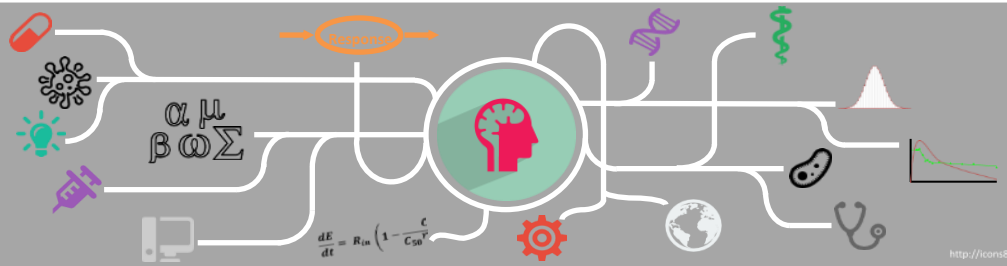
Modelling approaches in dose finding clinical trials: Simulation-based study comparing predictive performances of model averaging and model selection

S.Buatois[1,2], S.Ueckert[3], N.Frey[1], S.Retout[1], F.Mentré[2]

PAGE meeting, June 7 2017

1. Roche Pharma Research & Early Development, Clinical Pharmacology, Roche Innovation Center Basel
2. IAME, INSERM UMR 1137, University Paris Diderot, F-75018 Paris, France
3. Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden



Background

- Finding the right dose is a critical step in clinical drug development[1,2]
- Between 2000 and 2012, one of the highest causes of phase 3 submission failure was due to uncertainties related to dose selection[3]
- Increased interest in model based approaches to characterize the dose response relationship[4,5]

[1] Cross J. *et al*, Pharmacoepidemiol Drug Saf, 2002
[2] Heerdink E.R. *et al*, Pharmacoepidemiol Drug Saf, 2002
[3] Sacks L.V. *et al*, JAMA, 2014
[4] Bornkamp B. *et al*, J Biopharm Stat, 2007
[5] Pinheiro J. *et al*, Stat Med, 2014

Model based approaches

- Model selection (MS):
  - Most commonly used approach
  - Relies on selection of the model that best describes the data according to an information criterion *(e.g. AIC)*
  - Making inferences on the basis of the selected model ignores model uncertainty which could impair predictive performance[6,7]

[6] Buckland S.T. *et al*, Biometrics, 1997
[7] Mould D.R. *et al* , CPT Pharmacometrics Syst Pharmacol, 2012
[8] Aoki Y. *et al*, PAGE 23, 2014
[9] Schorning K. *et al*, Stat Med, 2016

Model based approaches

- Model selection (MS):
  - Most commonly used approach
  - Relies on selection of the model that best describes the data according to an information criterion *(e.g. AIC)*
  - Making inferences on the basis of the selected model ignores model uncertainty which could impair predictive performance[6,7]
- Model averaging (MA):
  - Allows measuring the uncertainty across a set of candidate models $l = 1, \ldots, L$ by weighting them in function of an IC[8,9] *(e.g. AIC)*

$$w_l = \frac{e^{\frac{-AIC_l}{2}}}{\sum_{i=1}^{L} e^{\frac{-AICi}{2}}}$$

[6] Buckland S.T. *et al*, Biometrics, 1997
[7] Mould D.R. *et al* , CPT Pharmacometrics Syst Pharmacol, 2012
[8] Aoki Y. *et al*, PAGE 23, 2014
[9] Schorning K. *et al*, Stat Med, 2016

Model Averaging

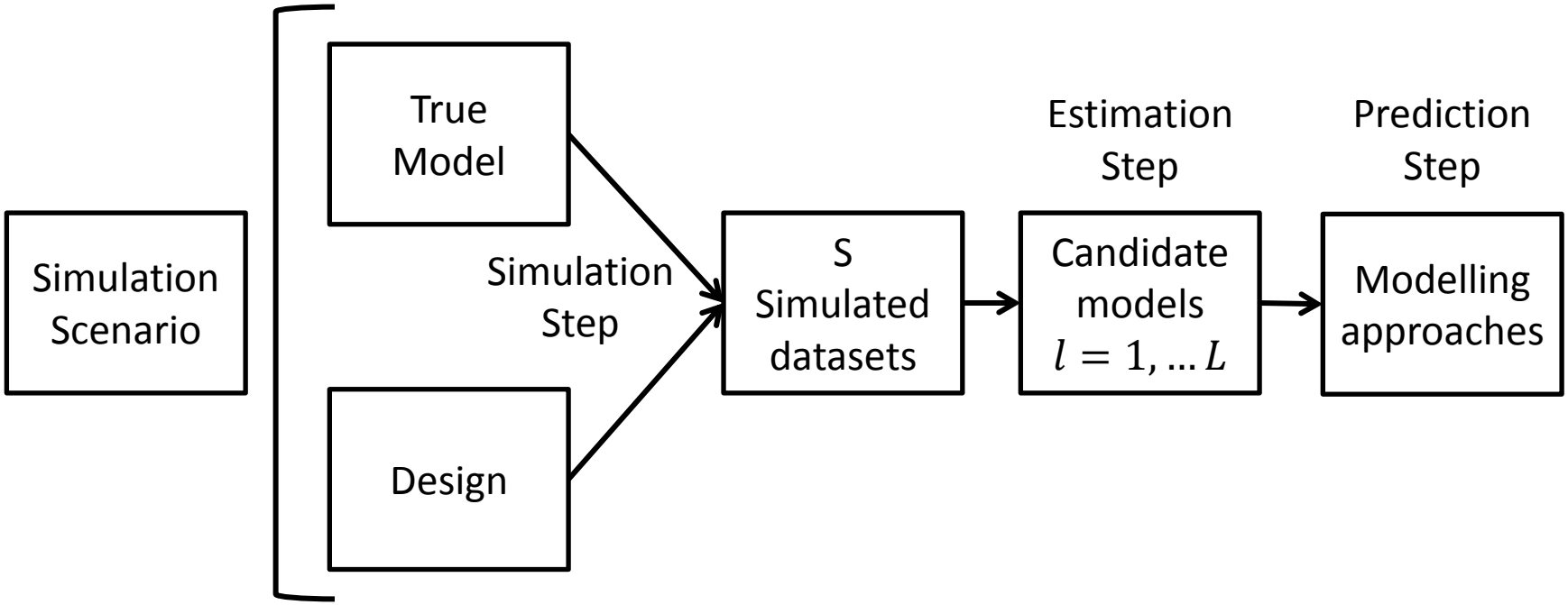
- Main applications in dose finding studies:
  - Aoki Y. *et al*, "Incorporate both the model parameter estimation uncertainty and the model structure uncertainty in dose selection"[8]
  - Schorning K. *et al*, "Model selection versus model averaging in dose finding studies"[9]

[8] Aoki Y. *et al*, PAGE 23, 2014
[9] Schorning K. *et al*, Stat Med, 2016

Objective:

- To compare predictive performances of model averaging (MA) and model selection (MS) based on a predefined set of NLMEMs with similar disease progression model and different dose-effect relationships

Workflow:


Simulation Scenario
True Model
Design
Simulation Step
S Simulated datasets
Estimation Step
Candidate models $l = 1, \ldots L$
Prediction Step
Modelling approaches

Case Study:

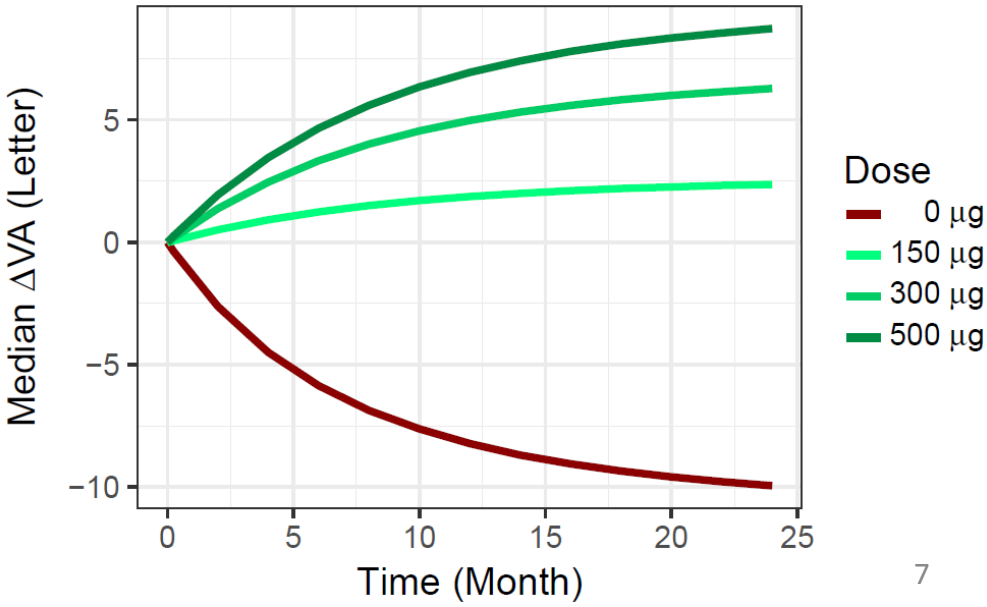
- Neovascular age-related macular degeneration (wet AMD)
- Biomarker: Visual Acuity (VA)
- True Model:

$$f(d_i, t_j, \Phi_i) = VA_{0,i} + \left(1 - e^{-k_{pr,i} \cdot t_j}\right) \cdot \left(\frac{Emax_i \cdot d}{ED_{50} + d} - \beta_i \cdot VA_{0,i} \right)$$

Log-normal distribution: VA_0, k_{pr}, β

Normal distribution: $Emax$

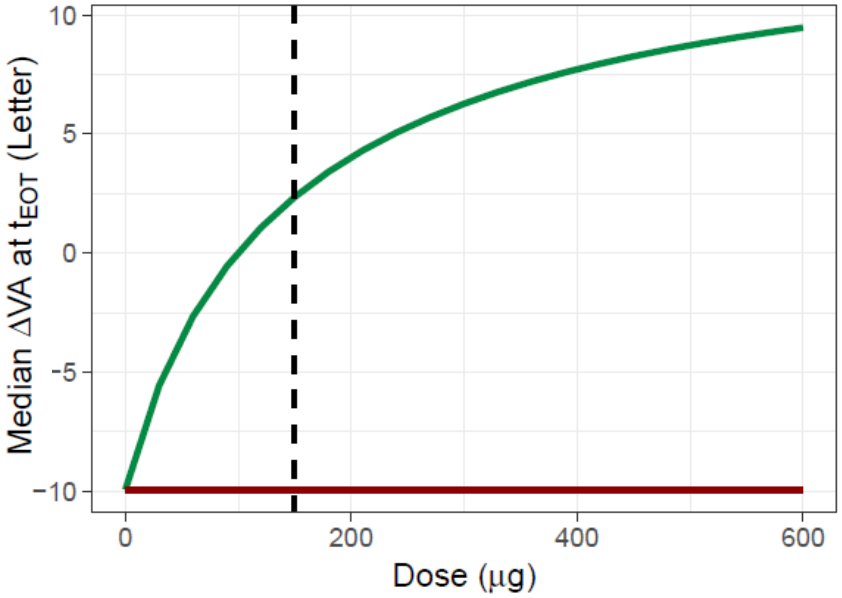
Parameter	μ	ω
VA_0 *(letter)*	55	0.26
k_{pr} (Day^{-1})	0.005	0.70
β	0.2	1.0
$Emax$ *(letter)*	30	12.2
ED_{50} (μg)	150	-



