

Three overlapping circles in shades of orange and yellow are positioned on the left side of the slide, serving as a decorative background for the text.

# Use of Informative Priors in Model-Informed Drug Development

**Nicky Best**

Advanced Biostatistics and Data Analytics, GSK

Acknowledgements:

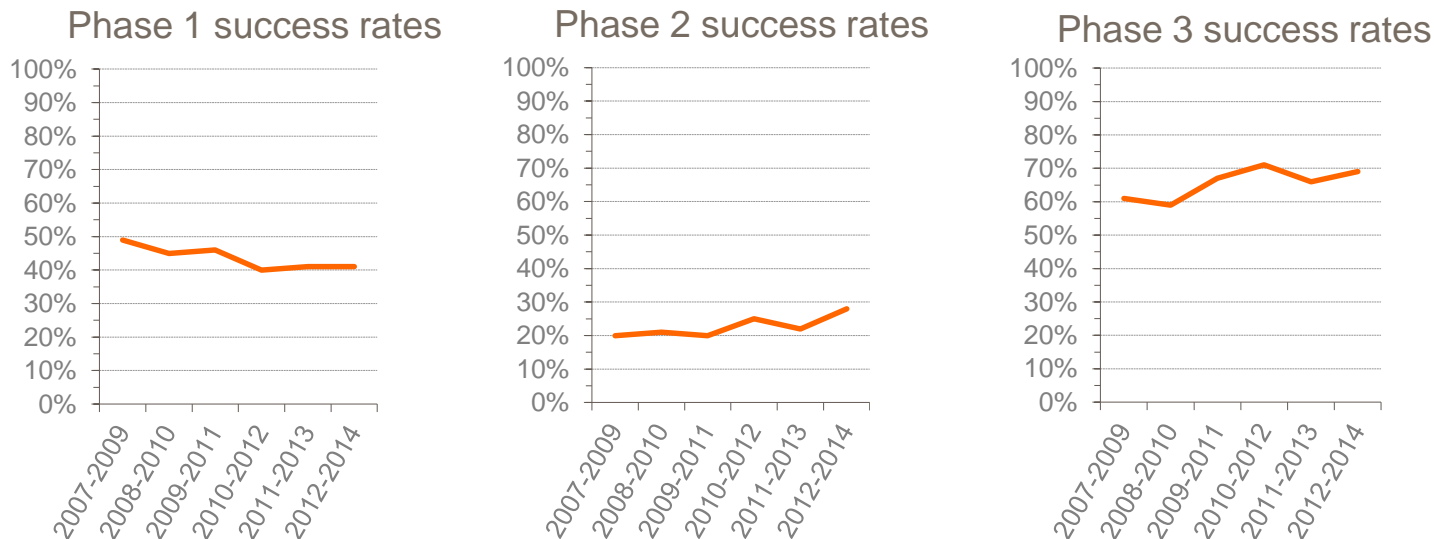
Tim Montague, Nigel Dallow,  
Tony O'Hagan, Fabio Rigat,  
Chiara Zecchin, Stefano Zamuner

- 
- Bayesian approaches to drug development are becoming more common
    - for trial design
    - for trial analysis
  - Key benefit (and challenge) is ability to use prior information
  - Prior knowledge exists on every project in some form, e.g.
    - actual data (e.g. prior clinical studies, animal data, PK data etc.)
    - scientific knowledge of the molecule/mechanism
    - clinical experience of treating patients
  - Different levels of uncertainty in predictability/relevance of the prior information
    - Often a translational gap between historical and current settings

- 
- Introduce methods for constructing informative priors from
    - historical data
    - elicitation from experts
  - Discuss methods for weighting priors in relation to data and for assessing and handling conflicts between prior and data
  - Share examples of how such priors are implemented in models to inform different stages of drug development

# Using informative priors in drug development

# Using priors to inform design of clinical development programmes



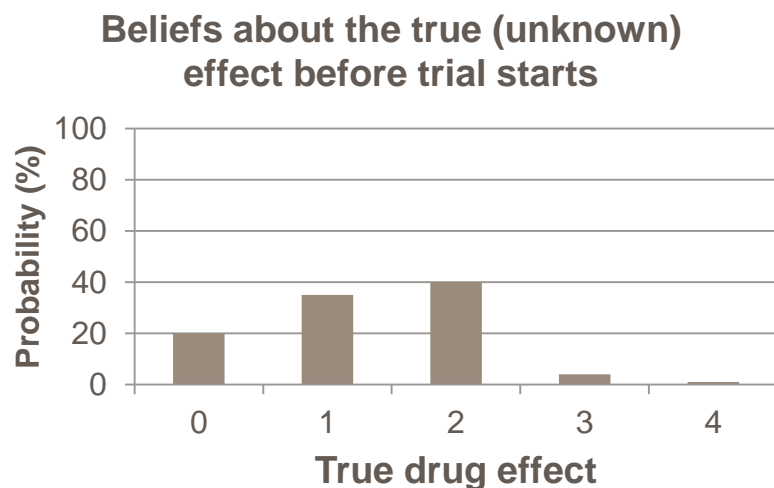
**Most late phase clinical trials  
are conducted with 90%  
power, but the success rate is  
much less than 90%**



# Power is not the Probability of Success (PoS)



- Example:
  - 400-pt trial, target superiority: > 2 point difference
  - 90% nominal power for assumed SD



True effect size	Power	Prior Belief	Power x Belief
0	2.5%	20%	0.5%
1	36%	35%	12.5%
2	90%	40%	36%
3	99.8%	4%	4%
4	99.9%	1%	1%
PoS			54%

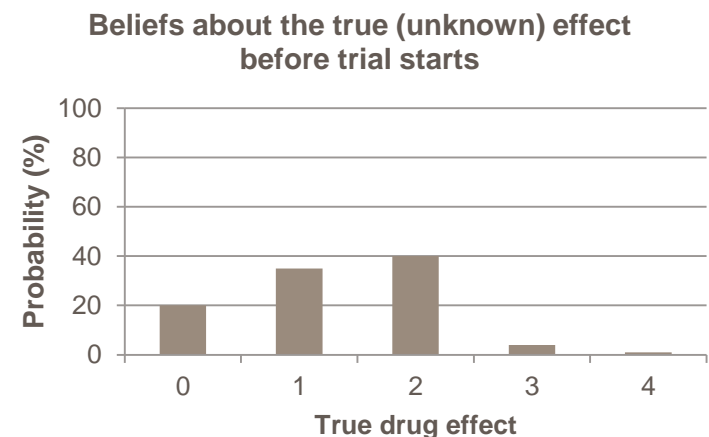
- PoS for a significant p-value outcome is known as Assurance\*
- Concept of Assurance can be extended to include criteria for magnitude of clinical effect
- Can also be extended to **Probability of Pharmacological Success** (see later)

\*O'Hagan et al (2005). "Assurance in clinical trial design," *Pharmaceutical Statistics*

# Assurance PoS supports trial design and portfolio decision-making



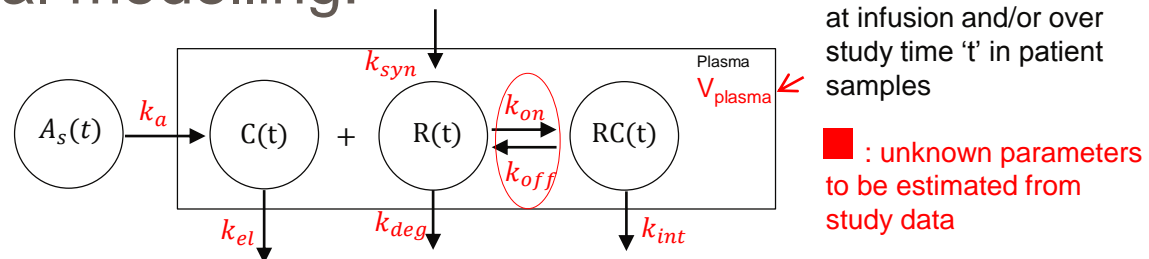
- Low PoS? ---► Consider futility interims
- Significant constraints on sample size? ---► Evaluate relationship between Trial Sample Size and PoS
- Portfolio decision-making
  - If this trial is a Go ---► what is the PoS for the next study?
  - Interim analysis to trigger spend on enabling activities ---► what is the probability of incurring the spend
- **Elements necessary to estimate PoS:**
  - Trial design
  - Definition of success
  - **Prior distribution** representing current beliefs about primary endpoint



Several opportunities:

- Increasing number of trials take place in “small populations”
  - Rare diseases; paediatrics; sub-groups; difficult-to-recruit populations..
- Designing trials to meet conventional evidentiary standards may not be feasible
  - Balance is needed between what is necessary and what is possible
  - Can using external data/prior information help?

– Pharmacological modelling:



- Estimation of model parameters often unstable when sample size small
- Can prior information on model parameters help?



