Evaluation of bootstrap methods in nonlinear mixed-effects models

Hoai -Thu Thai[1,2],

France Mentré[1], Nicholas N.G Holford[3], Christine Veyrat-Follet[2], Emmanuelle Comets[1]

[1]INSERM, UMR 738, Univ Paris Diderot, Sorbonne Paris Cité, F-75018 Paris, France

[2]Drug Disposition Department, Sanofi, France

[3]Department of Pharmacology and Clinical Pharmacology, University of Auckland, Auckland, New Zealand

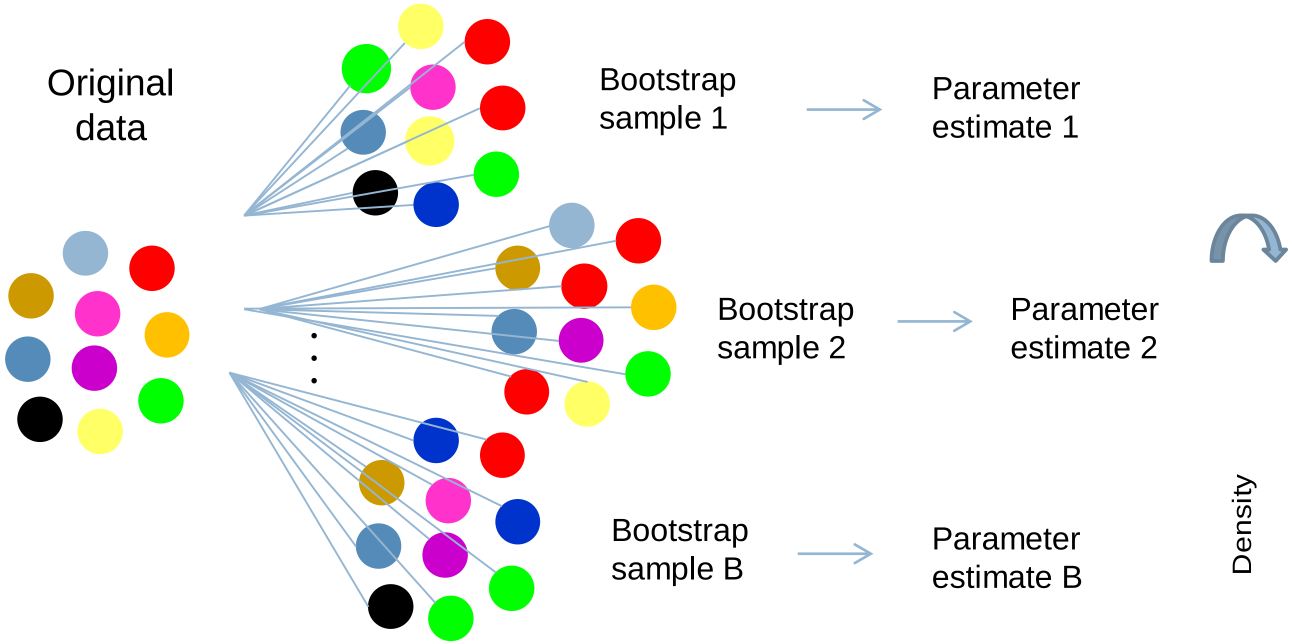
PAGE MEETING

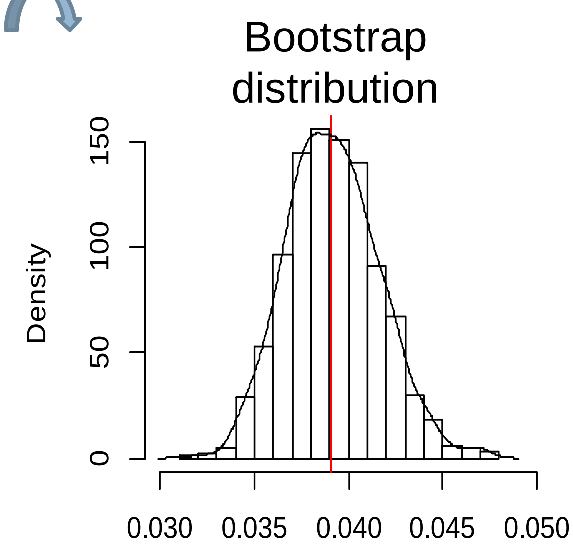
Glasgow, 14 June 2013

Principle of bootstrap (Efron 1979)[1]

- Resample with replacement the observed data to construct the distribution of an estimate or a statistic of interest
- A general bootstrap procedure in regression:

«Pull one self up with your own bootstrap»





- "Main" bootstrap methods
  - case bootstrap[2]: resample the pairs
  - residual bootstrap[1]: resample the residuals after model fitting

1. Efron B . *Annal Stat* 1979; 7: 1-26
2. Freedman D.A . *Annal Stat* 1981; 9: 1218-1228

Bootstrap in mixed-effects models (MEM)

Inserm
Institut national
de la santé et de la recherche médicale

MEM model

$$y_i = f(\xi_i, \mu, \eta_i) + g(\xi_i, \mu, \eta_i, \sigma)\varepsilon_i$$

μ : fixed parameters

$\eta_i \sim N(0, \Omega)$

$\varepsilon_i \sim N(0,1)$

- Take into account the characteristics of MEM[1]: repeated measures, heteroscedasticity, nonlinearity
- Respect two levels of variability: interindividual and residual variabilities[2,3]

- Previous simulation study in linear MEM to compare bootstrap methods (THAI HT et al. Pharm Stat 2013; 12 (3): 129-140)
  - poor performance of bootstraps resampling only one level of variability
  - good performance of three bootstrap methods resampling two levels of variability
  - some differences between bootstrap methods when applied to a real unbalanced dataset

1. Das S, Krishen A. *J Stat Plan Inference* 1999; 75: 237-245
2. Ocana J, Halimi E. *Mathematics Preprint Series 2005;* 367
3. Leeuw J, Meijer E. *Handbook of multilevel analysis*. Springer 2008

