UPPSALA UNIVERSITET

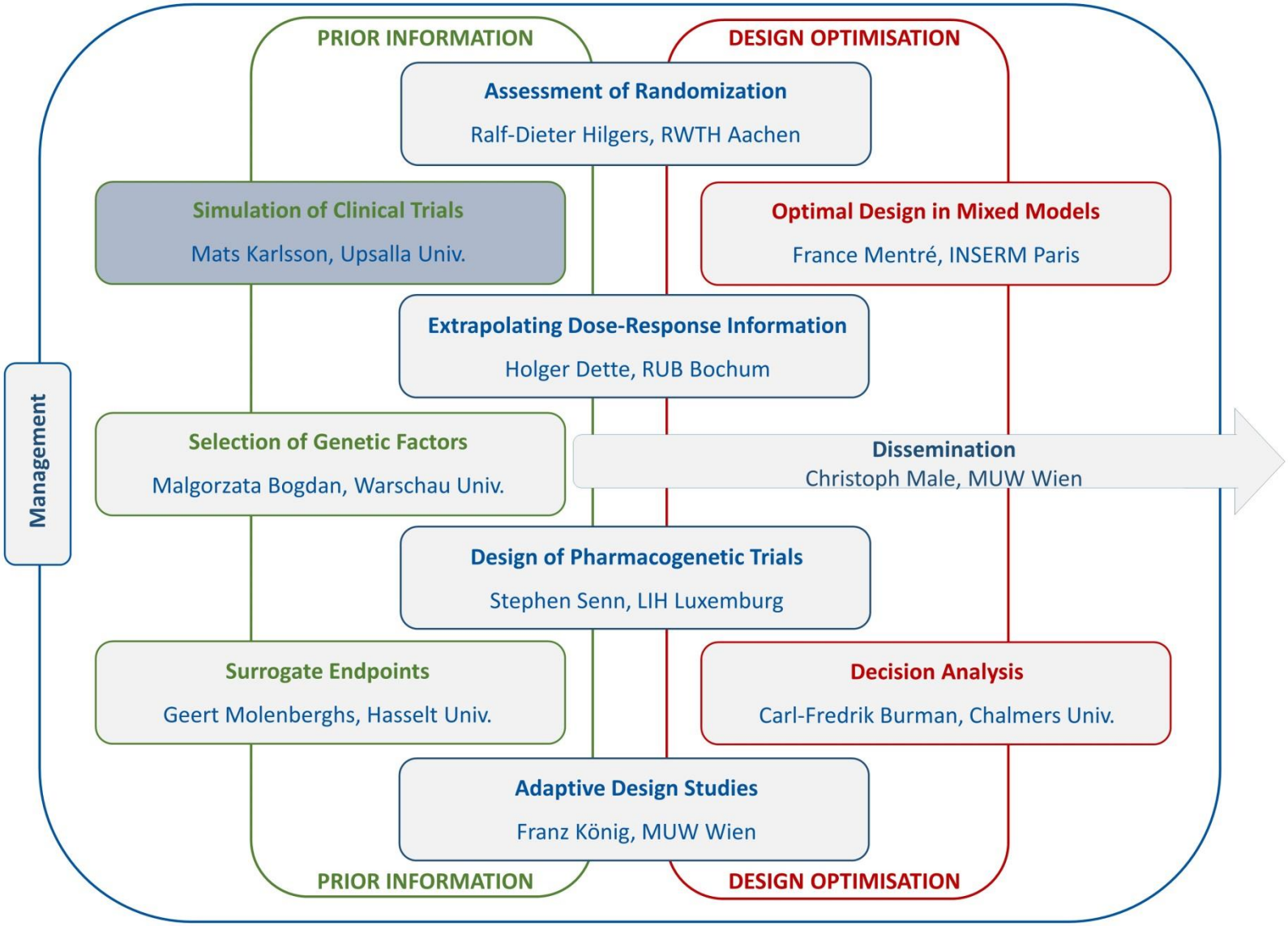
WP7: Simulation of clinical trials in small population groups

A Sampling Importance Resampling Procedure for Estimating Parameter Uncertainty

Oct 11, 2016

Mats Karlsson
Uppsala University

Workpackage structure



Considerations in small population groups

- Difficult to run trials with many subjects
- How can all relevant information be utilized in making decisions?
  - Nonlinear mixed effects models (**NLMEM**) incorporating drug and disease characteristics offer an attractive alternative

NLMEM – why attractive?

- Integrate information in data across
  - subjects
  - time (longitudinal analysis)
  - variables
  - covariates/predictors
- Allow prior knowledge to be incorporated
  - Drug/Disease driven structural models
  - Parameter constraints from biological/pharmacological knowledge
  - Other knowledge/assumptions as appropriate

Trial/treatment decisions using NLMEM

- Informed by
  - Model contrasts (hypothesis tests)
  - Parameter uncertainty distributions
  - Prediction distributions with uncertainty

Decisions using NLMEM – model contrasts

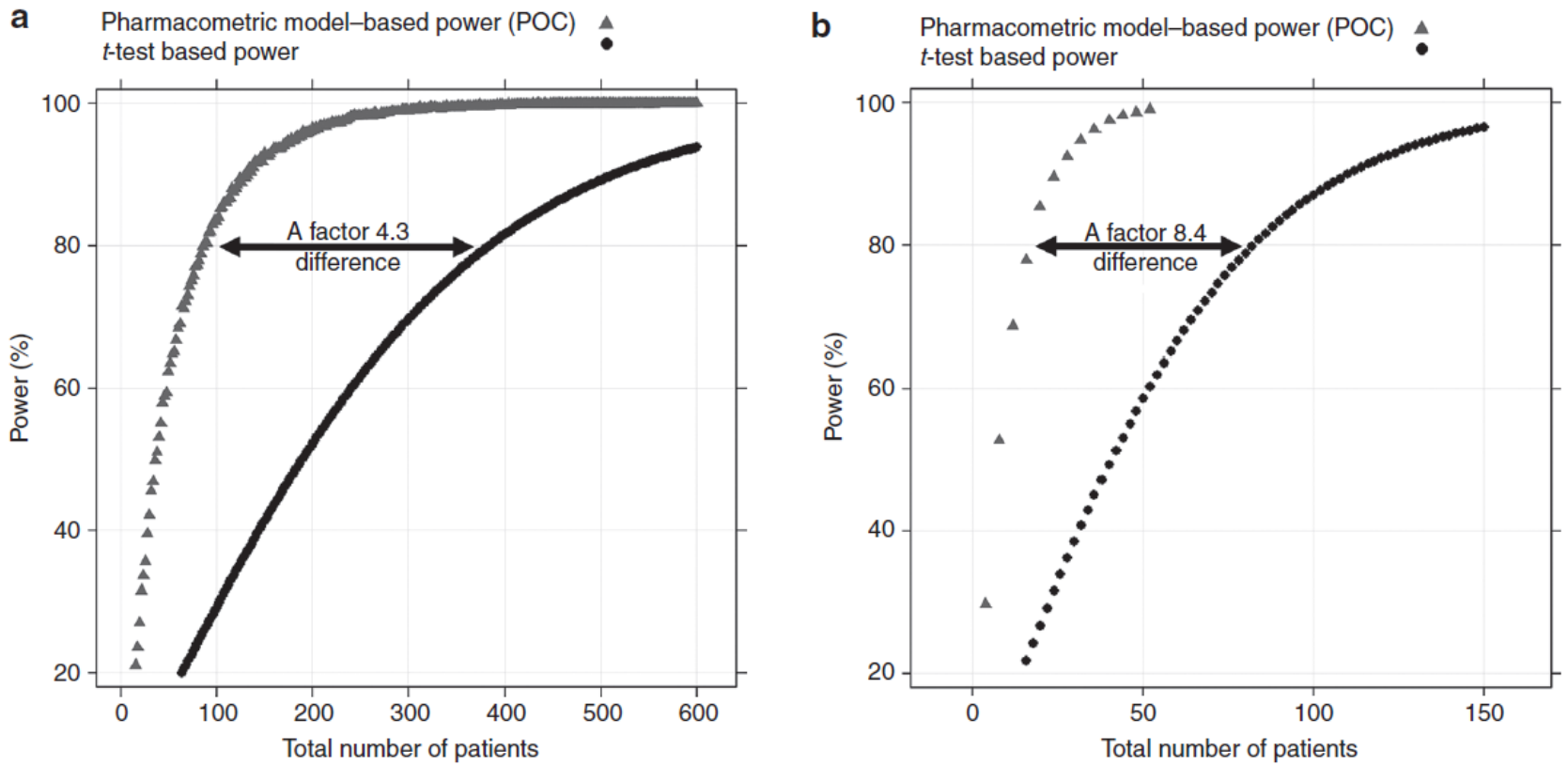


Figure 3 Power curve comparison between the pharmacometric model–based power (gray triangles) and the *t*-test based power (black diamonds), for the proof-of-concept scenario. (**a**) The power curves for the stroke example in which the difference in study size is a factor of 4.3 (90 vs. 388 total number of patients) is displayed. (**b**) In the diabetes example, the difference in study size was 8.4-fold (10 vs. 84 total number of patients) in favor of the pharmacometric approach.