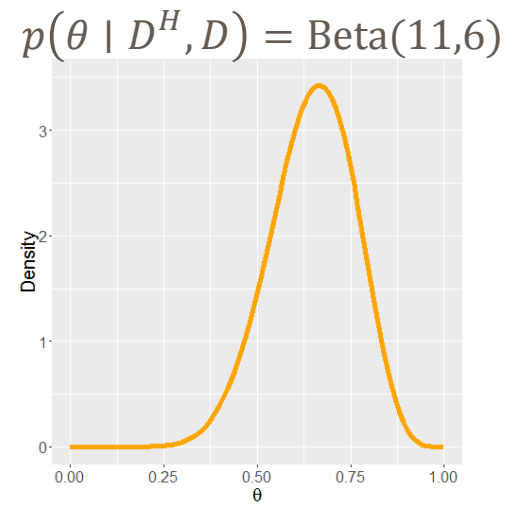
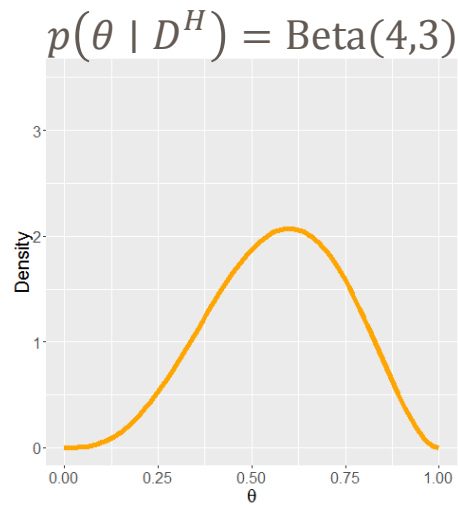
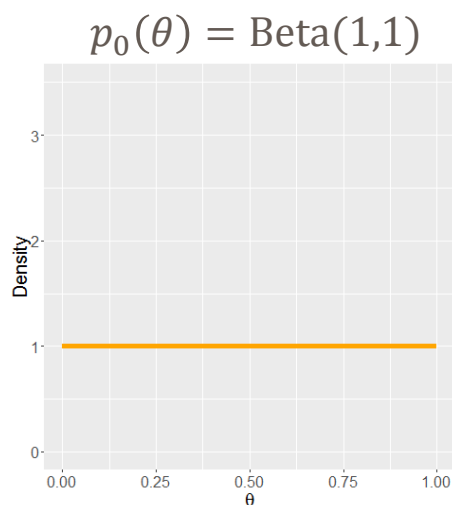
# Constructing priors from historical data

# Constructing priors from historical data



Bayesian mantra: today's posterior = tomorrow's prior

$D^H = \{3 \text{ responders in } 5 \text{ subjects}\}$      $D = \{7 \text{ responders in } 10 \text{ subjects}\}$



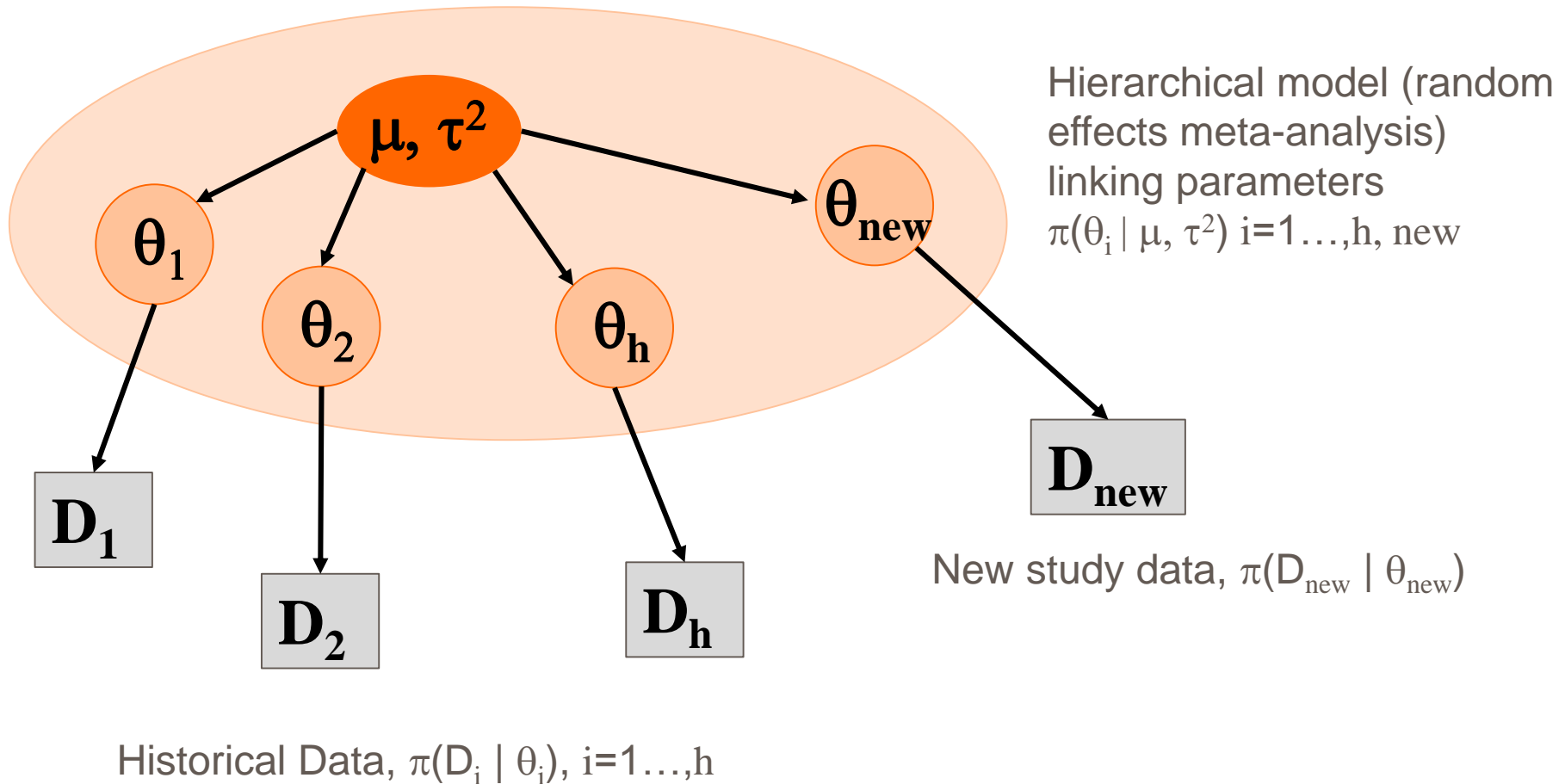
$D^H + D = \{10 \text{ responders in } 15 \text{ subjects}\}$

# Constructing priors from historical data

## Multiple historical studies



Meta-Analytic Predictive (MAP) Prior (Schmidli et al 2014)

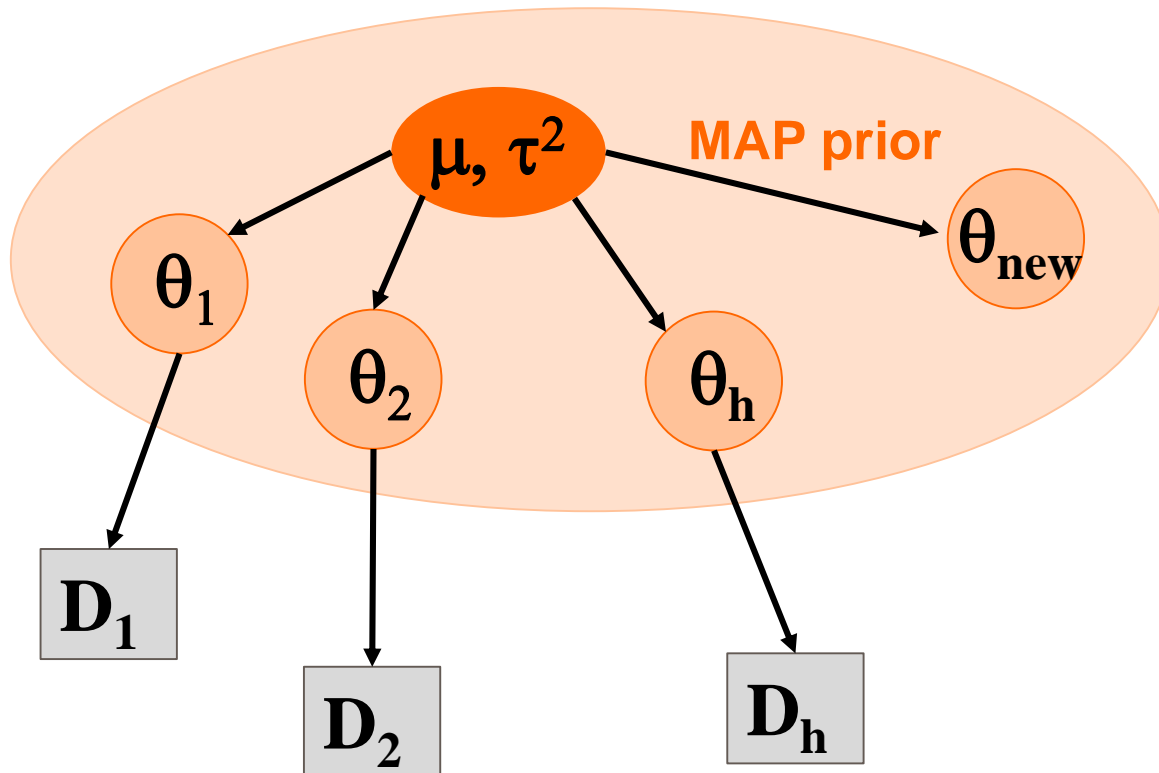


# Constructing priors from historical data

## Multiple historical studies



Meta-Analytic Predictive (MAP) Prior (Schmidli et al 2014)