Objective

Inserm
Institut national
de la santé et de la recherche médicale

Evaluate the performance of bootstrap methods for estimating uncertainty of parameters in nonlinear mixed-effects models (NLMEM) using a simulation study

Bootstrap methods

Inserm
Institut national
de la santé et de la recherche médicale

- Case bootstrap (B_{case})
- Nonparametric random effect and residual bootstrap ($B^{NP}_{\eta,\varepsilon}$)
- Parametric random effect and residual bootstrap ($B^{P}_{\eta,\varepsilon}$)

Resample the individuals

$$(\xi_i, y_i) \rightarrow (\xi_i^*, y_i^*)$$

Bootstrap methods

- Case bootstrap (B_{case})
- Nonparametric random effect and residual bootstrap ($B^{NP}_{\eta,\varepsilon}$)
- Parametric random effect and residual bootstrap ($B^{P}_{\eta,\varepsilon}$)

Fit the model to the data and estimate random effects $\hat{\eta}_i$ and standardized residuals $\hat{\varepsilon}_{ij}$

↓

Correct for shrinkage[1,2] in random effects and residuals

↓

Resample the random effects
Resample the residuals

↓

Generate the bootstrap observations

$$y_i^* = f(\xi_i, \hat{\mu}, \hat{\eta}_i^*) + g(\xi_i, \hat{\mu}, \hat{\eta}_i^*, \hat{\sigma})\varepsilon_i^*$$

1. Carpenter JR. *Appl Statist* 2003, 52: 431-443
2. Wang J et el.*Comput Methods Programs Biomed 2006*; 82: 130-143.

Bootstrap methods

Inserm
Institut national
de la santé et de la recherche médicale

- Case bootstrap (B_{case})
- Nonparametric random effect and residual bootstrap ($B^{NP}_{\eta,\varepsilon}$)
- Parametric random effect and residual bootstrap ($B^{P}_{\eta,\varepsilon}$)

Fit the model to the data

↓

Simulate the random effects from $N(0,\hat{\Omega})$
Simulate the residuals from N(0,1)

↓

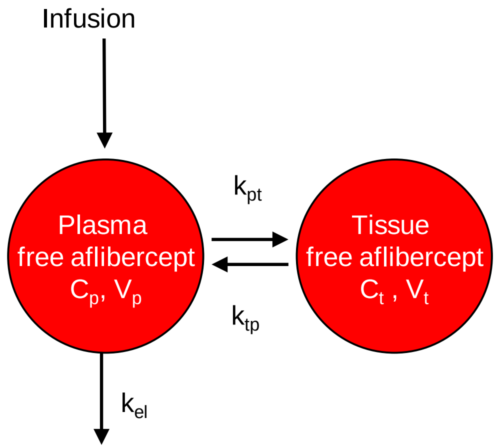
Generate the bootstrap observations

$$y_i^* = f(\xi_i, \hat{\mu}, \hat{\eta}_i^*) + g(\xi_i, \hat{\mu}, \hat{\eta}_i^*, \hat{\sigma})\varepsilon_i^*$$

Motivating example

Inserm
Institut national
de la santé et de la recherche médicale

- Pharmacokinetic data: aflibercept[1], an anti-angiogenic agent
  - 2 clinical trials[2]
    - phase I TCD 6120 trial (first dose) (N_1=53, n_1=median of 9)
    - phase III VITAL trial (first two doses) (N_2=291, n_2=2)
  - two - compartment PK model with first-order elimination
    - exponential model for IIV
    - proportional model for residual error



- Model fit to the data: SAEM (MONOLIX 4.1.2)[1]

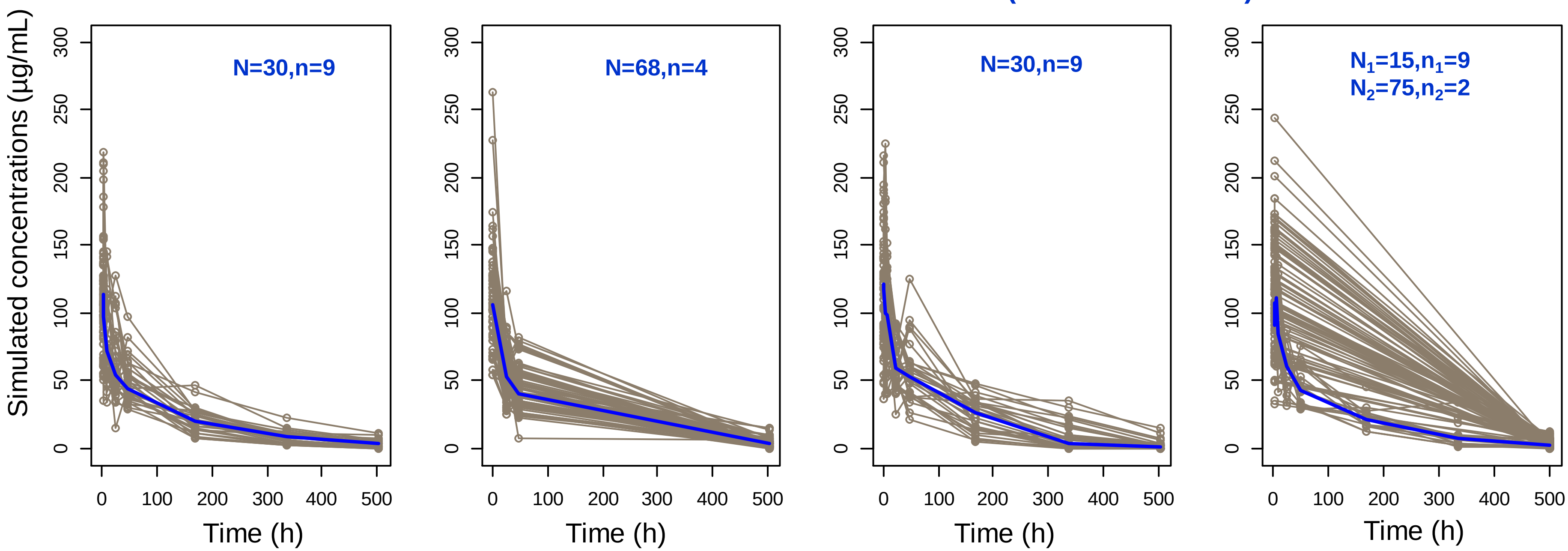
Parameter	Estimate (RSE)	IIV (%)
CL (L/hr)	0.04 (2)	28.8 (5)
V1 (L)	3.62 (2)	19.7 (10)
Q (L/hr)	0.14 (15)	111(12)
V2 (L)	2.9 (5)	-
Corr(CL,Q) (ρ)	0.90 (8)	
σ_p (%)	24.8 (4)	