Karlsson et al. CPT:PSP 2:e23 (2013)

Decisions using NLMEM – parameter uncertainty

Clarification on Precision Criteria to Derive Sample Size When Designing Pediatric Pharmacokinetic Studies

J Clin Pharmacol 2012 52: 1601

Yaning Wang, PhD, Pravin R. Jadhav, PhD, Mallika Lala, PhD, and Jogarao V. Gobburu, PhD

One of the important goals of the pediatric PK study is to ensure the precise estimate of important PK parameters, such as clearance and volume of distribution, to justify the choice of a safe and effective dose from a PK perspective. To achieve this goal, a standard regulatory requirement has been drafted and communicated to the sponsors, where applicable, as follows:

> The study must be prospectively powered to target a 95% CI [confidence interval] within 60% and 140% of the geometric mean estimates of clearance and volume of distribution for DRUG NAME in each pediatric sub-group with at least 80% power.

Internal decision making – predictive distributions

Model-Based Drug Development: A Rational Approach to Efficiently Accelerate Drug Development

PA Milligan[1], MJ Brown[2], B Marchant[3,10], SW Martin[1], PH van der Graaf[4,1], N Benson[4,11], G Nucci[5], DJ Nichols[5], RA Boyd[6], JW Mandema[7], S Krishnaswami[6], S Zwillich[8], D Gruben[2], RJ Anziano[2], TC Stock[9] and RL Lalonde[6]

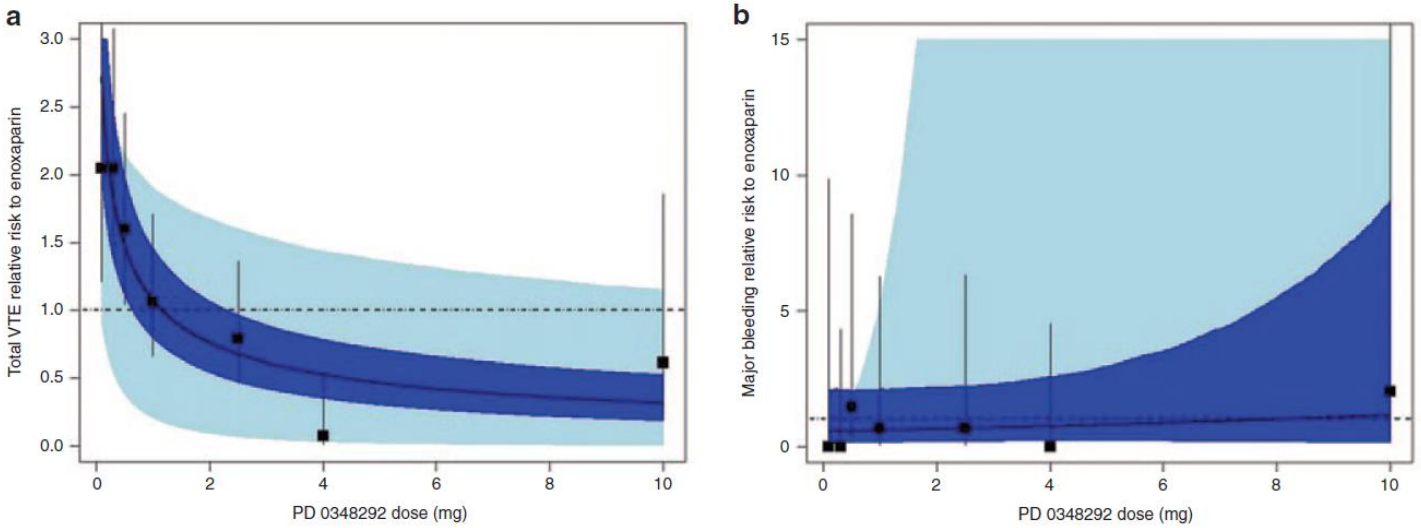


Figure 6 Observed relative risk of PD 0348292 vs. enoxaparin (symbols with 95% confidence intervals (CIs)) for (**a**) VTE and (**b**) MB and logistic regression model fit (solid line with dark blue area covering the 90% CI) in an adaptive phase II study. The light blue area covers the 90% CI before the trial based on the PK–PD model for inhibition of thrombin generation. MB, major bleeding; PK–PD, pharmacokinetics–pharmacodynamics; VTE, venous thromboembolism.

Clin Pharmacol Ther 2013

Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation

EFPIA MID3 Workgroup: SF Marshall[1]*, R Burghaus[2], V Cosson[3], SYA Cheung[4], M Chenel[5], O DellaPasqua[6], N Frey[3], B Hamrén[7], L Harnisch[1], F Ivanow[8], T Kerbusch[9], J Lippert[2], PA Milligan[1], S Rohou[10], A Staab[11], JL Steimer[12], C Tornøe[13] and SAG Visser[14]

: CPT Pharmacometrics Syst. Pharmacol. (2016) 5, 93–122;

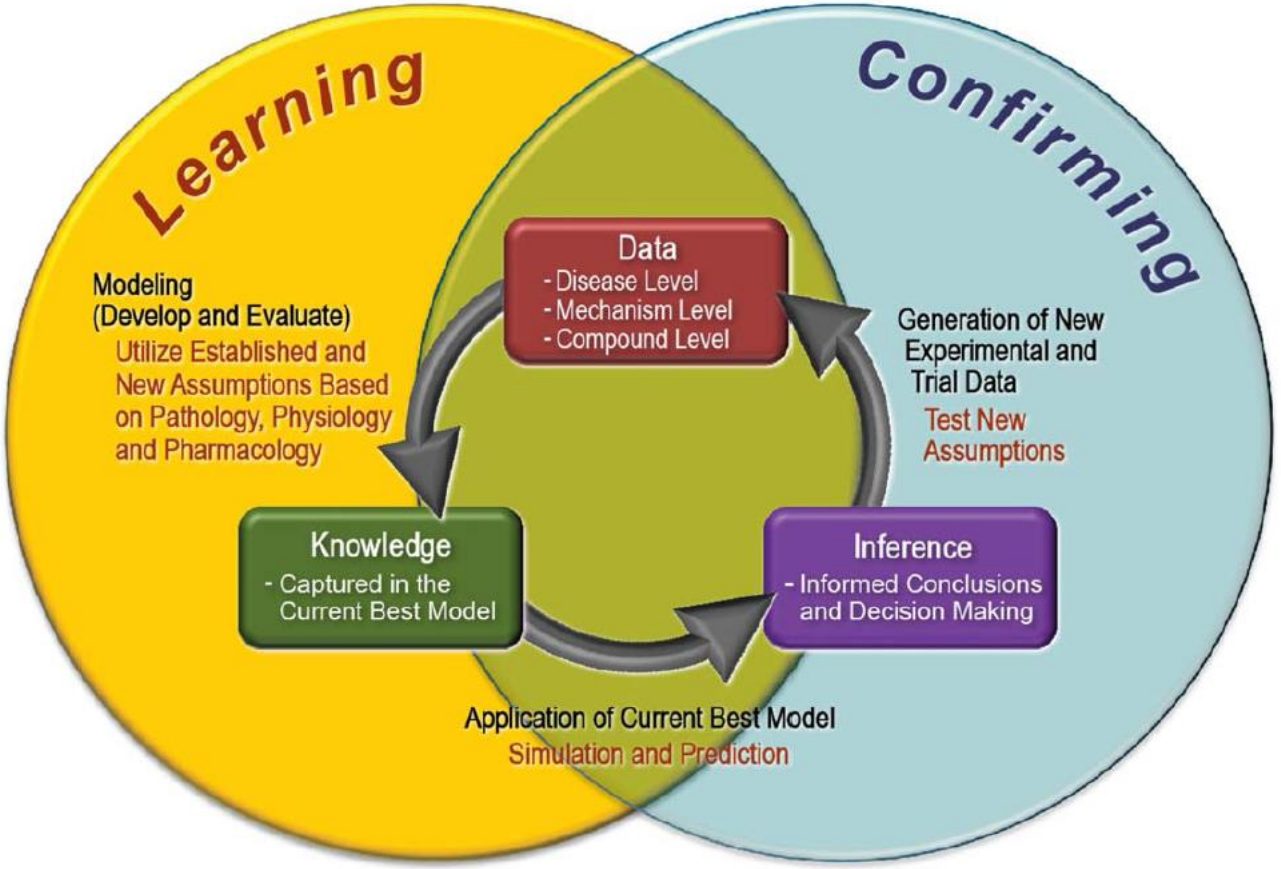
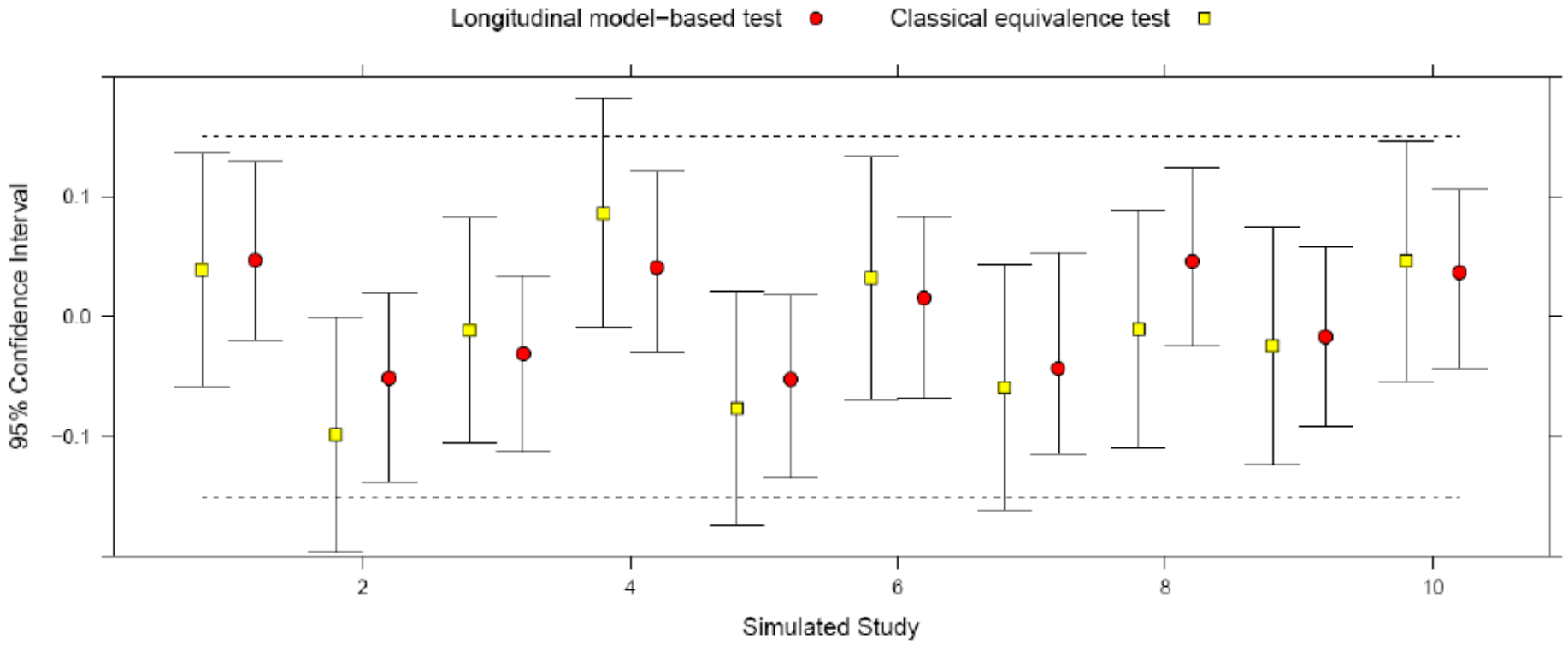


