

Case Study in the Use of Bayesian Hierarchical Modeling and Simulation for Design and Analysis of a Clinical Trial

William R. Gillespie
Pharsight Corporation
Cary, North Carolina, USA

2003 FDA/Industry Statistics Workshop
September 18-19, 2003
Hyatt Regency
Bethesda, Maryland

"Statistics: From Theory to Regulatory Acceptance"
Oral Session 1: Flexible Designs

Case Study in the Use of Bayesian Hierarchical Modeling and Simulation for Design and Analysis of a Clinical Trial

- Bayesian principles and methods provide a coherent framework for:
 - Quantifying uncertainty,
 - Making inferences in the presence of that uncertainty.
- Bayesian modeling and simulation are practical options for many applications due to recent advances in hardware, numerical methods and software.
- This presentation describes an approach used with a recent project to optimize the design and analysis of a Phase II proof-of-concept (PoC) trial.
 - Bayesian methods used throughout a model-based approach.
 - Model development
 - Trial simulation
 - Trial analysis
 - Focus on technical execution.

The scenario

- Simuzine is a NCE for treatment of a slowly progressive illness.
- Previous Phase II PoC study of simuzine:
 - Primary efficacy score measured at baseline and 6 months.
 - Results encouraging but inconclusive.
 - Longer duration treatment may be necessary to reach a decisive outcome.
- Additional longitudinal data available for model development:
 - Efficacy score for patients with observations at various times over durations up to 6 years.
 - Believed to be representative of the placebo group in the new trial.
- The new trial is already underway, so the M&S effort focuses on optimizing analysis of the trial results to support a PoC decision.

The scenario (cont.)

- New trial design:
 - Parallel, 2 treatment arm trial comparing simuzine 100 mg to placebo.
 - Primary endpoint = efficacy score at 2 years (LOCF imputation).
 - 100 patients per treatment arm.
- The example compares the use of 3 different trial analyses:
 - Conventional frequentist analysis of endpoint data (ANCOVA)
 - Bayesian longitudinal analysis with use of prior data.
 - Bayesian longitudinal analysis without prior data (non-informative priors).

Fully Bayesian approach to trial simulation and analysis

- The case study illustrates the following 3 applications of Bayesian modeling:
 - Model development
 - A Bayesian longitudinal model of an efficacy score is fitted using Markov chain Monte Carlo simulation (MCMC, WinBUGS).
 - Trial simulation
 - Uncertainty in the model parameters is considered by resampling the MCMC-generated samples from the joint posterior distribution of the model parameters (S-PLUS).
 - Trial analysis
 - The simulated trial data combined with prior data are analyzed with a Bayesian longitudinal model (WinBUGS).
 - The resulting inferences are compared with those obtained with a more conventional endpoint analysis (ANCOVA, S-PLUS) and use of the Bayesian longitudinal model without the prior data.

Model for the effect of simuzine on an efficacy score in the target patient population

- Selected model:
 - Log(score) changes linearly with time.
 - Sex, age and time from disease onset affect the intercept.
 - Dose of simuzine affects the slope.
 - Log(score) at the i^{th} observation time in the j^{th} patient is modeled as:

$$\log(\text{score}_{ij}) \sim N(\mathbf{a}_j + \mathbf{b}_j t_{ij}, \mathbf{s}^2)$$

$$\mathbf{a}_j = \mathbf{a}_{0j} + \mathbf{q}_{sex} I_{female,j} + \mathbf{q}_{age} (\text{age}_j - 55) + \mathbf{q}_b t_{onset,j}$$

$$\mathbf{b}_j = \mathbf{q}_b + \mathbf{q}_{drug} \text{dose}_j$$

$$\mathbf{a}_{0j} \sim N(\mathbf{q}_a, \mathbf{w}_a^2)$$

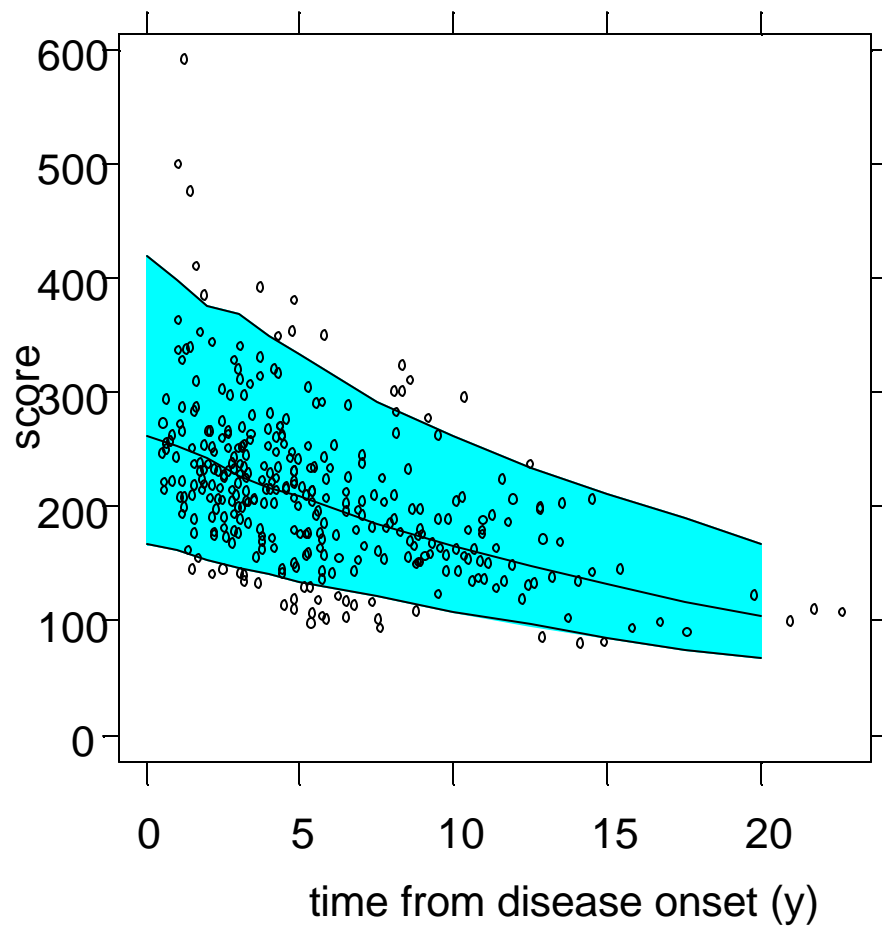
Prior distributions (chosen to be relatively non-informative)

$$\mathbf{q}_a \sim N(0, 10^6) \quad \mathbf{q}_b \sim N(0, 10^6) \quad \mathbf{q}_{drug} \sim N(0, 10^{12}) \quad \mathbf{q}_{age} \sim N(0, 10^6)$$

