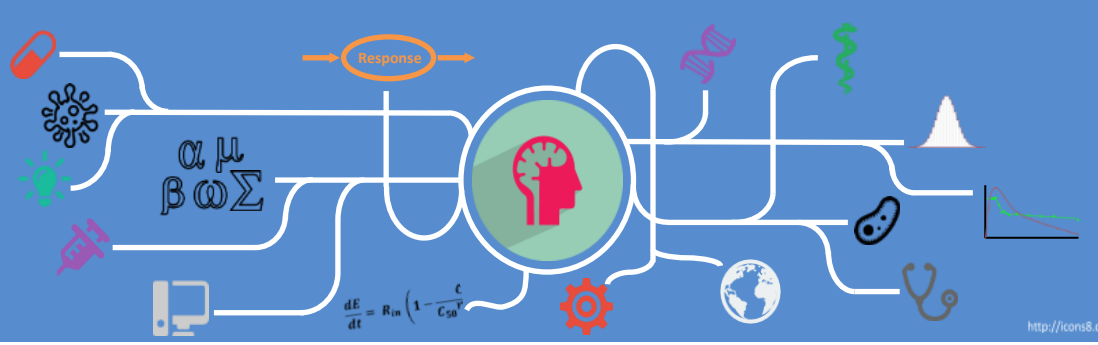


24th PAGE Meeting
Lewis Sheiner Student Session
4th June, 2015

MODELLING PHARMACOGENETIC DATA IN POPULATION STUDIES DURING DRUG DEVELOPMENT

ADRIEN TESSIER[1,2], JULIE BERTRAND[3], MARYLORE CHENEL[2], EMMANUELLE COMETS[1,4]

[1] IAME, INSERM UMR 1137, University Paris Diderot, Sorbonne Paris Cité
[2] Institut de Recherches Internationales Servier
[3] Genetic Institute, University College of London
[4] INSERM CIC 1414, University Rennes 1

iame UMR 1137
Infection • Antimicrobials • Modelling • Evolution

Inserm

université PARIS DIDEROT PARIS 7

SERVIER

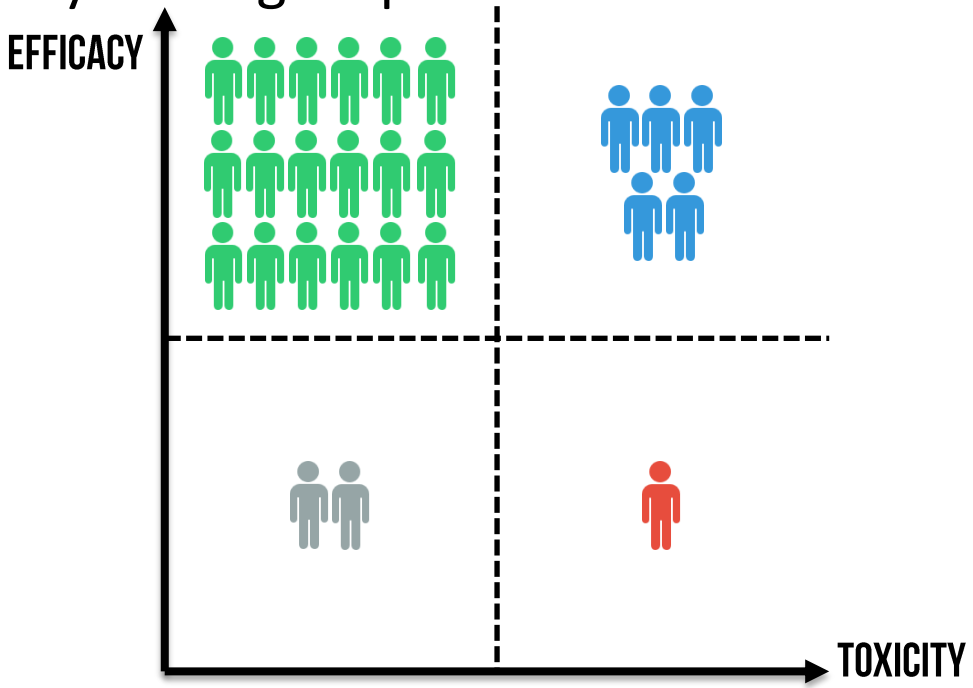
INTRODUCTION

CONTEXT

OBJECTIVES

PHARMACOGENETICS

Pharmacogenetics: study of the variation in genetics in relation to the interindividual variability in drug response[1]



Identifying the genetic variants related to response variability can significantly **increase drug efficacy**[2] and **reduce toxicity**[3]

1. Motulsky 1969
2. Rosell et al. 2009 (gefetinib)
3. Mallal et al. 2008 (abacavir)

PHARMACOGENETICS

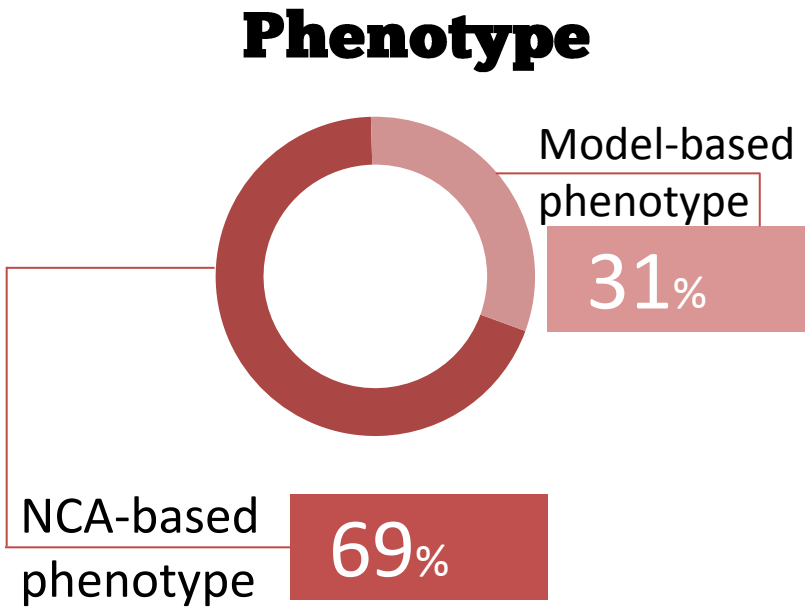
Pharmacogenetics: study of the variation in genetics in relation to the interindividual variability in drug response[1]

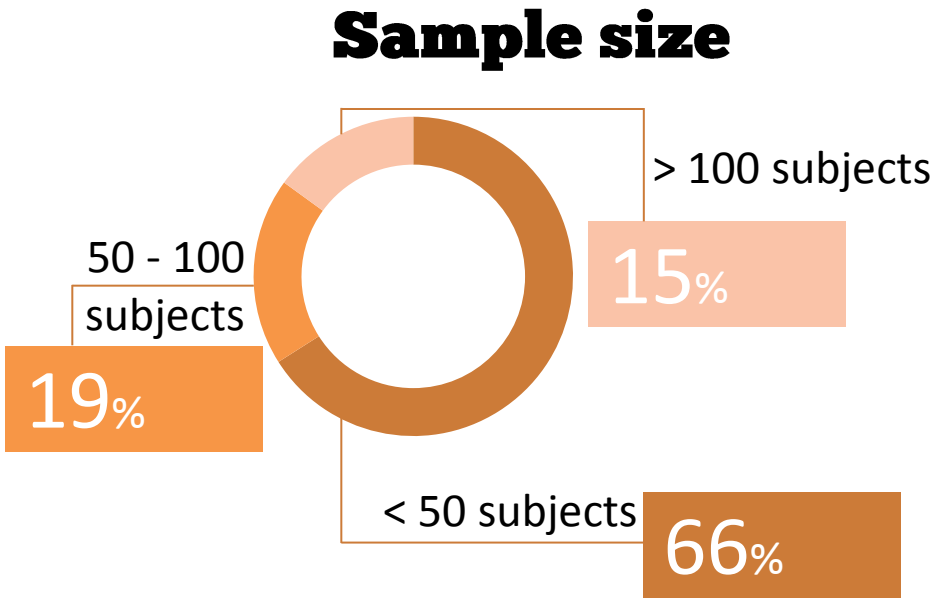
genetic variation: Single Nucleotide Polymorphisms (**SNPs**)
variation of only one base in genomic sequence
increasingly screened in clinical studies (DNA microarray)

interindividual variability in pharmacokinetics (**PK**): variation in the enzymatic activity (metabolism, transport)

1. Motulsky 1969

MULTIPLE METHODOLOGIES FOR PHARMACOGENETIC STUDIES





Survey on 85 pharmacogenetic studies published during the period 2010 - 2012 (Tessier et al. 2015)

Most analyses use NCA-based phenotype estimated from a limited number of subjects

MOTIVATING EXAMPLE
PHASE I CLINICAL STUDY

A drug S developed by Servier in phase I clinical development

PK:

N = 78 healthy volunteers
rich sampling design **(n = 16 samples per subject)**
nonlinear PK with doses

Pharmacogenetics:

all subjects genotyped for **176 SNPs** of genes known to be involved in the PK of drugs (metabolic enzymes, transporters, nuclear receptors)

- **challenging settings to compare NCA and model-based analysis**

THREE OBJECTIVES

PART 1: Assessment of pharmacogenetic analysis methods

To compare the ability of different **PK phenotypes** to detect genetic effects

To assess the performance of different **association tests**

PART 2: Enhance detection of genetic variants through combined designs

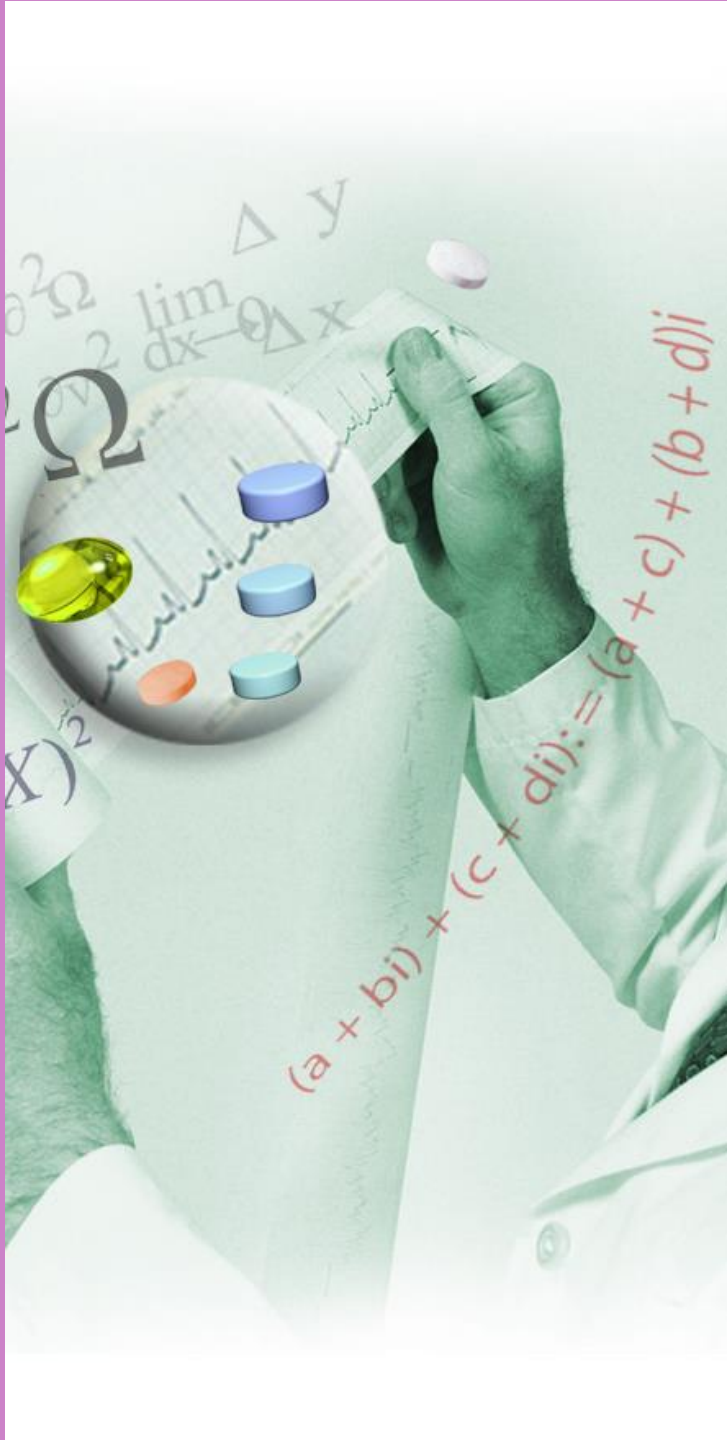
To assess **combined analysis** of phase I and II data