Hierarchical model (random effects meta-analysis)  
linking parameters  
 $\pi(\theta_i | \mu, \tau^2) \ i=1\dots, h, \text{ new}$

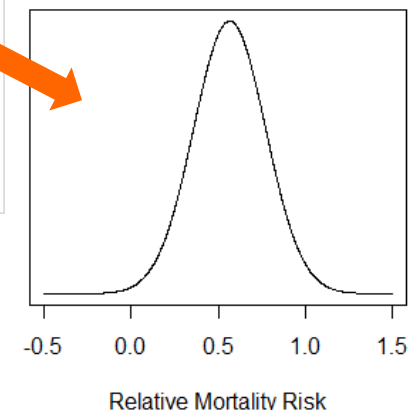
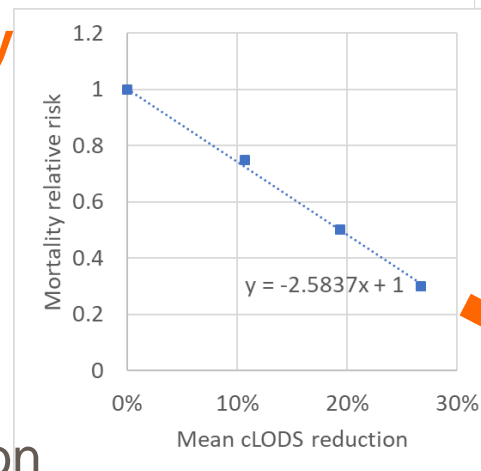
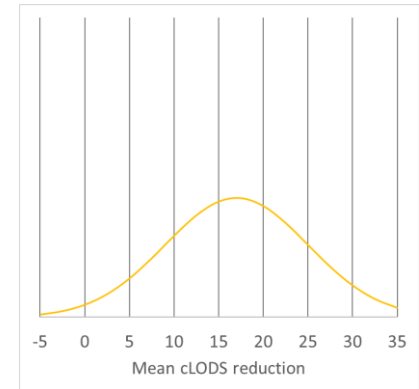
Historical Data,  $\pi(D_i | \theta_i), \ i=1\dots, h$

# Constructing priors from historical data

## More complex model-informed priors



- Phase II endpoint = **measure of organ dysfunction, LOD**
- Phase II data used to obtain posterior for  $\Delta$  in LOD score
- **Registration endpoint = mortality**
- Developed model of relationship between  $\Delta$  in LOD score and mortality RR based on published data
- Used model to extrapolate prior on mortality given prior on LOD



# Eliciting priors from experts

# Why Prior Elicitation?



- In principle, historical data can provide a prior probability distribution for a treatment effect or model parameter
- But
  - such data rarely completely match the precise parameter definition
  - expert judgement is often needed to bridge those gaps
- If historical evidence is substantial, this may be a simple judgement not requiring formal expert elicitation
  - e.g. an ad hoc increase in variance
- Alternatively, formal expert elicitation methods can be used to “translate” existing evidence and scientific knowledge to a new setting
- In 2014, GSK implemented a formal expert elicitation process to translate prior data and expert knowledge into quantitative prior distributions to support trial design and internal decision making\*
  - 60+ elicitations to date

# Subjective opinions about objective science?

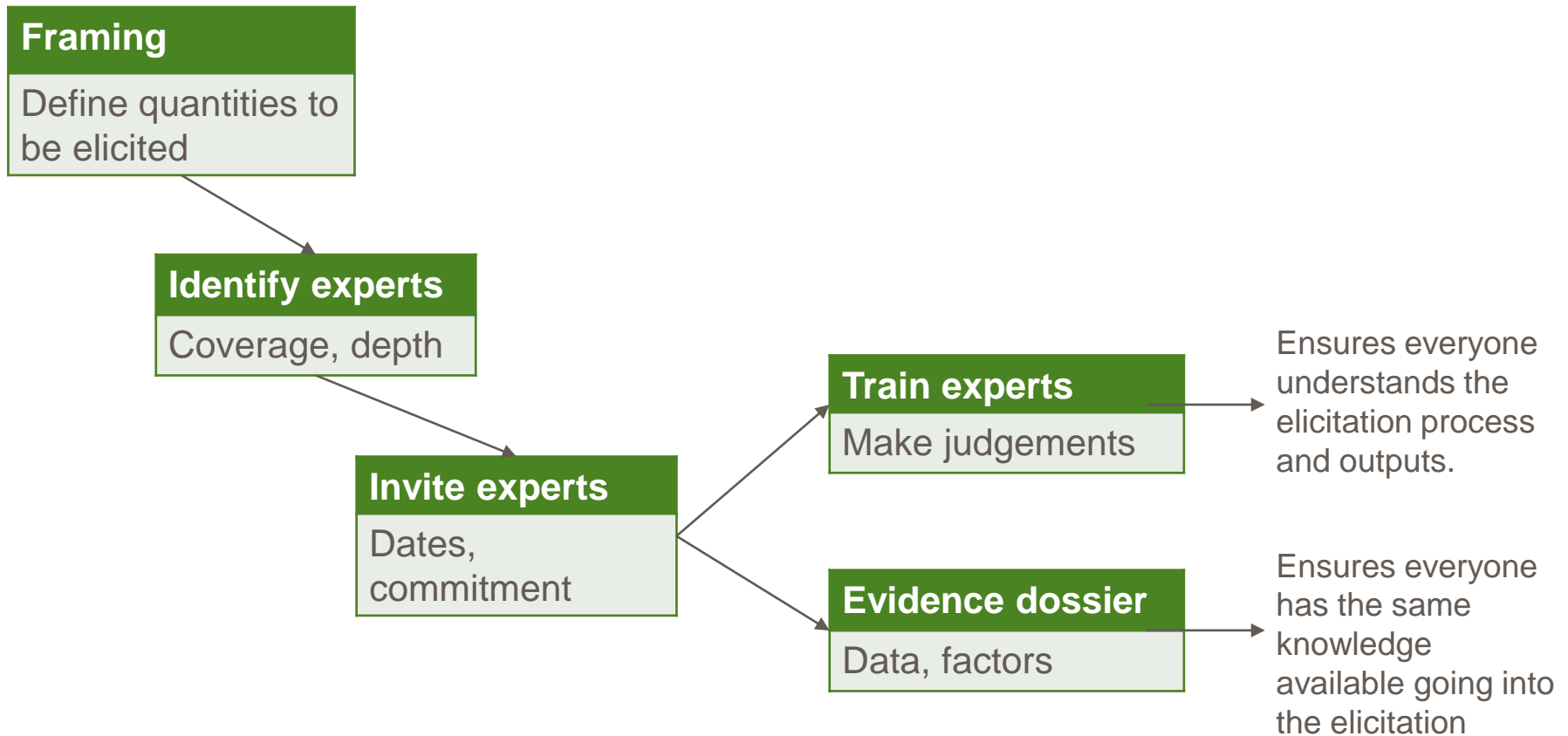


- Your belief about a drug effect or PK parameter is **subjective** (it's different to mine), whereas the attributes of the drug are **objective** (in truth, they are independent of our views)
- But scientists still have to make **judgements**
  - Take meta-analysis, for example....
- Important point is that subjectivity is minimized through basing scientific judgements on **defensible evidence** and **transparent reasoning**
- Prior elicitation process is designed to facilitate transparency and accountability and to enable any subjectivity to be **open to inspection and critique**
- Important to use the best judges/experts





# Prior Elicitation: A rigorous process – preparation\*



# Prior Elicitation: A rigorous process – preparation\*



**Decision problem:** Phase III planning for fixed dose combination (FDC) of two approved products.

**Relevant Data:** A positive Phase II study and a wealth of data and knowledge on individual components and other FDCs.

**Unknown:** How results from the phase II study (challenge model) translate to Phase III clinical study (real world situation).