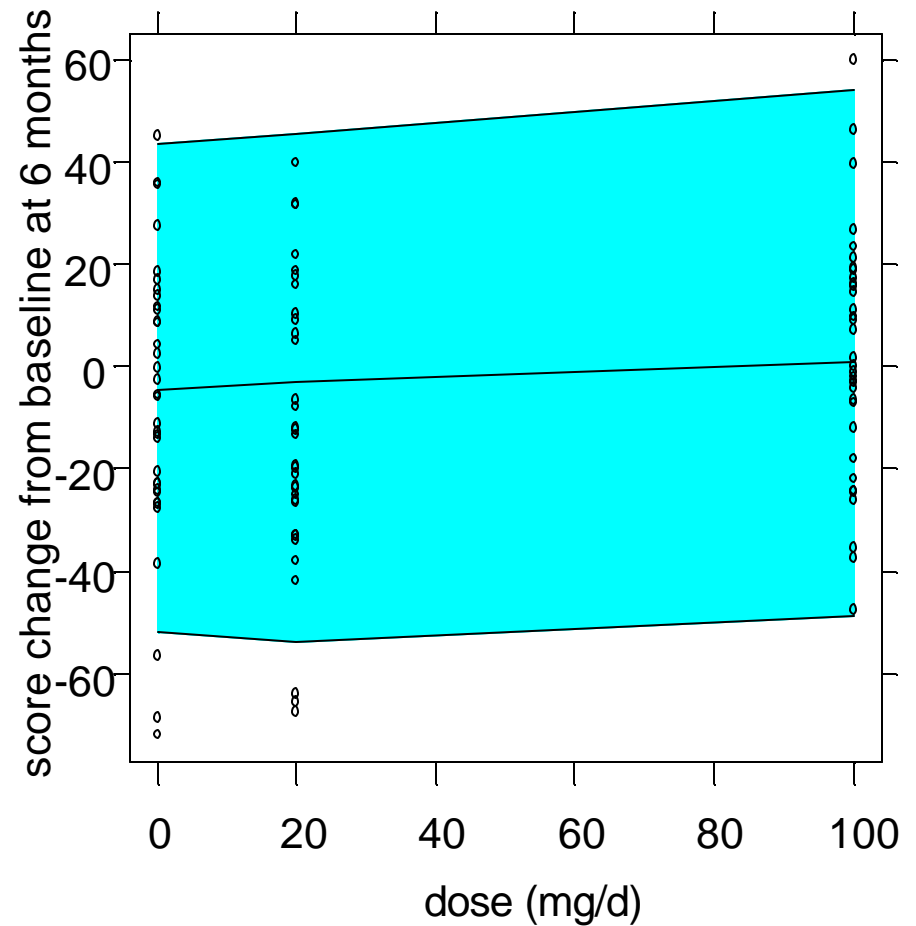
$$\mathbf{q}_{sex} \sim N(0, 10^6) \quad \frac{1}{\mathbf{s}^2} \sim \text{gamma}(10^{-4}, 10^{-4}) \quad \frac{1}{\mathbf{w}_a^2} \sim \text{gamma}(10^{-4}, 10^{-4})$$

Model fitting results: Posterior predictive intervals

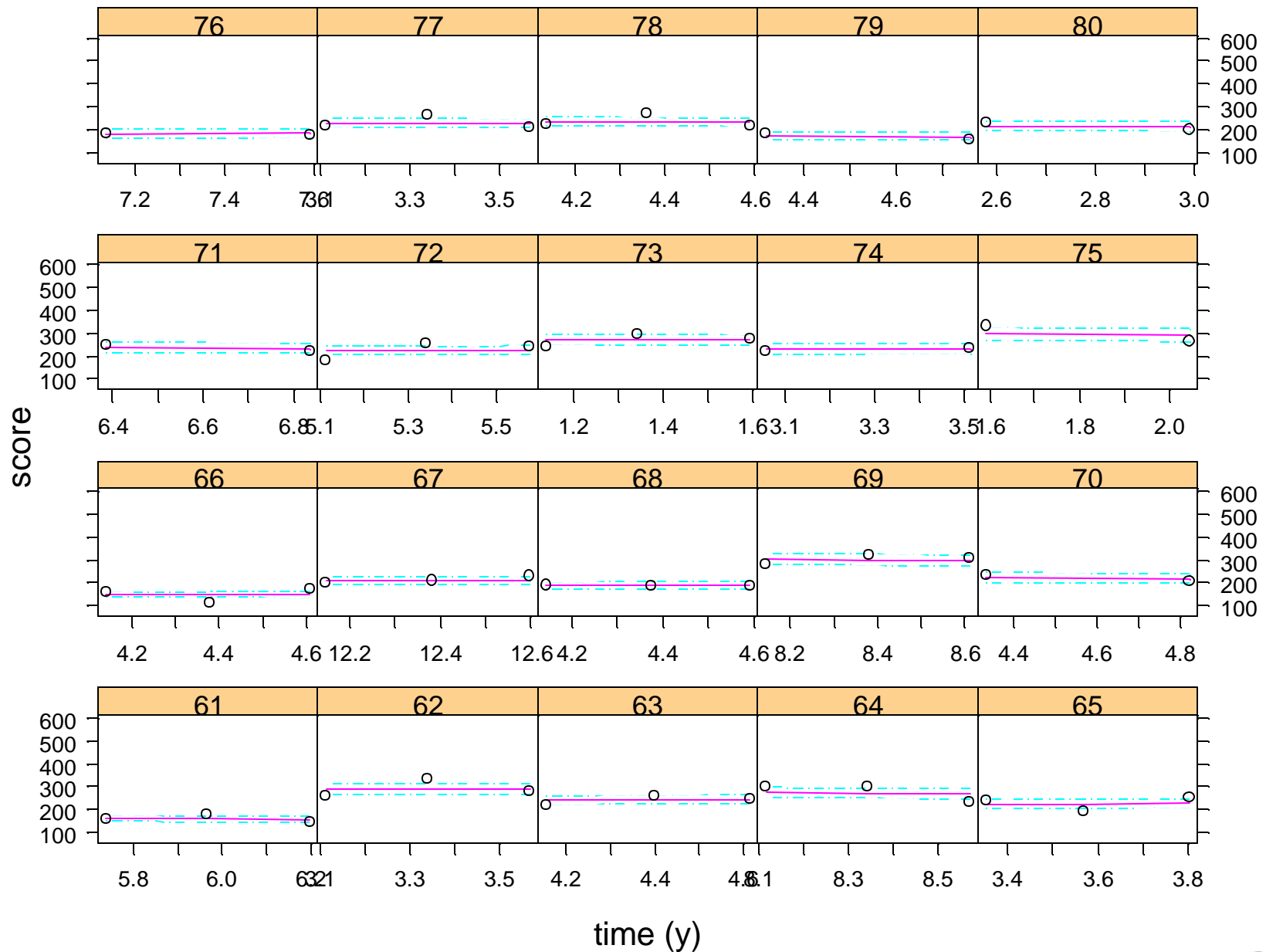
The model captures the decline in score after disease onset: Observed and model predicted (median and 90% prediction intervals) scores (placebo data)