Figure 3 MID3: a quantitative framework for prediction and extrapolation centered on knowledge and inference generated from integrated models of compound, mechanism, and disease level data aimed at improving the quality, efficiency, and cost-effectiveness of decision-making. The colored boxes represent essential components of the "Learn and Confirm Cycle". The arrows represent processes that link these components.

Regulatory decision making – predictive distributions

Model-based analyses for pivotal decisions,
with an application to equivalence testing for biosimilars
Bieth et al, PAGE 2012

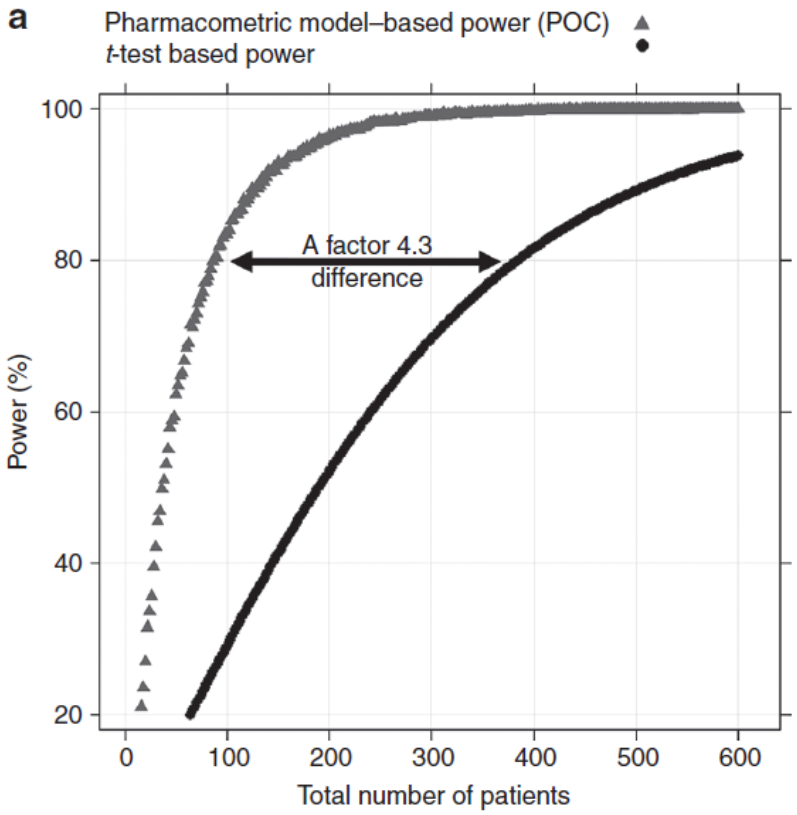


NLMEM in trial/treatment evaluations

- Power calculations
  - How to do timely power calculations?
- Hypothesis tests
  - How to achieve type 1 error control?
- Model uncertainty
  - What if the NLMEM is not appropriate?
- Adaptive designs for small populations
  - NLMEM-Based Adaptive Optimal Design
- Parameter uncertainty (PU)
  - Diagnostics for adequacy of PU
  - Sampling-Importance-Resampling (SIR)

NLMEM in trial/treatment evaluations

- **Power calculations**
  - **How to do timely power calculations?**
- Hypothesis tests
  - How to achieve type 1
- Model uncertainty
  - What if the NLMEM is r
- Adaptive designs for s
  - Model-Based Adaptive
- Parameter uncertainty
  - Diagnostics for adequa
  - Sampling-Importance-



Increased speed in power calculations

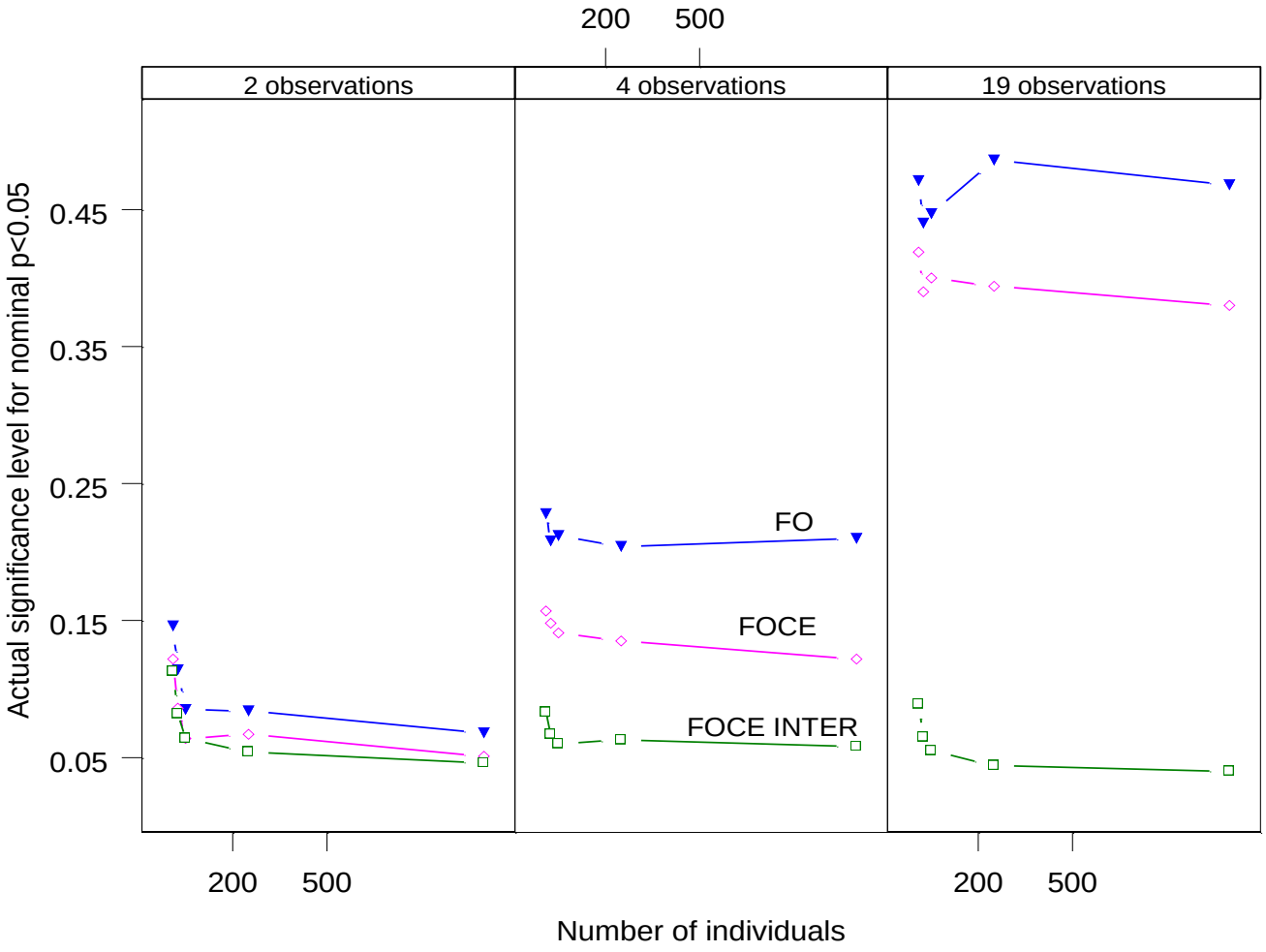
- **Monte Carlo Mapped Power (MCMP)**
- Simulate 1 data set with large N
- Fit full and reduced model
- Obtain dOFVi for each subject
- Resample dOFVi to obtain power for study size of interest
- Vong et al., AAPS J 2012