**Elicitation aim:** to elicit true mean treatment difference between FDC and monotherapy

## Framing

Define quantities to be elicited

## Identify experts

Coverage, depth

## Invite experts

Dates, commitment

## Evidence dossier

Data, factors

Ensures everyone has the same knowledge available going into the elicitation

# Prior Elicitation: A rigorous process – preparation\*



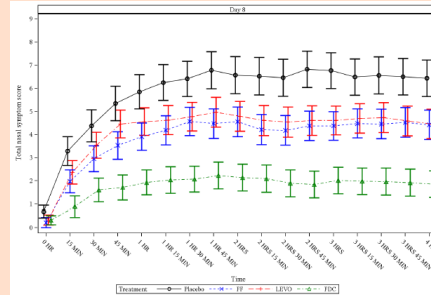
Fr  
De  
be

Clinical Review  
Jennifer Rodriguez Pippins, MD, MPH  
NDA 202-236  
Dymista (azelastine hydrochloride 0.1% / fluticasone propionate 0.037% nasal spray)

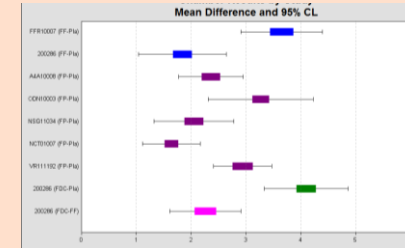
Table 9. rTNSS (AM and PM Combined), Change from Baseline over the 14-Day Treatment Period, ITT Population, Analysis using Imputed Scores

Treatment Arm	N	Baseline LS Mean (SD)	Change from baseline		Treatment Difference from MP29-02	
			LS Mean (SD)	LS Mean (SD)	LS Mean	95% CI P-value
<b>Trial MP-4002</b>						
MP29-02	207	18.27 (3.038)	-5.01 (3.225)	-	-	-
Azelastine hydrochloride	208	18.28 (3.038)	-4.23 (4.629)	-1.38	-2.22-0.54	0.001
Fluticasone propionate	207	18.22 (3.233)	-4.71 (4.678)	-0.9	-1.74-0.07	0.034
Vehicle	209	18.61 (3.175)	-2.92 (3.923)	-2.69	-3.48-1.91	<0.001
<b>Trial MP-4004</b>						
MP29-02	193	18.28 (3.341)	-5.54 (5.183)	-	-	-
Azelastine hydrochloride	193	18.64 (3.147)	-4.54 (4.621)	-1.60	-1.90-0.69	0.002
Fluticasone propionate	188	18.64 (2.918)	-4.55 (5.148)	-0.99	-1.91-0.05	0.038
Vehicle	199	18.24 (3.067)	-3.03 (3.933)	-2.51	-3.33-1.67	<0.001
<b>Trial MP-4006</b>						
MP29-02	448	19.34 (2.431)	-5.53 (5.180)	-	-	-
Azelastine hydrochloride	443	19.47 (2.520)	-4.62 (4.762)	-0.71	-1.30-0.13	0.016
Fluticasone propionate	450	19.41 (2.378)	-4.89 (4.665)	-0.64	-1.22-0.07	0.029
Vehicle	448	19.44 (2.383)	-3.40 (4.542)	-2.13	-2.70-1.57	<0.001

Source: Section 3.2, pg 17 (Table 9); Section 3.3, pg 31 (Table 10); Section 3.3, pg 35 (Table 11)