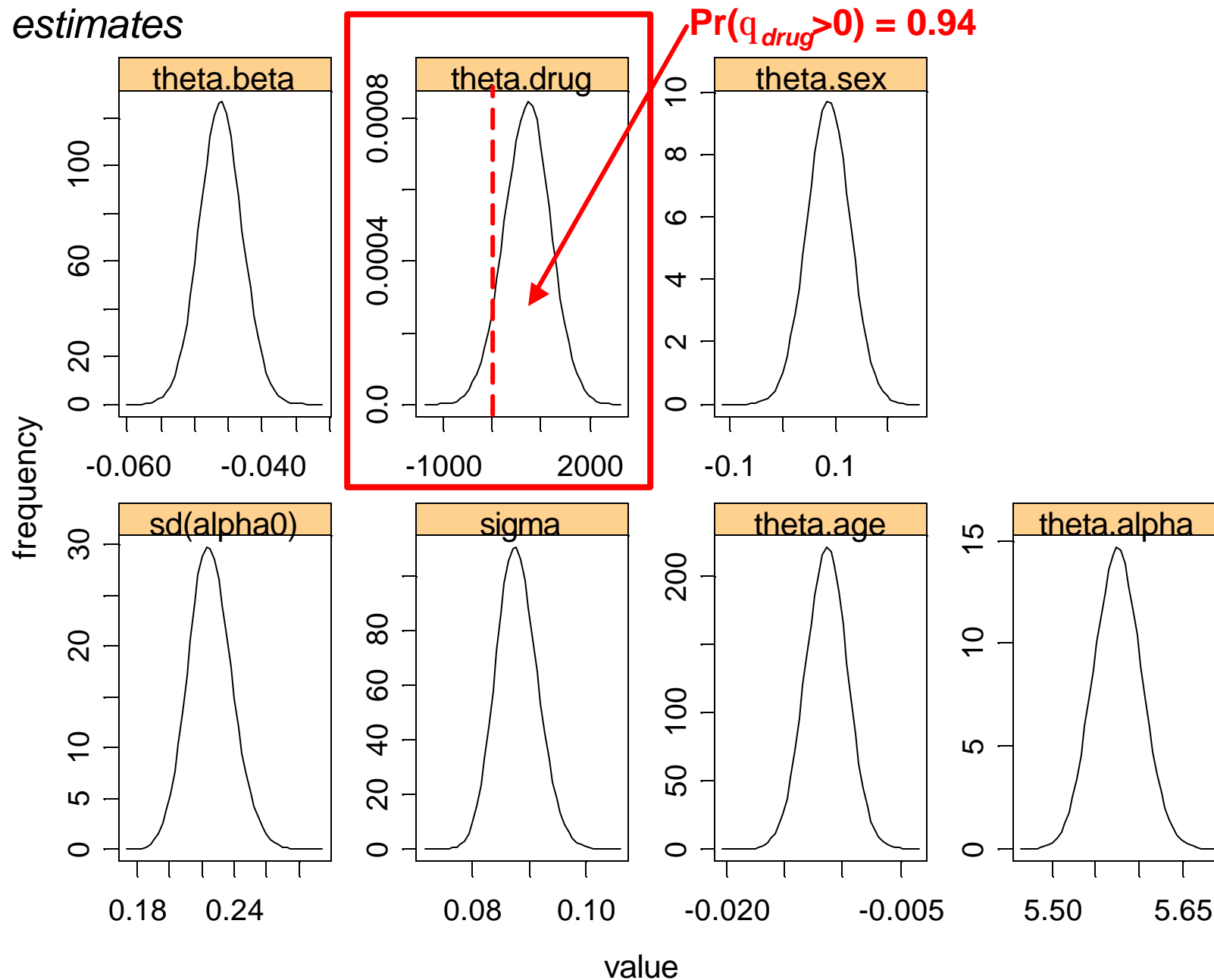
And the effect of simuzine: Observed and model predicted (median and 90% prediction intervals) change from baseline



Model fitting results: Examples of individual fits



Model fitting results: Posterior marginal distributions of parameter estimates



Extending the model for simulating 2 year outcomes

- The current model is based on data from only 6 months of simuzine administration.
 - Model predictions for 1 and 2 years represent major extrapolations from experience.
 - The magnitude of the response at 1 and 2 years is more uncertain than indicated by simple linear extrapolation.
- For example, the available clinical evidence is also consistent with a more pessimistic model in which the drug benefit is not sustained beyond 6 months.
- Extrapolation beyond 6 months is based on expert judgment:
 - Upper bound: Linear extrapolation (constant slope)
 - Lower bound: Slope changes to pretreatment value after 6 months
 - Uncertainty in the post-6 month slope is modeled as a uniform distribution between those extremes.

Model for simulating the effect of simuzine over 2 years

- Log(score) at the i^{th} observation time in the j^{th} patient is modeled as:

$$\log(\text{score}_{ij}) \sim N(\mathbf{m}_j, \mathbf{s}^2)$$

$$\mathbf{m}_j = \begin{cases} \mathbf{a}_j + \mathbf{b}_{1j} t_{ij}, & t_{ij} \leq 0.5 \\ \mathbf{a}_j + \mathbf{b}_{1j}(0.5) + \mathbf{b}_{2j}(t_{ij} - 0.5), & t_{ij} > 0.5 \end{cases}$$

$$\mathbf{a}_j = \mathbf{a}_{0j} + \mathbf{q}_{sex} I_{\text{female},j} + \mathbf{q}_{age} (\text{age}_j - 55) + \mathbf{q}_b t_{\text{onset},j}$$