1. Chu QS. *Expert Opin Biol Ther* 2009;9:263-271
2. Gaya A et al. *Cancer Treat Rev* 2012 38): 484-493
3. Lavielle M. MONOLIX. MONOLIX group, Orsay, France, 2008.

Simulation settings

Inserm
Institut national de la santé et de la recherche médicale

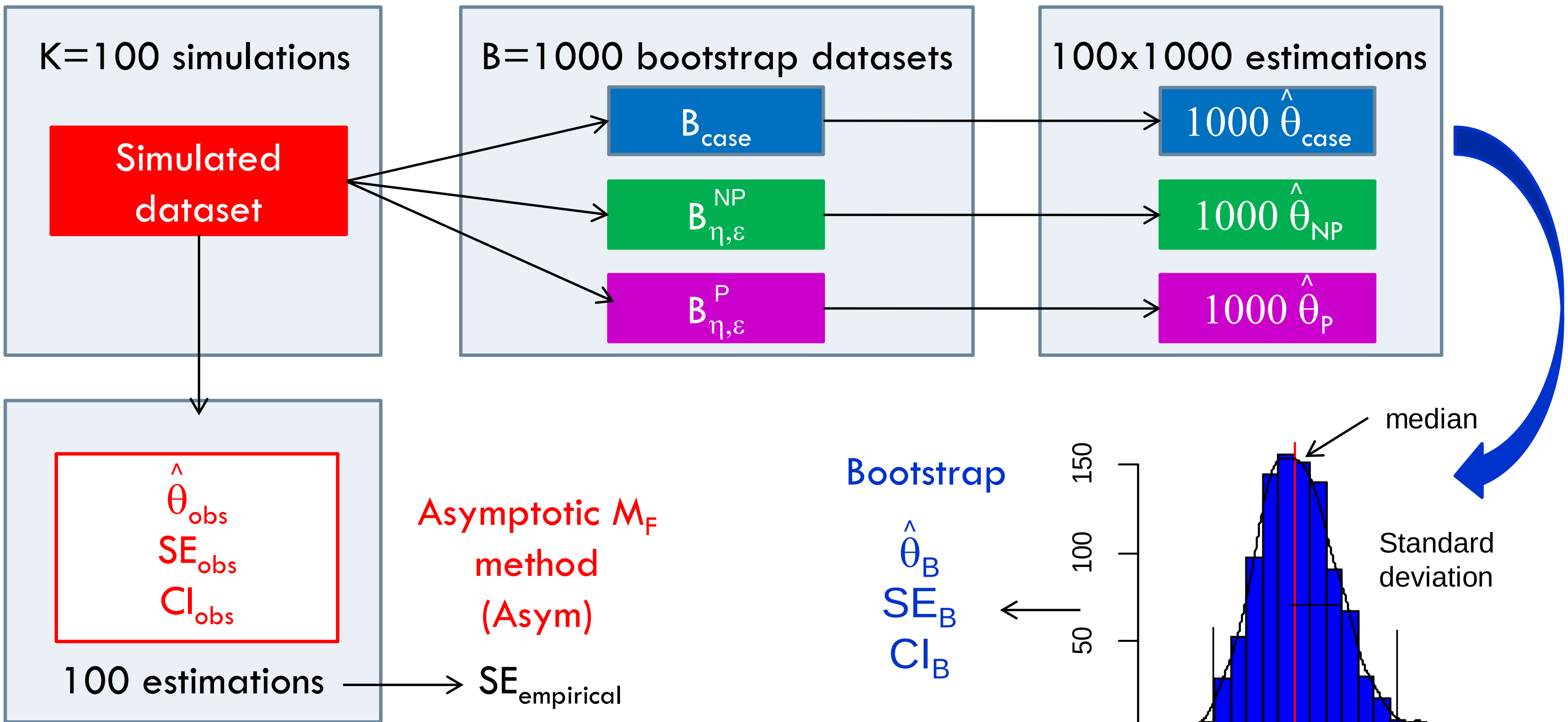
Frequent balanced design first-order | **Sparse balanced design first-order** | **Frequent balanced design mixed-order (Michaelis-Menten)** | **Unbalanced design first-order**