GSK Historical Data Sets

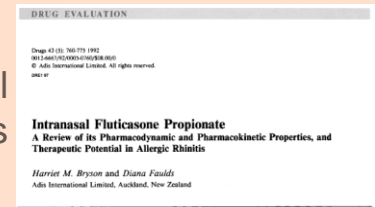


Data summaries from GSK reports and published competitor studies

Regulatory Reviews

**Evidence dossier**

Journal Articles

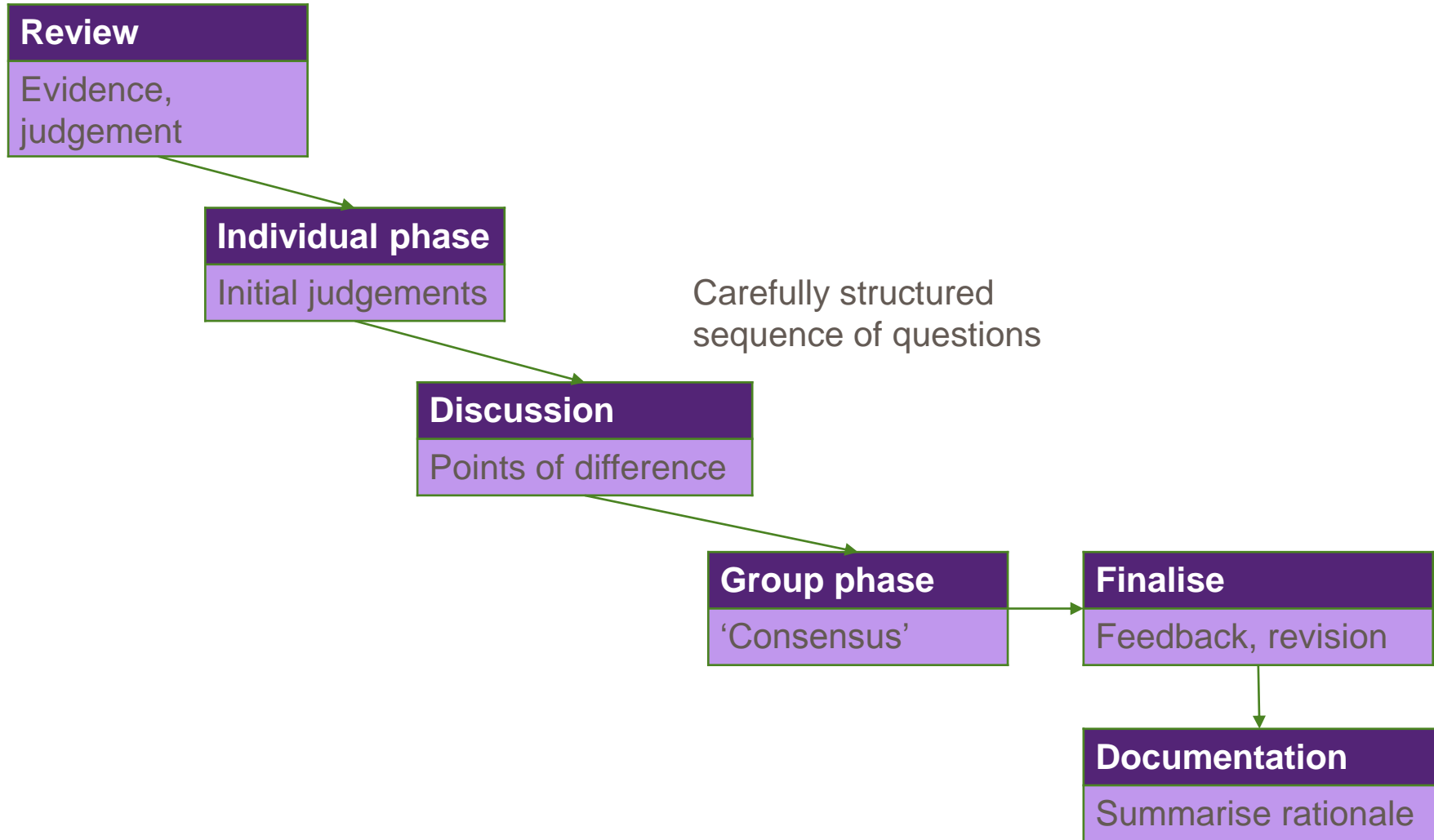


Dates, commitment

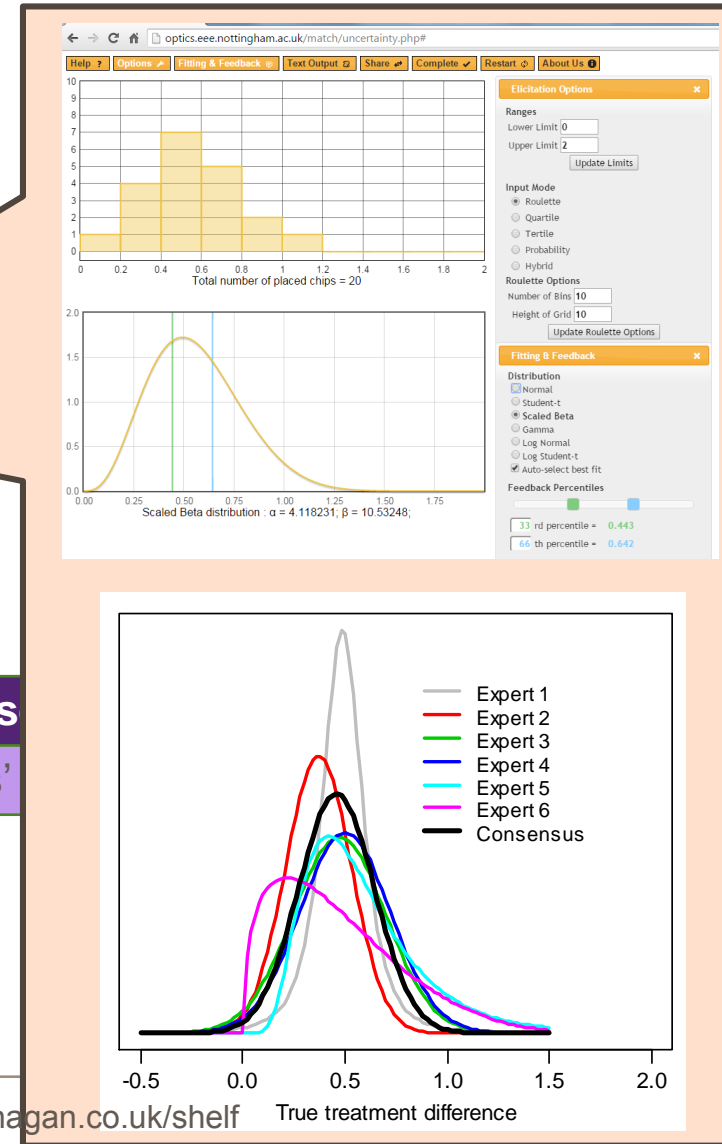
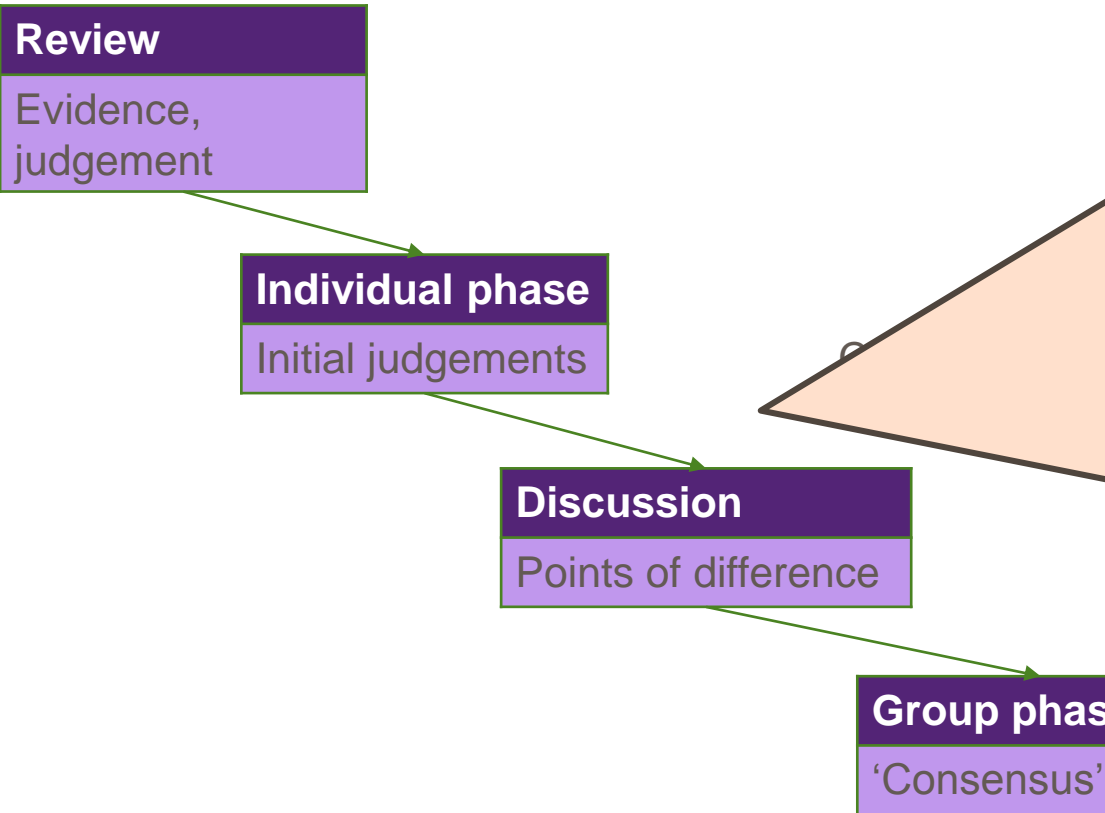
**Evidence dossier**  
Data, factors

Ensures everyone has the same knowledge available going into the elicitation

# Prior Elicitation: A rigorous process – workshop\*



# Prior Elicitation: A rigorous process – workshop\*



# Prior Elicitation: A rigorous process – workshop\*



## Review

Evidence,  
judgement

Ind

Init

The Sheffield Elicitation Framework

SHELF v2.0

### ELICITATION RECORD – Part 1 – Context

Ellicitation title	
Session	
Date	
Part 1 start time	
Attendance and roles	
Purpose of elicitation	
Orientation and training	
Participants' expertise	
Declarations of interests	
Strengths & weaknesses	
Evidence	
Structuring	
Definitions	

Structured  
questions

The Sheffield Elicitation Framework

SHELF v2.0

### ELICITATION RECORD – Part 2 – Distribution

#### Roulette Method

Definition	Define quantity to be elicited (X)
Evidence	Review of evidence relating to X
Plausible range	Record the range of plausible values for X elicited from each expert
Chips in bins	Each expert asked to create histogram representing his/her beliefs about X. Record histograms/chip placements here.
Fitting	Record distributions fitted to each of the experts' histograms
Group elicitation	Experts invited to discuss their different distributions and share knowledge and reasoning about differences. Record key points of this discussion, together with the consensus histogram.
Fitting and feedback	Record process of fitting, feedback and revision of the group consensus judgement.
Chosen distribution	Record and show the final fitted distribution
Discussion	Record experts' reactions to the process and to the final fitted distribution, plus any difficulties that arose during the elicitation.

## Finalise

Feedback, revision

## Documentation

Summarise rationale

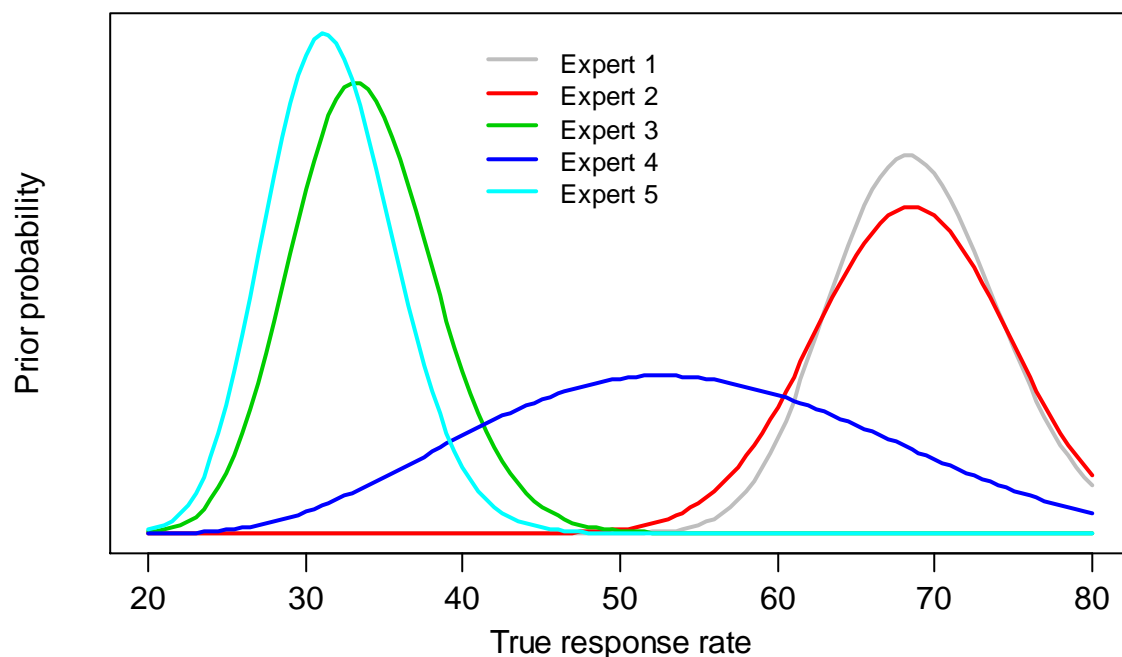
- Judgements elicited from several experts to cover range of scientific opinion and expertise
- But (ideally) a single prior is needed for decision-making
  - SHELF protocol uses behavioural aggregation for consensus prior
  - Alternative is mathematical aggregation (weighted average)