PART 3: Use of the conditional distribution to enhance genetic covariates analysis

To assess a PK **phenotype enrichment approach**

GENERAL METHODS

SIMULATION STUDY ASSOCIATION TESTS



SIMULATION STUDY
GENOTYPES

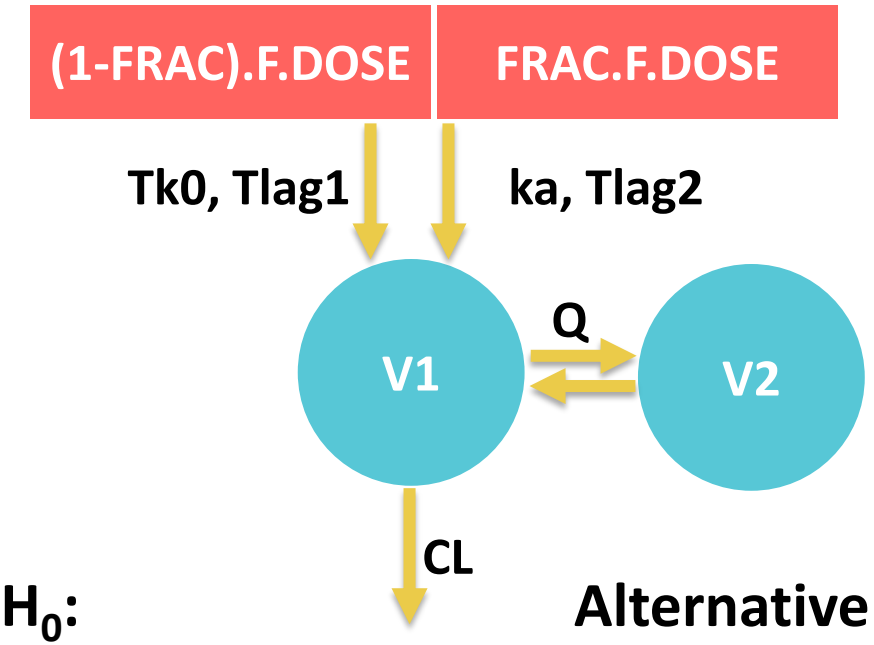
176 SNPs simulated based on the DNA microarray developed by Servier, retaining correlations between variants found in the human genome

using a reference panel of Hapmap genotypes data set[1] and a specialised software (Hapgen2)[2]

1. International HapMap Consortium 2003
2. Su et al. 2011

SIMULATION STUDY
POPULATION PHARMACOKINETIC MODEL

Nonlinearity on drug absorption (F and FRAC parameters)



Null hypothesis H_0:
No genetic effect

Alternative hypothesis H_1:
Genetic variants affect drug clearance

SIMULATION STUDY
GENETIC EFFECT

Under H_1, 6 SNPs drawn randomly affect the clearance:

$$log(CL_i) = log(\mu_{CL}) + \sum_{k=1}^{6} \beta_k \times SNP_{ik} + \eta_{i_{CL}}$$

$SNP_{ik} = \{0, 1, 2\}$: additive genetic model

β_k: effect size associated to the genotype SNP_{ik}, depends on:

p_k: the frequency of the minor allele

$\boldsymbol{R_{GC_k}}$: % of the interindividual variability in CL explained by the SNP[1]

SNP	1	2	3	4	5	6
$\mathbf{R_{GC_k}}$	1 %	2 %	3 %	5 %	7 %	12 %

1. Bertrand and Balding. 2013

SIMULATION STUDY
PHENOTYPES: NCA VS NLMEM

Individual PK profiles were simulated from PK model with genetic effects

- AUC estimated through **NCA**
  - normalised by the doses

- Individual clearances (EBE_{CL}) estimated by **NonLinear Mixed effects Models** (**NLMEM**) under H_0
  - Stochastic Approximation Expectation Maximisation algorithm (SAEM)[1]
  - Monolix software (v 4.2.2)[2]

All phenotypes were log-transformed

1. Kuhn and Lavielle. 2004
2. www.lixoft.eu

ASSOCIATION TESTS
GENETIC DATA CONSIDERATIONS

Posterior PG analysis in a PK study

No assumption of genetic mechanisms influencing drug PK

The 176 simulated SNPs are tested

Correlations between variants

Family Wise Error Rate (FWER) correction

FWER controlled around 20%

FWER correction: $1 - (1 - FWER)^{1/N_{SNP}}$

ASSOCIATION TESTS
USUAL STRATEGY

1. Stepwise procedure

Iterative method

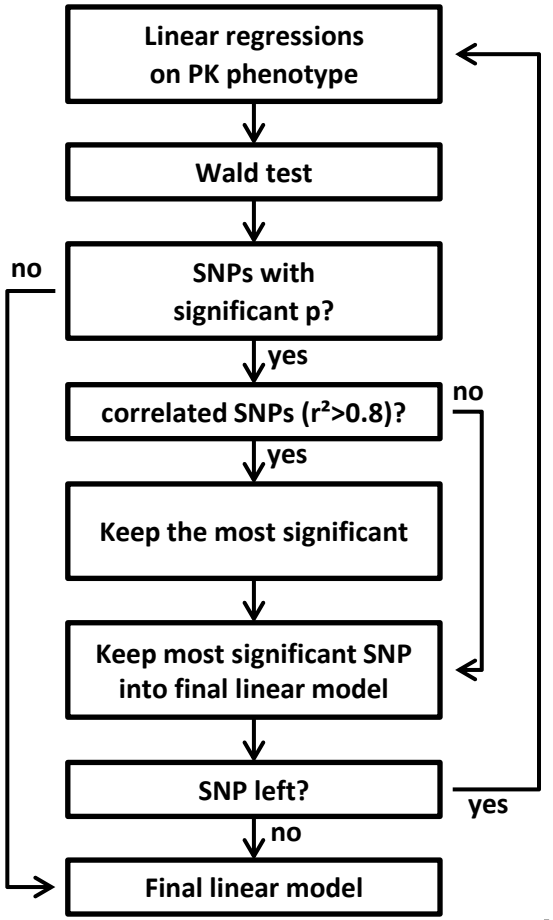
Effect sizes estimated through **univariate regressions**

$$\hat{\beta}_k = argmin(Y - X\beta)^2$$

Variants selected using a Wald test

FWER correction for significance threshold

Taking into account correlations between variants[1]



1. Inspired from Lehr et al. 2010

ASSOCIATION TESTS
PENALISED REGRESSIONS

Multivariate analysis

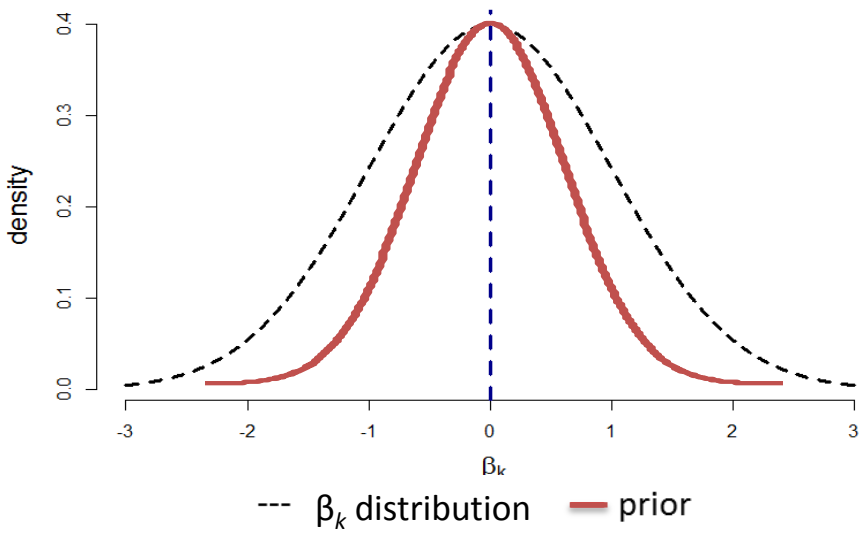
$$\hat{\beta}_k = argmin(Y - X\beta)^2 + penalty$$

2. Ridge regression[1]

Gaussian prior

Wald test for variable selection

FWER correction for significance threshold



1. Cule et al. 2011

ASSOCIATION TESTS
PENALISED REGRESSIONS

Multivariate analysis

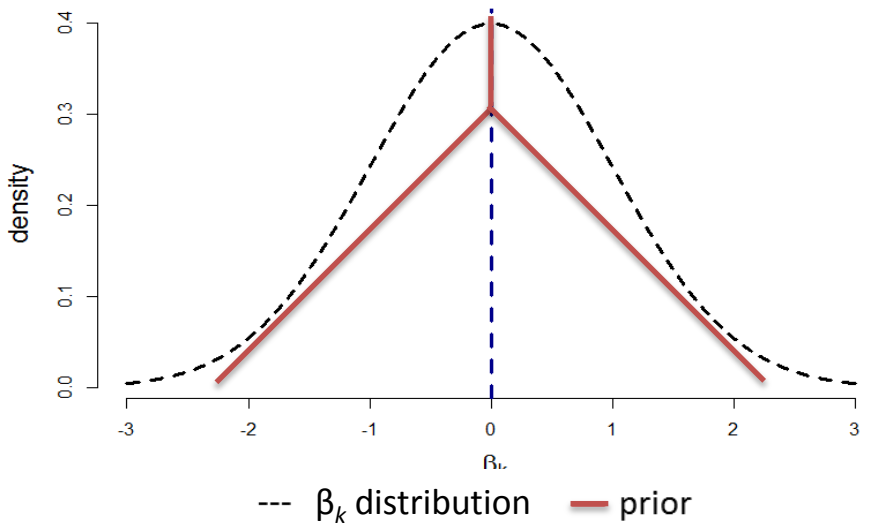
$$\hat{\beta}_k = argmin(Y - X\beta)^2 + penalty$$