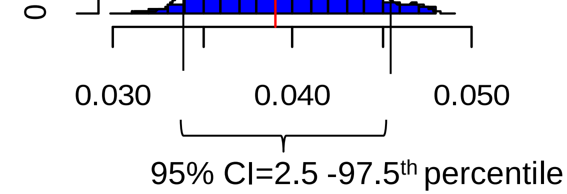


K=100 replications for each design

Bootstrap settings

Inserm
Institut national
de la santé et de la recherche médicale

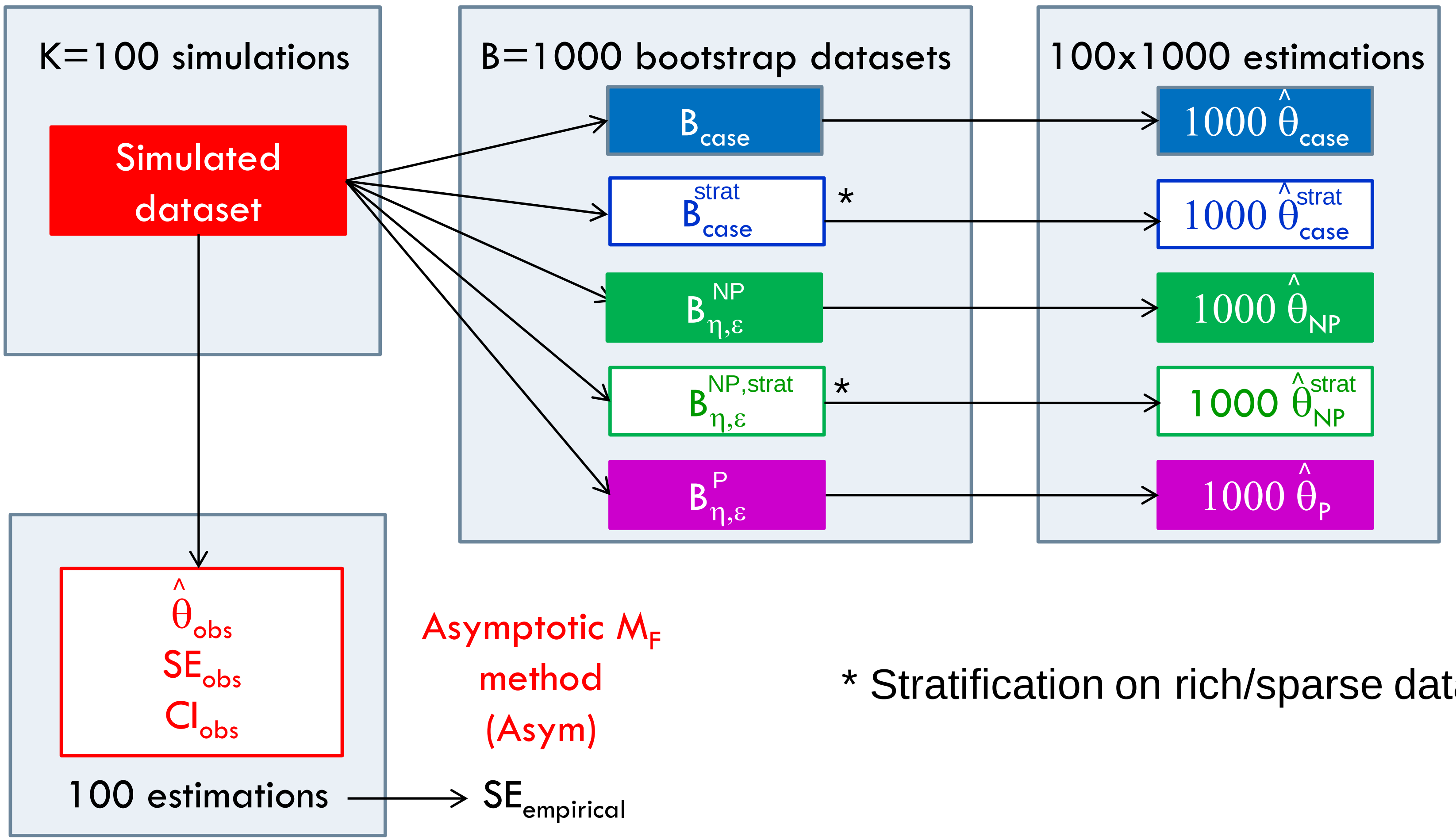




Estimation method: SAEM (Monolix 4.1.2)

Fisher information matrix (M_F) computed by linearisation

Bootstrap settings for the unbalanced design



* Stratification on rich/sparse data

Evaluation

- For each method (k=100 replicated datasets)
  - relative bias of parameter (%)

$$\mathrm{RBias}(\hat{\theta}_{\mathrm{B}}^{(l)}) = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{\hat{\theta}_{\mathrm{B};k}^{(l)} - \hat{\theta}_{k}^{(l)}}{\hat{\theta}_{k}^{(l)}} \times 100\right)$$

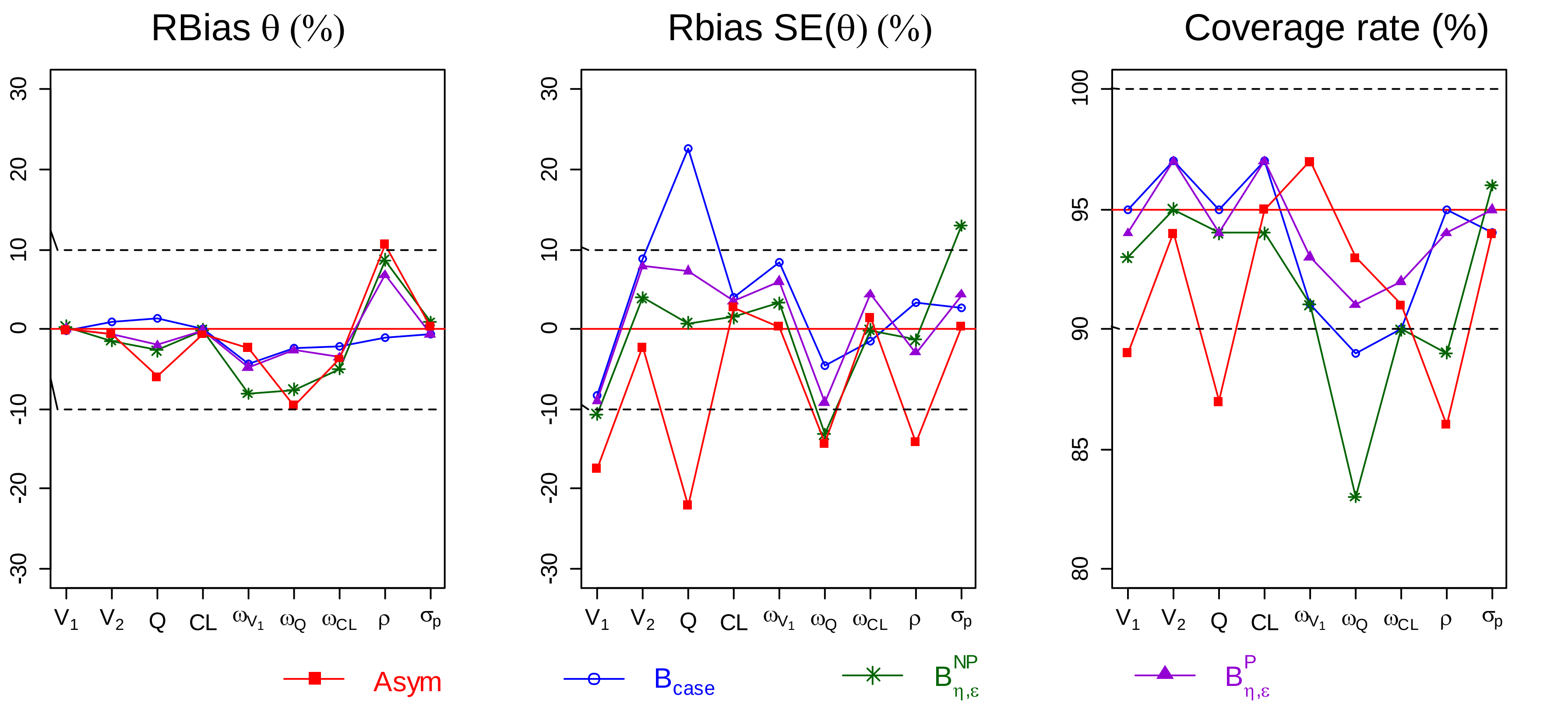
  - relative bias of SE (%)