Benefits of Behavioural Aggregation	Risks of Behavioural Aggregation
<ul style="list-style-type: none"><li>• Encourages sharing knowledge</li><li>• Avoids using an arbitrary mathematical rule</li><li>• Consensus prior intended to represent view of a Rational Impartial Observer</li></ul>	<ul style="list-style-type: none"><li>• Difficulty of managing the experts</li><li>• Difficulty of ensuring all opinions are treated on their merits</li><li>• Experts required to ‘put themselves in someone else’s shoes’</li></ul>

# Achieving an aggregate prior: Example



- Setting:
  - Planning for PIII trial of existing drug in new indication with unmet medical need
  - Heterogeneous patient population
  - No well-established disease-severity index
- Elicitation of response rate on Standard of Care (SOC)



## Individual expert priors

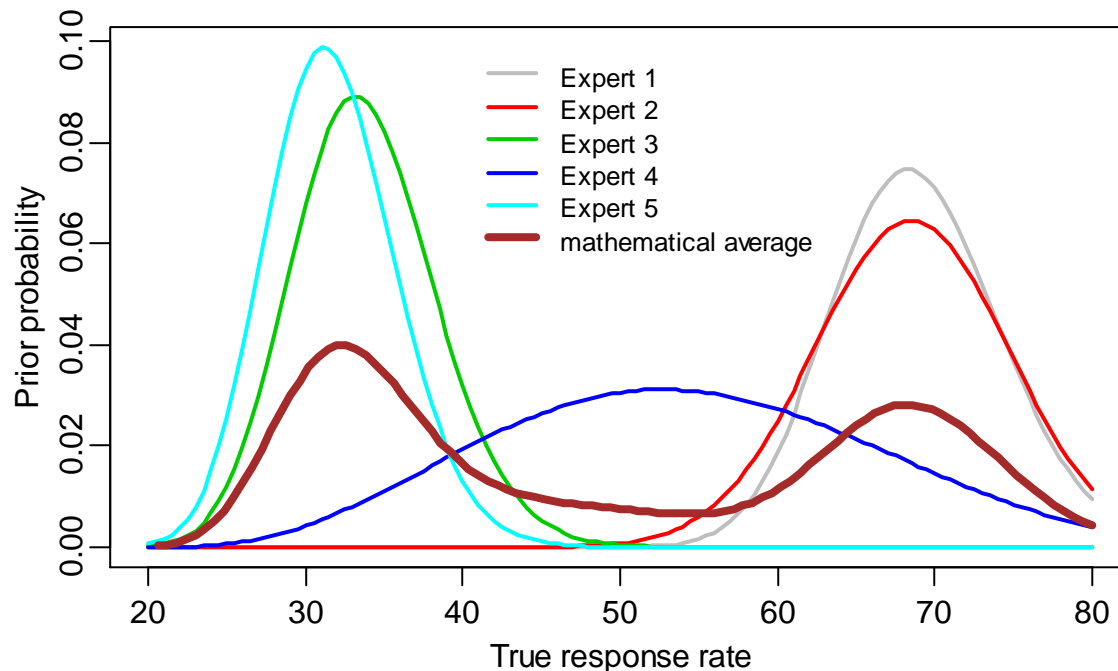
- Expert 1 & 2 based on clinical experience (primary care)
- Expert 3 & 5 based on literature and tertiary care experience
- Expert 4 based on literature allowing for heterogeneity



# Achieving an aggregate prior: Example



- Setting:
  - Planning for PIII trial of existing drug in new indication with unmet medical need
  - Heterogeneous patient population
  - No well-established disease-severity index
- Elicitation of response rate on Standard of Care (SOC)

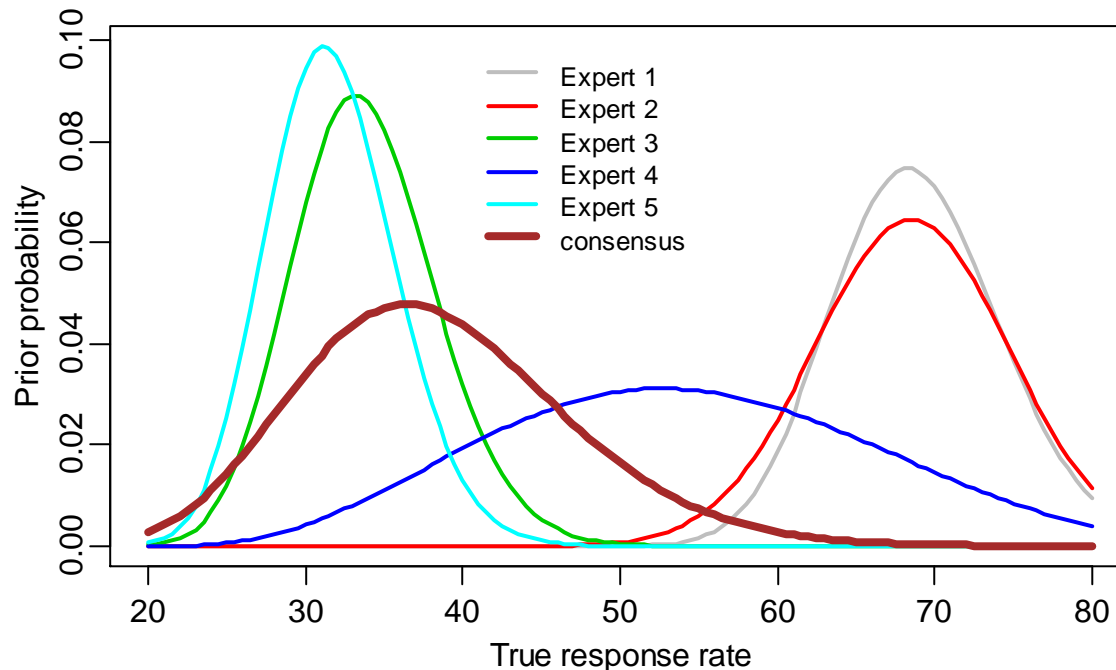


**Mathematical average**

# Achieving an aggregate prior: Example



- Setting:
  - Planning for PIII trial of existing drug in new indication with unmet medical need
  - Heterogeneous patient population
  - No well-established disease-severity index
- Elicitation of response rate on Standard of Care (SOC)



## Consensus

- Reflects discussion around patient heterogeneity and expectation that patient population for trial likely to be more severe

# Feedback from experts

---



“It is the process itself which is most valuable for the team, uncovering heterogeneity among expert views in a totally transparent way”

“The negotiation among experts and the exchange of rationale for probabilities was probably the most valuable part”

“Allowed internal team to have a clear and honest discussion with external experts without either side trying to say what other side wants to hear”

“It challenges your views - often entrenched and biased.”

# **Weighting prior information and assessing and dealing with prior-data conflict**

# Combining prior information and new data: Standard Bayesian updating

- Historical data used to generate predictive prior distribution for response rate in new trial

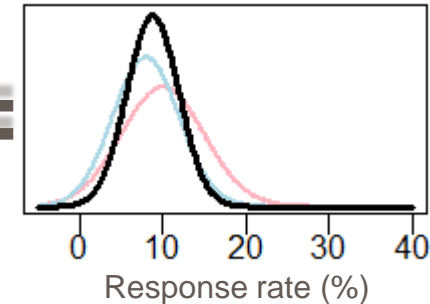
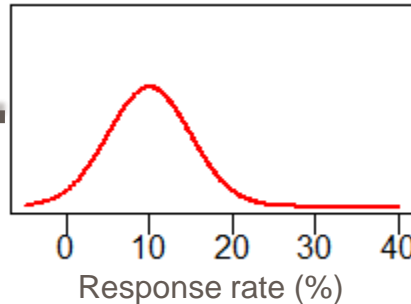
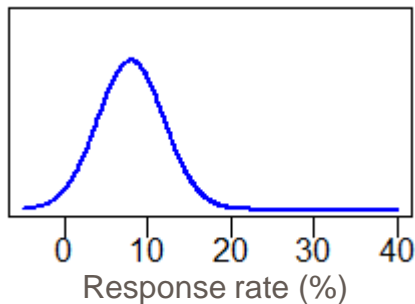
Predictive distribution for  
what we believe about future  
responses based on the  
historical studies (“prior”)

What we see in  
the new study  
(“sampling distribution”)