3. Lasso[1]

Double Exponential (DE) prior

Set some coefficients to 0 (no test)

FWER correction to compute penalty



1. Tibshirani 1994

ASSOCIATION TESTS
PENALISED REGRESSIONS

Multivariate analysis

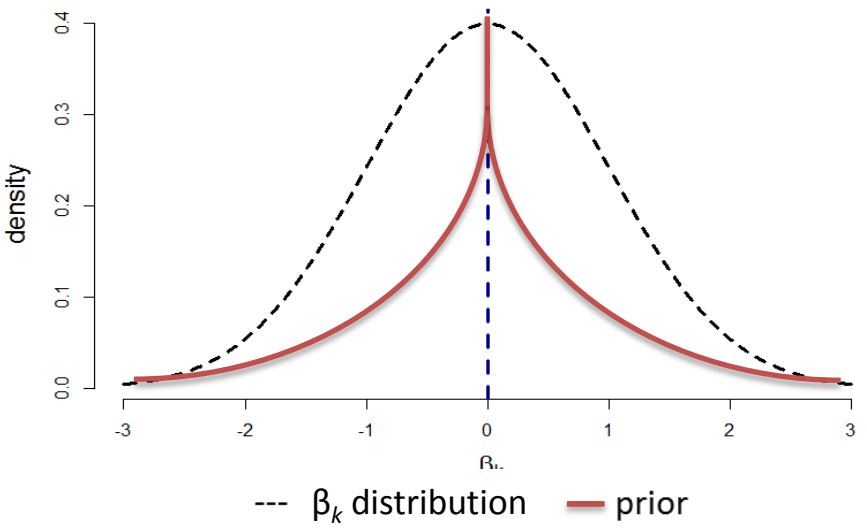
$$\hat{\beta}_k = argmin(Y - X\beta)^2 + penalty$$

4. HyperLasso[1]

Normal exponential gamma (NEG) prior
- shape and scale parameters
- sparser solutions

Set some coefficients to 0 (no test)

FWER correction to compute penalty



1. Hoggart et al. 2008

PART 1

ASSESSMENT OF PHARMACOGENETIC ANALYSIS METHODS

Tessier A, Bertrand J, Chenel M, Comets E. AAPS J. 2015

SCENARIOS

S_{real}

Phase I design

Administration	Single dose
Number of subjects (N_1)	78 subjects
Elementary design (ξ_{ij})	16 sampling times[a]

[a] from 0.5 to 192h

S_{large}

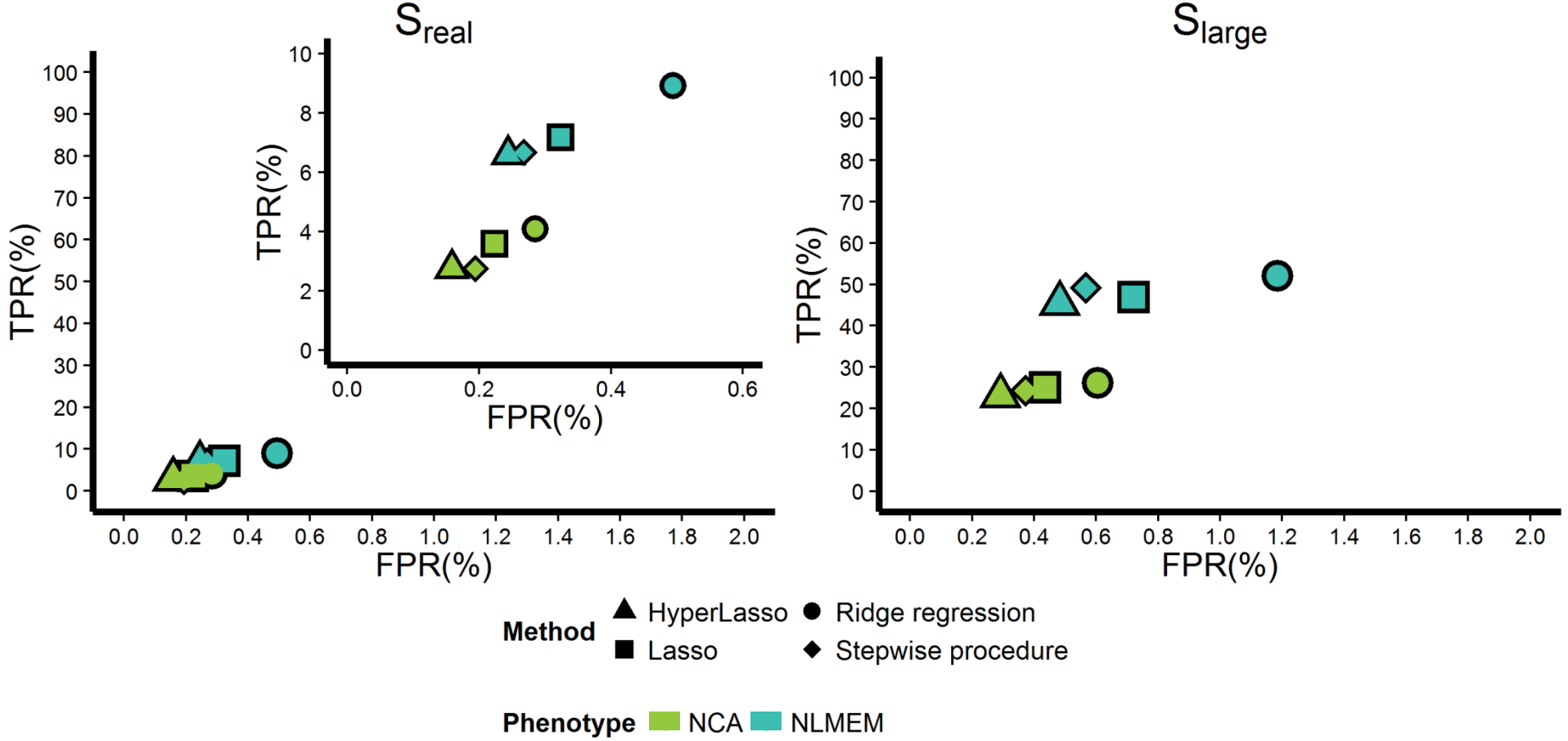
Asymptotic conditions for N

Administration	Single dose
Number of subjects (N_1)	384 subjects
Elementary design (ξ_{ij})	16 sampling times[a]

200 data sets were simulated for each scenario

COMPARISON OF PHENOTYPE ESTIMATION METHODS

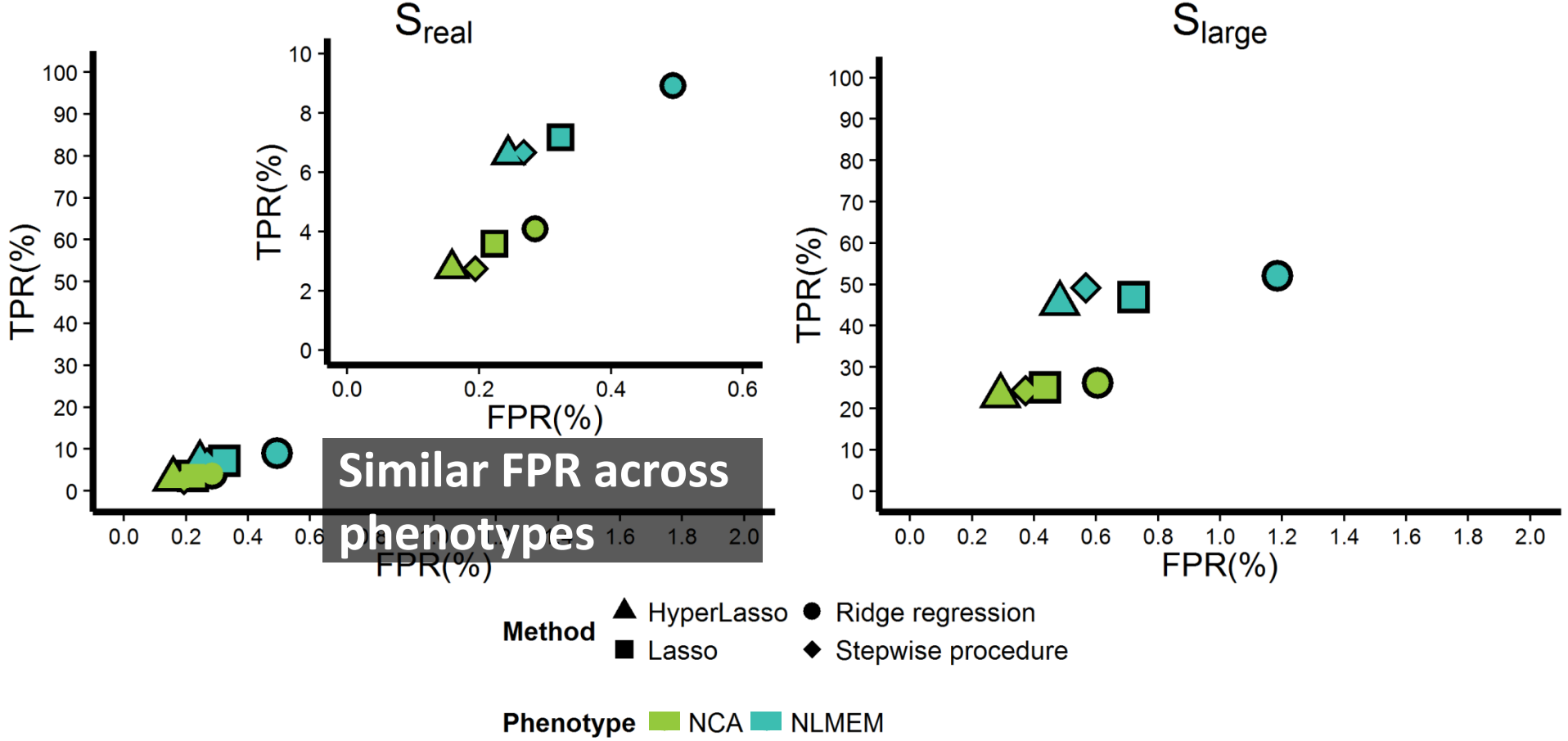
NCA vs NLMEM



True positives (TP): SNP selected in the model and indeed associated to CL in the simulation
False positives (FP): SNP selected in the model but not present in the simulation