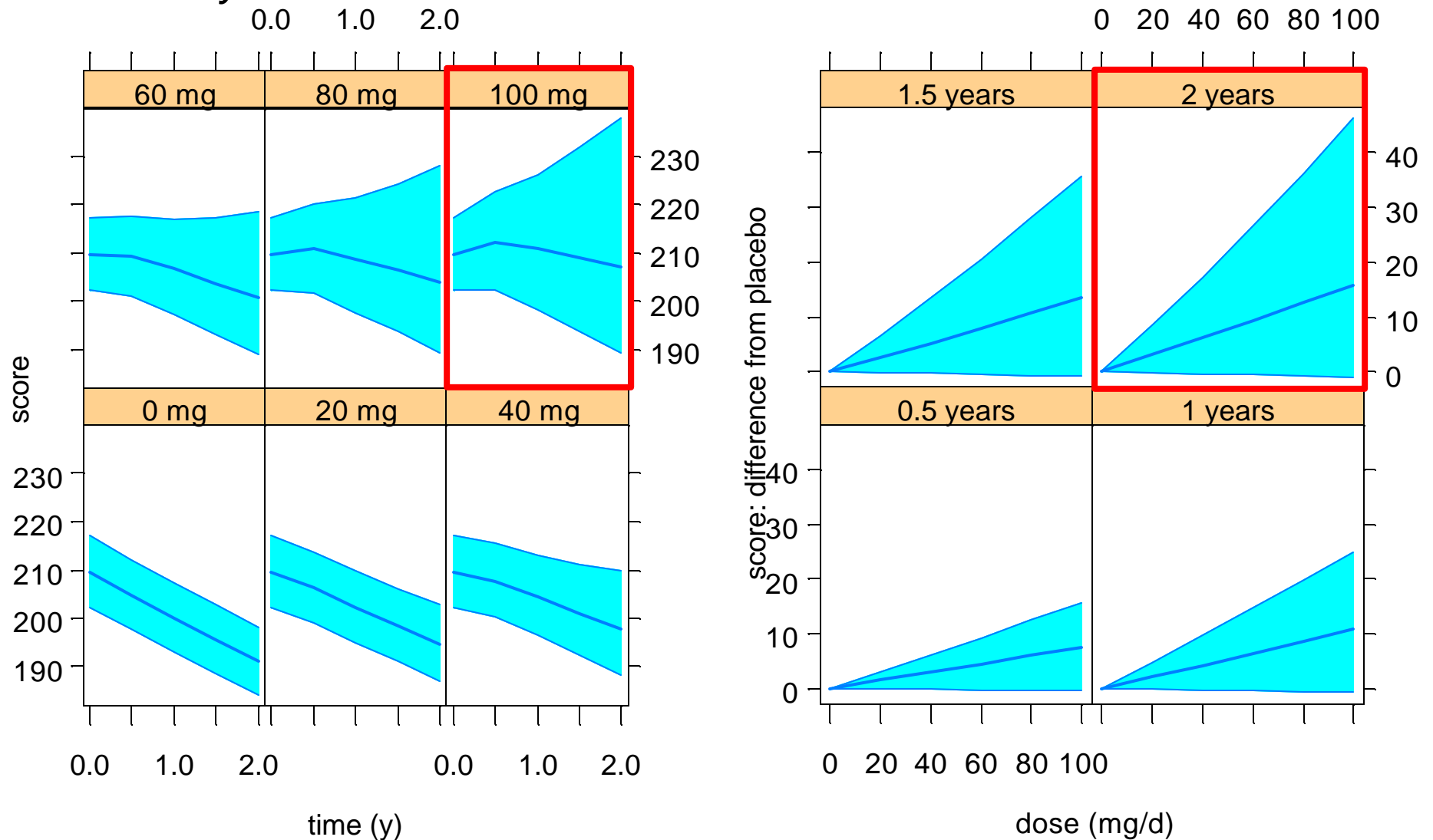
$$\mathbf{b}_{1j} = \mathbf{q}_b + \mathbf{q}_{drug} \text{dose}_j \quad \mathbf{b}_{2j} = \mathbf{q}_b + \mathbf{q}_{extrap} \mathbf{q}_{drug} \text{dose}_j$$

$$\mathbf{a}_{0j} \sim N(\mathbf{q}_a, \mathbf{w}_a^2)$$

$$\mathbf{q}_{extrap} \sim \text{Uniform}(0,1)$$

$$(\mathbf{q}_a, \mathbf{q}_b, \mathbf{q}_{drug}, \mathbf{q}_{age}, \mathbf{q}_{sex}, \mathbf{s}^2, \mathbf{w}_a^2) \sim \text{MCMC estimated joint posterior distribution}$$

Model results indicate a highly uncertain but potentially large drug effect on the efficacy score



Model-predicted population mean score (median & 90% probability intervals) as a function of dose and time