- **Parametric Power Estimation (PPE)**
- Simulate X data sets with N subjects
- Fit full and reduced model
- Estimate λ from dOFV assuming non-central chi-square distribution
- Extrapolate to other study sizes using λ
- Ueckert et al., JPKPD 2016

NLMEM in trial/treatment evaluations

- Power calculations
  - How to do timely power calculations?
- **Hypothesis tests**
  - **How to achieve type 1 error control?**
- Model uncertainty
  - What if the NLMEM is not appropriate?
- Adaptive designs for small populations
  - Model-Based Adaptive Optimal Design
- Parameter uncertainty (PU)
  - Diagnostics for adequacy of PU
  - Sampling-Importance-Resampling (SIR)

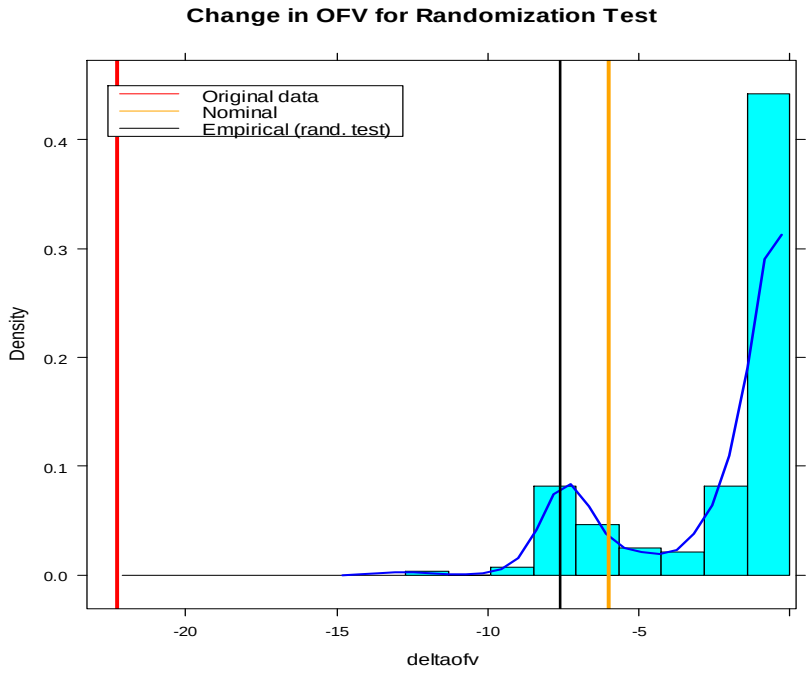
NLMEM – Type 1 error control

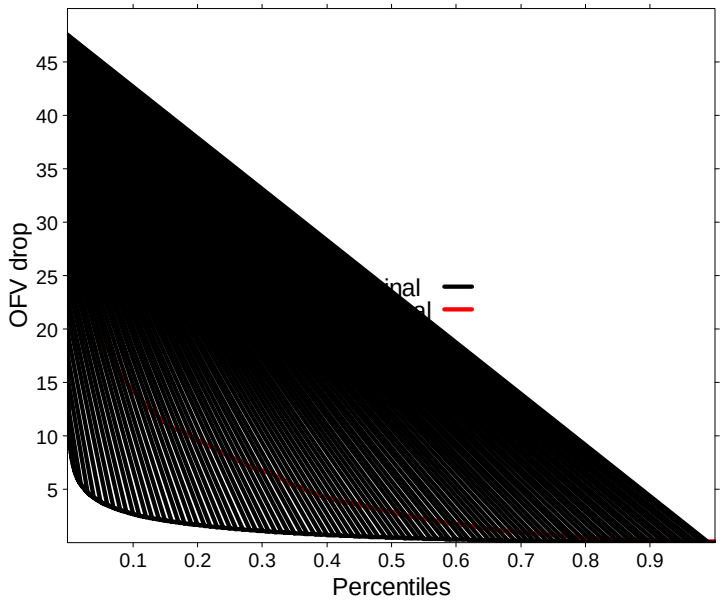


Wählby U et al., J Pharmacokinet Pharmacodyn 28:231-52 (2001)

Permutation (Randomisation) tests for NLMEM

- Permutation test for
  - prespecified NLMEM model
  - (mixture) model built using blinded data





Deng et al. PAGE 2015, Harling et al. PAGE 2016,

NLMEM in trial/treatment evaluations

- Power calculations
  - How to do timely power calculations?
- Hypothesis tests
  - How to achieve type 1 error control?
- **Model uncertainty**
  - **What if the NLMEM is not appropriate?**
- Adaptive designs for small populations
  - Model-Based Adaptive Optimal Design
- Parameter uncertainty (PU)
  - Diagnostics for adequacy of PU
  - Sampling-Importance-Resampling (SIR)

Model-averaging

- Model-averaging for
  - longitudinal dose–response*
  - biosimilar superiority testing**
  - confidence interval-based QT-test***

*Aoki et al., PAGE 2014, PAGE 2016
**Dosne et al., in manuscript
***Dosne et al, PAGE 2016

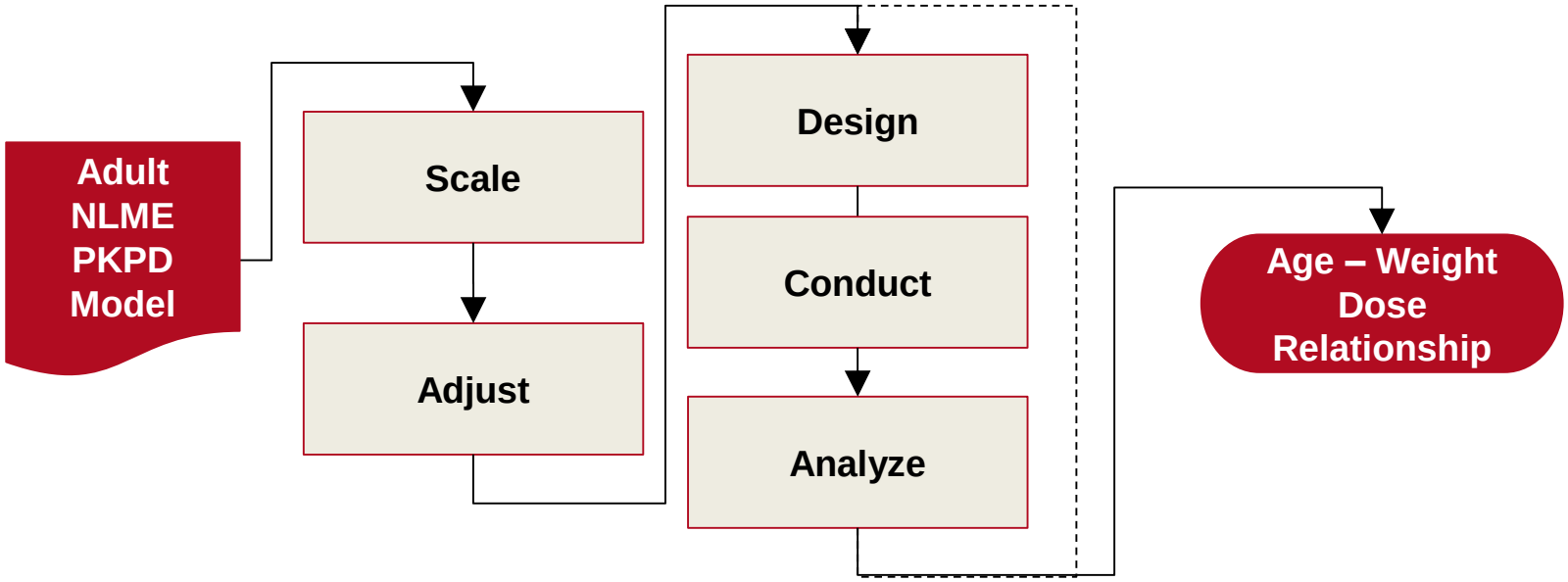
NLMEM in treatment evaluations

- Power calculations
  - How to do timely power calculations?
- Hypothesis tests
  - How to achieve type 1 error control?
- Model uncertainty
  - What if the NLMEM is not appropriate?
- **Adaptive designs for small populations**
  - **Model-Based Adaptive Optimal Design**
- Parameter uncertainty (PU)
  - Diagnostics for adequacy of PU
  - Sampling-Importance-Resampling (SIR)