“Posterior” distribution:  
weighted average of prior  
and new trial data

## Scenario 1

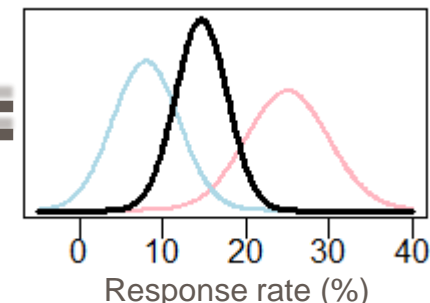
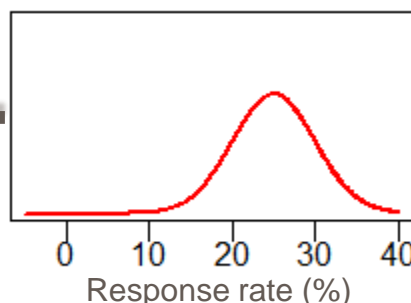
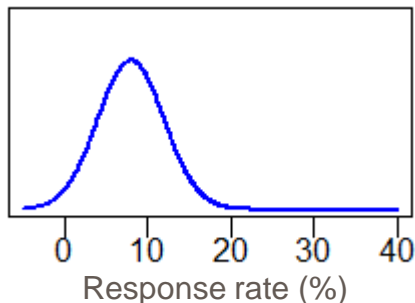
Historical  
and new  
data are  
consistent



- But, can result in potentially unrealistic estimates if historical data **conflicts** with new data

## Scenario 2

Historical  
and new  
data in  
conflict



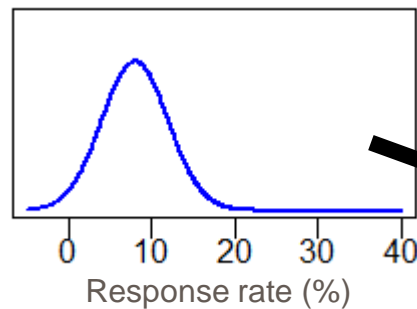
# Robust mixture priors to address prior-data conflict



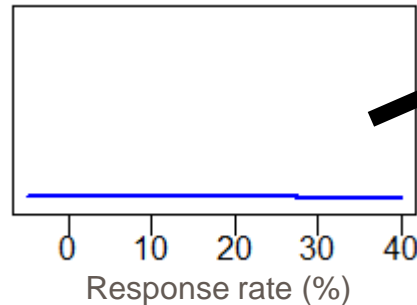
## – Cromwell's rule (after Dennis Lindley):

"I beseech you, in the bowels of Christ, think it possible that you may be mistaken"

Prior assuming historical data are relevant



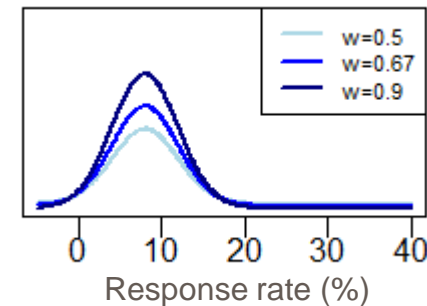
"In case we are mistaken" prior i.e. assuming historical data are not relevant



$W$

$1-W$

Robust prior = weighted mixture of these 2 priors



Schmidli et al. Biometrics (2014)

# Robust (dynamic) Bayesian models to deal with potential prior-data conflict



Mixture prior with 50% weight on historical data and 50% weight on flat prior

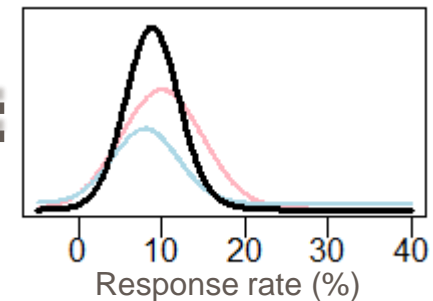
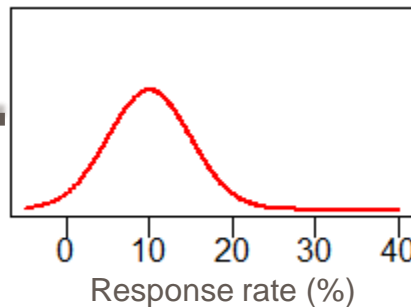
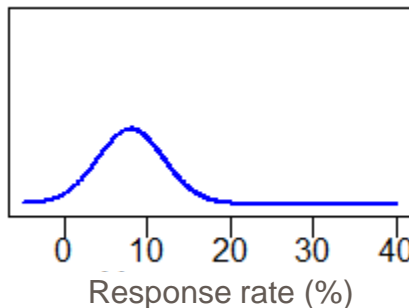
What we see in the new study ("sampling distribution")

"Posterior" distribution: weighted average of prior and new trial data



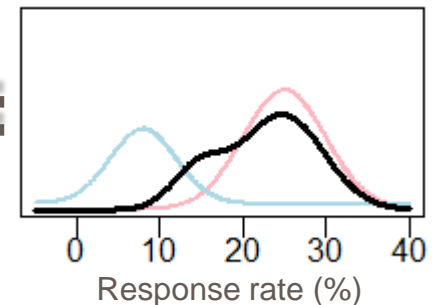
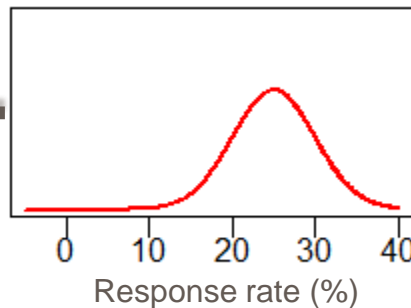
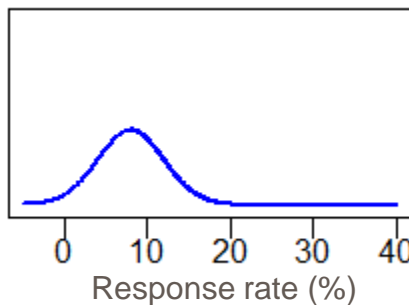
## Scenario 1

Historical and new data are consistent



## Scenario 2

Historical and new data in conflict



- Informative and vague components can be thought of as **two alternative Bayesian models** for new study:
  - M1: parameters of historical and new study are the same
  - M2: parameters of historical and new study are unrelated
- Models M1 and M2 differ by the assumed prior for the parameters of the new data likelihood:
  - $\pi(\theta|M_1)$  = informative prior based on historical data
  - $\pi(\theta|M_2)$  = vague prior
- **Weight** on each component = **prior probability for each model**
- Marginal prior for  $\theta$  is equivalent to robust mixture:

$$\pi(\theta) = \Pr(M_1) \pi(\theta|M_1) + \Pr(M_2) \pi(\theta|M_2)$$



- Given observed data  $D$  for new trial
  - Conditional posteriors for each model can be updated separately to give  $\pi(\theta|M_1, D)$  and  $\pi(\theta|M_2, D)$

- Posterior model probabilities are updated via Bayes theorem

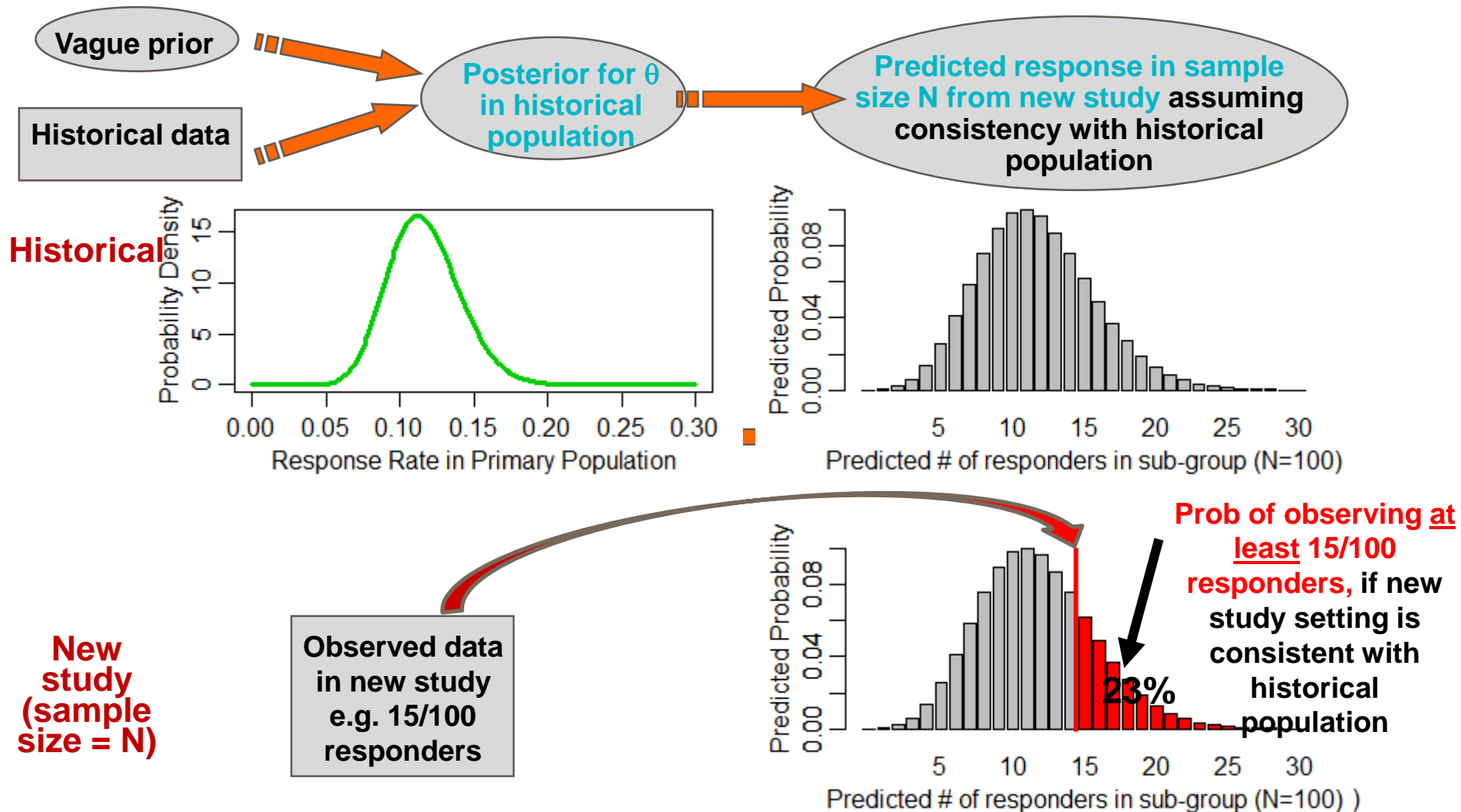
$$\pi(M_1|D) = \frac{\pi(M_1)f(D|M_1)}{\pi(M_1)f(D|M_1) + \pi(M_2)f(D|M_2)}$$

