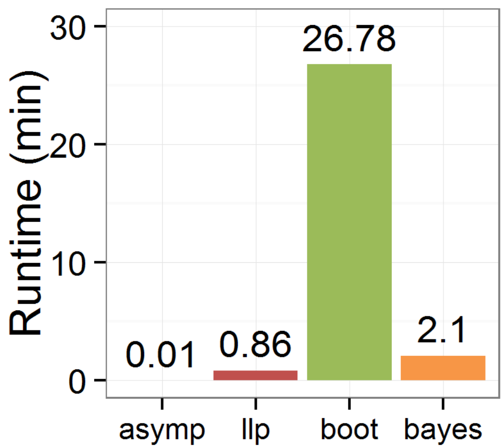




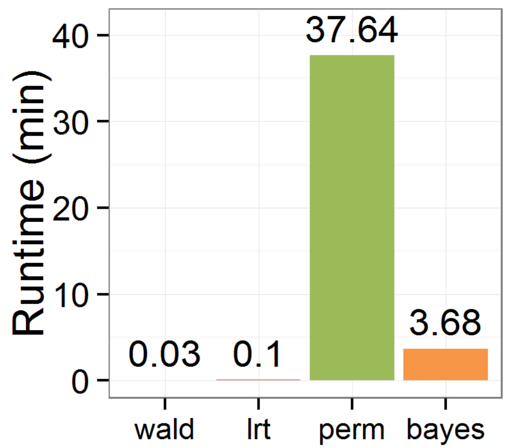
	WT on Cl	APGR on Cl
wald	$< 10^{-3}$	0.68
lrt	$< 10^{-3}$	0.43
perm	$< 10^{-3}$	0.48
bayes	$< 10^{-3}$	0.65