COMPARISON OF PHENOTYPE ESTIMATION METHODS

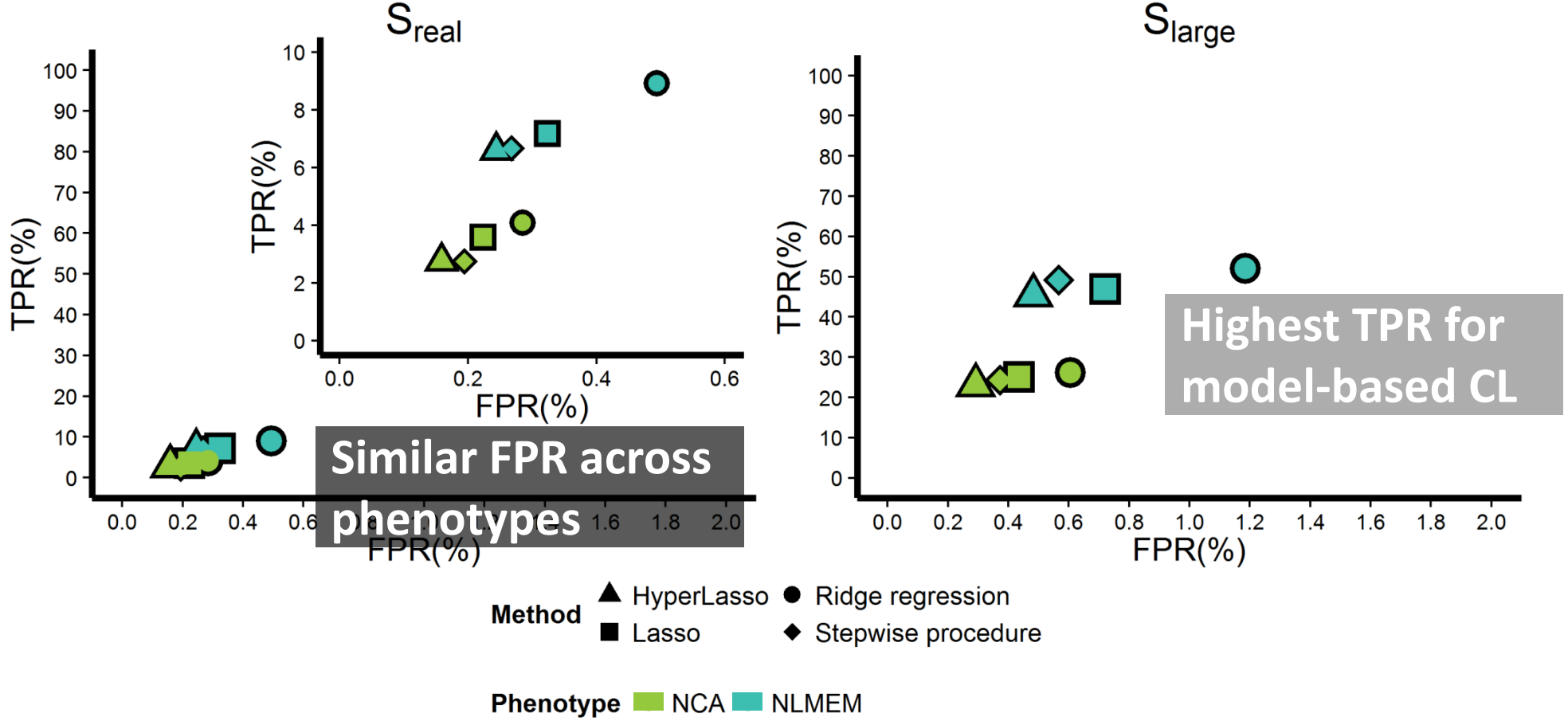
NCA vs NLMEM



True positives (TP): SNP selected in the model and indeed associated to CL in the simulation
False positives (FP): SNP selected in the model but not present in the simulation

COMPARISON OF PHENOTYPE ESTIMATION METHODS

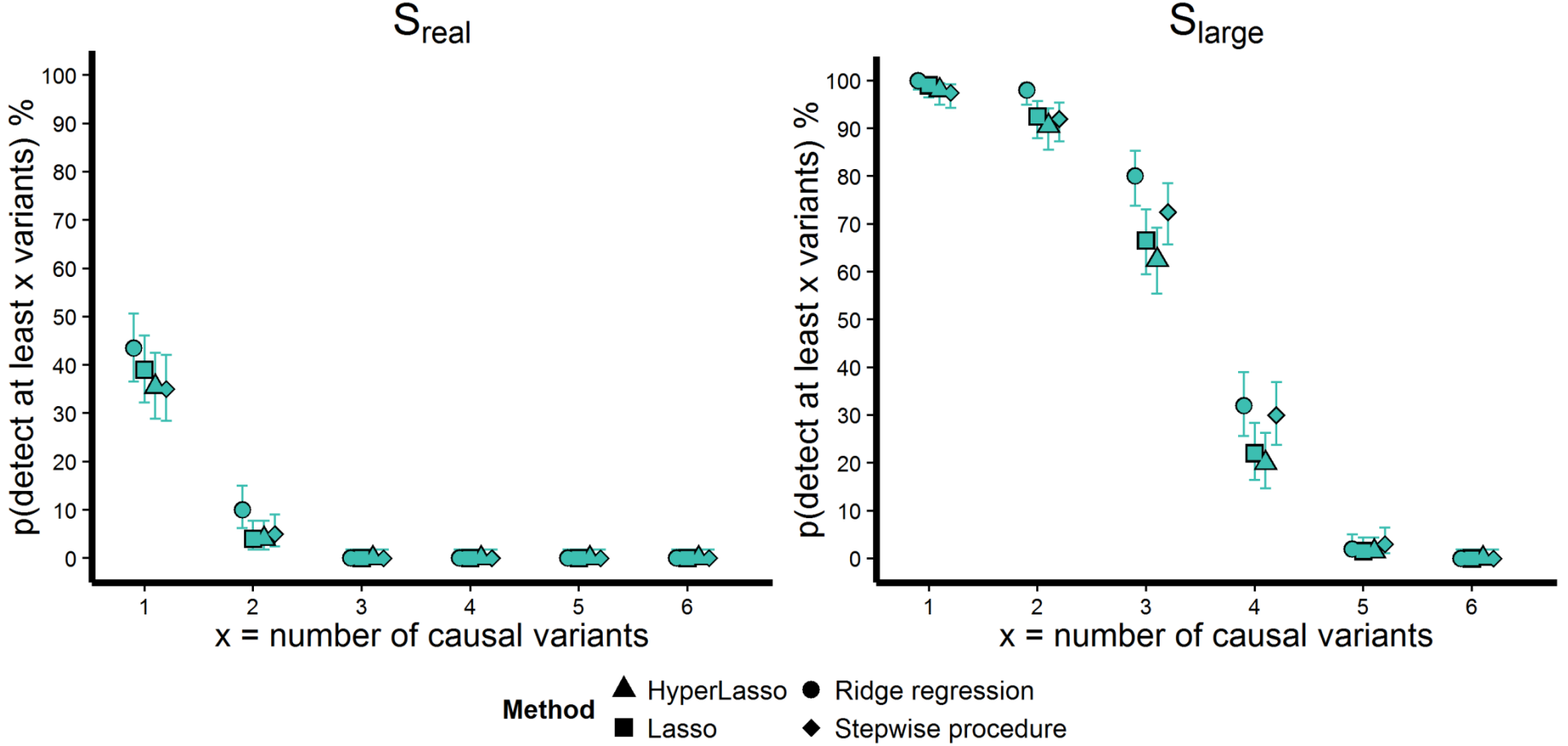
NCA vs NLMEM



True positives (TP): SNP selected in the model and indeed associated to CL in the simulation
False positives (FP): SNP selected in the model but not present in the simulation

COMPARISON OF ASSOCIATION TESTS

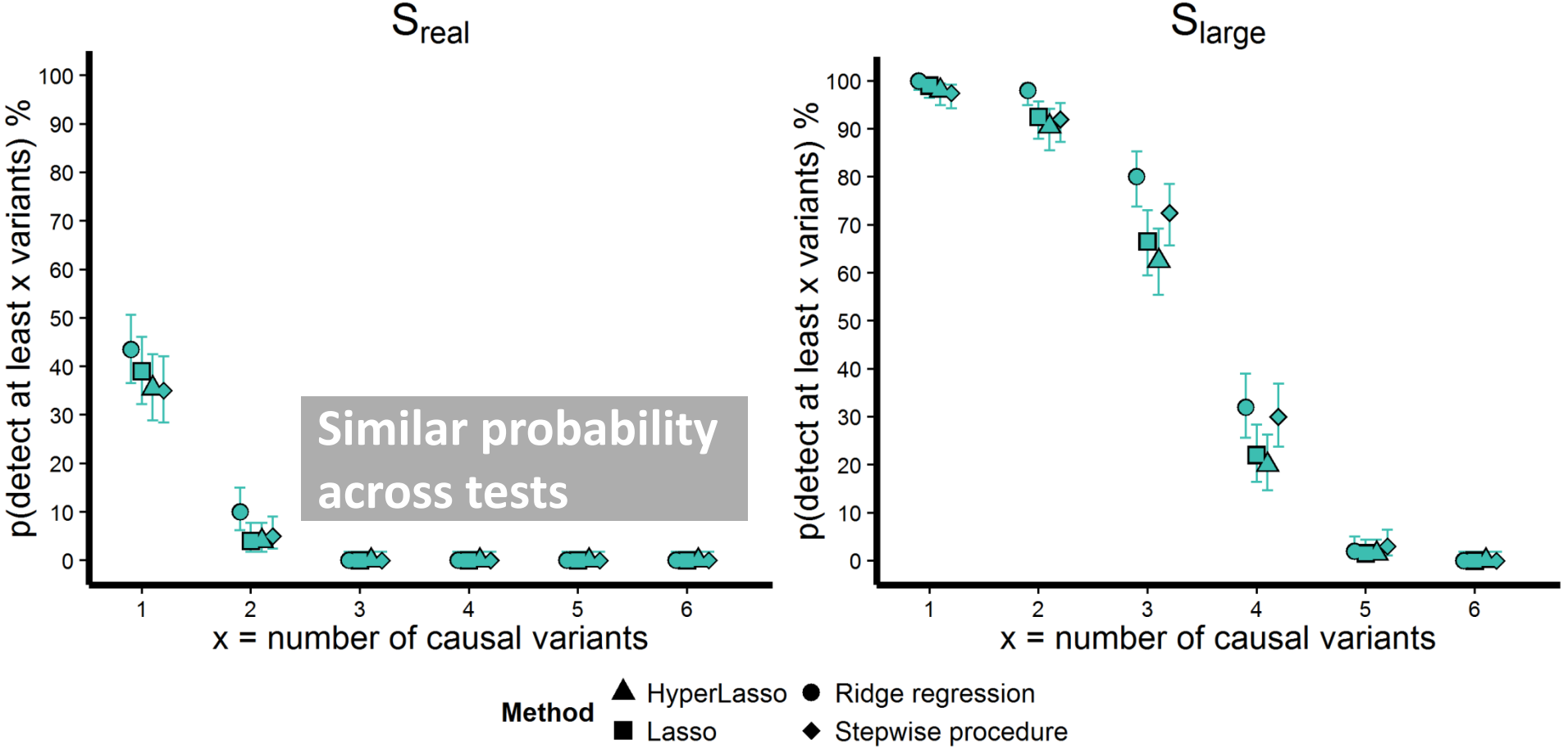
PROBABILITY TO DETECT GENETIC EFFECTS ON NLMEM PHENOTYPE



Probability to detect genetic effects: % of data sets simulated where at least x SNPs (x=1, …, 6) of the 6 SNPs are selected (200 simulated data set)