Study design:

- 300 patients equally distributed across the different dose levels
- 4 arms
- 26 observations per patient: baseline, day 7 & every month during 24 months
- End of trial (EOT): 24 months

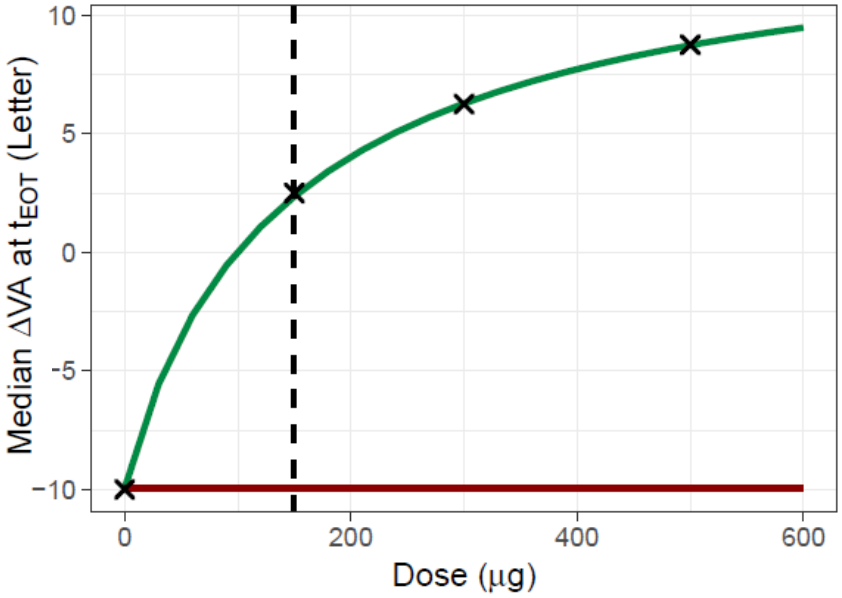
Simulation scenarios:



Study design:

- 300 patients equally distributed across the different dose levels
- 4 arms
- 26 observations per patient: baseline, day 7 & every month during 24 months
- End of trial (EOT): 24 months

Simulation scenarios:

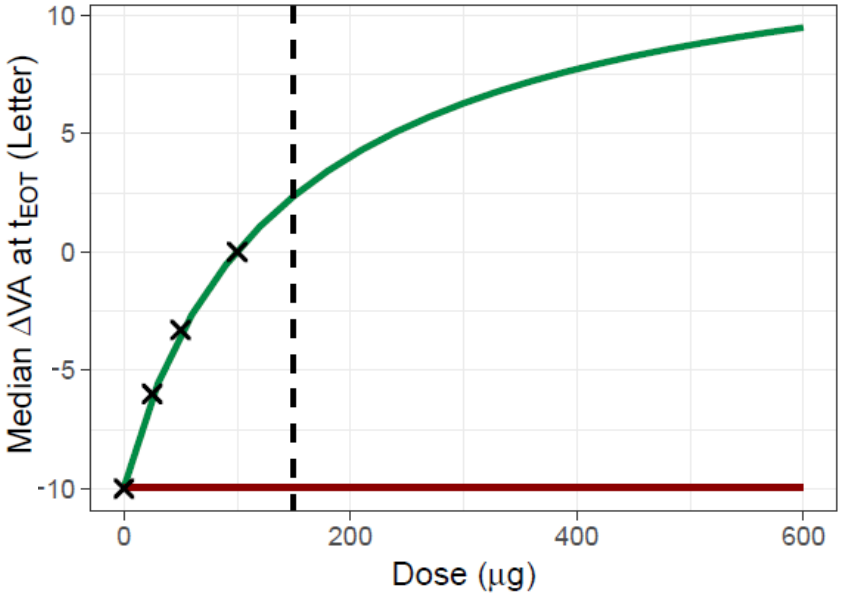


Scenario	Doses	Model
I	$0, 150, 300, 500\ \mu g$	Emax

Study design:

- 300 patients equally distributed across the different dose levels
- 4 arms
- 26 observations per patient: baseline, day 7 & every month during 24 months
- End of trial (EOT): 24 months

Simulation scenarios:

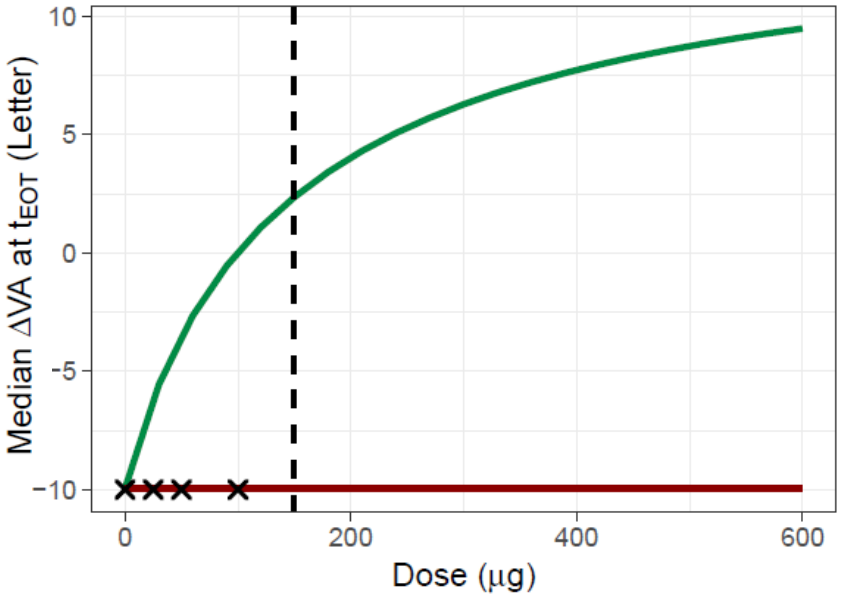


Scenario	Doses	Model
I	$0, 150, 300, 500\ \mu g$	Emax
II	$0, 25, 50, 100\ \mu g$	Emax

Study design:

- 300 patients equally distributed across the different dose levels
- 4 arms
- 26 observations per patient: baseline, day 7 & every month during 24 months
- End of trial (EOT): 24 months

Simulation scenarios:

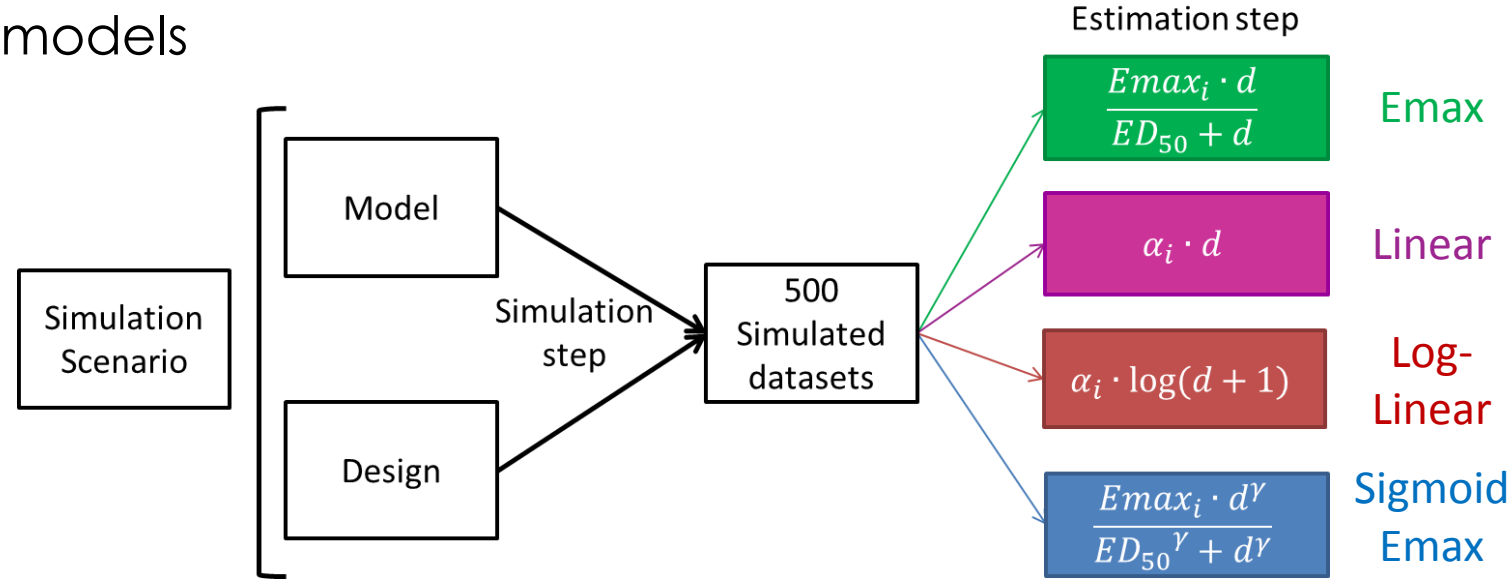


Scenario	Doses	Model
I	$0, 150, 300, 500\ \mu g$	Emax
II	$0, 25, 50, 100\ \mu g$	Emax
III	$0, 25, 50, 100\ \mu g$	No drug effect

Simulations & Estimations:

Simulations:

- For a given simulation scenario, $s = 1,..,500$ datasets were simulated and re-estimated using the $l = 1,..,4$ candidate models



Estimations:

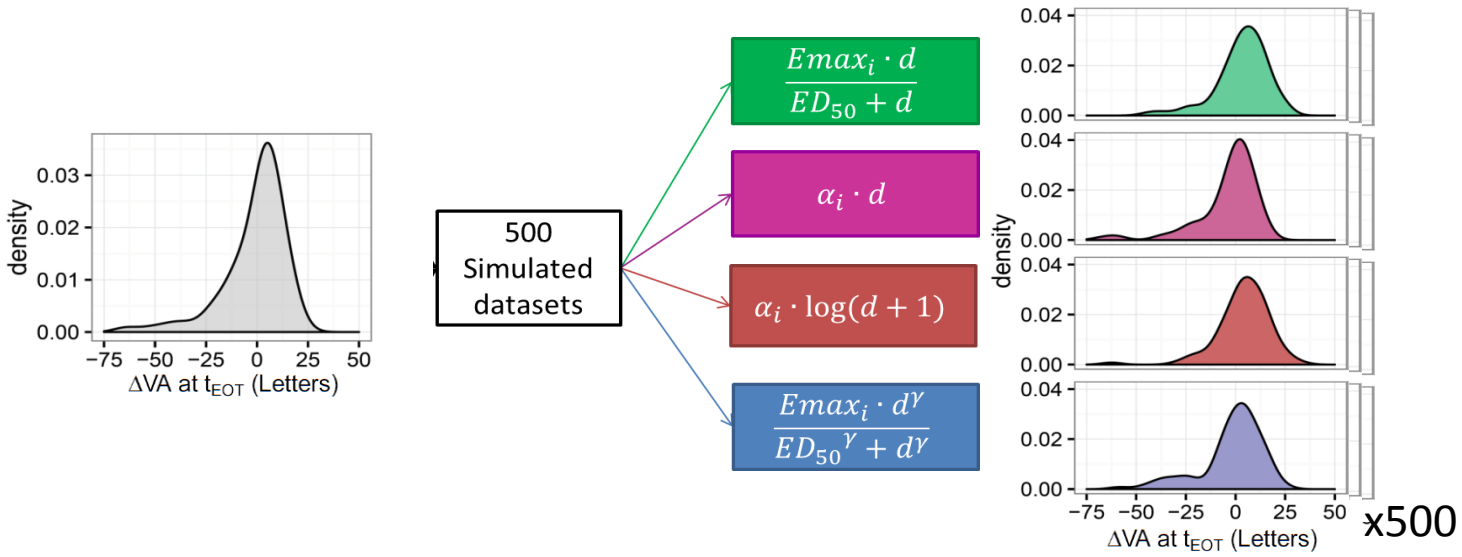
- Estimation of $\hat{\Psi}_{s,l}$ by maximizing the likelihood function
  - Expectation maximization method using importance sampling
- Software NONMEM 7.3

Model Predictions

- Model predictions were used to compute the true and estimated probability distribution of the VA change from baseline (ΔVA) at end of trial (t_{EOT})
- Predictions were computed for each dose $d^k = \{0, 150, 300, 500\}$
- Modelling approaches: MS

True probability distribution:

Model selection

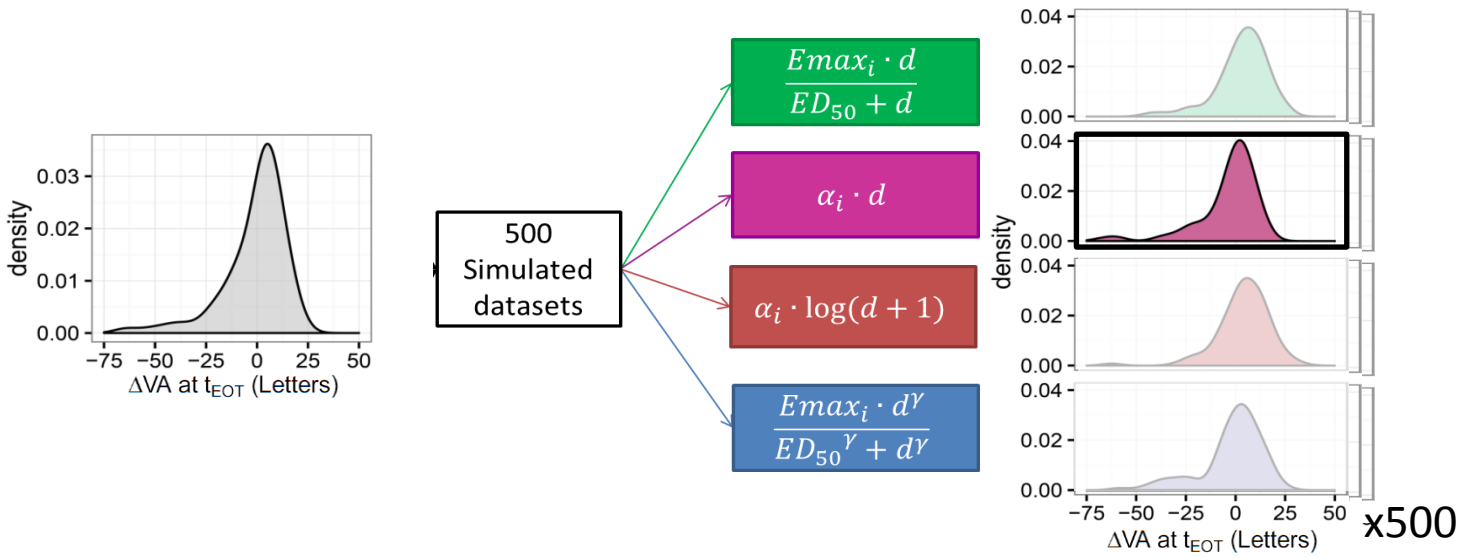


Model Predictions

- Model predictions were used to compute the true and estimated probability distribution of the VA change from baseline (ΔVA) at end of trial (t_{EOT})
- Predictions were computed for each dose $d^k = \{0, 150, 300, 500\}$
- Modelling approaches: MS

True probability distribution:

Model selection

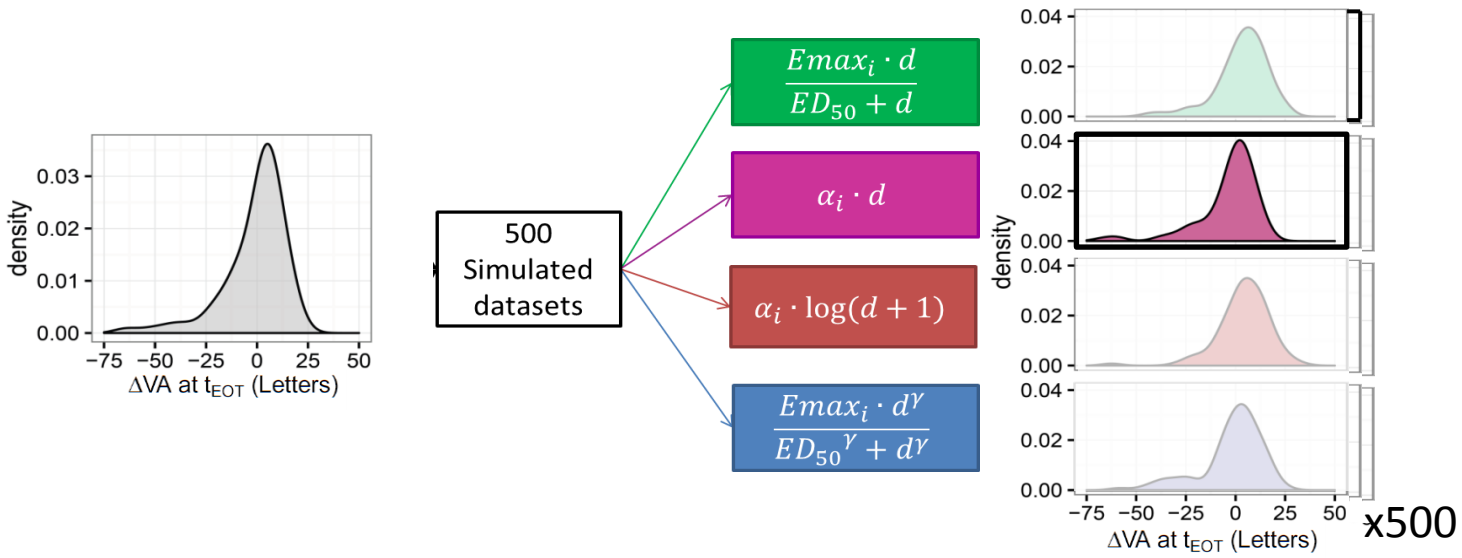


Model Predictions

- Model predictions were used to compute the true and estimated probability distribution of the VA change from baseline (ΔVA) at end of trial (t_{EOT})
- Predictions were computed for each dose $d^k = \{0, 150, 300, 500\}$
- Modelling approaches: MS

True probability distribution:

Model selection

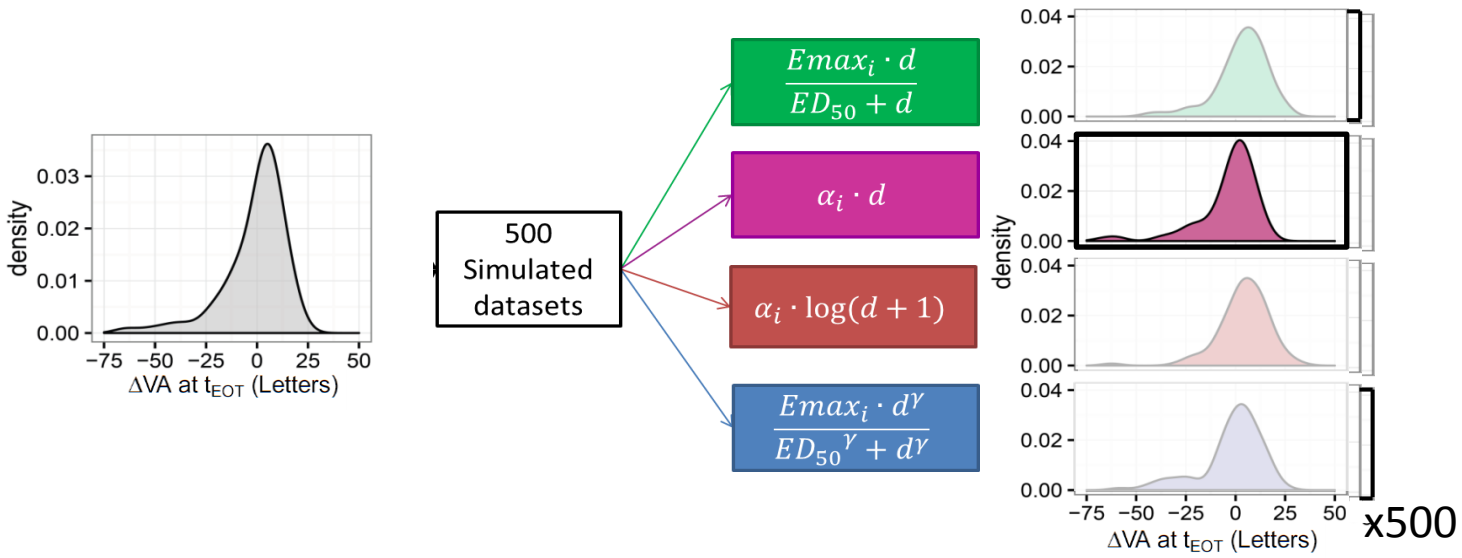


Model Predictions

- Model predictions were used to compute the true and estimated probability distribution of the VA change from baseline (ΔVA) at end of trial (t_{EOT})
- Predictions were computed for each dose $d^k = \{0, 150, 300, 500\}$
- Modelling approaches: MS