Runtimes

CIs



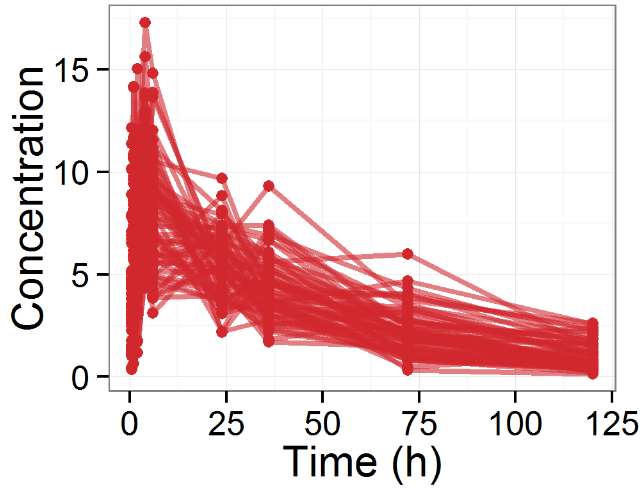
p-value

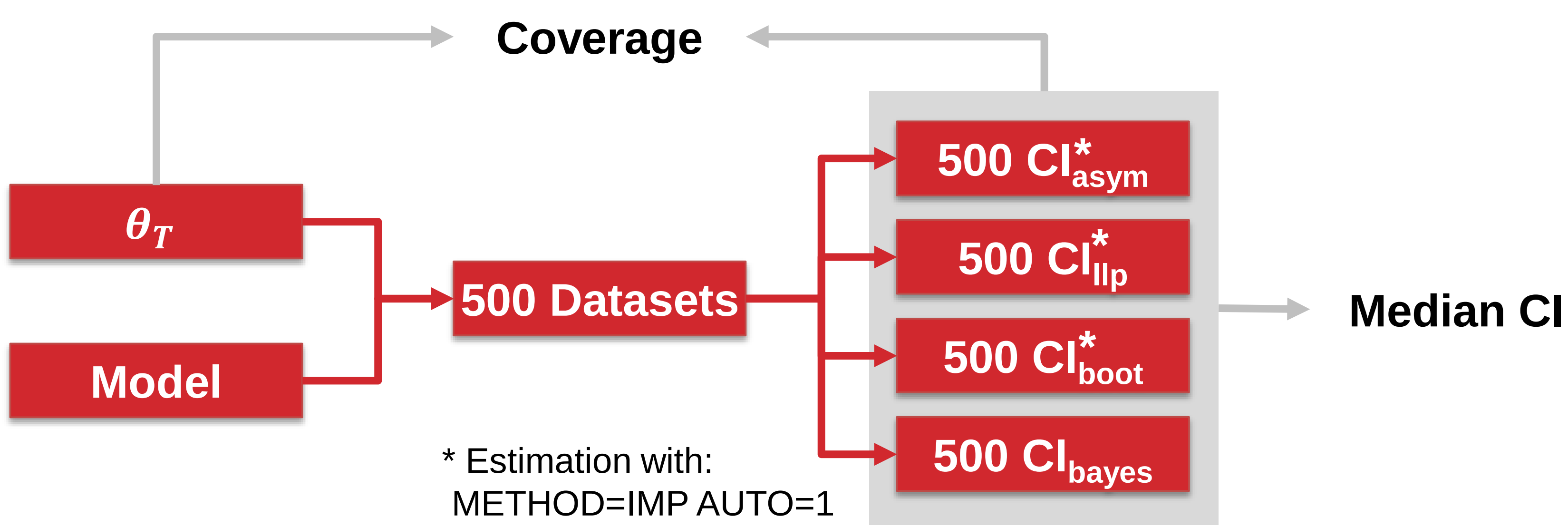


Promising results, but don't know the truth…

Simulation study – CIs

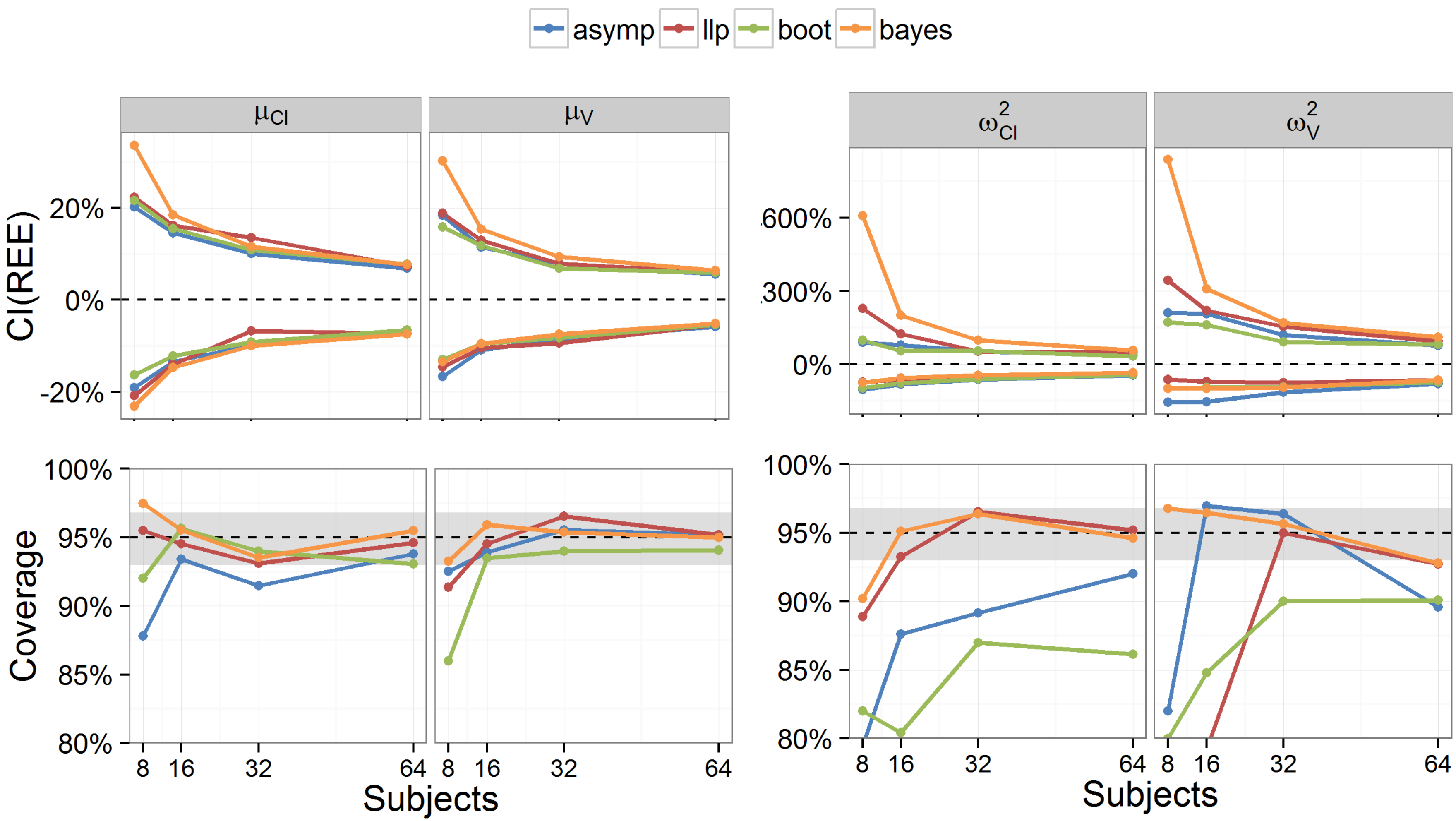
- **Warfarin example**[1]
  - 1 compartment model
  - 1st order absorption
  - Proportional RUV
  - Log-normal IIV
  - Single dose 9 obs/subject
- **95% CI evaluation** (8, 16, 32, 64 subjects)