Model-based adaptive optimal design

Simulated model based adaptive optimal design of adult to children bridging study using FDA stopping criteria

- Interim analysis after every cohort
- Update of design for next cohort
- Stopping if precision is sufficient



Strömberg and Hooker. PAGE 2015; Karlsson EMA Workshop 2016

NLMEM in treatment evaluations

- Power calculations
  - How to do timely power calculations?
- Hypothesis tests
  - How to achieve type 1 error control?
- Model uncertainty
  - What if the NLMEM is not appropriate?
- Adaptive designs for small populations
  - Model-Based Adaptive Optimal Design
- **Parameter uncertainty (PU)**
  - Diagnostics for adequacy of PU
  - Sampling-Importance-Resampling (SIR)

Parameter uncertainty (PU)

- Parameter uncertainty distributions provide decision basis for probability and confidence interval (CI)-based decisions
- Several ways to estimate PU
  - Cov-matrix, bootstrap, …
- Different methods provide different PU and have different properties
  - Which one to use?

Parameter uncertainty – covariance matrix

- Covariance matrix
  - Not always retrievable or suitable
  - Assumes symmetry & linear correlations
- NLMEM CIs often assymetric
  - Non-linear model
  - (Interindividual) Variability parameters
  - Context-driven parameter boundaries

Parameter uncertainty – bootstrap

- Bootstrap: sensitivity to sample size
  - For simple models, robust down to small sample sizes (N≈10-12)
  - For NLME models, sample size dependence less well explored/understood

Bootstrap of NLME models

- Factors likely to increase sample size demand
  - Simultaneous estimation of multiple parameters
  - Hierarchical models with ≥2 levels of random-effects
  - Heterogeneous designs including covariate distributions
  - Data-driven model development
  - Model misspecification

dOFV distribution - a diagnostic for PU

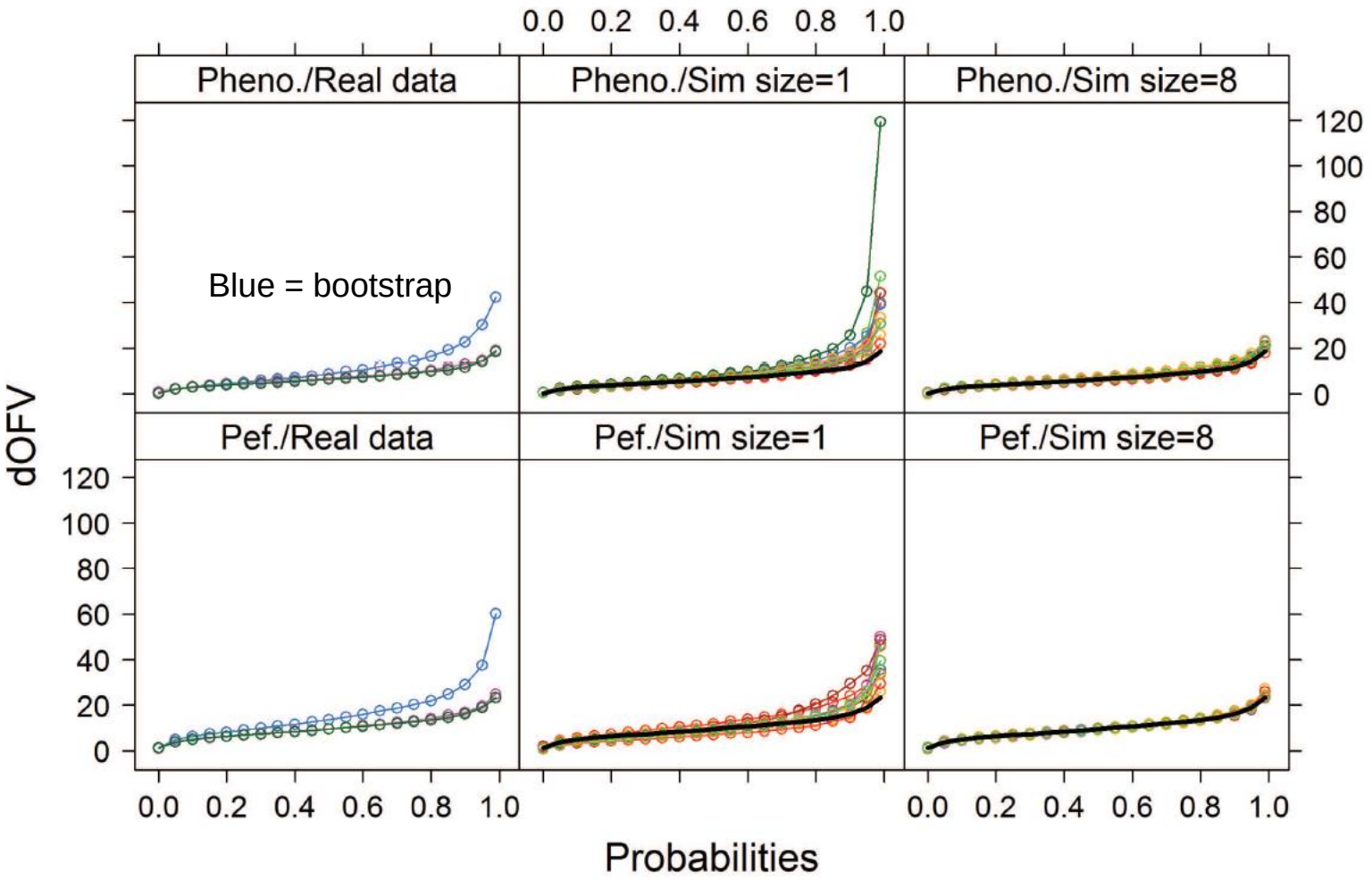
- Objective:
  - Provide a diagnostic for the adequacy of an estimate of parameter uncertainty

Dosne et al. JPKPD 2016 (in press)

dOFV distribution

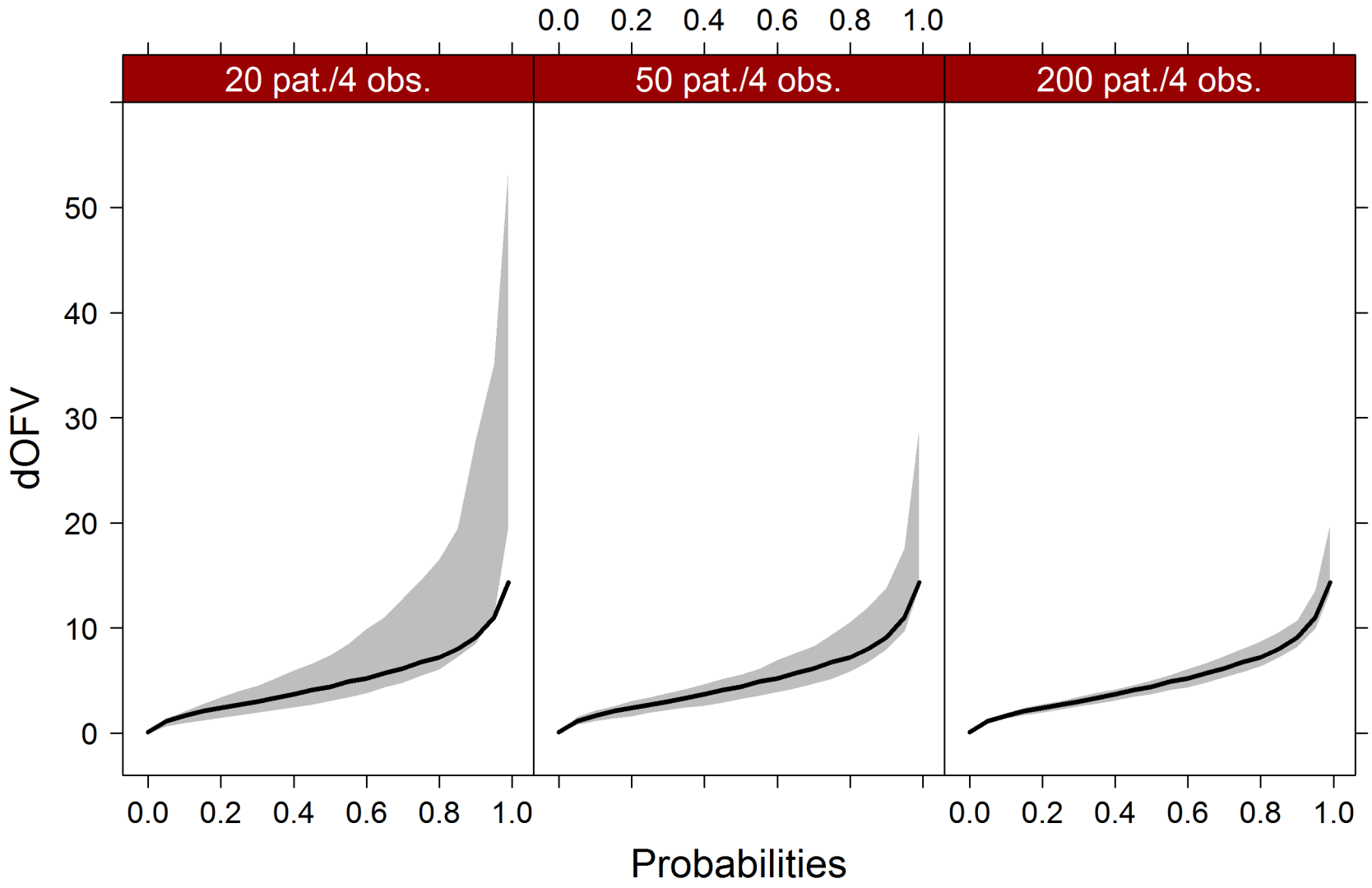
1. Evaluate parameter vectors sampled from PU distribution on original dataset
2. Subtract OFV of the final model for original data set
3. Compare bootstrap dOFV distribution with reference (chisq) dOFV distribution