True probability distribution:

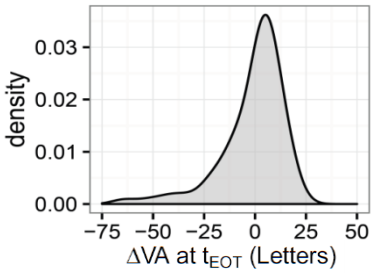
Model selection



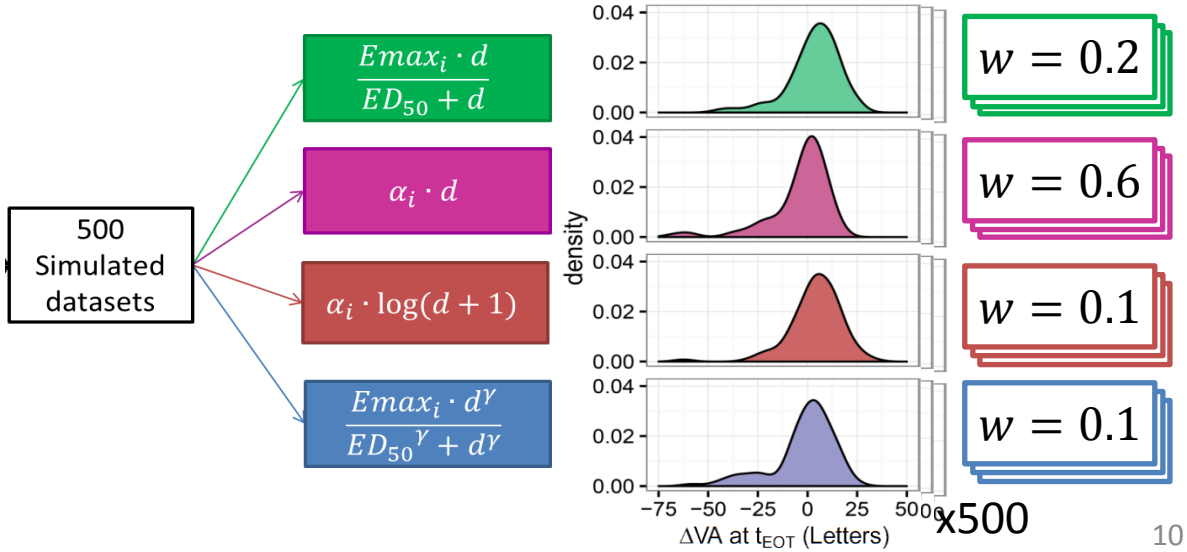
Model Predictions

- Model predictions were used to compute the true and estimated probability distribution of the VA change from baseline (ΔVA) at end of trial (t_{EOT})
- Predictions were computed for each dose $d^k = \{0, 150, 300, 500\}$
- Modelling approaches: MS, MA

True probability distribution:



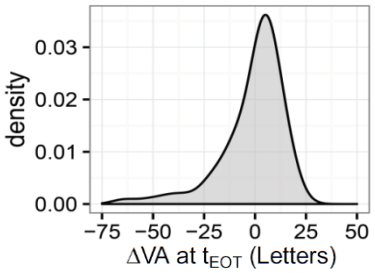
Model averaging



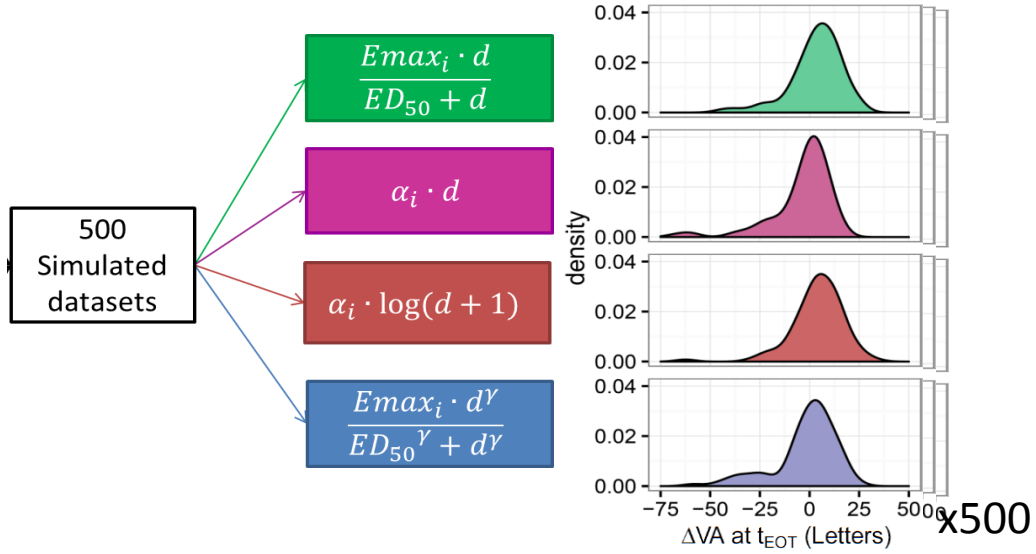
Model Predictions

- Model predictions were used to compute the true and estimated probability distribution of the VA change from baseline (ΔVA) at end of trial (t_{EOT})
- Predictions were computed for each dose $d^k = \{0, 150, 300, 500\}$
- Modelling approaches: MS, MA, Candidate models

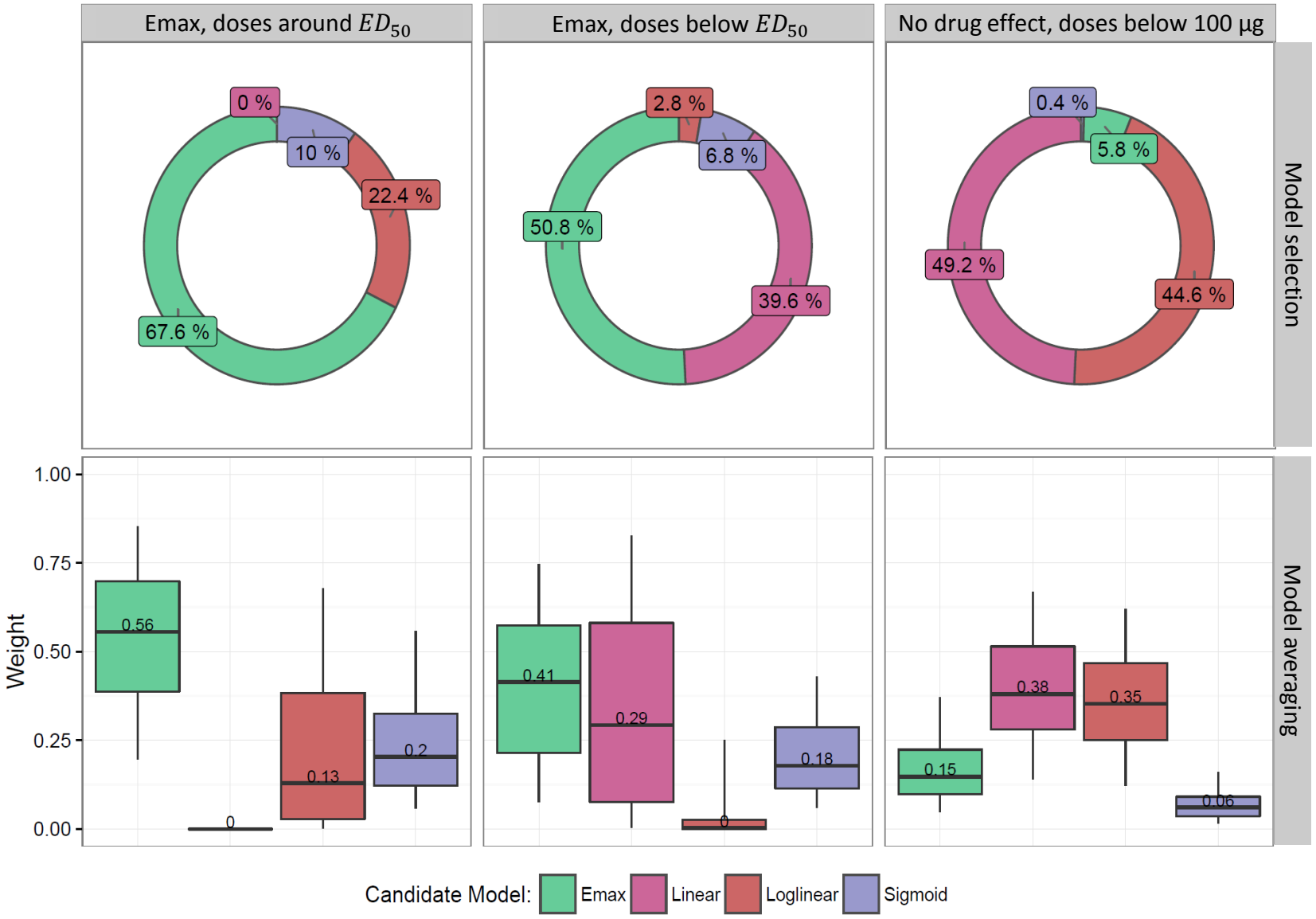
True probability distribution:



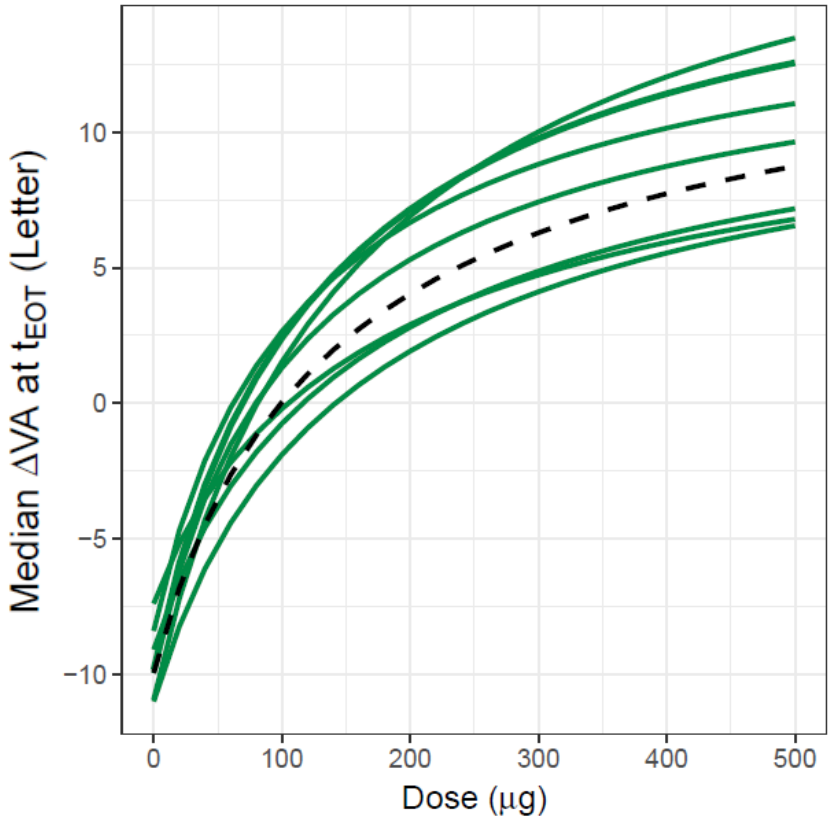
Candidate models



Model selection & Model averaging:



Performance criteria



<u>Primary end point:</u> median Δ_{VA} at t_{EOT}

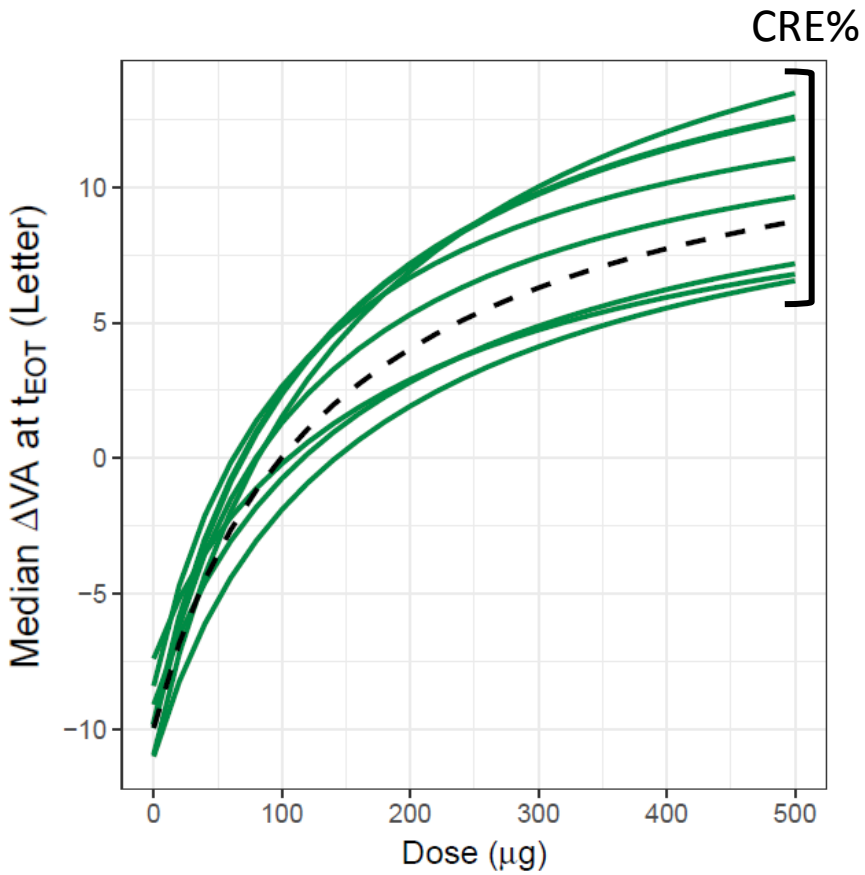
<u>Clinically relevant effect:</u> increase of the median Δ_{VA} at t_{EOT} of at least 15 points compared to placebo patients

- - - True

— Predicted

Performance criteria

1) Percentage of trials concluding to a clinically relevant effect at the highest simulated dose d^k



Primary end point: median Δ_{VA} at t_{EOT}

Clinically relevant effect: increase of the median Δ_{VA} at t_{EOT} of at least 15 points compared to placebo patients

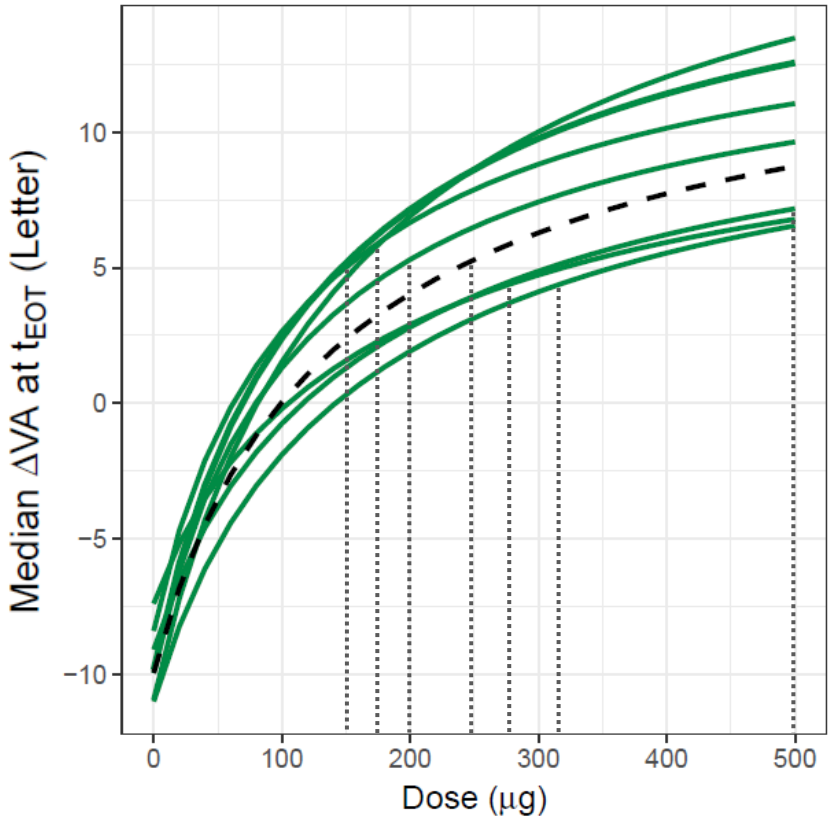
True

Predicted

Performance criteria

1) Percentage of trials concluding to a clinically relevant effect at the highest simulated dose d^k

2) Minimum dose at which a clinically relevant effect is achieved



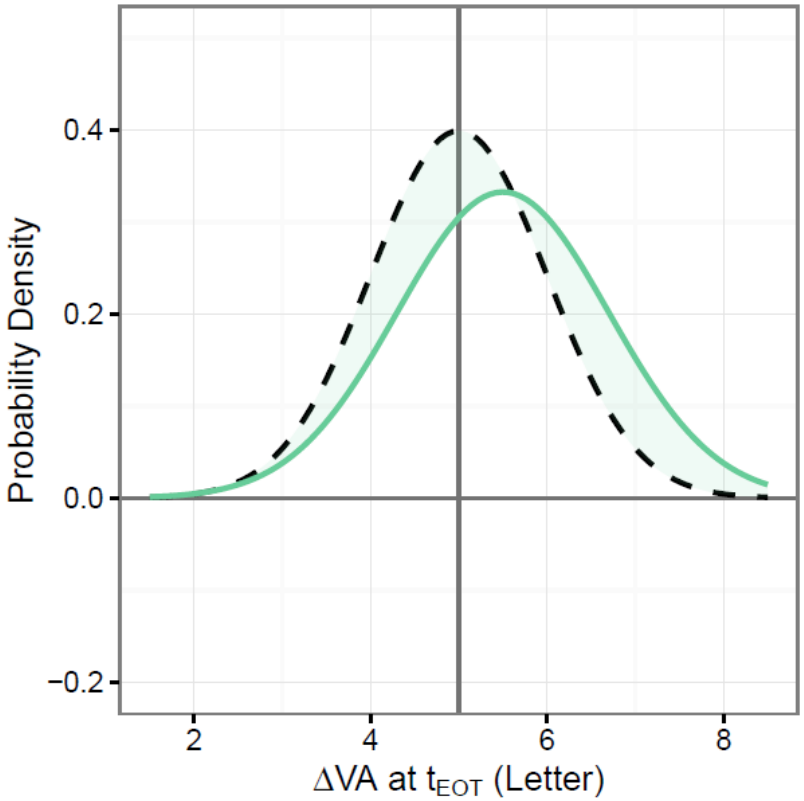
Primary end point: median Δ_{VA} at t_{EOT}

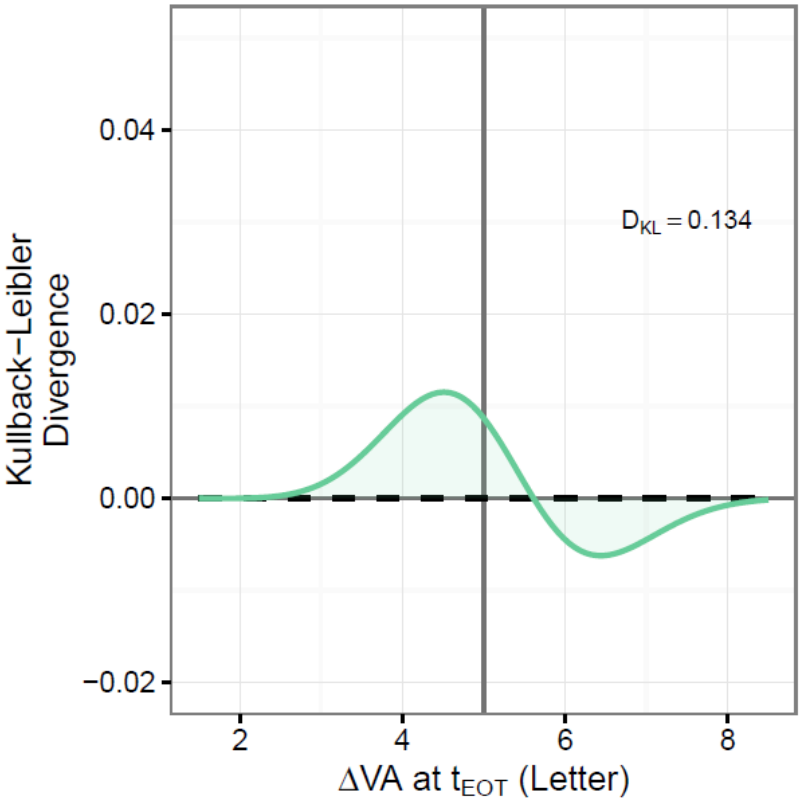
Clinically relevant effect: increase of the median Δ_{VA} at t_{EOT} of at least 15 points compared to placebo patients