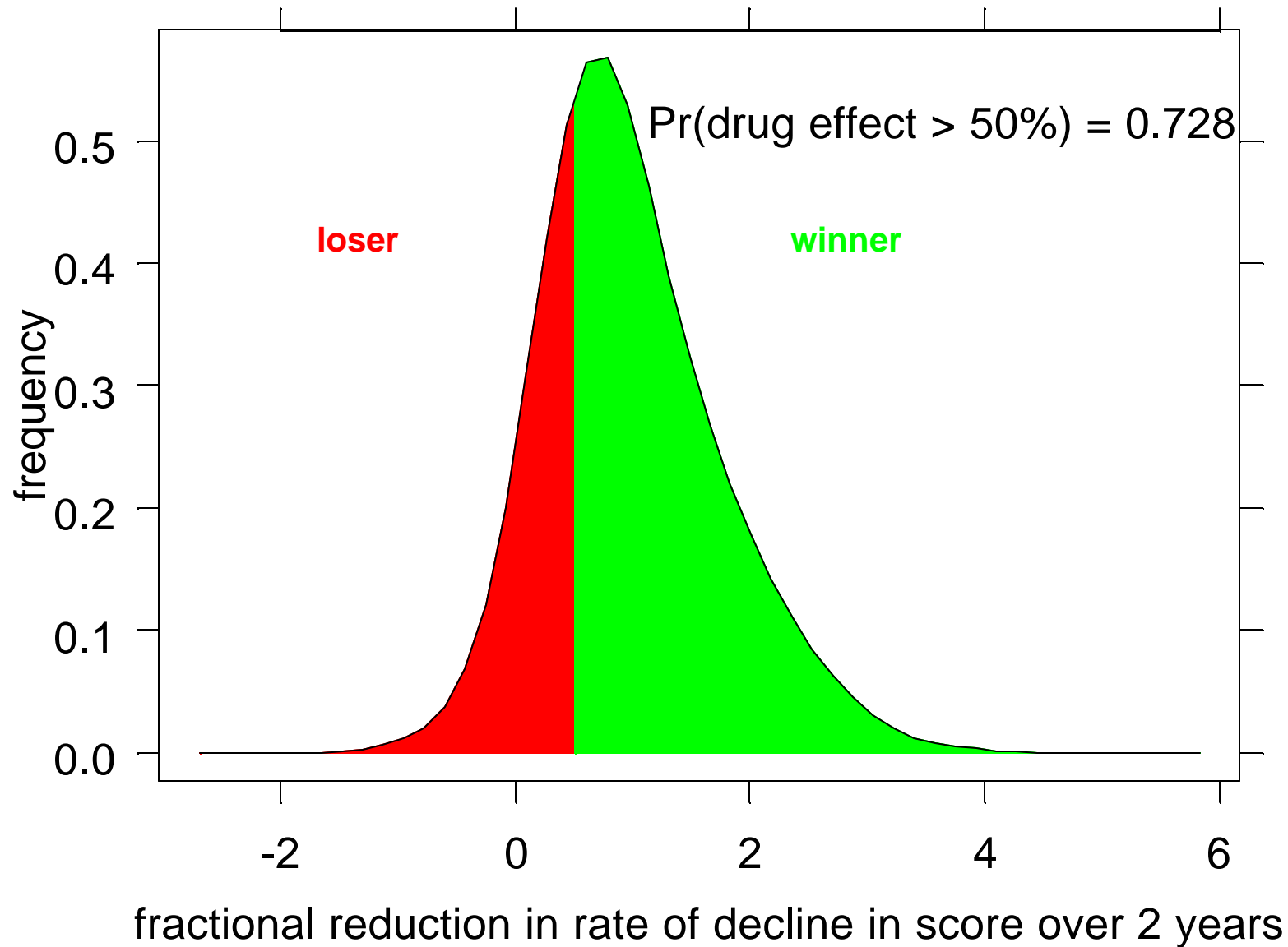
Trial simulation with uncertainty

- Uncertainty is modeled as inter-trial variation in the model parameters
 - Each trial is simulated using one random draw from the joint posterior distribution of the model parameters.
 - Those parameter values represent the “truth” for that simulated trial. Each simulated trial outcome is compared to its own unique “truth” under the model.
 - The basic notion is that you don’t know the real truth, so you would like to explore the performance of the trial design over a range of possibilities consistent with your uncertainty.
- Algorithm:
 - For $j = 1$ to n .trials
 - Sample parameters from the joint posterior distribution.
 - Simulate the trial.
 - Calculate statistic(s) of interest (e.g., treatment means, hypothesis test results, go/no-go decision, choice of treatment regimen further development, etc.).
 - Assess performance by comparison to model “truth”.
 - Implemented in S-PLUS

Trial performance is measured by the quality of the proof-of-concept decision

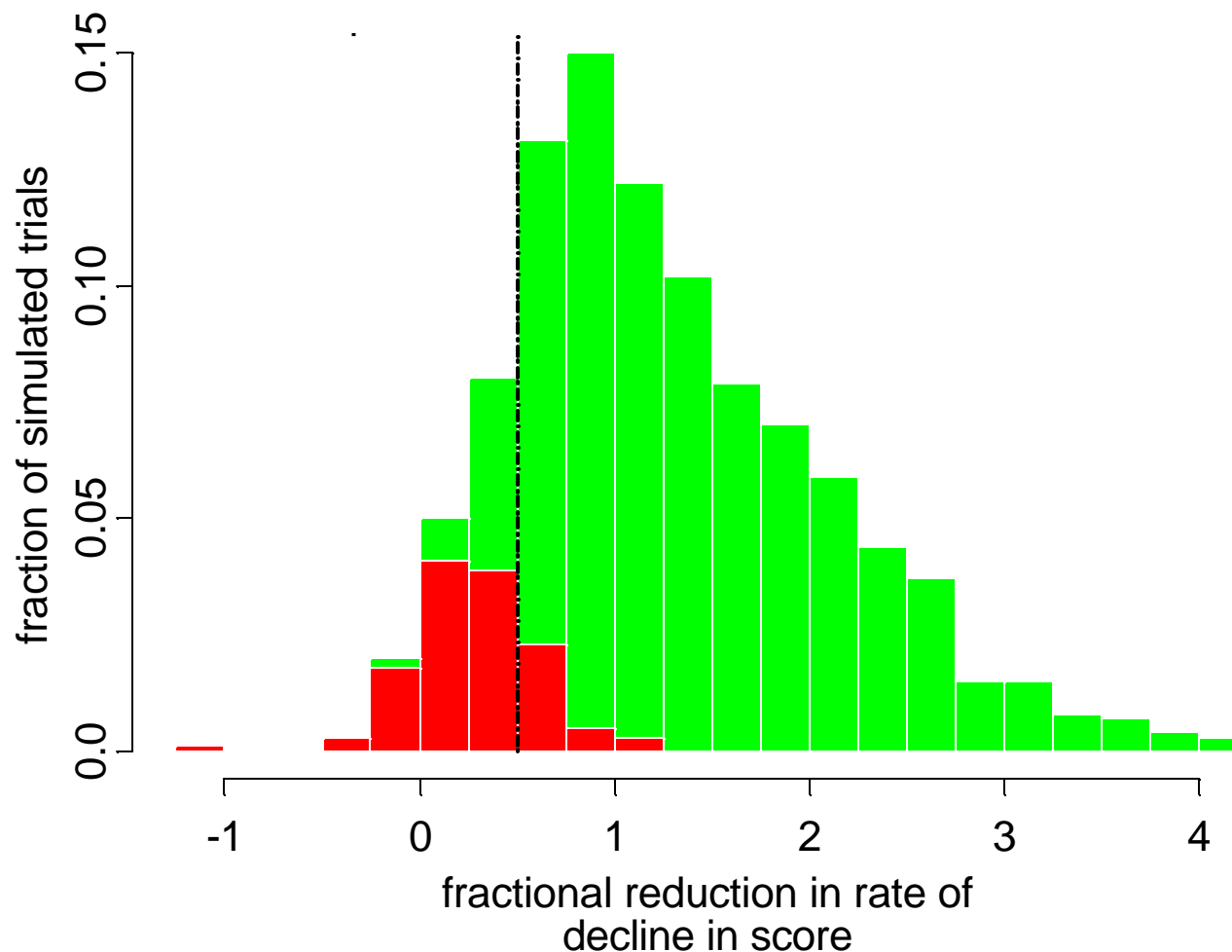
- Probability of reaching the correct (highest value) decision, i.e., go for a “winner” drug and no-go for a “loser” drug.
- You want to choose a trial design and a go/no-go decision method and criteria that minimizes:
 - $\Pr(\text{go}|\text{loser})$: probability of an incorrect go decision.
 - $\Pr(\text{stop}|\text{winner})$: probability of a lost opportunity.
- What is a “winner” or “loser” drug treatment?
 - The working definition of a “winner” used for the analyses presented here is a drug treatment that results in at least a 50% reduction in the rate of decline of the efficacy score over 2 years.

Prior information indicates a 73% probability that simuzine 100 mg is a “winner”



Current trial design and per protocol analysis results in too many go decisions for “losers”

- $\Pr(\text{go}|\text{loser}) = 0.34$ high probability of an incorrect go decision.
- $\Pr(\text{stop}|\text{winner}) = 0.037$ low probability of a lost opportunity.



Results of Bayesian longitudinal analyses

Go criteria: $\Pr(\geq 50\% \text{ reduction in the rate of decline over 2 years}) \geq p_{\text{crit}}$

Bayesian longitudinal analysis (without prior information) can be calibrated to markedly improve the PoC decision by reducing incorrect go decisions.

Analysis criteria	Simulated trial results	
p_{crit}	Pr(stop winner)	Pr(go loser)
Bayesian longitudinal analysis without prior information		
0.5	0.065	0.152
0.6	0.085	0.122
0.7	0.097	0.084
0.8	0.126	0.057
0.9	0.182	0.041
0.95	0.216	0.020
ANCOVA results	0.037	0.338

Results of Bayesian longitudinal analyses

Go criteria: $\Pr(\geq 50\% \text{ reduction in the rate of decline over 2 years}) \geq p_{\text{crit}}$

Incorporation of the prior data offers little or no additional improvement in the quality of the PoC decision.