$$\mathrm{RBias}(\widehat{SE}_{\mathrm{B}}^{(l)}) = \frac{\frac{1}{K}\sum_{k=1}^{K}\widehat{SE}_{\mathrm{B;k}}^{(l)} - \widehat{SE}_{\mathrm{empirical}}^{(l)}}{\widehat{SE}_{\mathrm{empirical}}^{(l)}} \times 100$$

  - coverage rate of 95% CI: % bootstrap CI contain the true value of parameter

- "Good" bootstrap
  - relative bias of parameters and SE: unbiased (±10%)
  - coverage rate of 95% CI: good (90-100%)

Frequent balanced design with first-order elimination



✓ Parametric bootstrap had the best performance

✓ Asymptotic approach performed slightly less well than the bootstrap methods

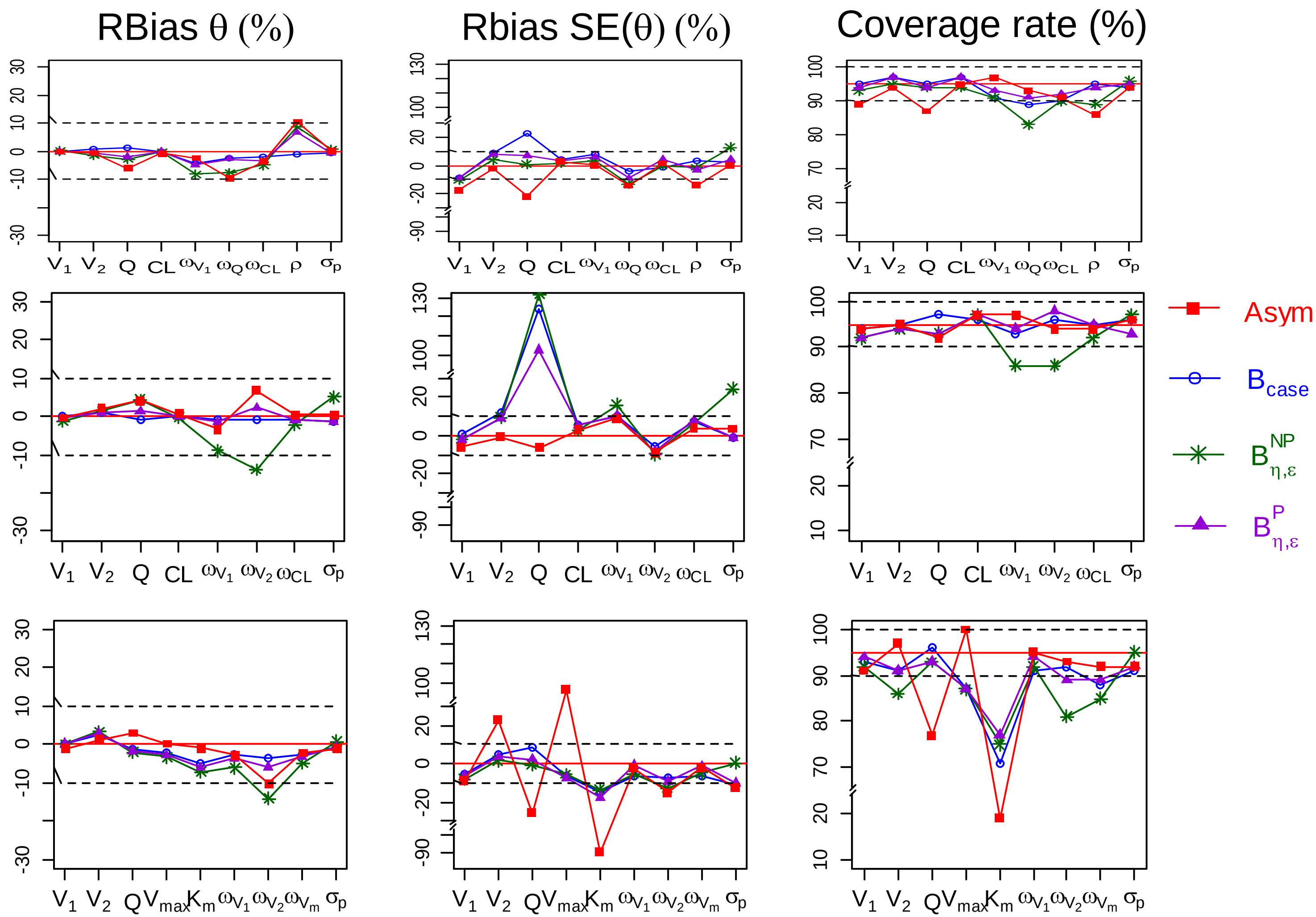
Three balanced designs

Inserm
Institut national de la santé et de la recherche médicale

Frequent design
first order
(N=30, n=9)