True
Predicted
MED

Performance criteria

1) Percentage of trials concluding to a clinically relevant effect at the highest simulated dose d^k
2) Minimum dose at which a clinically relevant effect is achieved
3) Kullback–Leibler divergence (D_{KL})[10]: for a given dose





[10] MacKay DJC. New York: Cambridge University Press; 2003

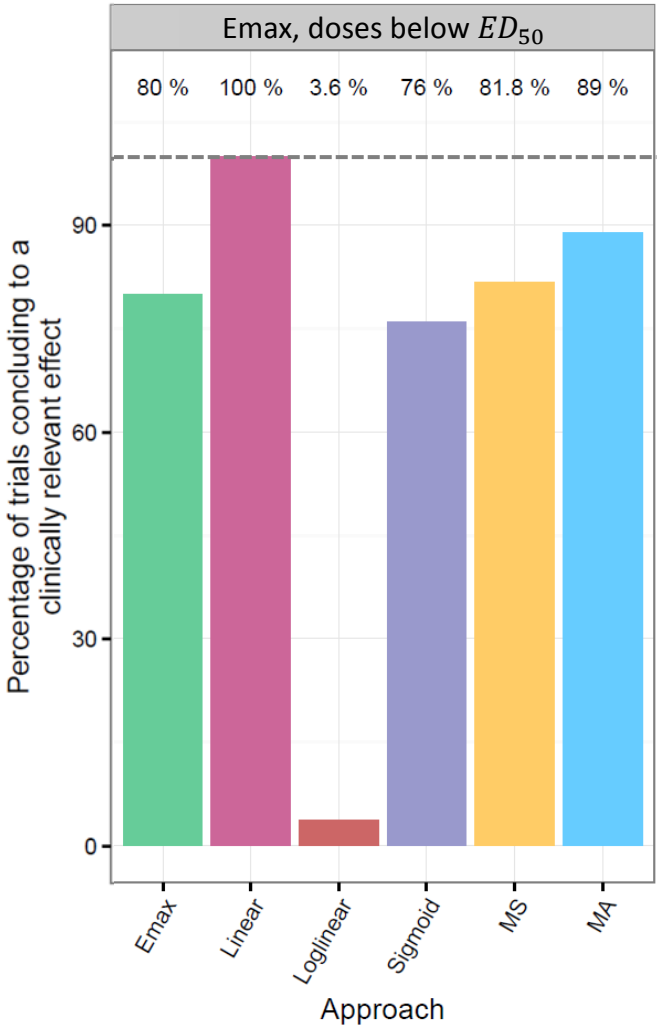
Performance criteria

1) Percentage of trials concluding to a clinically relevant effect at the highest simulated dose d^k
2) Minimum dose at which a clinically relevant effect is achieved
3) Kullback–Leibler divergence (D_{KL})[10]: for a given dose
   - Total D_{KL}: over the set of doses d^k at t_{EOT}

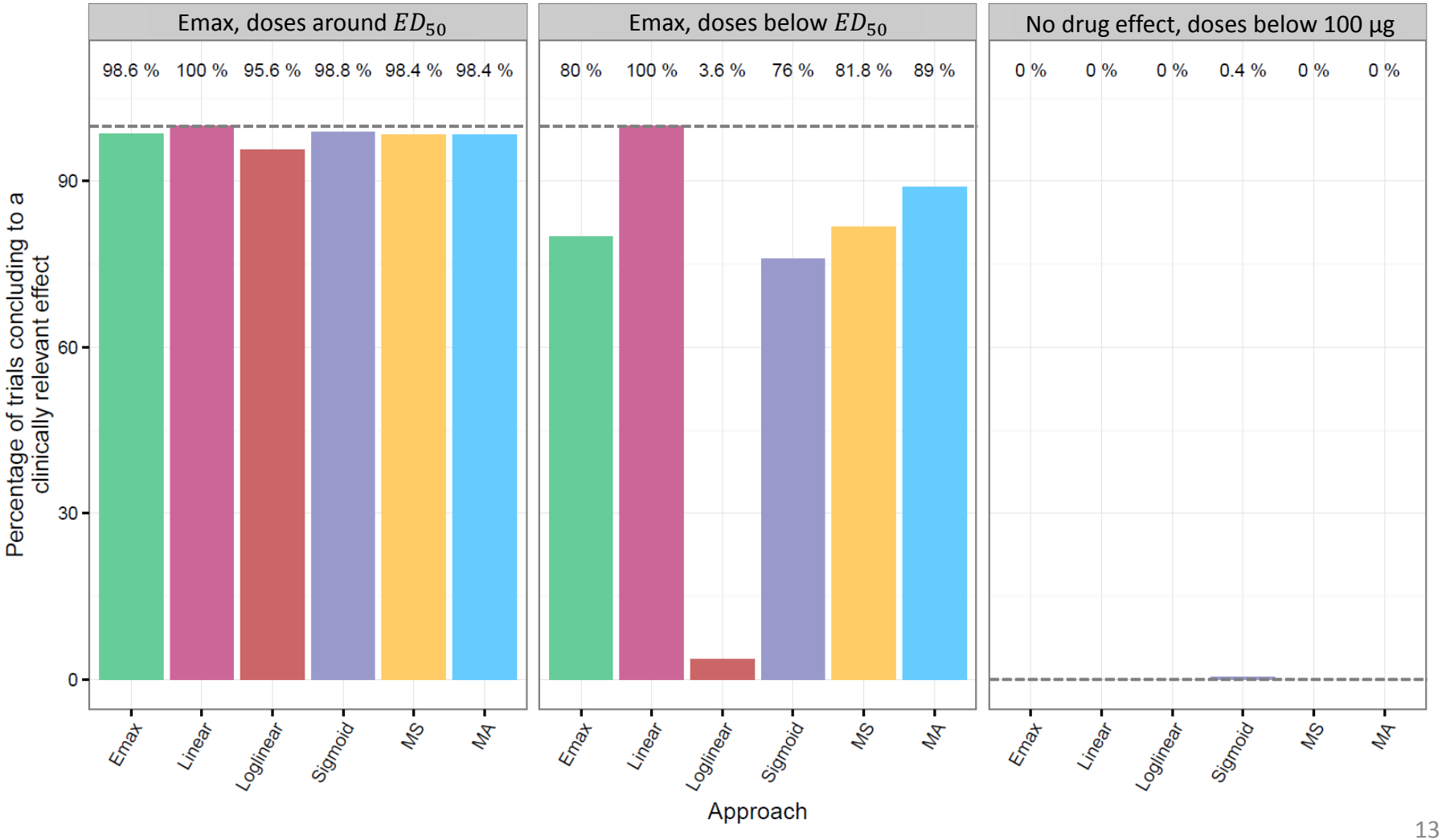
$$Total\ D_{KL}(p^*|p) = \sum_{k=1}^{K} D_{KL_k}(p^*|p)$$

[10] MacKay DJC. New York: Cambridge University Press; 2003

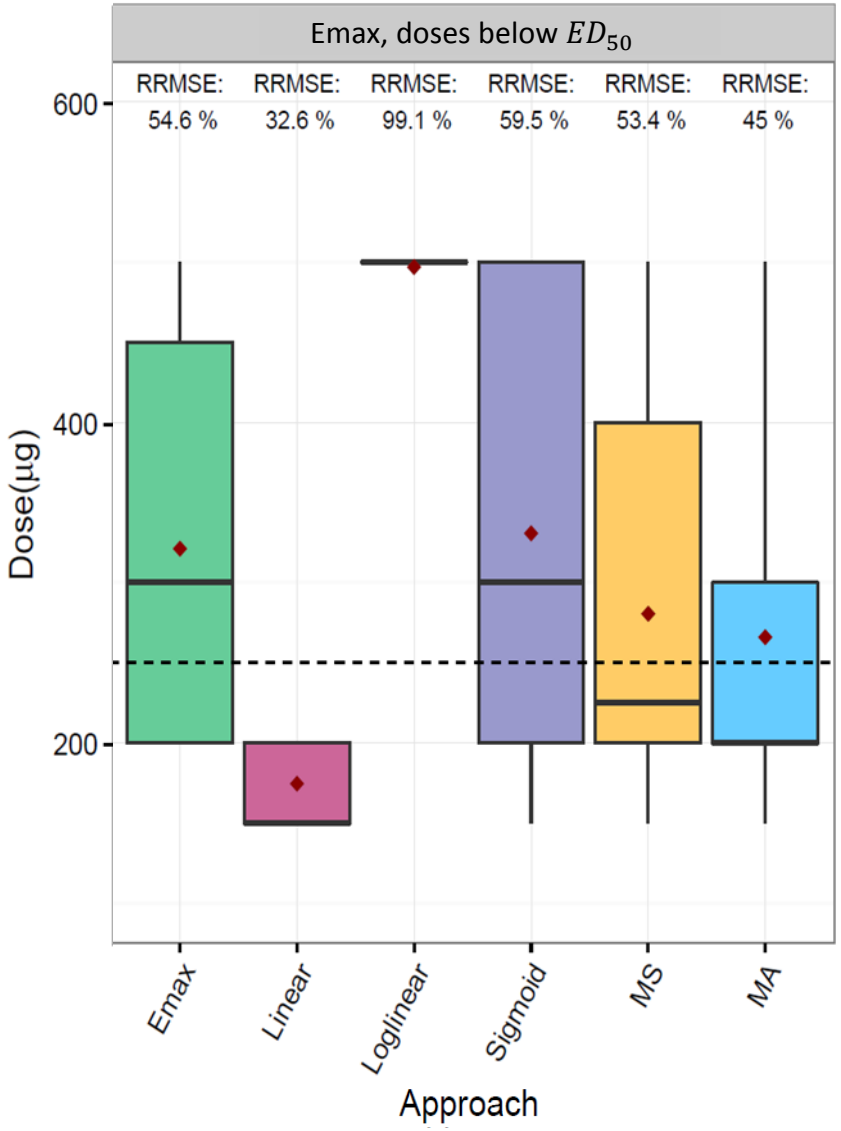
1) Clinically relevant effect: (CRE)