COMPARISON OF ASSOCIATION TESTS

PROBABILITY TO DETECT GENETIC EFFECTS ON NLMEM PHENOTYPE



Probability to detect genetic effects: % of data sets simulated where at least x SNPs (x=1, …, 6) of the 6 SNPs are selected (200 simulated data set)

PART 1:

CONCLUSIONS

With a nonlinear PK, higher probability to detect genetic effects with a phenotype estimated through NLMEM

Similar power for the 4 association tests (penalised regressions or the stepwise procedure)

PART 2

ENHANCE DETECTION OF GENETIC VARIANTS THROUGH COMBINED DESIGNS

PHOTOGRAPH BY ADAM VOORHES

SCENARIOS

Three phase II designs:

N_2 = 306 subjects at steady-state

	$PII_{3s.96h}$[a]	$PII_{3s.24h}$[a]	$PII_{1s.24h}$
Elementary design (ξ_{ij})	**2, 22, 96h**	**2, 22, 24h**	**24h**

[a]designs optimised using PFIM[1,2]

Four analysis scenarios:

SPI: Phase I data only (N_1 = 78)

SPI/PII$_{3s.96h}$: Phase I + Phase II data {2, 22, 96h}

SPI/PII$_{3s.24h}$: Phase I + Phase II data {2, 22, 24h}

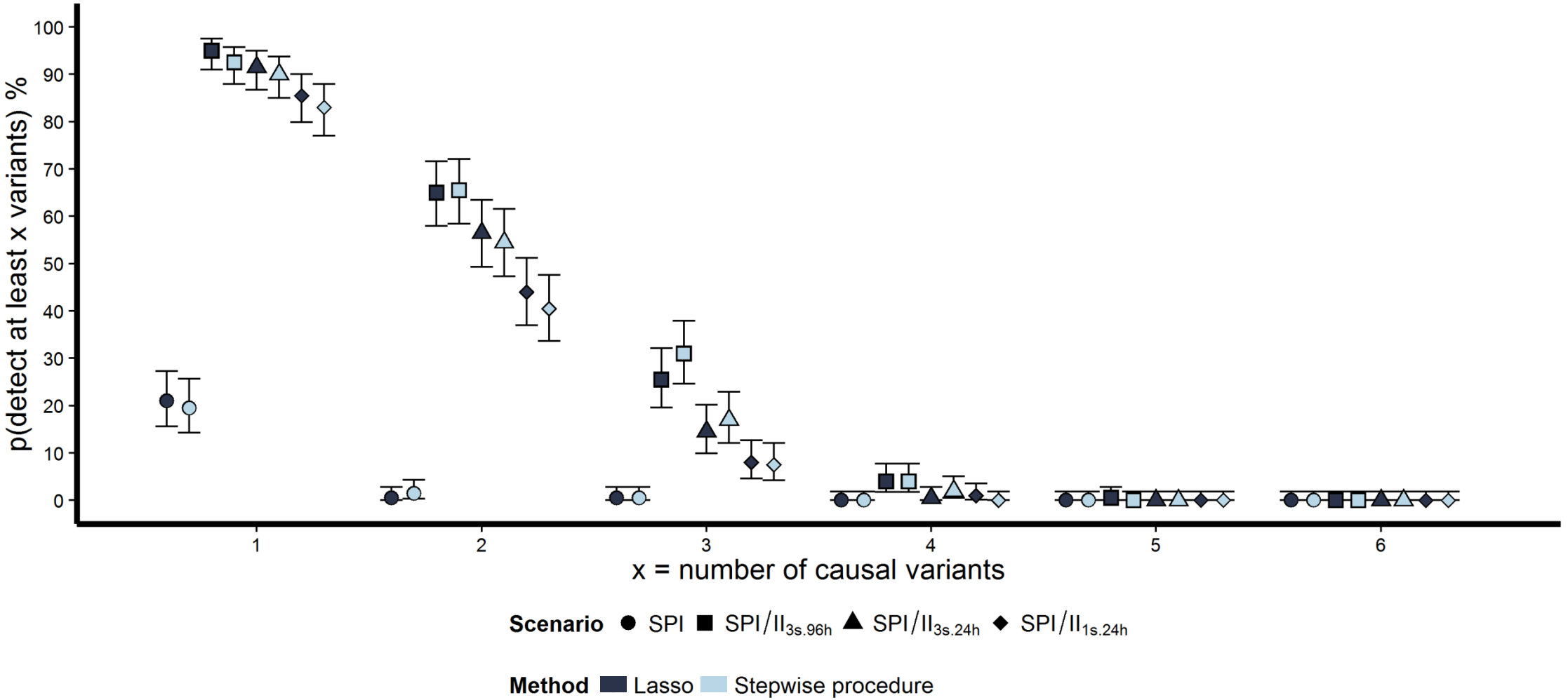
SPI/PII$_{1s.24h}$: Phase I + Phase II data {24h}