Comparison with expected distribution

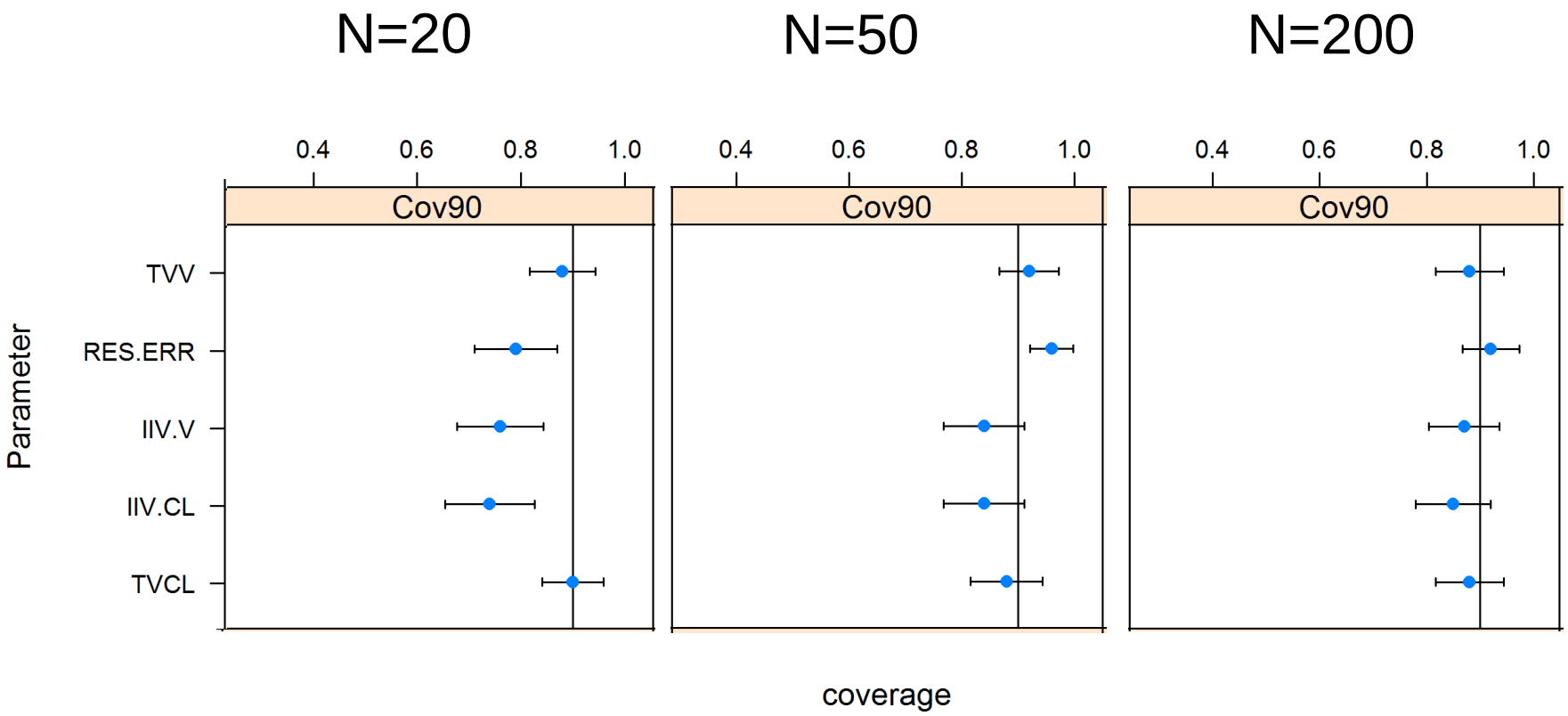


Dosne et al. JPKPD 2016 (in press)

Simulation example 1



Simulation example 1



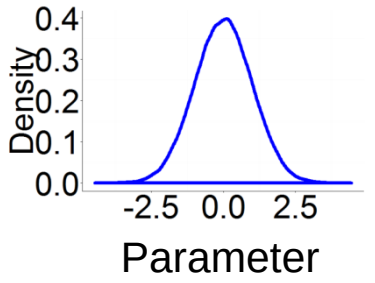
Parameter uncertainty estimation

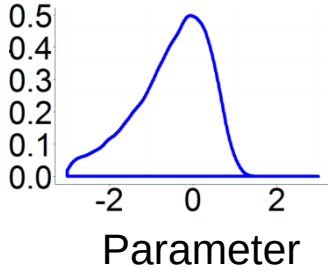
- Covariance matrix
  - Not always retrievable or suitable
  - Assumes symmetry & linear correlations
- Bootstrap
  - Empirically shown to be inadequate for small/medium-sized data
  - Computationally problematic (time & stability)
- Need for additional PU estimation methods
  - **S**ampling – **I**mportance – **R**esampling (**SIR**)

SIR principle:
Sampling, Importance weighting, Resampling

- → Approximate unknown posterior distribution by weighted known distribution[1]

S	**SAMPLING** *Step 1*	• **Sample** *p* parameter vectors from covariance matrix
I	**IMPORTANCE WEIGTHING** *Step 2*	• **Calculate weights** based on fit to original data
R	**RESAMPLING** *Step 3*	• **Resample** *M* vectors based on weights from step 2 • Compute confidence intervals





[1] Rubin DB, Bayesian Statistics. 1988;3:395-402

Importance ratios (IR)

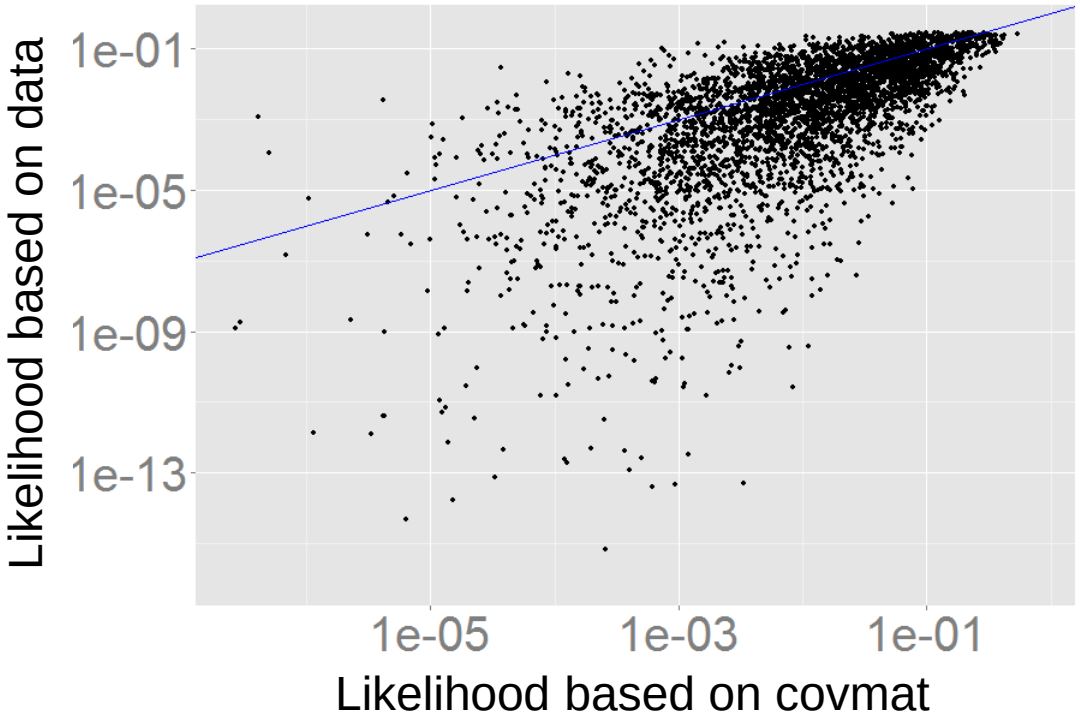
- **Resampling probabilities:**

$$IR = \frac{lik(Y|\theta)}{h(\theta)}$$

- Likelihood of data given parameter vector divided by likelihood of vector in proposal
- How well vector fits data compared to how well it should fit data
- $IR = 1$: as expected → not reweighted in resampling
- $IR > 1$: better than expected → upweighted
- $IR < 1$: worse than expected → downweighted