1. Nyberg et al. 2015

Simulation study – CIs

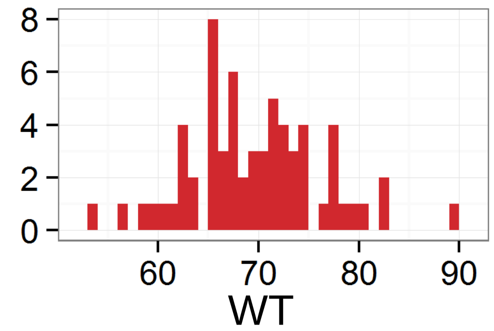


Simulation study – p-values

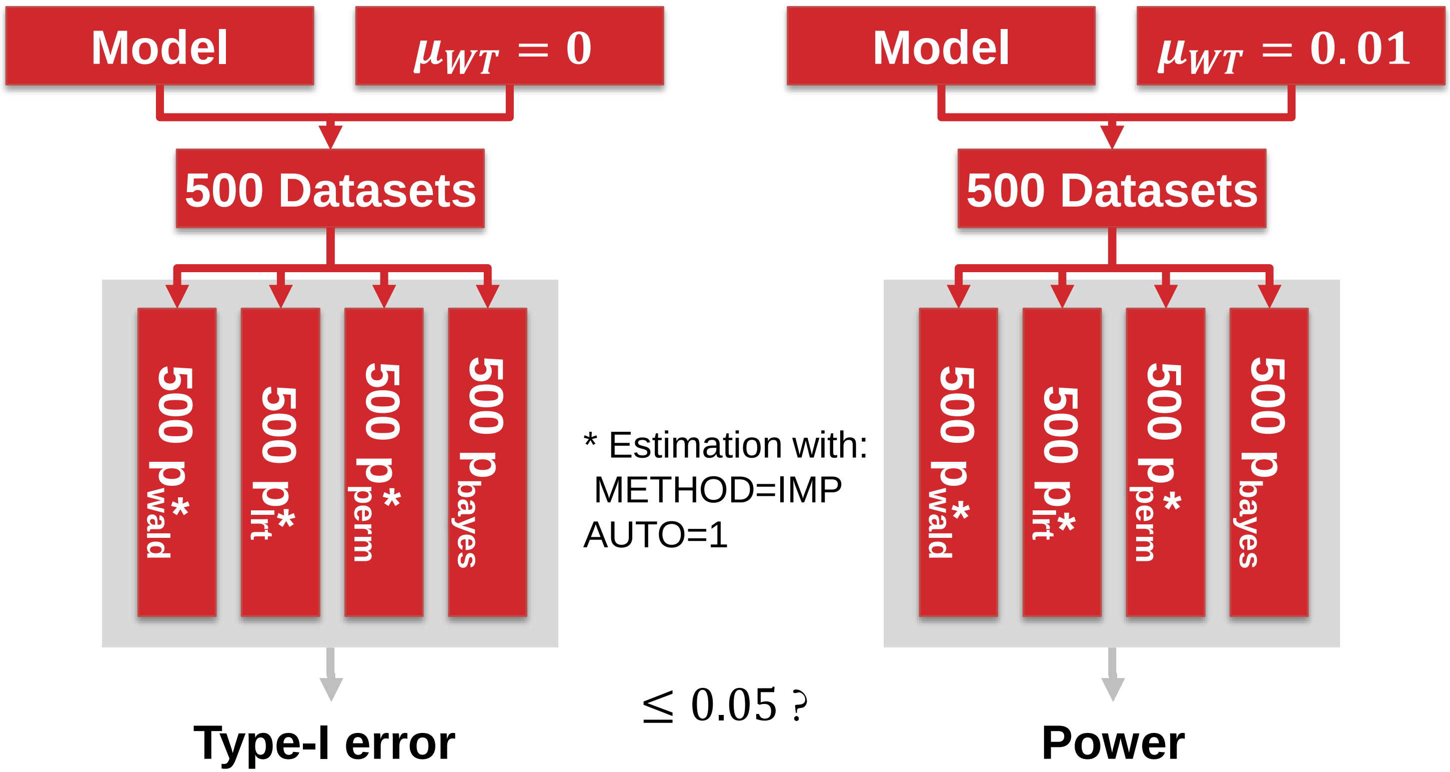
- **Warfarin example**

$$\log Cl = \log \mu_{Cl} + \mathrm{b}_{i,Cl} + \mu_{WT}(1 + (WT - 70))$$

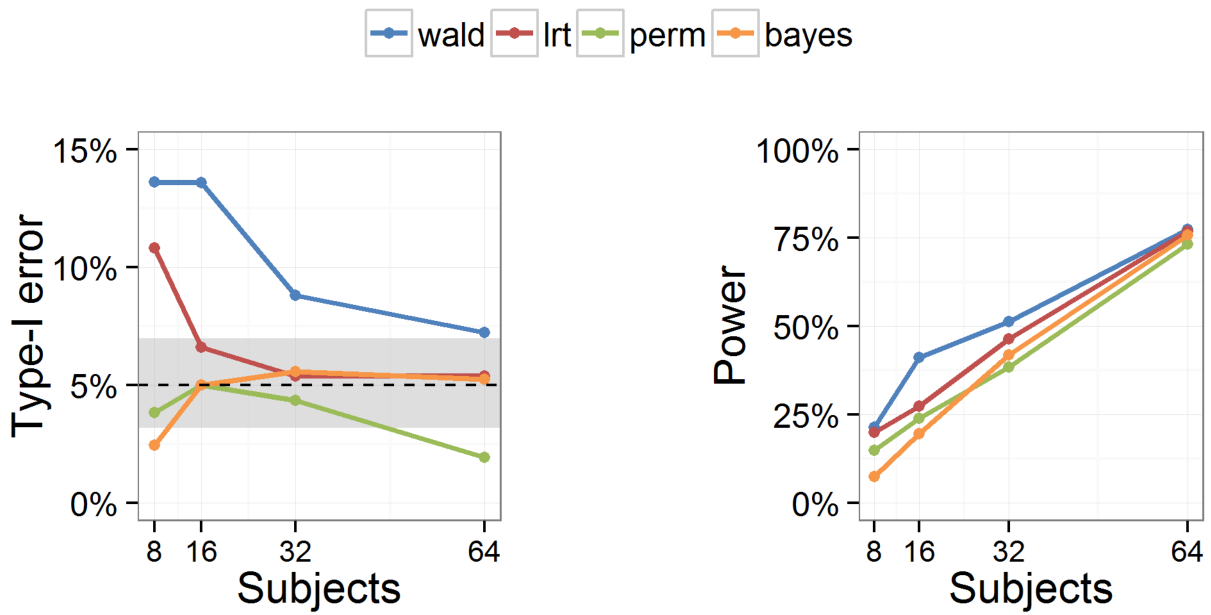
$$WT \sim \mathcal{N}(70{,}7)$$



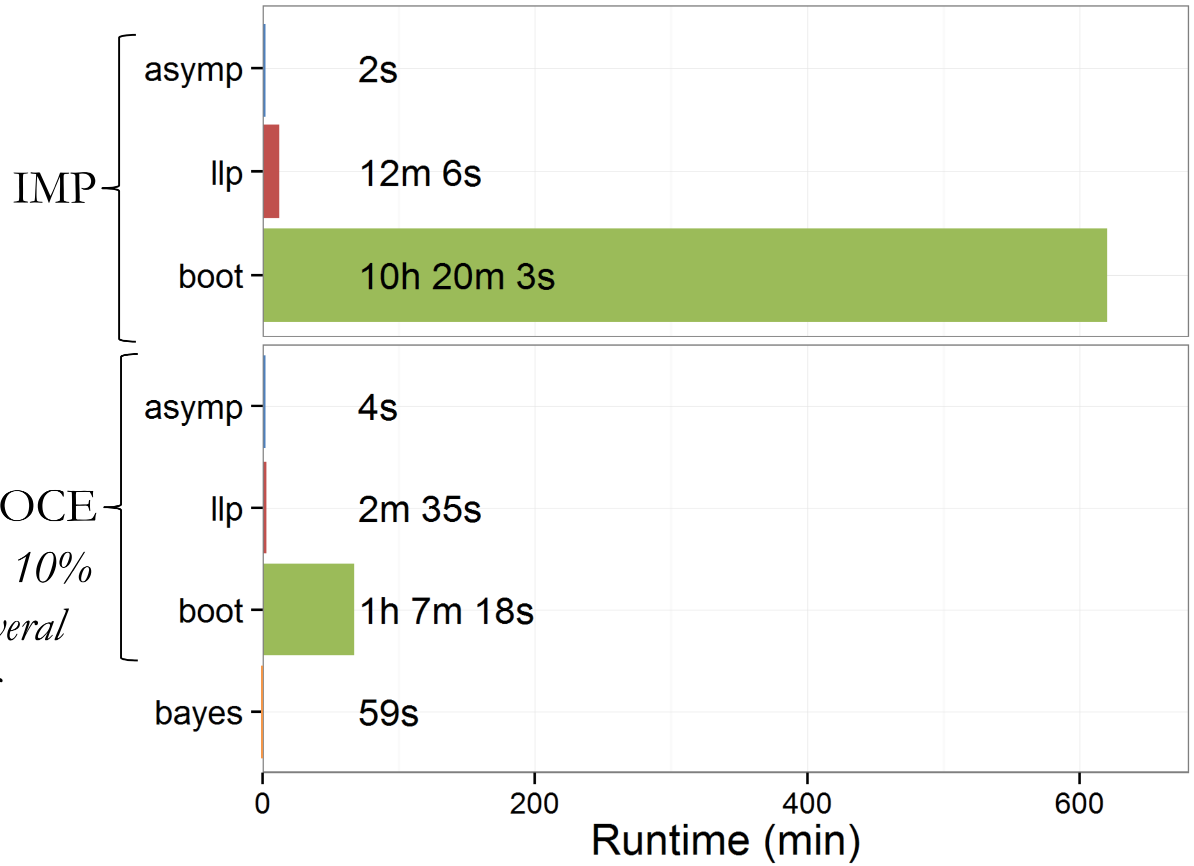
- **P-value evaluation** (8, 16, 32, 64 subjects)



Simulation study – p-values



Runtime – CI calculation



More than 10% bias for several parameters

Use in practice

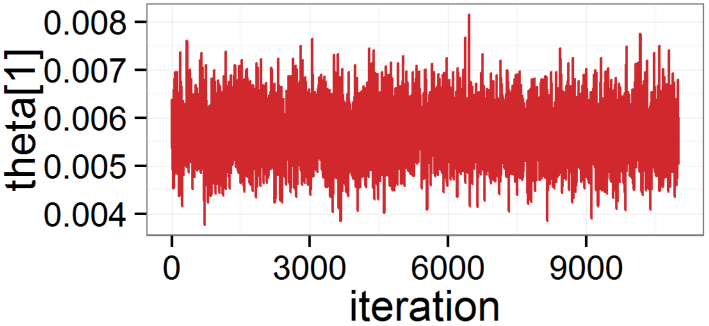
Sampling from posterior distribution for CIs and p-values attractive alternative

- No model linearization
- No large samples assumptions (in contrast to all other methods except perm)
- Good theoretical properties (coverage & type-I error)