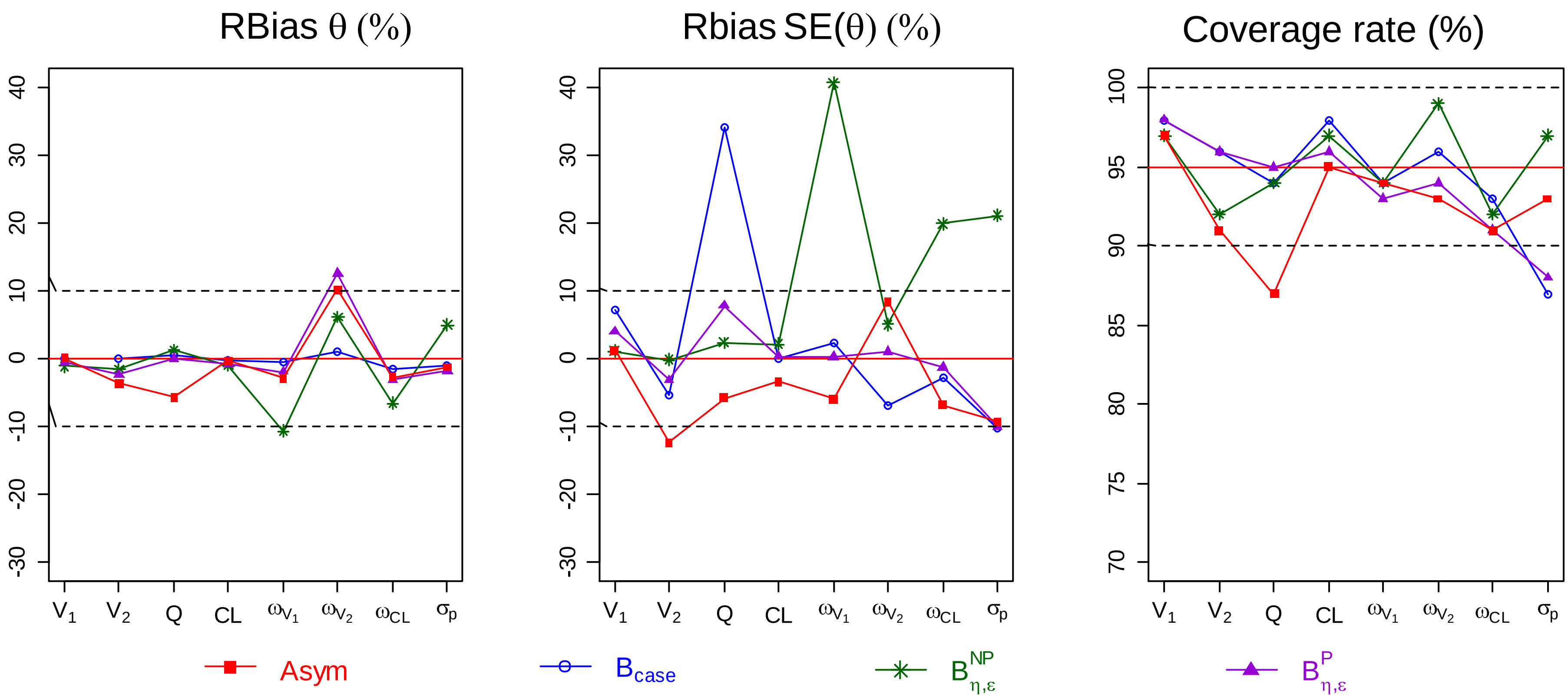
Sparse design
first order
(N=68, n=4)

Frequent design
mixed-order
(N=30, n=9)



Unbalanced design with first-order elimination

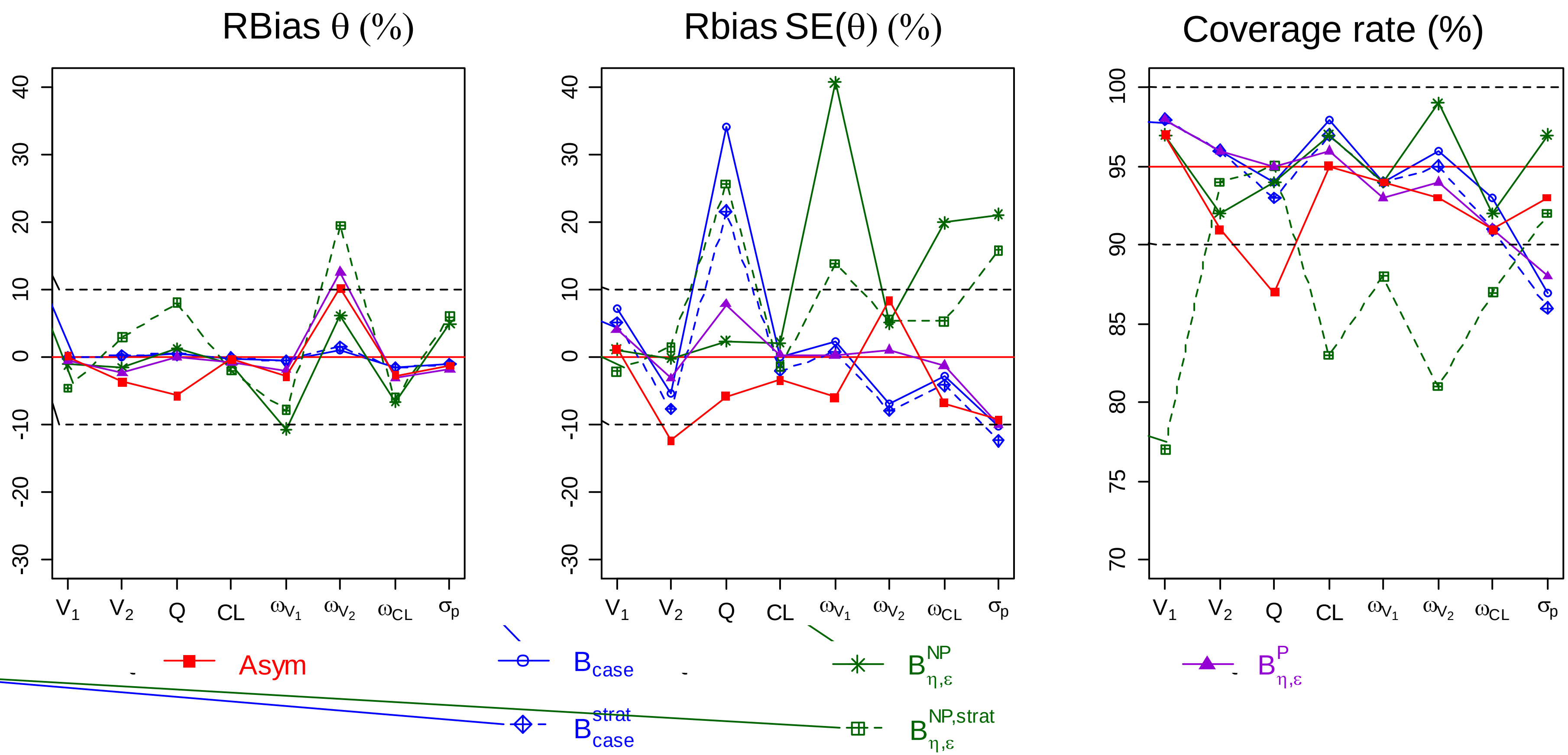
Inserm
Institut national de la santé et de la recherche médicale