200 data sets were simulated for each scenario

1. Bazzoli et al. 2010 2. www.pfim.biostat.fr

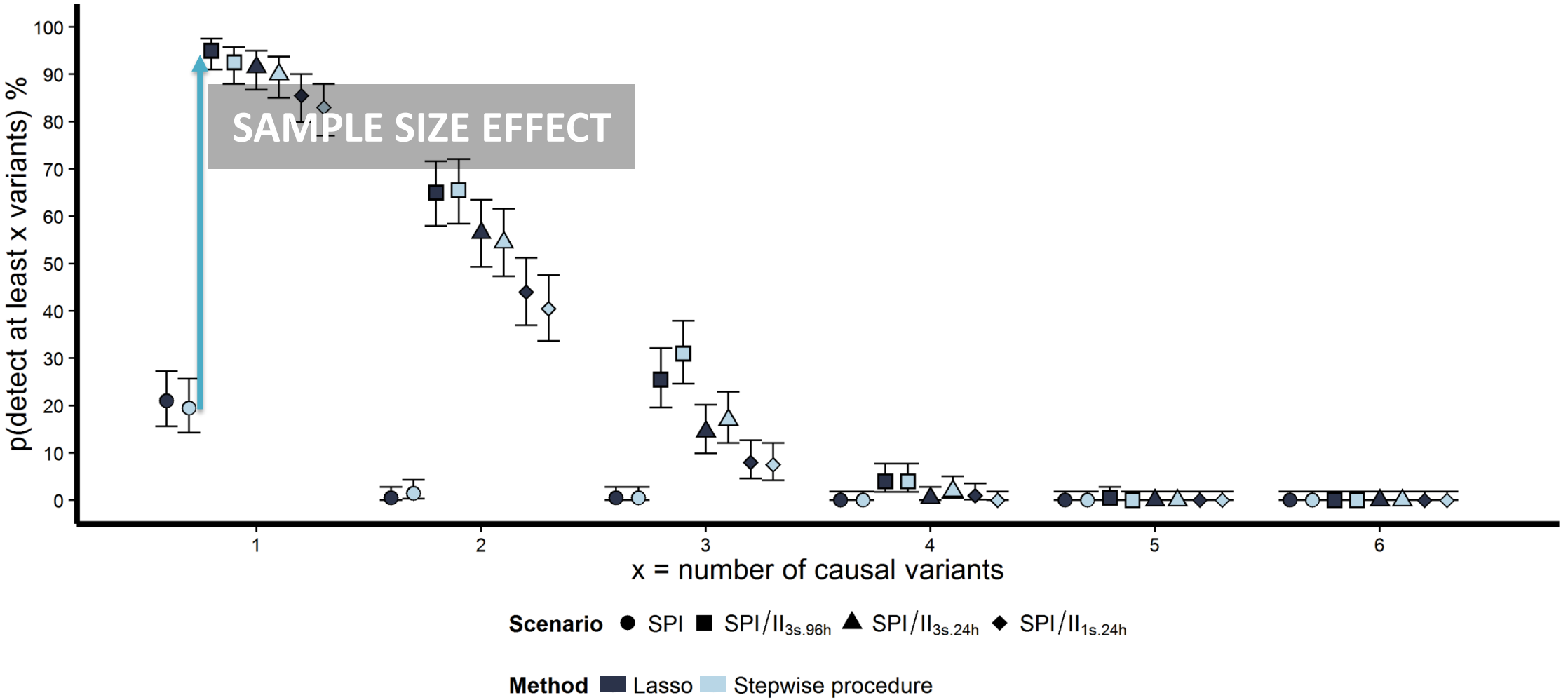
COMPARISON OF DESIGNS

PROBABILITY TO DETECT GENETIC EFFECTS ON NLMEM PHENOTYPE



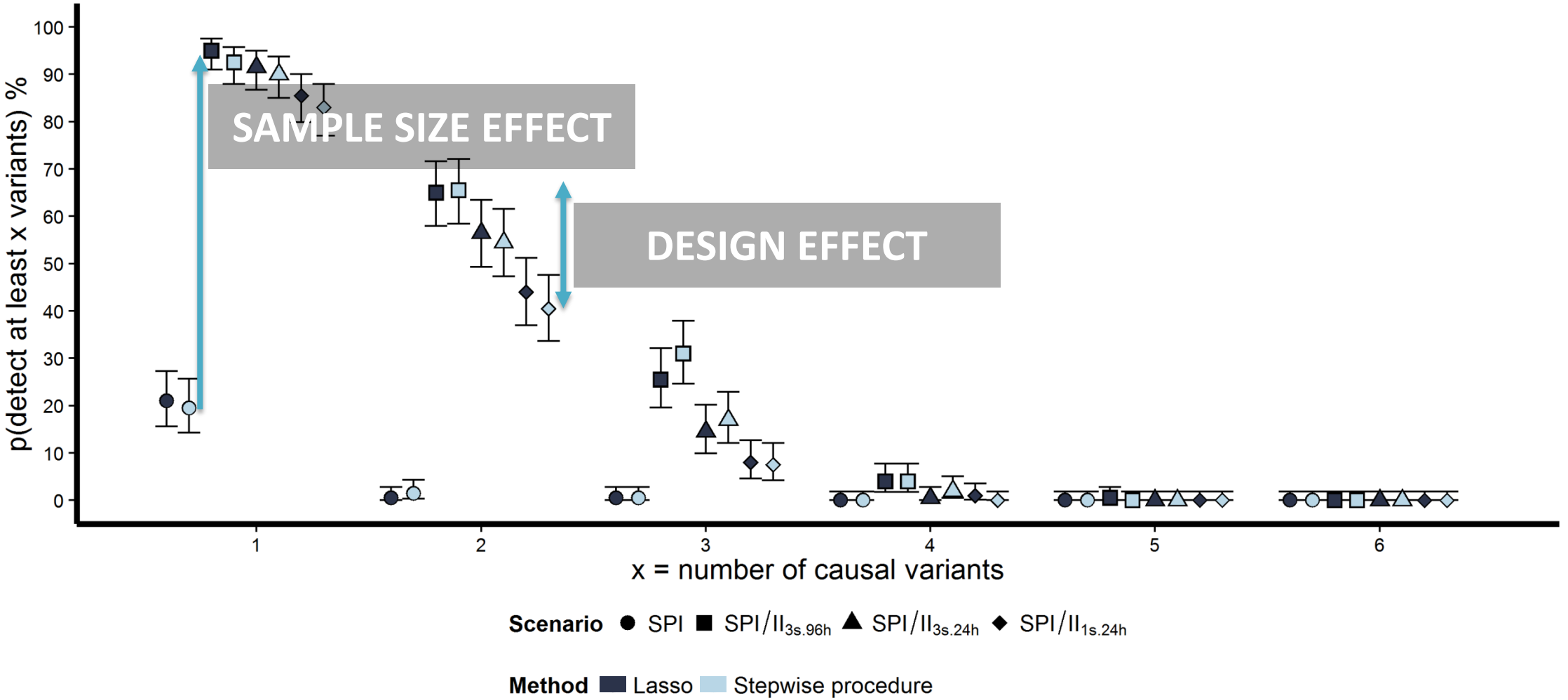
COMPARISON OF DESIGNS

PROBABILITY TO DETECT GENETIC EFFECTS ON NLMEM PHENOTYPE



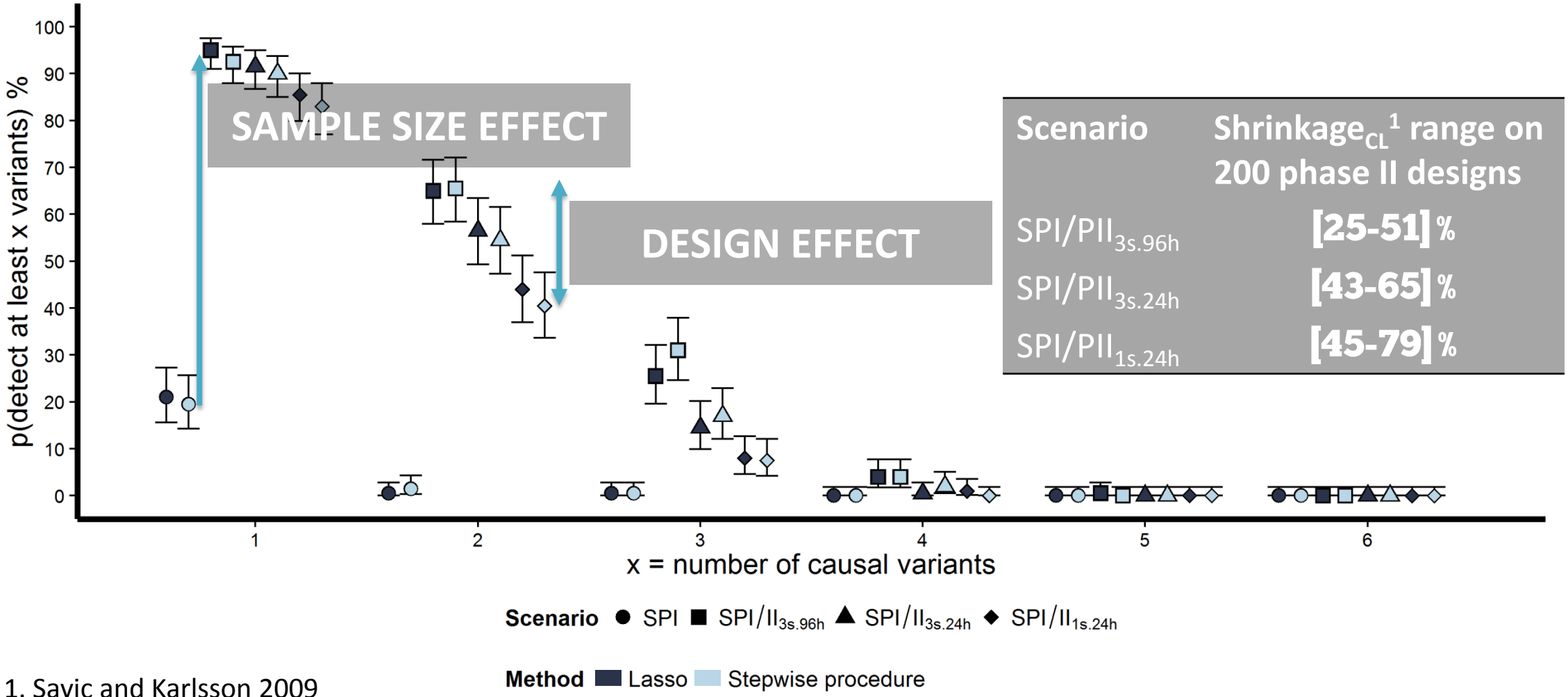
COMPARISON OF DESIGNS

PROBABILITY TO DETECT GENETIC EFFECTS ON NLMEM PHENOTYPE



COMPARISON OF DESIGNS

PROBABILITY TO DETECT GENETIC EFFECTS ON NLMEM PHENOTYPE

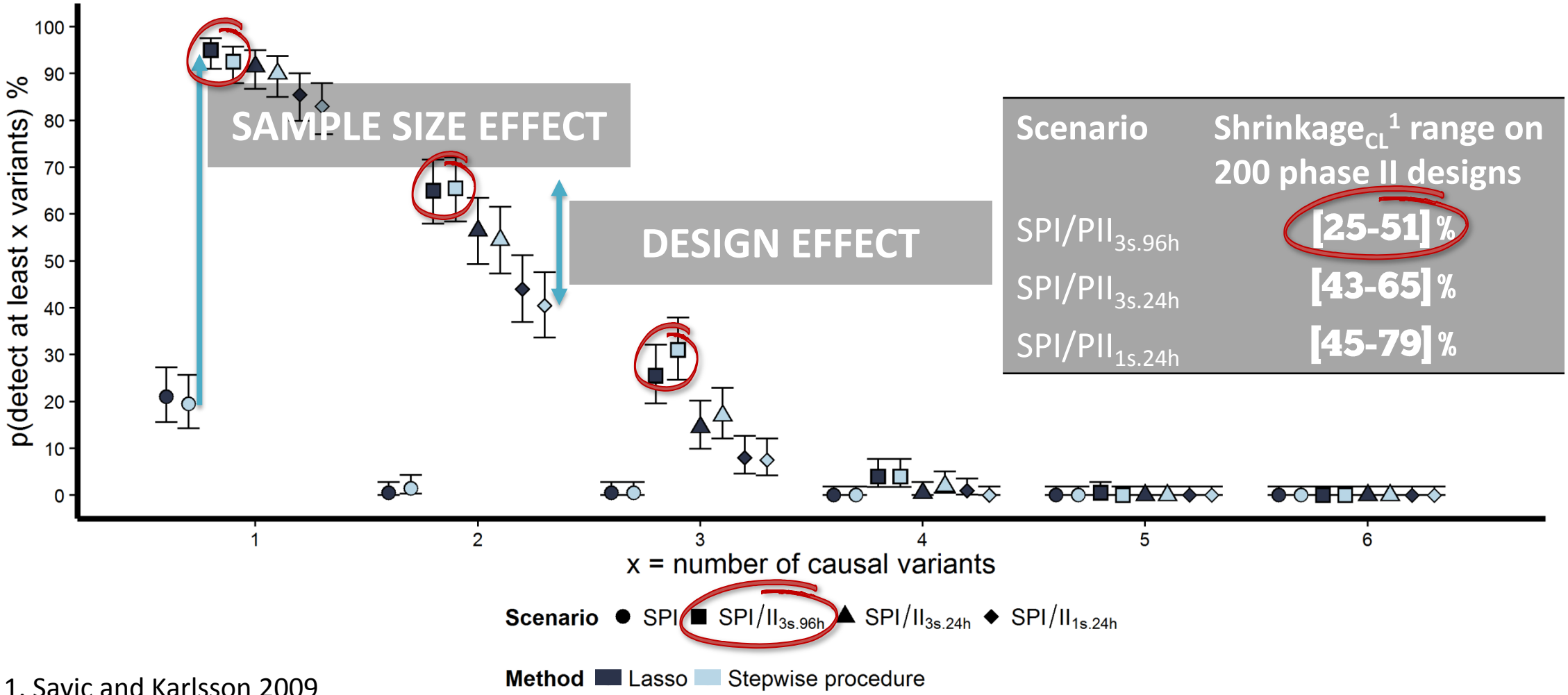


Scenario	Shrinkage$_{CL}$[1] range on 200 phase II designs
SPI/PII$_{3s.96h}$	[25-51] %
SPI/PII$_{3s.24h}$	[43-65] %
SPI/PII$_{1s.24h}$	[45-79] %

1. Savic and Karlsson 2009

COMPARISON OF DESIGNS

PROBABILITY TO DETECT GENETIC EFFECTS ON NLMEM PHENOTYPE



Scenario	Shrinkage$_{CL}$[1] range on 200 phase II designs
SPI/PII$_{3s.96h}$	[25-51] %
SPI/PII$_{3s.24h}$	[43-65] %
SPI/PII$_{1s.24h}$	[45-79] %

1. Savic and Karlsson 2009

PART2:

CONCLUSIONS

Even with advanced methods (NLMEM, stepwise procedure or penalised regressions) results showed poor performance using phase I data only

due to the small sample size

Significant improvement of the probability to detect realistic polymorphisms combining data from phase I and phase II in NLMEM

Mixing subsets with rich and sparse designs could also be performed within a phase II study

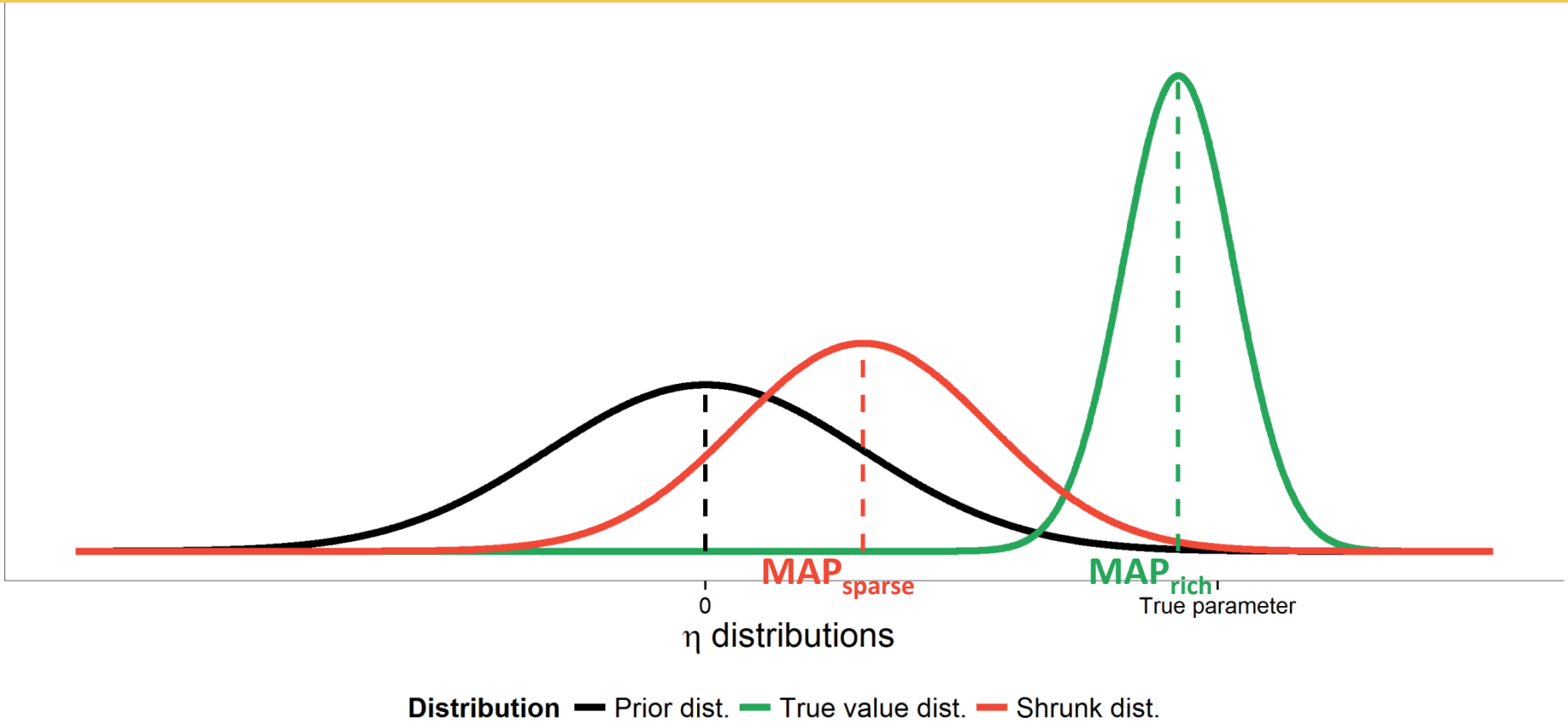
The design of phase II studies benefits from optimisation