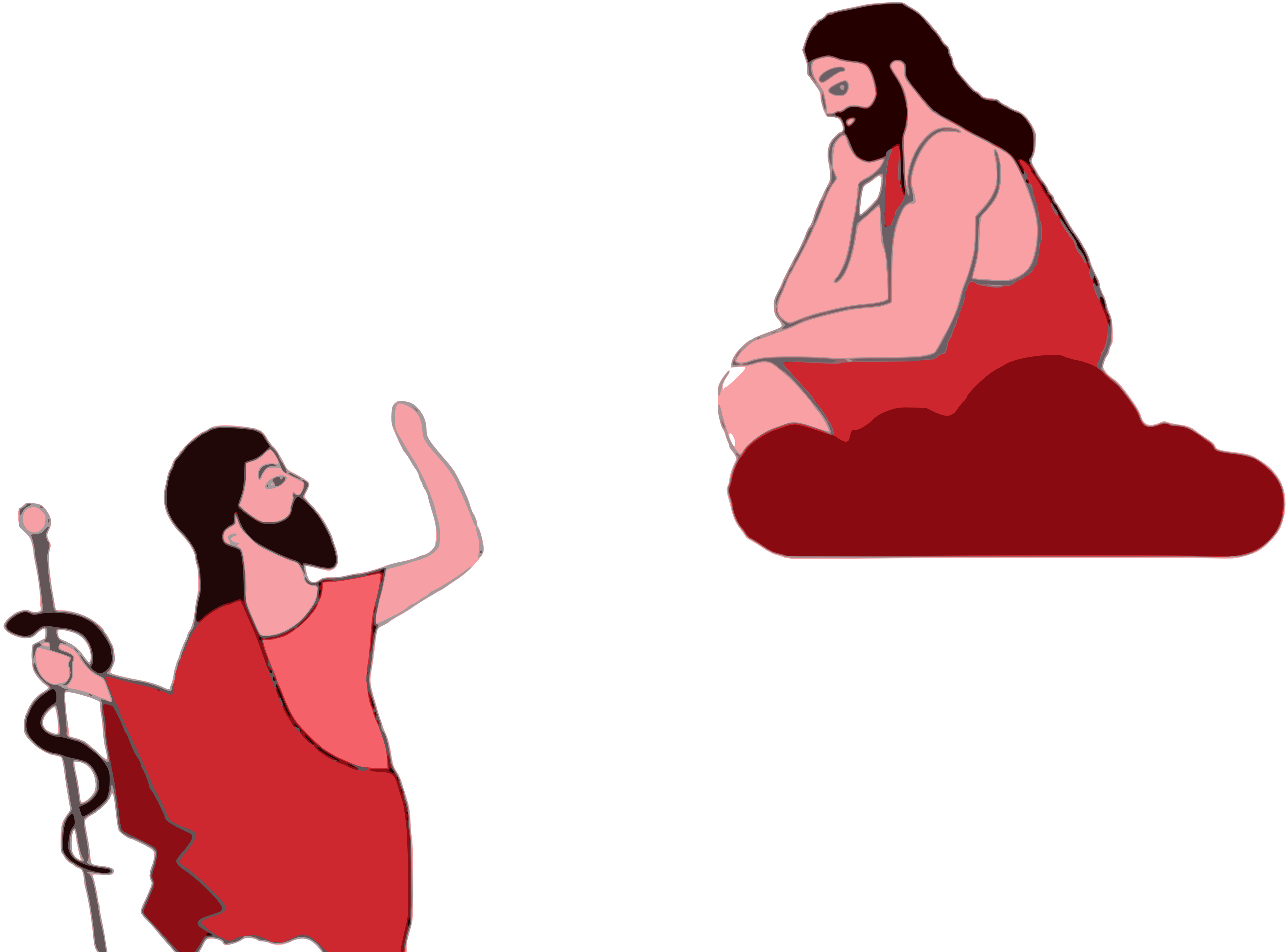
- Much faster than resampling based methods (boot & perm)
  - Factor 10-60 when using FOCE
  - Multiple orders of magnitude when using IMP

Use in practice (2)

- Reimplementation of model impractical
- Use Bayesian sampler in NONMEM (linear mu-referencing if possible)
  - Phenobarbital
    - Good agreement for CIs & short runtime
  - Warfarin
    - Sampling terminated
- Verify:
  - Mixing & convergence of chains
    - Potential scale reduction statistic $\hat{R}$[1]
    - Trace plot
  - Sufficient number of samples
    - Effective sample size



1.Gelman & Rubin 1992



Thanks to

- You for listening
- Paris and Uppsala colleagues for fruitful discussions
- DDMoRe initiative for funding

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n° 115156, resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. The DDMoRe project is also supported by financial contribution from Academic and SME partners. This work does not necessarily represent the view of all DDMoRe partners.