✓ Asymptotic approach and parametric bootstrap performed reasonably well and better than other bootstrap methods

Unbalanced design with first-order elimination

Inserm
Institut national de la santé et de la recherche médicale



✓Stratification improved slightly the performance of case bootstrap and but degraded that of nonparametric random effect and residual bootstrap in term of coverage rate

Conclusions

Inserm
Institut national de la santé et de la recherche médicale

- Better estimation of uncertainty by the bootstrap methods than the asymptotic method in NLMEM with high nonlinearity
- Caution with bootstrap methods in presence of large fluctuations in parameter estimates between bootstrap samples
- The choice of bootstrap methods in NLMEM
  - parametric bootstrap: best description of uncertainty

    ⟶ study robustness in case of model misspecification
  - case bootstrap: fast, simple and robust (e.g heteroscedasticity, missing data)

    ⟶ evaluate stratification for complex designs
  - nonparametric random effect and residual bootstrap: maintain the same design as the original dataset

    ⟶ improve correction for shrinkage in unbalanced designs

Acknowlegements

- Sanofi for financial support

Annexes

Correction for variance underestimation

- Correction for random effects[1,2]
  - center: $\tilde{\eta}_i = \hat{\eta}_i - \bar{\eta}_i$
  - transform: by the ratio between empricial vs estimated variance-covariance (A_η)
    - Cholesky decomposition: matrice Ω_R positive
    - EVD (*Eigen Value Decomposition*): matrice Ω_R semi-positive

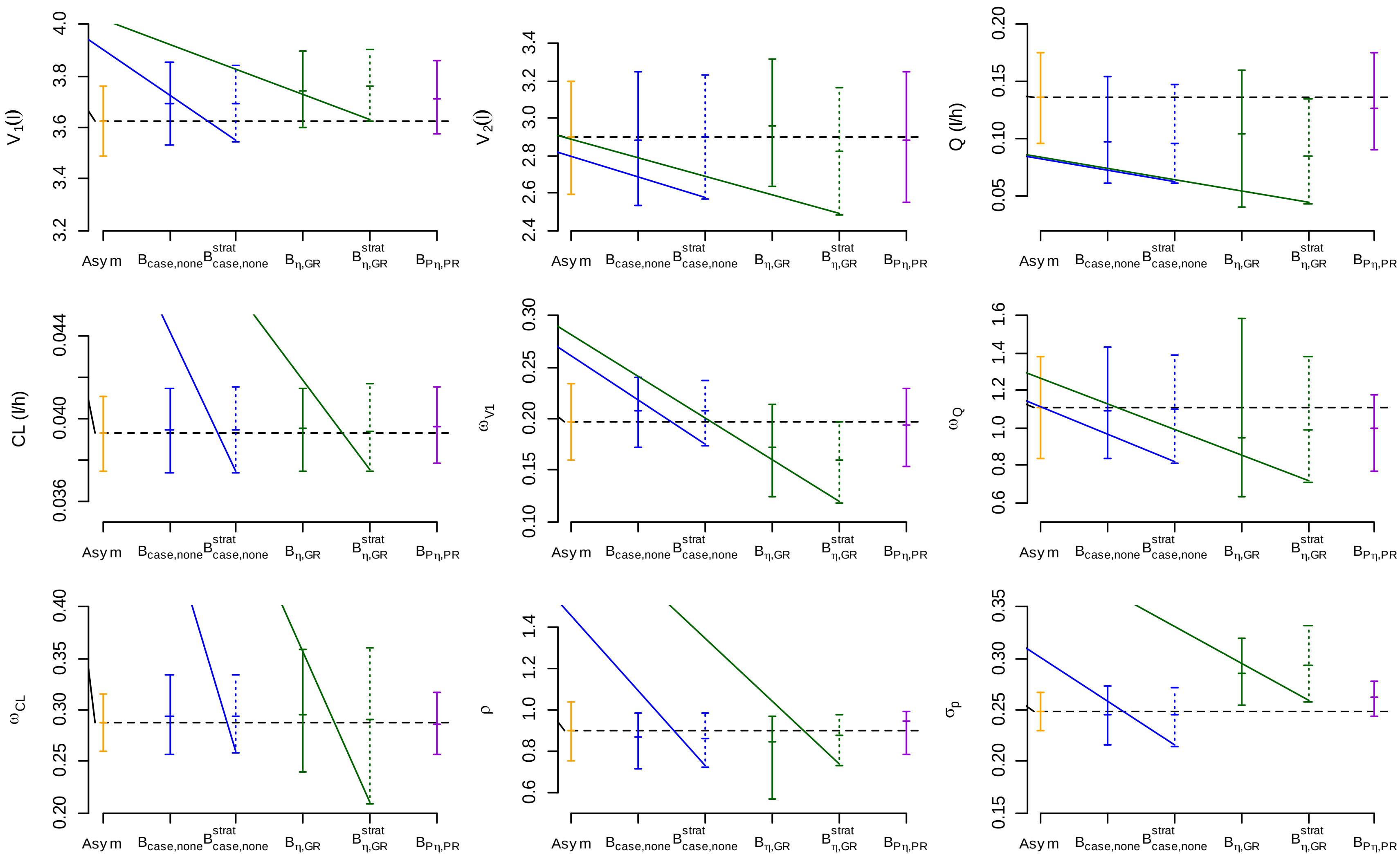
$$\hat{\eta}_i' = \tilde{\eta}_i \times A_\eta$$