Analysis criteria	Simulated trial results	
p_{crit}	Pr(stop winner)	Pr(go loser)
Bayesian longitudinal analysis without prior information		
0.5	0.065	0.152
0.6	0.085	0.122
0.7	0.097	0.084
0.8	0.126	0.057
0.9	0.182	0.041
0.95	0.216	0.020
Bayesian longitudinal analysis with prior information		
0.5	0.078	0.149
0.6	0.111	0.115
0.7	0.126	0.088
0.8	0.145	0.057
0.9	0.180	0.027
0.95	0.223	0.017
ANCOVA results	0.037	0.338

What next?

- The presented approach for assessing trial design performance does not explicitly optimize the tradeoffs between false positives (go|loser) and false negatives (stop|winner).
- That may be addressed by associating values (possibly economic) to the losses due to those competing errors and using Bayesian decision analysis to optimize the choice of analysis criteria (p_{crit}).
- Alternatively, the go/no-go decision method could be based on Bayesian decision analysis rather than the approach shown here.

Take home message:

Bayesian modeling and simulation can be done NOW!

- Recent advances in computer hardware, numerical methods and software make fully Bayesian approaches a practical option for many modeling, simulation and decision analysis applications.
- Bayesian principles and methods provide a coherent framework for:
 - Quantifying uncertainty,
 - Making inferences in the presence of that uncertainty.

Additional slides

Model development

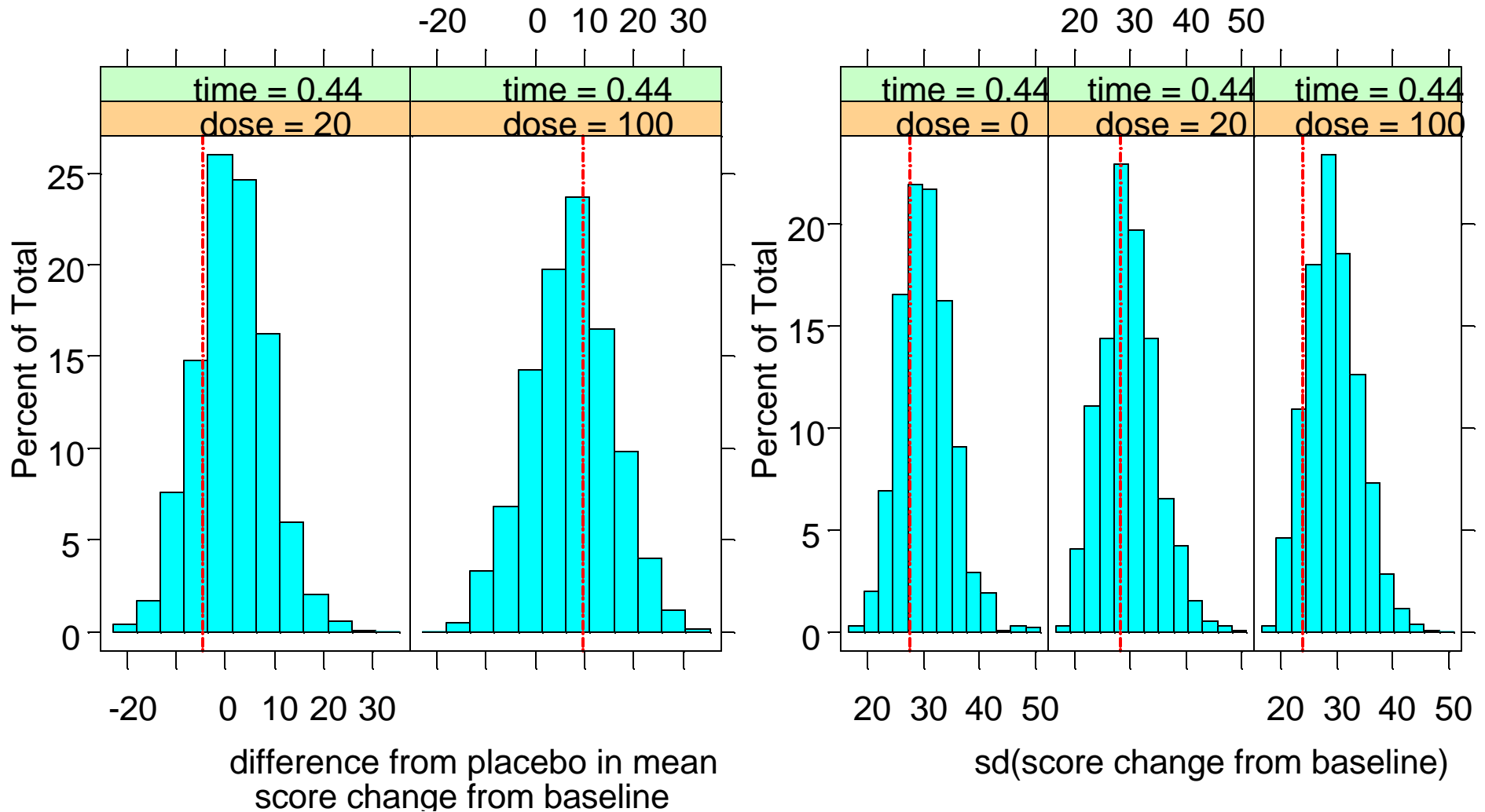
- General approach
 - Exploratory data analysis
 - Graphical exploration plus crude regression analyses (primarily S-PLUS).
 - Model exploration and selection
 - Linear and nonlinear mixed effects models fitted by maximum likelihood methods (NONMEM or S-PLUS)
 - Final model parameter estimation
 - Mixed effects models fitted by Bayesian methods (WinBUGS).
 - More rigorously characterizes the correlated uncertainties in the parameter estimates.
 - Posterior distributions of the parameters are used in subsequent clinical trial simulations.

Simulations to optimize the Phase II PoC trial analysis

- Focuses on methods of trial analysis, i.e., something that may still be influenced now that the trial is underway.
 - Conventional endpoint analysis (ANCOVA) versus Bayesian longitudinal analysis with or without use of prior data.
- Implementation:
 - Simulations performed using S-PLUS.
 - Simulated trials analyzed with S-PLUS (conventional ANCOVA) or WinBUGS (Bayesian longitudinal analyses).