Components of Importance Ratios

- Many vectors do not fit as well as expected



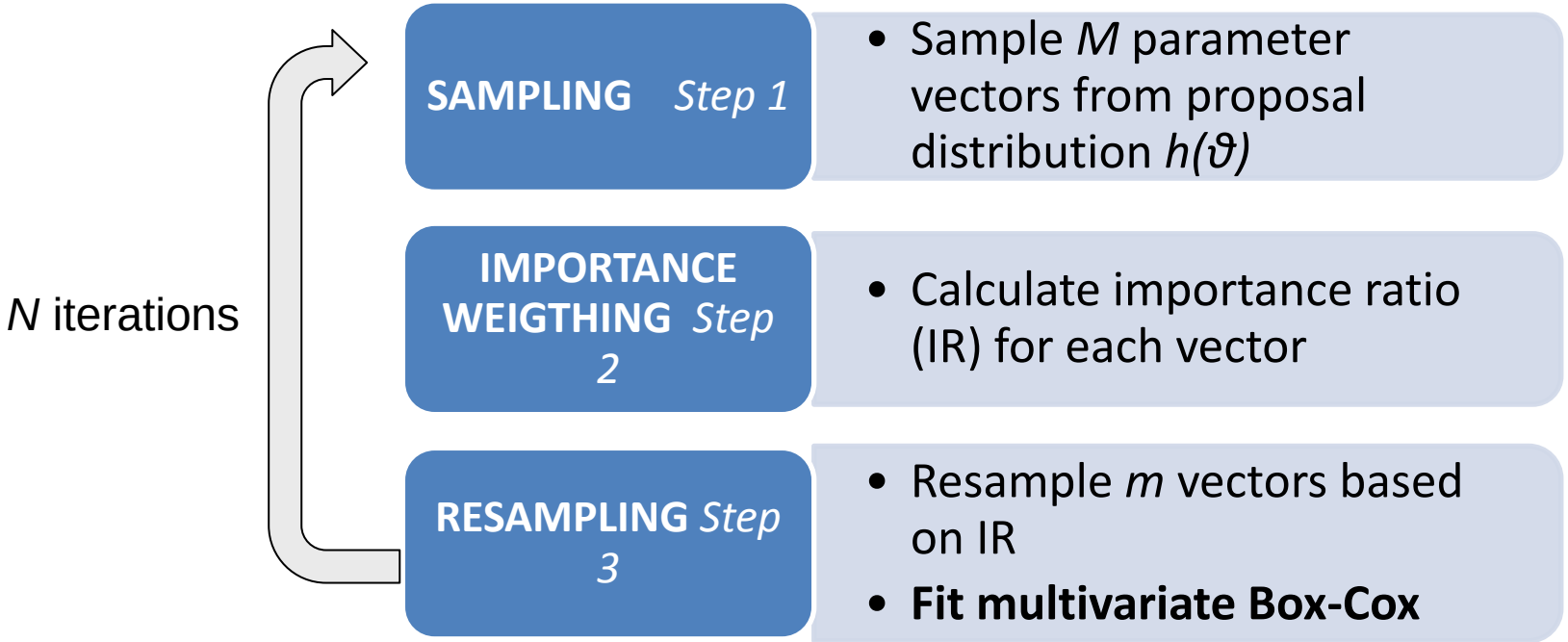
SIR optimization I

- SIR is a procedure with options

Number of initial samples	• The higher number the better • A costly way of increasing precision
Resampling	• Resampling can improve efficiency – but also decrease performance • With or without replacement?
Inflation of sampling distribution	• A too wide proposal is better than a too constrained – basis for inflation?

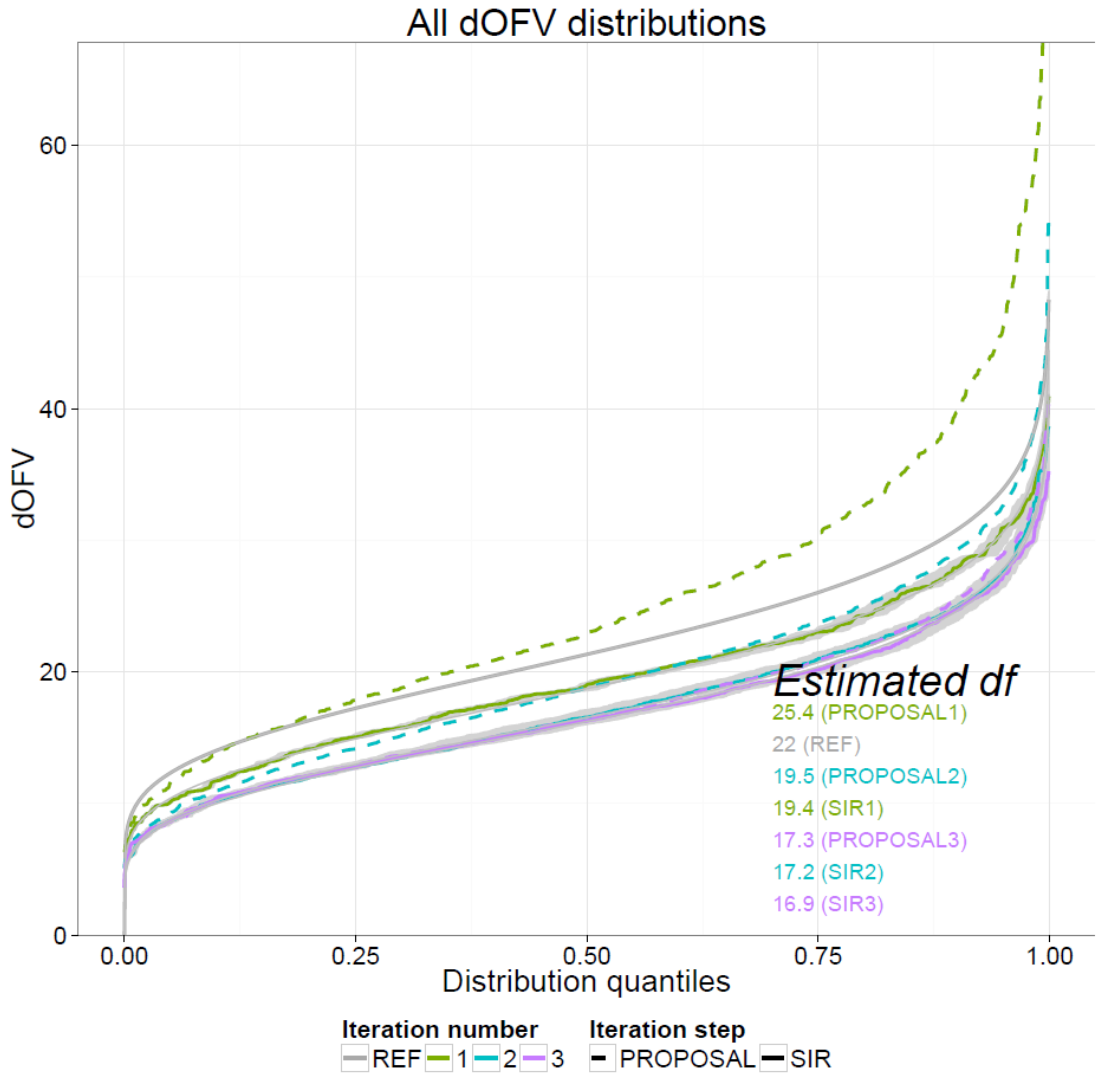
SIR optimization II - make SIR iterative

- Updating of proposal is more efficient than increasing initial sampling
- Fit multivariate Box-Cox to SIR output and use as new proposal



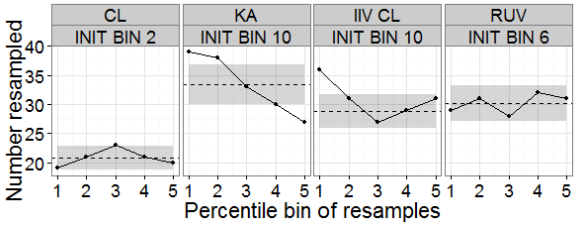
Dosne et al., PAGE 2016

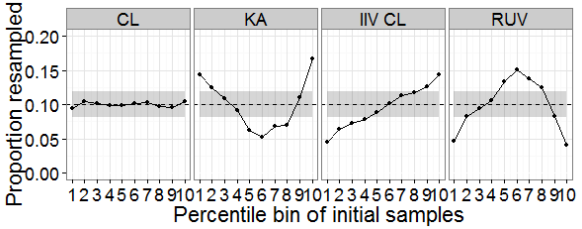
dOFV plot iterative SIR – convergence check