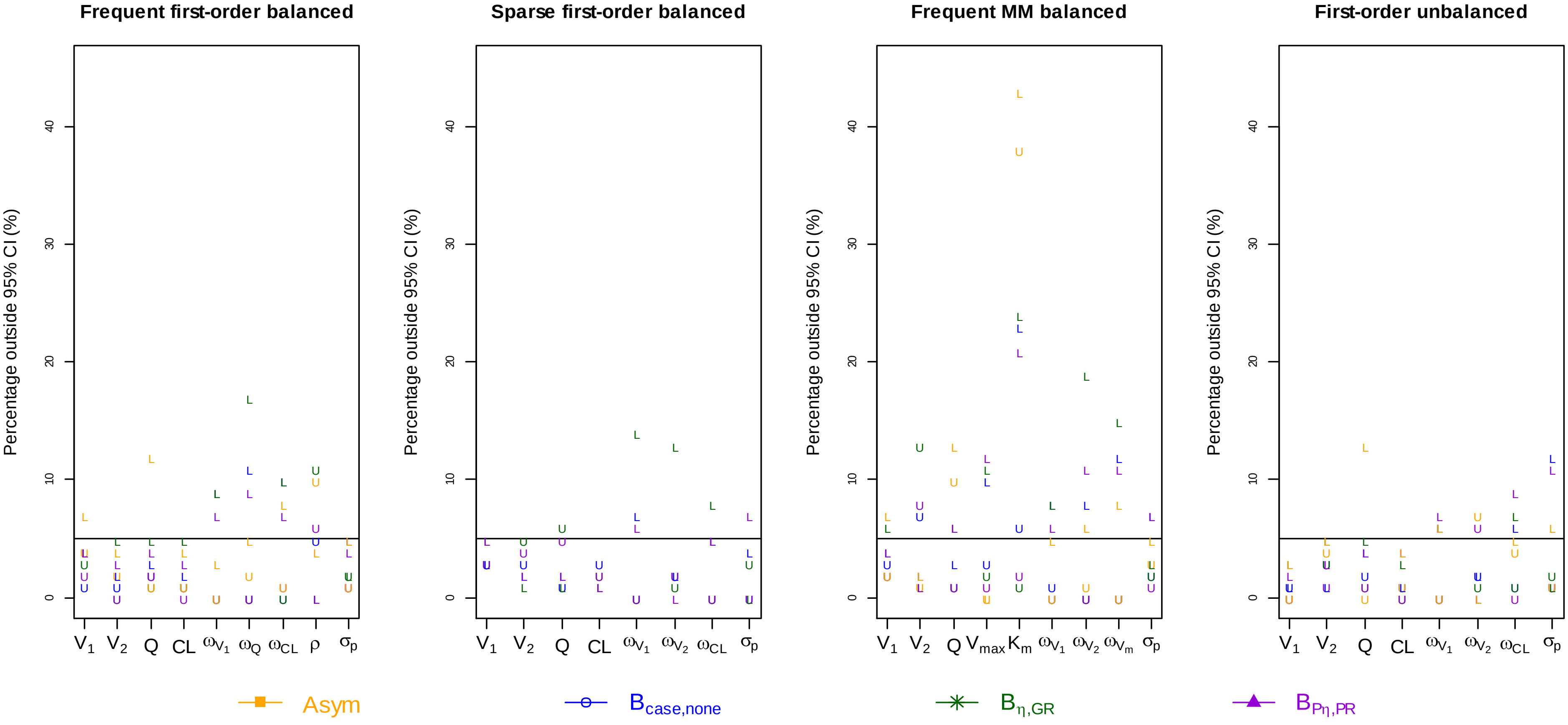
- Correction for residuals[1,2]
  - center: $\tilde{\epsilon}_{ij} = \hat{\epsilon}_{ij} - \bar{\epsilon}_{ij}$
  - transform: by the ratio between empirical vs estimated variance-covariance(A_σ)
    - homoscedastic error: $A_\sigma = \hat{\sigma} / sd(\tilde{\epsilon}_{ij})$
    - heteroscedastic error: $A_\sigma = 1$ / SD(standardized residuals)

$$\hat{\epsilon}_{ij}' = \tilde{\epsilon}_{ij} \times A_\sigma$$

1. Carpenter JR. *Appl Statist* 2003, 52: 431-443
2. Wang J et al.*Comput Methods Programs Biomed 2006*; 82: 130-143.

Application to aflibercept data





Frequent first-order balanced
Sparse first-order balanced
Frequent MM balanced
First-order unbalanced
Percentage outside 95% CI (%)
0
10
20
30
40
V_1 V_2 Q CL ω_{V_1} ω_Q ω_{CL} ρ σ_p
V_1 V_2 Q CL ω_{V_1} ω_{V_2} ω_{CL} σ_p
V_1 V_2 Q V_{max} K_m ω_{V_1} ω_{V_2} ω_{V_m} σ_p
V_1 V_2 Q CL ω_{V_1} ω_{V_2} ω_{CL} σ_p
Asym
$B_{case,none}$
$B_{\eta,GR}$
$B_{P\eta,PR}$