to a lower extent

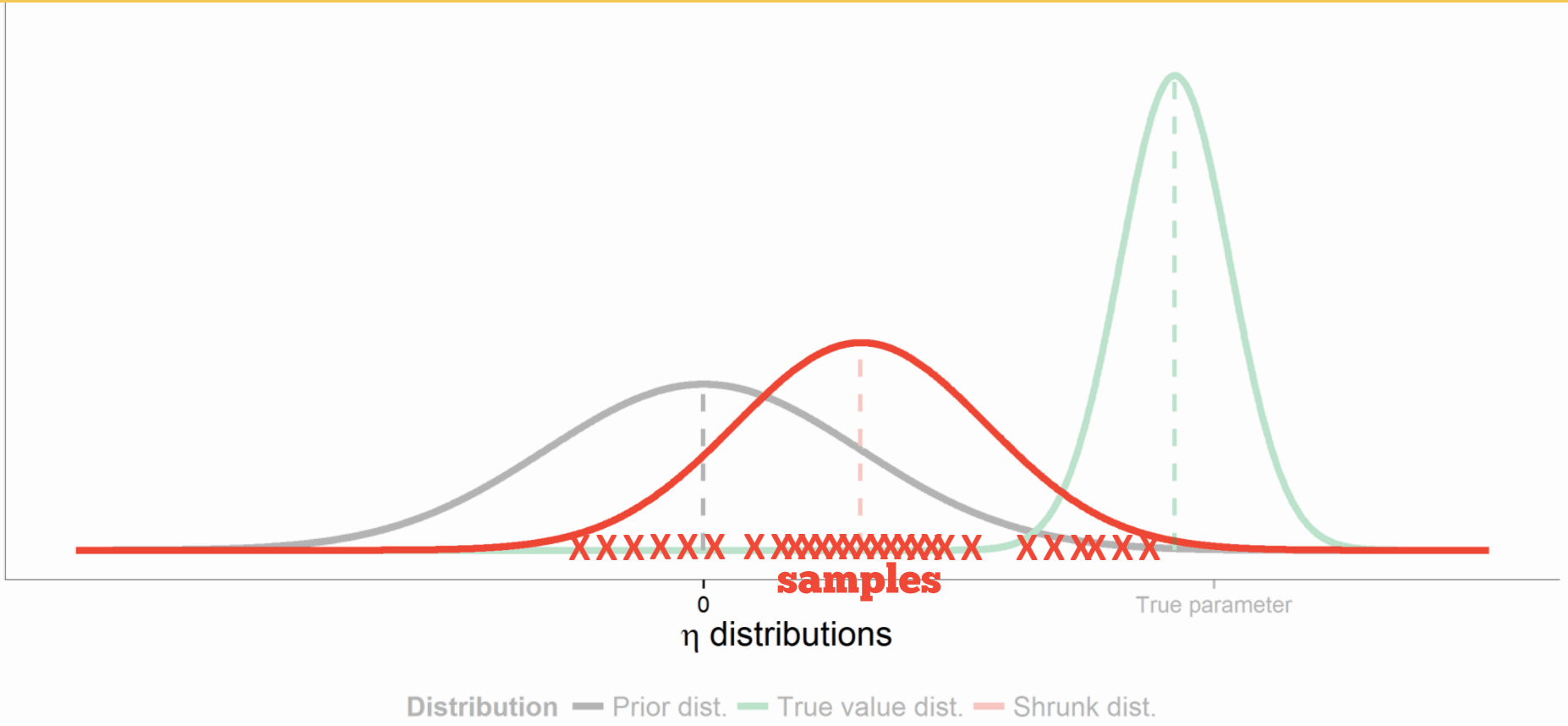
PART 3

USE OF THE CONDITIONAL DISTRIBUTION TO ENHANCE GENETIC COVARIATES ANALYSIS

BAYESIAN APPROACH FOR INDIVIDUAL ESTIMATIONS



BAYESIAN APPROACH FOR INDIVIDUAL ESTIMATIONS



Samples from conditional distribution computed using Monolix software[1]

Proposed in diagnostic plots in last versions

1. www.lixoft.eu

SAMPLES IN THE CONDITIONAL DISTRIBUTION SCENARIOS

Individual PK profiles were simulated from PK model with genetic effects

1 SNP drawn randomly to affect the clearance:

R_{GC} = {5, 12, 20%}

N = 78 subjects

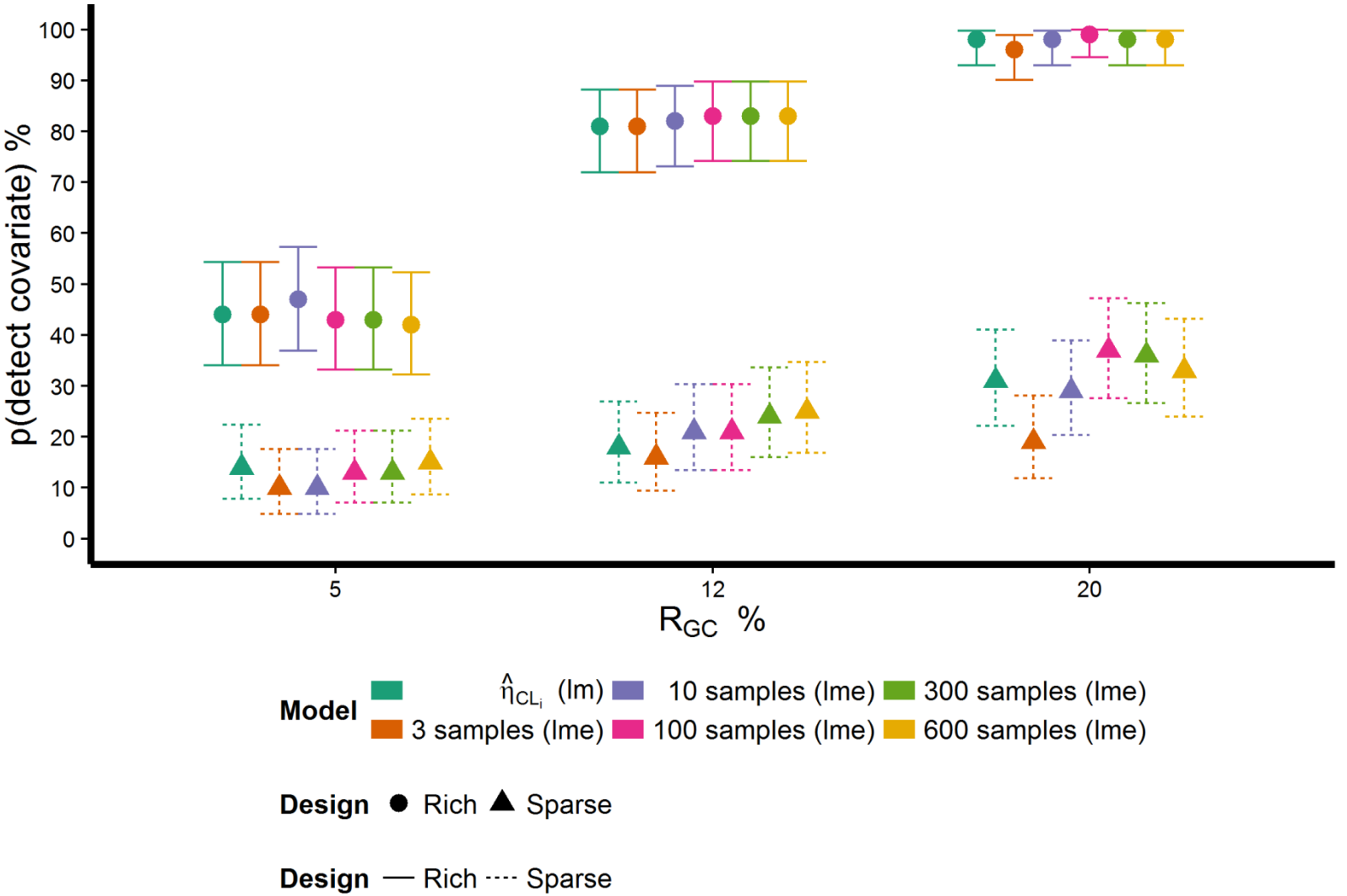
2 scenarios: Rich design (n = 16) or Sparse design (n = 1 with 3 groups)