Model fitting results: examples of posterior predictive checks

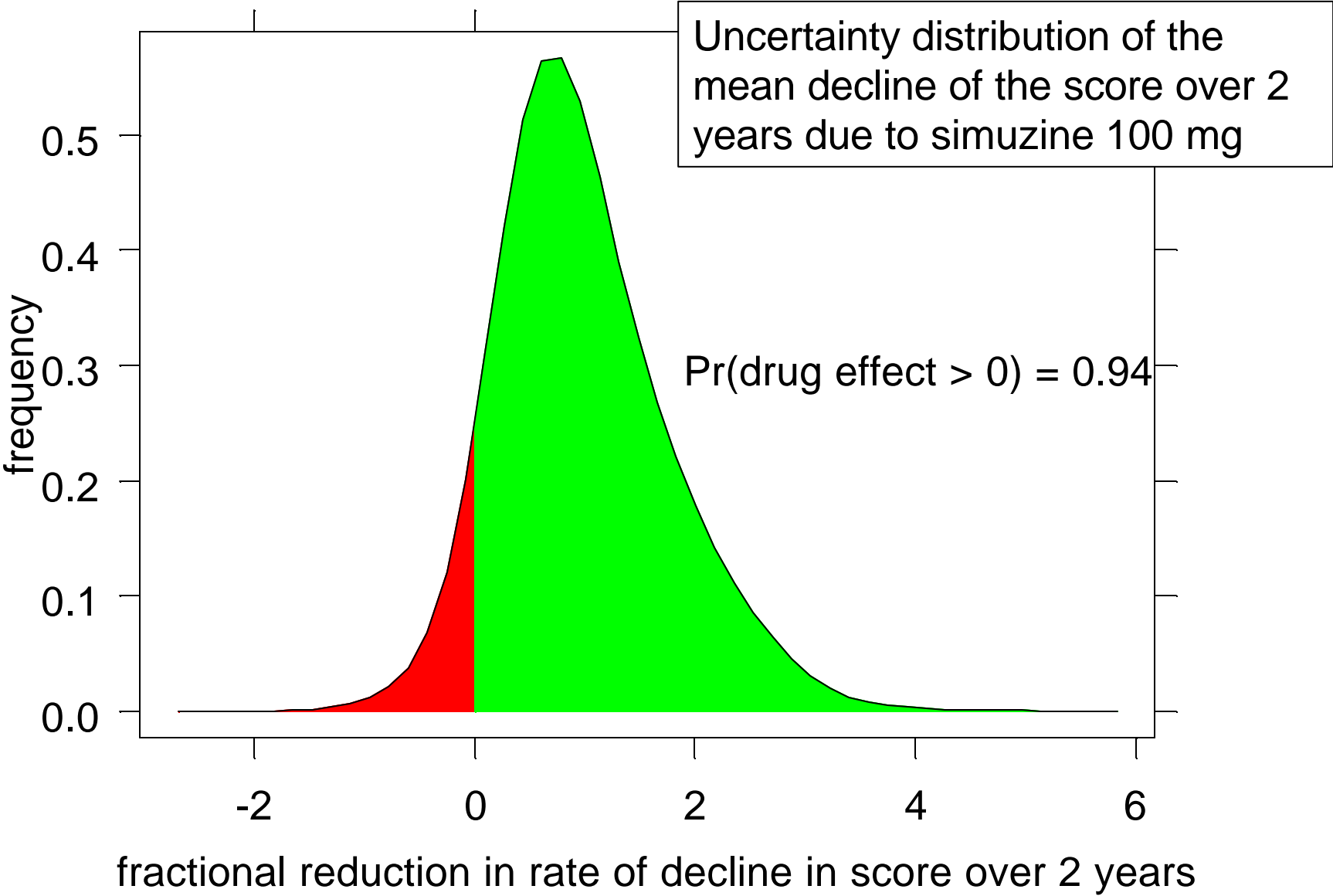
Simulated trial results are consistent with the results of the previous trial.



Histograms depict the distribution of simulated trial outcomes using the same design and patient covariates as the previous trial.

Observed values are shown as vertical dashed lines.

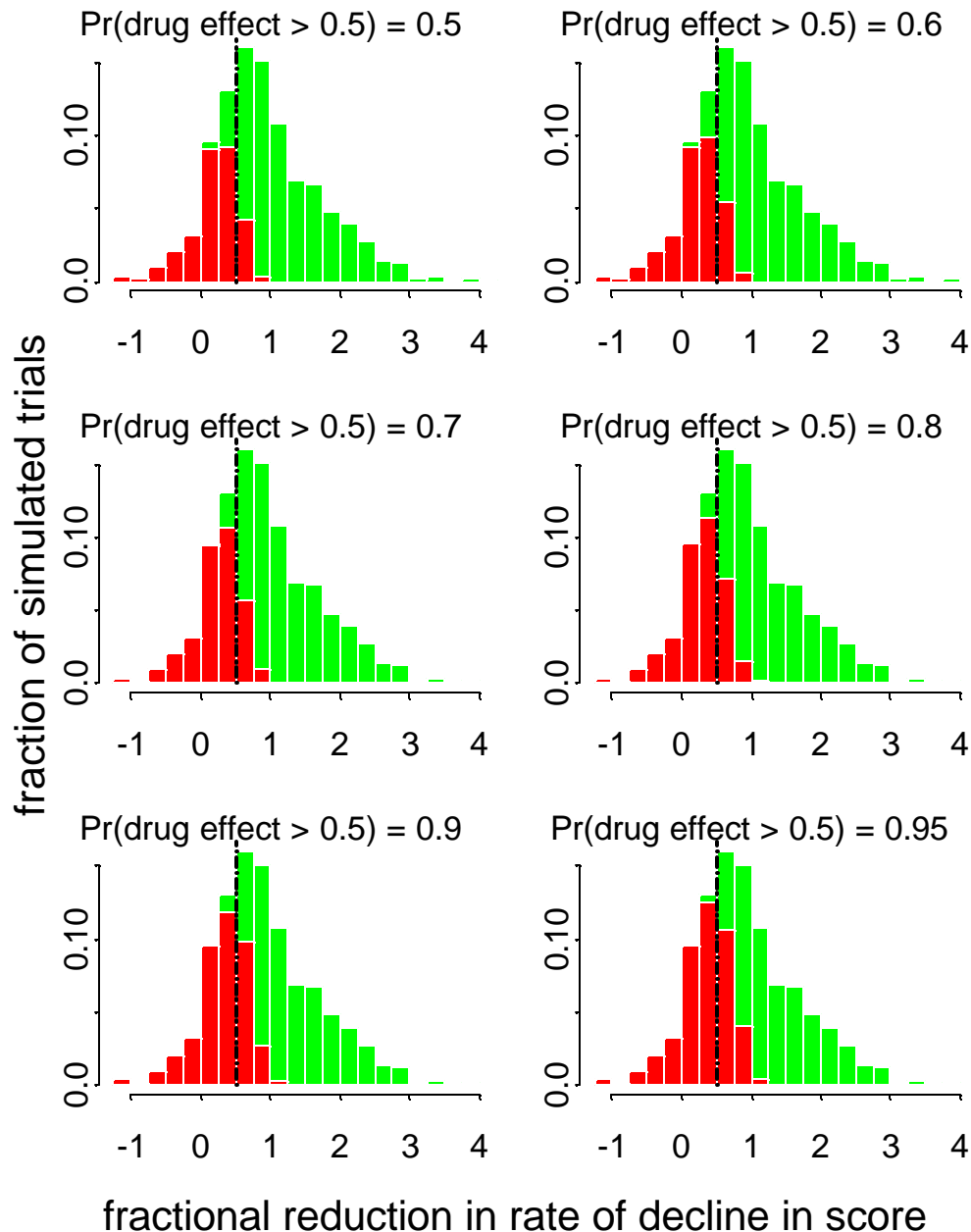
Model results indicate a highly uncertain but potentially large drug effect on the efficacy score