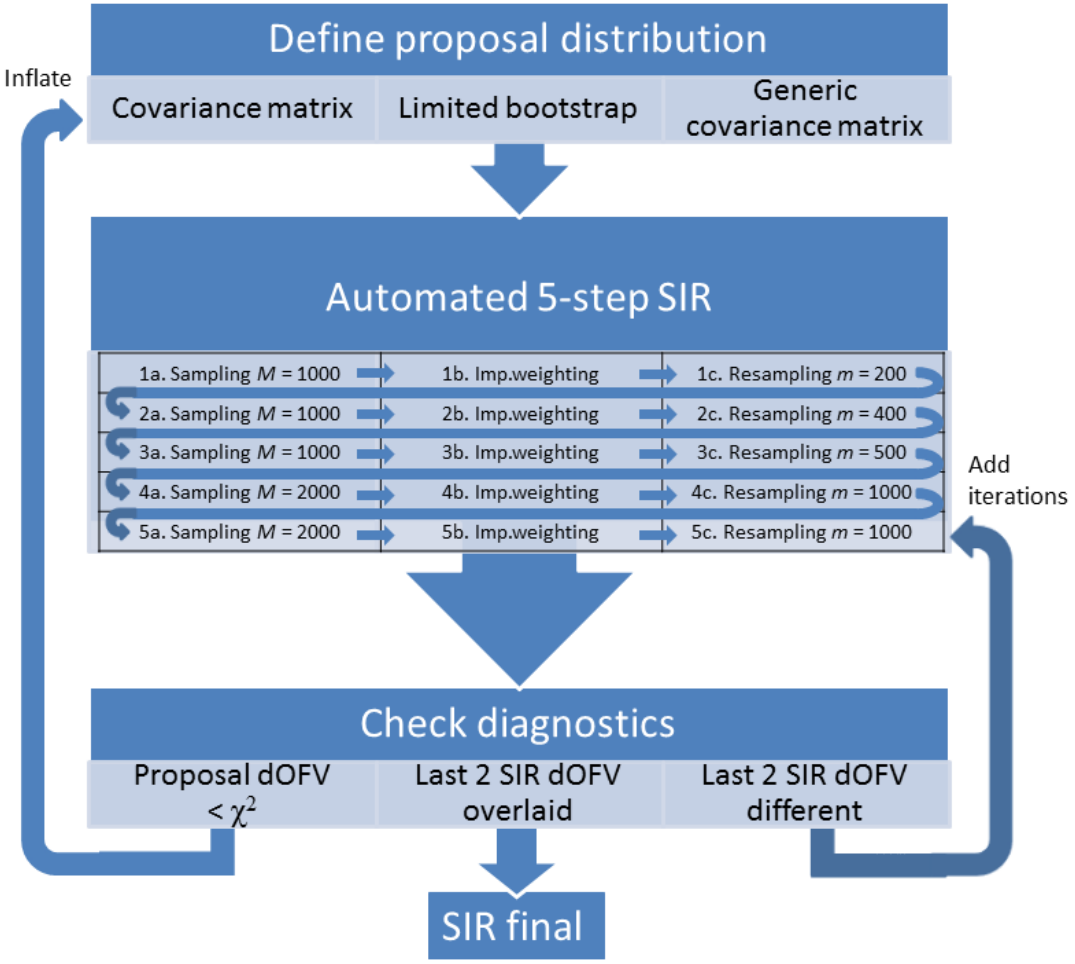


- ✓ dOFV distribution $\sim \chi^2$
- ✓ $df \leq n\,params$
- ✓ dOFV distribution stable over last 2 iterations

Implementation of SIR in PsN/NONMEM







Dosne et al., PAGE 2016

Conclusion SIR

- SIR

✓ allows for asymmetry in uncertainty distribution

✓ does not require parameter re-estimation

- "Fast and stable" method to assess parameter uncertainty, in particular if:
  - ✓ long estimation times
  - ✓ bootstrap convergence issues
  - ✓ unbalanced/small study designs
  - ✓ model-based meta-analysis
  - ✓ informative priors in model

References I

Ueckert S., Karlsson MO., Hooker AC. "Accelerating Monte Carlo power studies through parametric power estimation", Journal of Pharmacokinetics and Pharmacodynamics, vol. 43 (2), 223—234, 2016.

Vong C, Bergstrand M, Nyberg J, Karlsson MO. "Rapid sample size calculations for a defined likelihood ratio test-based power in mixed-effects models", AAPS Journal, vol. 14 (2), 176—186, 2012.

Dosne A.G., Bergstrand M., Karlsson MO. "Determination of Appropriate Settings in the Assessment of Parameter Uncertainty Distributions using Sampling Importance Resampling (SIR)", PAGE 24 (2015), ISSN 1871-6032.

Dosne A.G., Bergstrand M., Karlsson MO. "Application of Sampling Importance Resampling to estimate parameter uncertainty distributions", PAGE 22 (2013), ISSN 1871-6032.

Niebecker R., Karlsson MO. "Are datasets for NLME models large enough for a bootstrap to provide reliable parameter uncertainty distributions?", PAGE 22 (2013), ISSN 1871-6032.

Hooker AC., van Hasselt JGC. "Platform for adaptive optimal design of nonlinear mixed effect models", PAGE 22 (2013), ISSN 1871-6032.

Deng C., Plan EL., Karlsson MO. "Influence of clinical trial design to detect drug effect in systems with within subject variability", PAGE 24 (2015), ISSN 1871-6032.

Dosne A.G., Bergstrand M., Karlsson MO. "Improving The Estimation Of Parameter Uncertainty Distributions In Nonlinear Mixed Effects Models Using Sampling Importance Resampling" J Pharmacokinet Pharmacodyn (in press)

References II

Dosne A.G., Niebecker R., Karlsson MO. ” dOFV Distributions: A New Diagnostic For The Adequacy Of Parameter Uncertainty In Nonlinear Mixed-Effects Models Applied To The Bootstrap” J Pharmacokinet Pharmacodyn (in press)

Aoki Y., Nordgren R., Hooker AC. “Preconditioning of Nonlinear Mixed Effects Models for Stabilisation of Variance-Covariance Matrix Computations”, AAPS Journal, vol. 18 (2), 505--518, 2016.

Hooker A., “Model based adaptive design for nonlinear mixed effect models”, Oral presentation at ACOP 2013, http://www.acop7.org/previous-acop-meetings-acop-2013-program

Hooker AC and van Hasselt JGC , “Platform for adaptive optimal design of nonlinear mixed effect models”, Population Optimum Design of Experiments (PODE), Windelsham, UK, 2013 http://www.maths.qmul.ac.uk/~bb/PODE/PODE2013.html

Ryeznik Y, Sverdlov O and Hooker AC, “Adaptive dose finding for time-to-event outcomes with adaptive choice of patient number based on response rate”, Oral presentation at PAGE 2016,http://www.page-meeting.org/?abstract=5932

Ryeznik Y, Hooker AC Sverdlov A, “Adaptive designs for dose finding clinical trials with time-to-event outcomes”, Poster at PAGE 2015, http://www.page-meeting.org/default.asp?abstract=3608

References III

Strömberg EA, Hooker AC, "Simulated model based adaptive optimal design of adult to children bridging study using FDA stopping criteria.", Oral presentation at PAGE 2015, http://www.page-meeting.org/default.asp?abstract=3614

Aoki Y, Hamrén B, Röshammar D, Hooker AC. "Averaged Model Based Decision Making for Dose Selection Studies". www.page-meeting.org/?abstract=3121. 23rd meeting of the Population Approach Group in Europe (PAGE), Alicante, Spain, 2014.

Aoki Y, Hamrén B, Röshammar D, Hooker AC. "Averaged model based decision making for dose selection studies", European Medicines Agency/European Federation of PharmaceuticalIndustries and Associations workshop on the importance of dose finding and dose selection for the successful development, licensing and lifecycle management of medicinal products, Dec. 2014, European Medicines Agency, London UK

Aoki Y, Hooker AC, "Model Averaging and Selection methods for model structure and parameter uncertainty quantification", PAGE 2016 Poster, http://www.page-meeting.org/?abstract=5951

Dosne AG, Bergstrand M, Karlsson MO, Renard D, Heimann G, "Robust Inference in Thorough QT Studies", AcoP 2013, http://www.acop7.org/previous-acop-meetings-acop5-posters

Dosne AG, Bergstrand M, Karlsson MO, Renard D, Heimann G, "Robust Inference in Thorough QT Studies", PAGE 2016,http://www.page-meeting.org/?abstract=6021

Team @ UU

- Mats Karlsson – Prof
- Andrew Hooker – Assoc Prof
- Kristin Karlsson – Researcher
- Sebastian Ueckert – Researcher
- Yasunori Aoki – Postdoc
- Ronald Niebecker – Postdoc
- Chenhui Deng - Postdoc
- Anne-Gaelle Dosne – PhD student
- Eric Strömberg – PhD student
- Kajsa Harling – System developer
- Rikard Nordgren – System developer