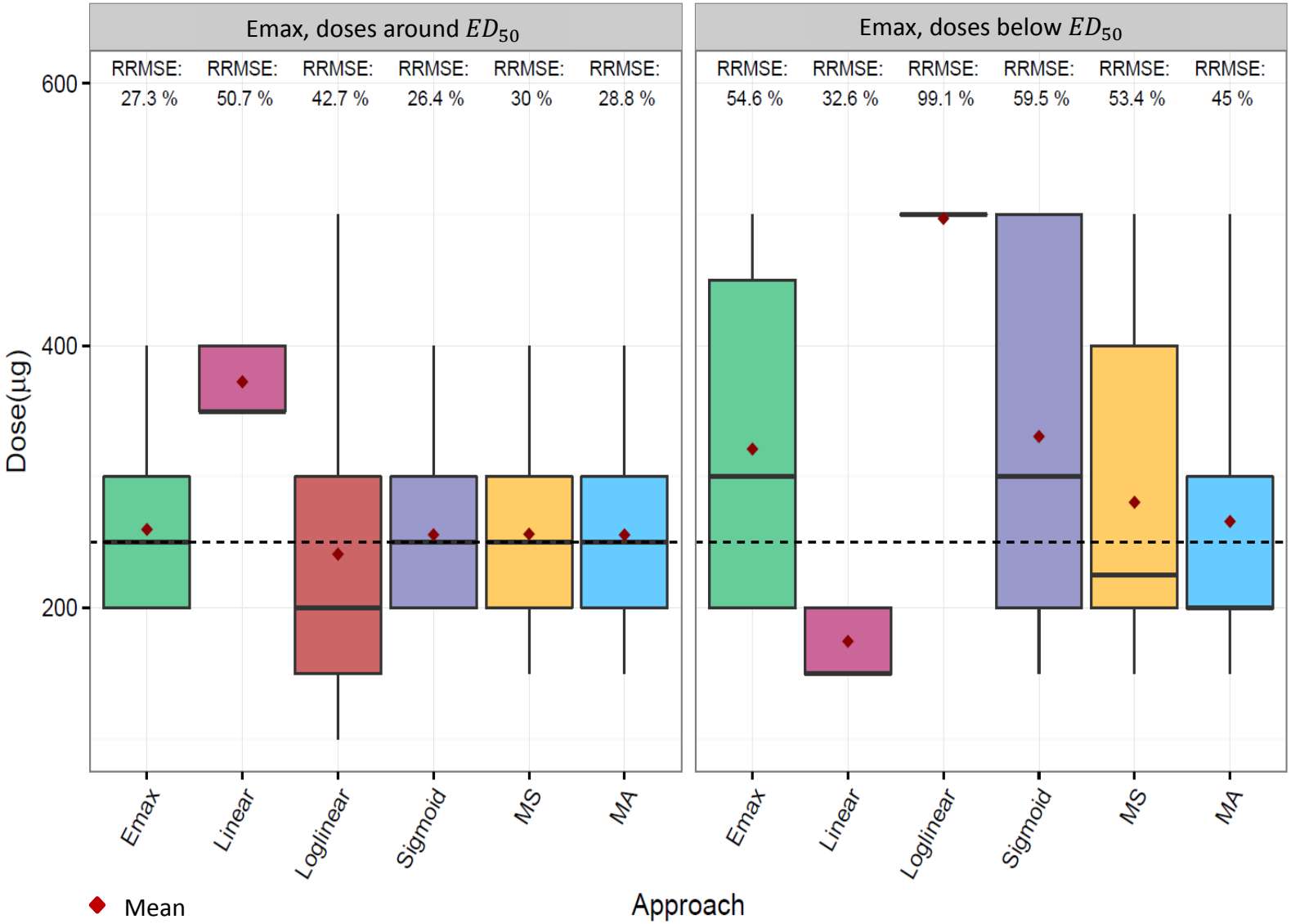
1) Clinically relevant effect: (CRE)



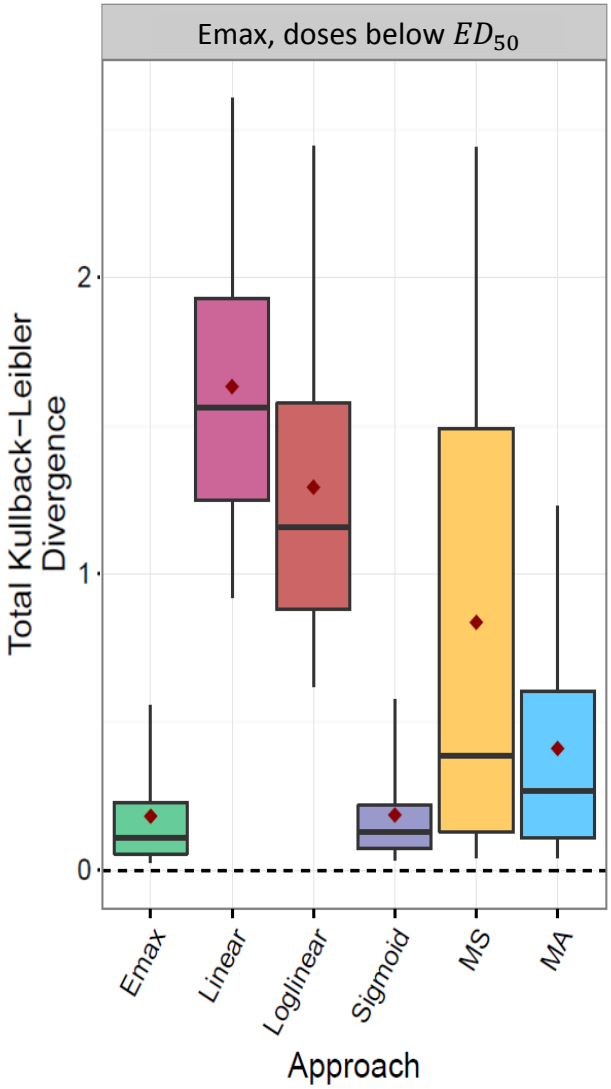
2) *Target dose d: Minimum effective dose (MED)*



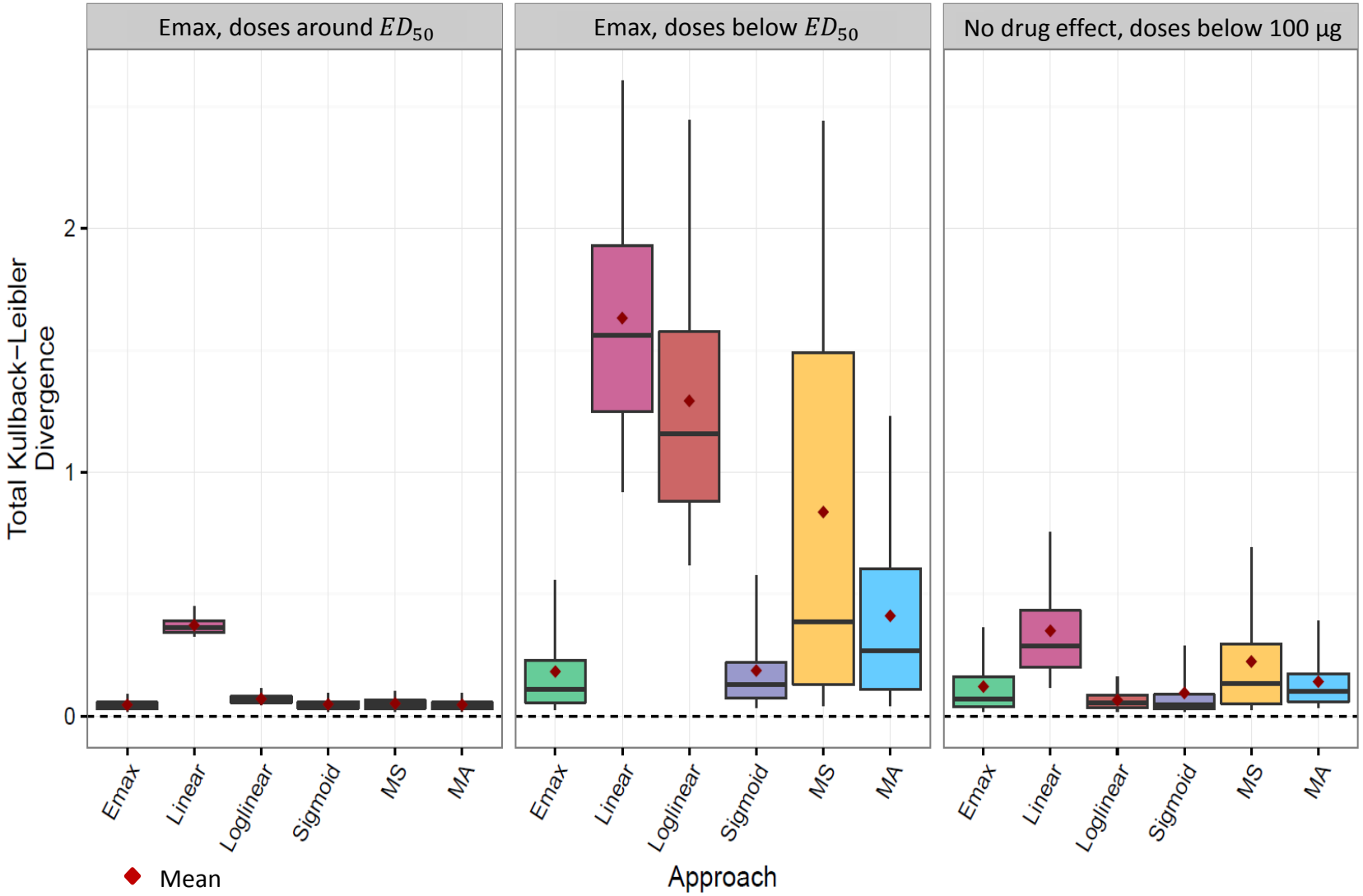
2) *Target dose d: Minimum effective dose (MED)*



3) *Kullback–Leibler divergence*



3) *Kullback–Leibler divergence*



Conclusions:

- Under an informative design, MA & MS provided similar predictive performances and led to an accurate prediction of the target dose
- Under less informative designs, by estimating weights on a predefine set of NLMEMs, MA showed relatively better predictive performance than MS increasing the likelihood to accurately characterize the dose response relationship

Perspectives:

- Include parameter uncertainties in the predictions
  - Compare coverage performances of MS and MA
- Explore the case where the true model is not in the set of candidate models
- Include different disease progression models in the set of candidate models

Thanks to:

- Inserm Colleagues:

- Roche Colleagues:

Backup

5) Model selection & Model averaging:

- Both approaches rely on an information criterion I[8]
- The value I_l was calculated under each candidate model

$$I = -2LL(y, \Psi) + 2pen$$

I	Penalty (pen) term for model l
AIC	p
BIC_N	$0.5 \times p \times \log(\mathrm{N})$
BIC_{nt}	$0.5 \times p \times \log(n_{tot})$
$CAIC_N$	$0.5 \times p \times (\log(\mathrm{N}) + 1)$
$CAIC_{nt}$	$0.5 \times p \times (\log(n_{tot}) + 1)$

Model selection:

Predictions are obtain using the model with the lowest I_l value among the L candidate models

Model averaging: [9]

Weights are associated with each of the candidate models w_l

$$w_l = \frac{e^{\frac{-I_l}{2}}}{\sum_{i=1}^{L} e^{\frac{-I_i}{2}}}$$

[8] Bertrand J. *et al*, J Biopharm Stat. 2008
[9] Claeskens G . *et al*, New York: Cambridge University Press, 2008