Univariate regression (lm) on **individual parameters** $\hat{\eta}_{CL_i}$

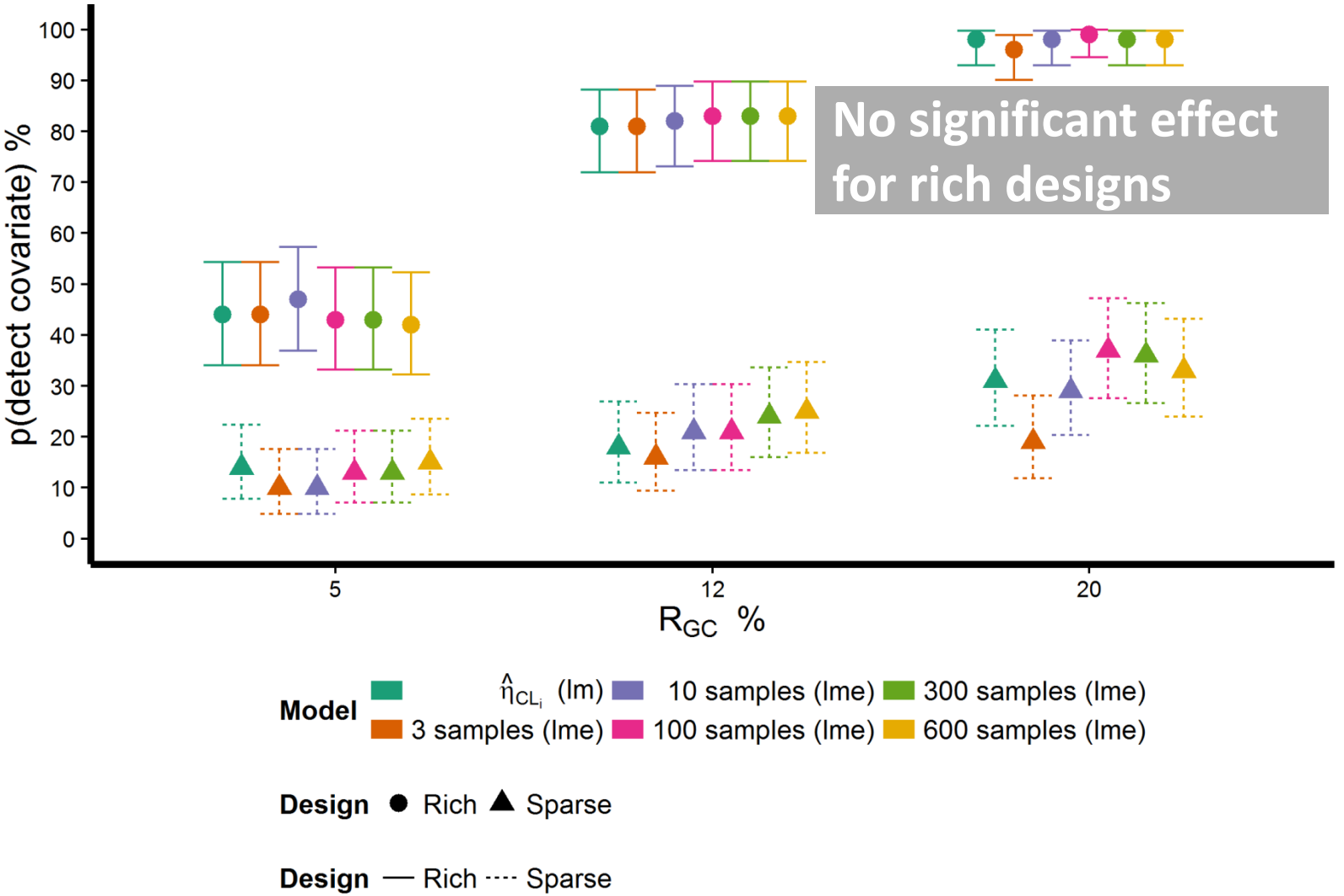
Linear mixed effects model (lme) on **samples from conditional distribution**

- random effect to handle correlations between the different samples for the same subject
- with different numbers of samples (from 3 to 600)

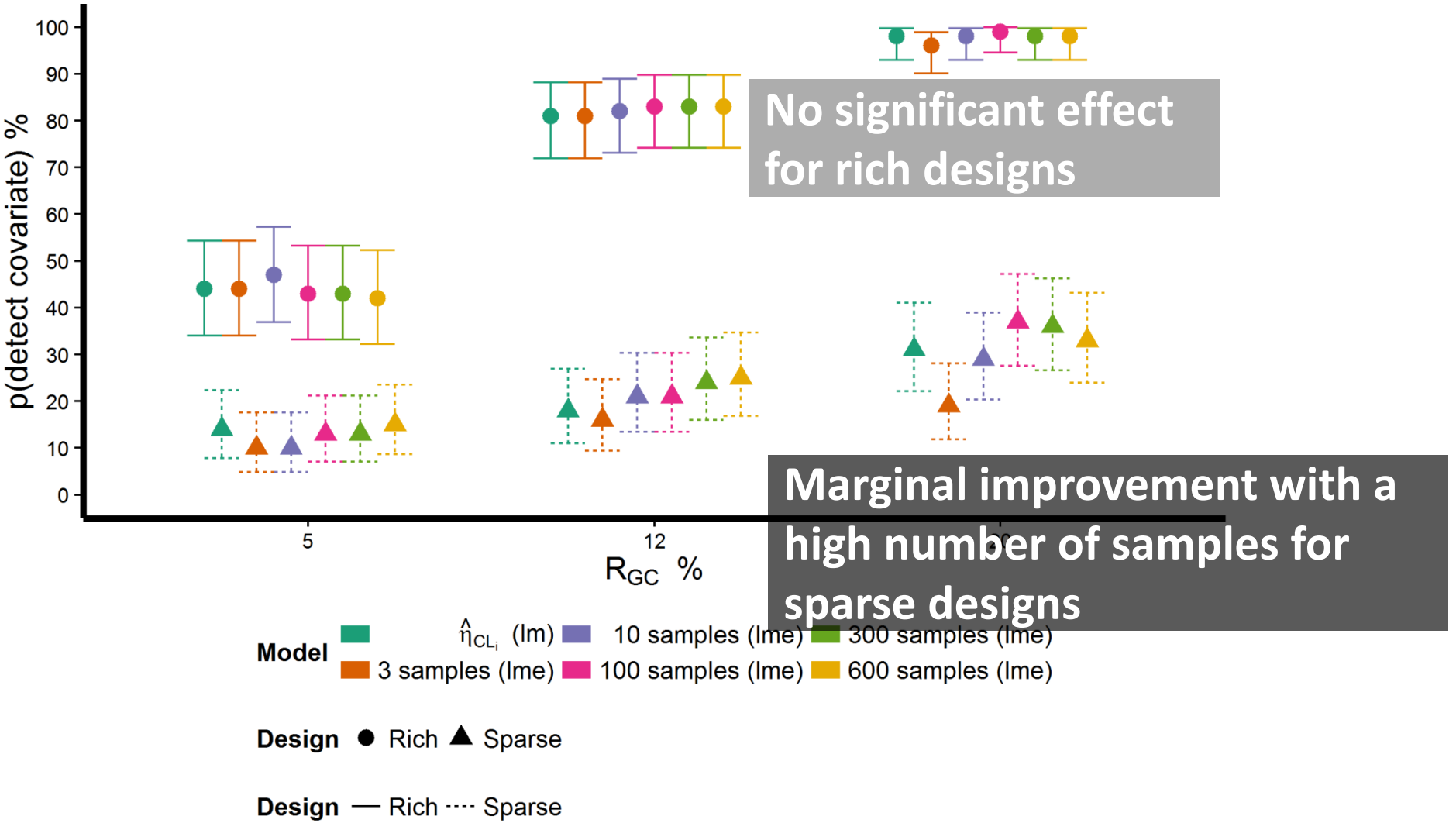
PROBABILITY TO DETECT THE GENETIC EFFECT



PROBABILITY TO DETECT THE GENETIC EFFECT



PROBABILITY TO DETECT THE GENETIC EFFECT



PART 3:

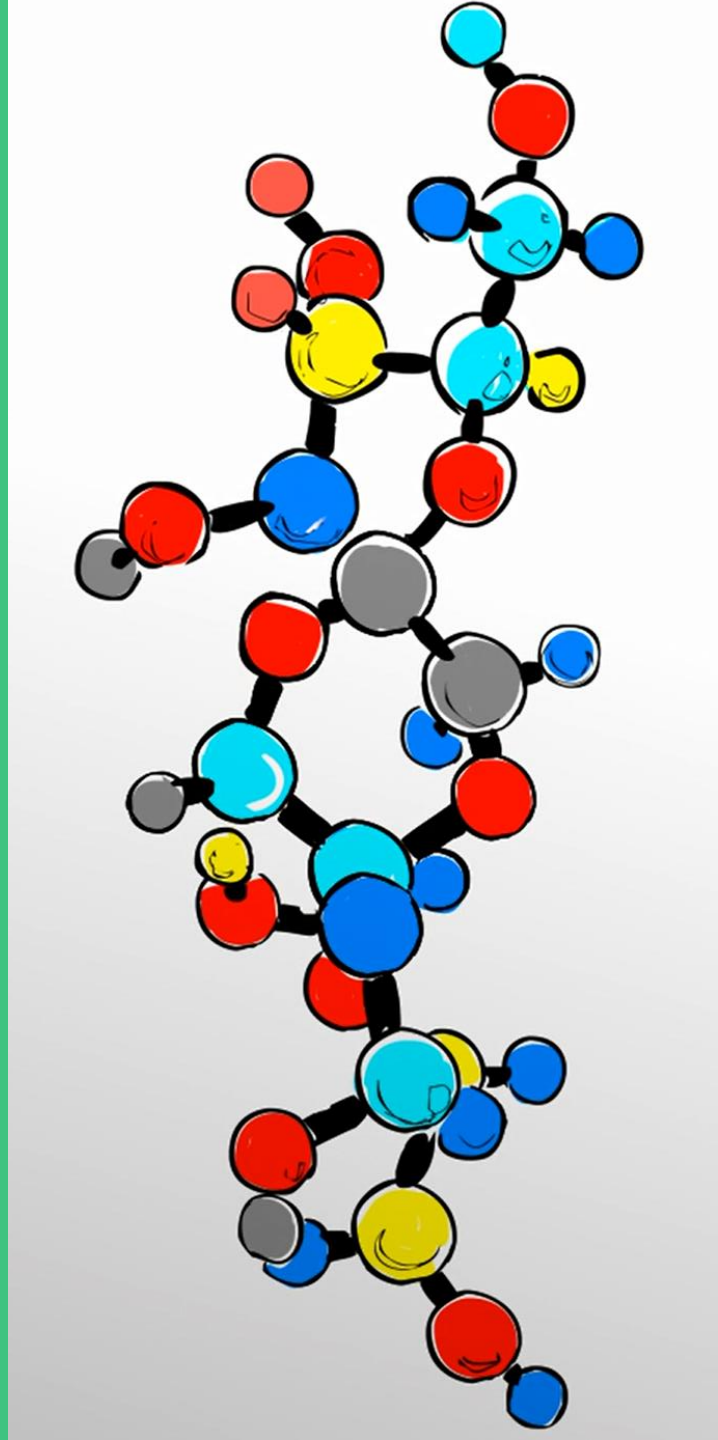
CONCLUSIONS

No benefit of using samples from the conditional distribution with rich designs

MAP from posterior distribution close to the true parameter

Marginal improvement in the probability to detect the genetic variant for sparse designs requires a large number of samples

CONCLUSION



CONCLUSION

RECOMMENDATIONS FOR PHARMACOGENETIC ANALYSIS

Phenotypes

Modelling approaches should be preferred to estimate the PK phenotype

Association methods

No benefit of penalised regressions

Design

We recommend

- to combine analysis of phase I and phase II data for the exploration of genetic associations
- to prospectively optimise the phase II study design

LIMITS AND PERSPECTIVES

Limits

Only one setting investigated in simulations

- additive genetic model
- nonlinear PK

No re estimation of EBE conditional to the genetic variants in stepwise procedure

Perspectives

Develop joint estimation/selection with penalised regression in NLMEM[1]

Investigate effect of model misspecifications on detection probability

- correlation between model parameters

1. Bertrand et al. PAGE 2013

Acknowledgments

Emmanuelle Comets

Marylore Chenel

Julie Bertrand

Servier Clinical PK & PMx team

Bernard Walther
Laurent Ripoll

IAME team

France Mentré

PAGE committee

We are grateful to Marc Lavielle for help implementing the imputations in the conditional distribution