Trial analysis methods

- ANCOVA
 - Dependent variable: percent change in score at endpoint
 - Covariates: baseline score and simuzine dose (as a categorical variable)
 - Go criteria: $p < 0.05$ for dose effect and percent change in score greater for simuzine 100 mg than for placebo
- Bayesian longitudinal analysis
 - Dependent variable: score (all observation times)
 - The model used for simulation is fit to the trial data
 - Trial data alone
 - Trial data + prior data
 - Relatively non-informative prior distributions used for the model parameters
 - Go criteria:
 - $\Pr(\geq 50\% \text{ reduction in the rate of decline over 2 years}) \geq p_{\text{crit}}$
 - A range of p_{crit} values are explored (0.5, 0.6, 0.7, 0.8, 0.9, 0.95).

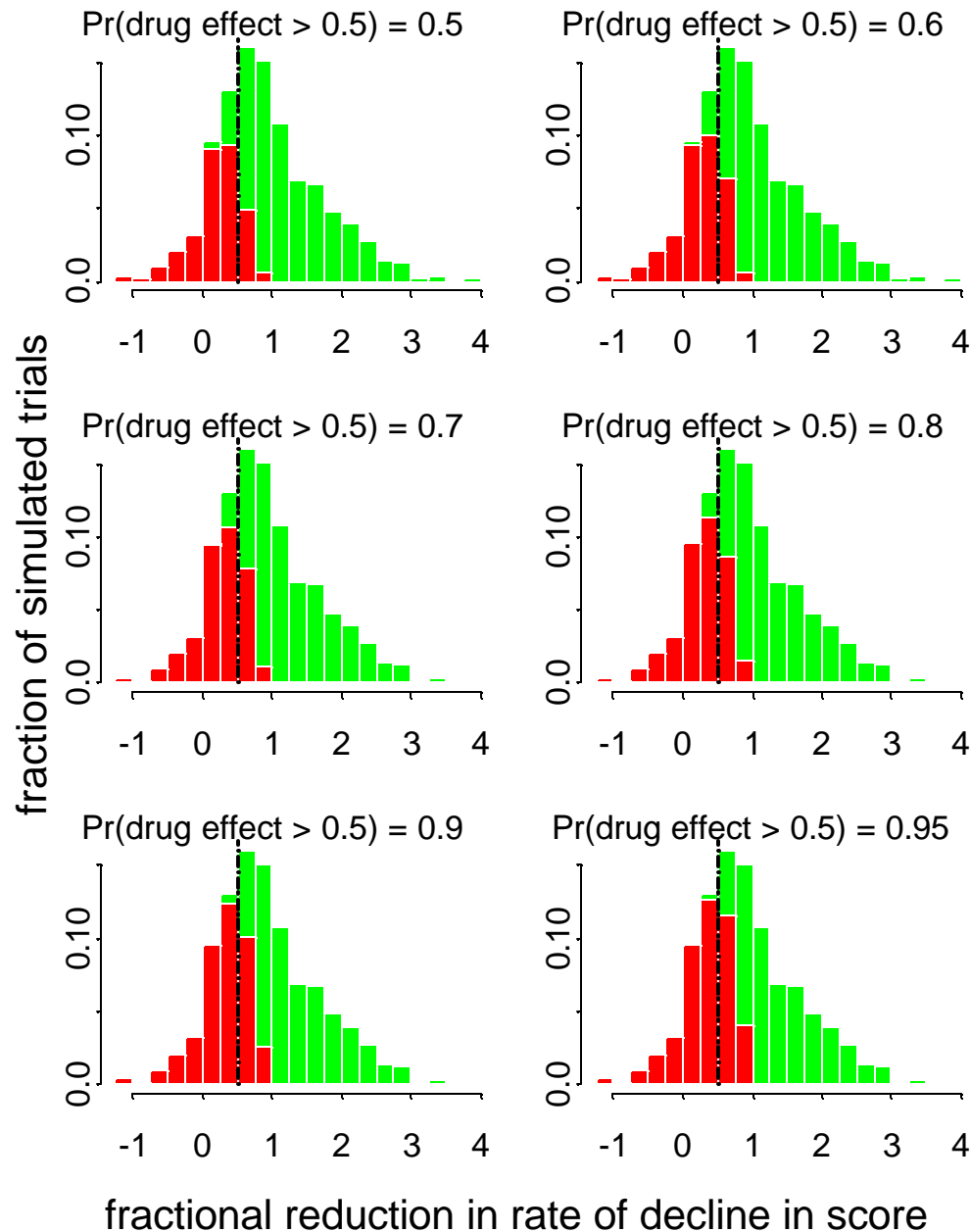
A Bayesian longitudinal analysis (without prior information) can be calibrated to markedly improve the PoC decision by reducing incorrect go decisions



Go criteria: $\Pr(\geq 50\% \text{ reduction in the rate of decline over 2 years}) \geq p_{\text{crit}}$

Analysis criteria	Simulated trial results		
	p_{crit}	Pr(stop winner)	Pr(go loser)
	0.5	0.065	0.152
	0.6	0.085	0.122
	0.7	0.097	0.084
	0.8	0.126	0.057
	0.9	0.182	0.041
	0.95	0.216	0.020
ANCOVA results		0.037	0.338

Incorporation of the prior data offers little or no additional improvement in the quality of the PoC decision (Bayesian longitudinal analysis method)



Go criteria: $\Pr(\geq 50\% \text{ reduction in the rate of decline over 2 years}) \geq p_{\text{crit}}$

Analysis criteria	Simulated trial results	
p_{crit}	$\Pr(\text{stop} \text{winner})$	$\Pr(\text{go} \text{loser})$
Bayesian longitudinal analysis without prior information		
0.5	0.065	0.152
0.6	0.085	0.122
0.7	0.097	0.084
0.8	0.126	0.057
0.9	0.182	0.041
0.95	0.216	0.020
Bayesian longitudinal analysis with prior information		
0.5	0.078	0.149
0.6	0.111	0.115
0.7	0.126	0.088
0.8	0.145	0.057
0.9	0.180	0.027
0.95	0.223	0.017
ANCOVA results	0.037	0.338

Summary of key inferences

- Current trial design and per protocol analysis results in too many go decisions for “losers”.
- A Bayesian longitudinal analysis (without prior information) can be calibrated to markedly improve the PoC decision by reducing incorrect go decisions.
- Incorporation of the prior data offers little or no additional improvement in the quality of the PoC decision (Bayesian longitudinal analysis method).