Information criteria:

- Scenario: Emax, doses around ED50

Scenario	Doses	Model
1	0,150,300,500 µg	*Emax*

Model selection:

I	Emax	Linear	Log-Linear	Sigmoid Emax
AIC	57%	0%	21%	22%
BIC_N	41%	0%	56%	3%
BIC_{nt}	22%	0%	77%	1%
$CAIC_N$	34%	0%	64%	2%
$CAIC_{nt}$	18%	0%	81%	1%

Selected percentage per candidate model in the S dataset for the *I* information criteria

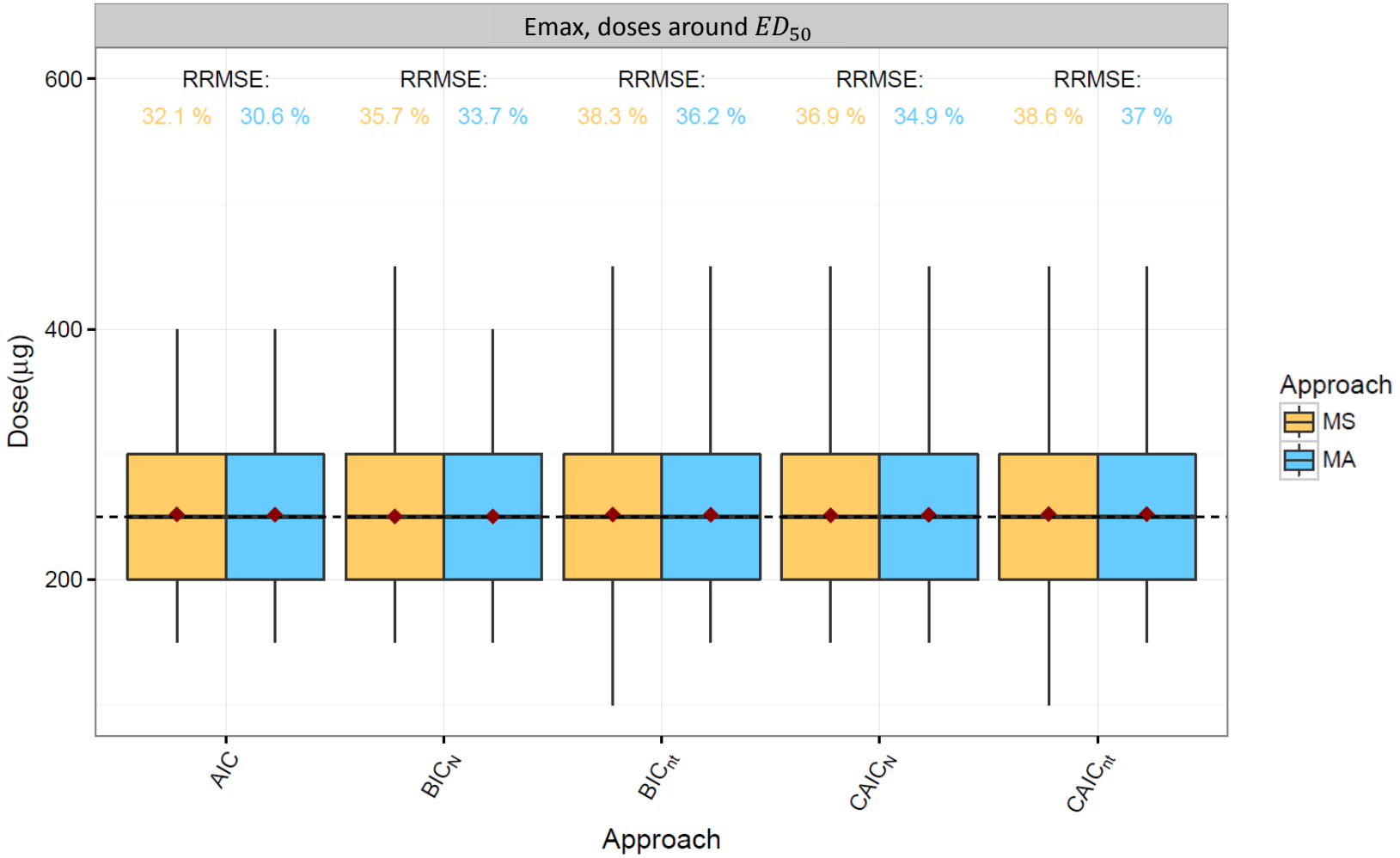
Model averaging:

I	Emax	Linear	Log-Linear	Sigmoid Emax
AIC	0,50	0,00	0,13	0,22
BIC_N	0,37	0,00	0,57	0,02
BIC_{nt}	0,12	0,00	0,88	0,00
$CAIC_N$	0,27	0,00	0,69	0,01
$CAIC_{nt}$	0,07	0,00	0,92	0,00

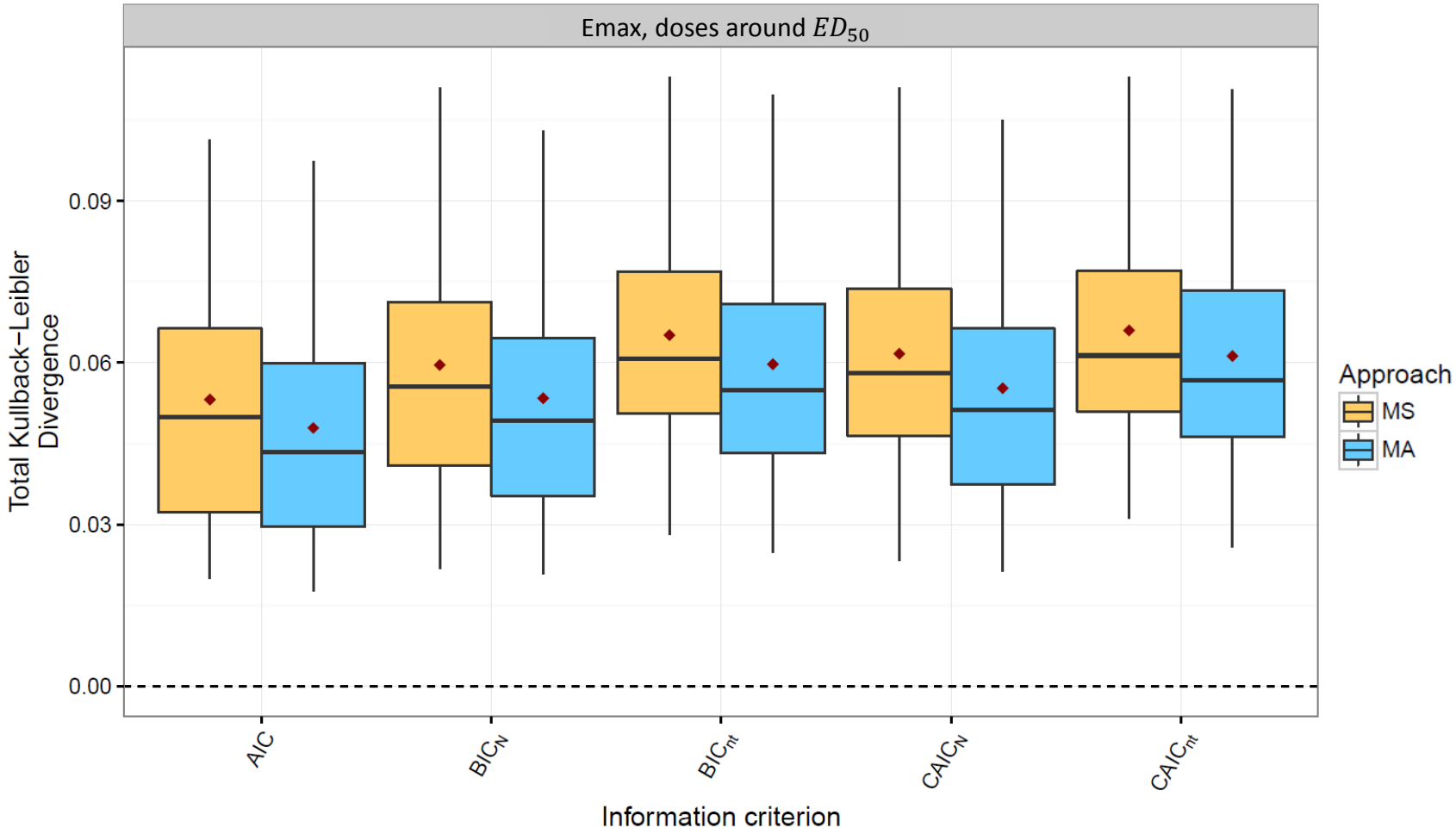
Median of the estimated weight per candidate model in the S dataset for the *I* information criteria

24

Target dose: Boxplot representation of the predicted MED for the
l information criteria. The dashed line represents the reference
and the diamonds the mean values



Dose response profile: Boxplot representation of the total D_{KL} for the I information criteria. The dashed line represents the reference and the diamonds the mean values



Information criteria:

- Scenario: Emax, doses below ED50

Model selection:

I	Emax	Linear	Log-Linear	Sigmoid Emax
AIC	50%	37%	4%	9%
BIC_N	14%	72%	14%	0%
BIC_{nt}	2%	80%	18%	0%
$CAIC_N$	9%	75%	16%	0%
$CAIC_{nt}$	1%	80%	19%	0%

Selected percentage per candidate model in the S dataset for the I information criteria

Model averaging:

I	Emax	Linear	Log-Linear	Sigmoid Emax
AIC	0,40	0,26	0,00	0,18
BIC_N	0,15	0,77	0,01	0,01
BIC_{nt}	0,03	0,94	0,01	0,00
$CAIC_N$	0,10	0,84	0,01	0,00
$CAIC_{nt}$	0,02	0,96	0,01	0,00

Median of the estimated weights per candidate model in the S dataset for the I information criteria

Information criteria:

- Scenario: No drug effect

Model selection:

I	Emax	Linear	Log-Linear	Sigmoid Emax
AIC	2%	48%	50%	0%
BIC_N	0%	51%	49%	0%
BIC_{nt}	0%	51%	49%	0%
$CAIC_N$	0%	51%	49%	0%
$CAIC_{nt}$	0%	51%	49%	0%

Selected percentage per candidate model in the S dataset for the I information criteria

Model averaging:

I	Emax	Linear	Log-Linear	Sigmoid Emax
AIC	0,00	0,45	0,45	0,00
BIC_N	0,00	0,48	0,50	0,00
BIC_{nt}	0,00	0,49	0,51	0,00
$CAIC_N$	0,00	0,49	0,50	0,00
$CAIC_{nt}$	0,00	0,49	0,51	0,00

Median of the estimated weights per candidate model in the S dataset for the I information criteria