where  $f(D|M_i)$  is the **marginal likelihood** of the data,  $D$ , under model  $M_i$

- Either select model with highest posterior probability (**test then pool**)...
- ...or adopt model averaging approach (**robust mixture**)

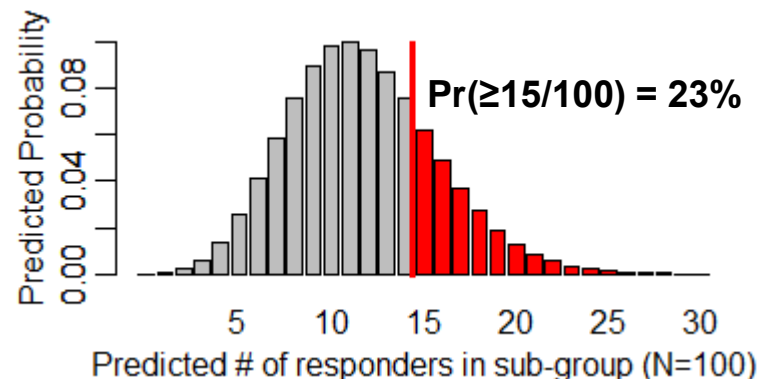
$$\pi(\theta|D) = \Pr(M_1|D) \pi(\theta|M_1, D) + \Pr(M_2|D) \pi(\theta|M_2, D)$$

# Assessing consistency between prior and data



- Predicted distribution and tail-area probability can be used as descriptive measures of consistency between sub-groups

- Does the observed treatment effect look consistent with what is predicted?



- Tail-area probability (Box's p-value) = predictive probability of data at least as extreme as that observed, assuming the treatment effect in the new population is consistent with that in the historical population
  - Extreme values of Box's p-value indicate evidence of conflict between prior and data

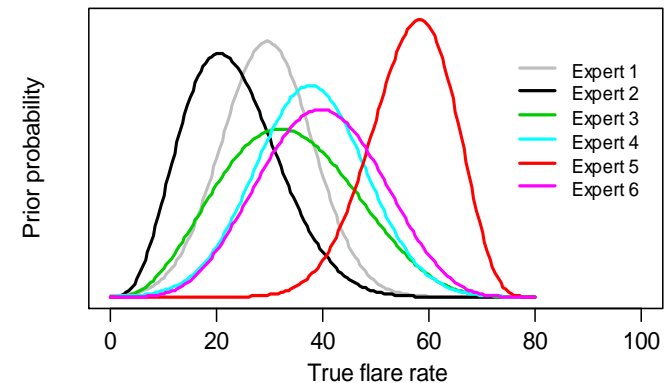
**Example: Use of elicited priors to calculate PoS (assurance) to inform clinical development planning**

# 1. PoS for Ph3 study for rare disease



## Elicited prior for placebo outcome

- Setting:
  - Planning for PIII trial in rare disease with high unmet medical need
  - Novel clinical endpoint
  - No historical data
- Elicitation of flare rate on placebo
- Individual expert priors
  - Some experts believed incl/ excl criteria would lead to **stable patients** being enrolled
  - Expert 5 (red) assumed stable patients **wouldn't be enrolled** and so a much higher flare rate on placebo
- Clear **rationale for differences** in prior beliefs which can be addressed as part of study design

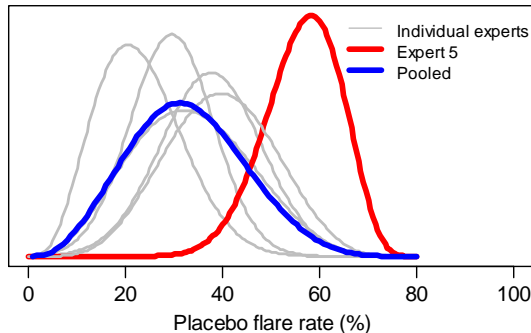


# 1. PoS for Ph3 study for rare disease

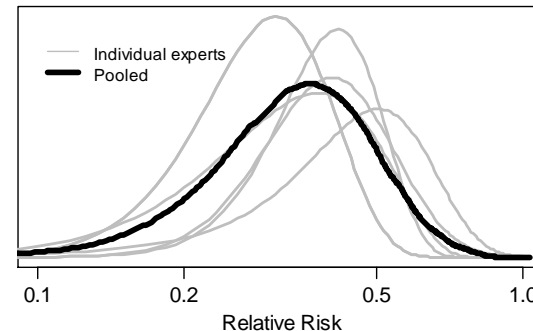


## Transparency in expert opinions and impact on study design

Expert beliefs about placebo flare rate



Expert beliefs about relative risk of flare on active vs placebo



Probability of success for planned Ph3 trial

	N=40	N=60
<b>Expert 5 placebo prior</b>	<b>78%</b>	<b>89%</b>
<b>Pooled placebo prior</b>	<b>43%</b>	<b>57%</b>

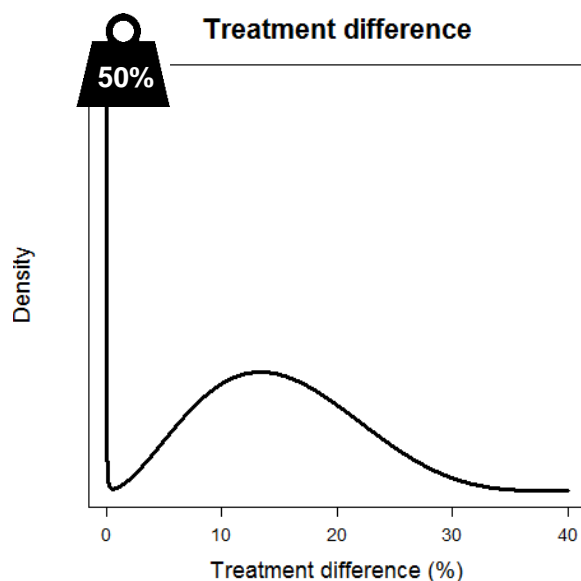
- Placebo flare rate is a key determinant of PoS
- Expert 5 had a higher expectation than other experts due to different beliefs about stability of patients recruited
- Inc/Excl criteria **modified** and **blinded sample size readjustment planned** to mitigate risk of low placebo flare rate

## 2. PoS for early phase asset with novel MOA

### Elicited prior for treatment effect



Setting – planning development of asset with novel MOA at start of Ph2



#### Prior

- Elicited from internal experts, including 3 non-project experts
- **50% weight on no difference due** to unknowns with the translation of the mechanism and low portfolio success rates in this disease area

#### Key points:

- Good biology package.
- Unknown how biology package will translate into humans.
- Potential redundancy of mechanism
- Refractory population

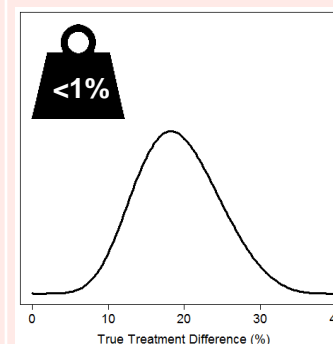
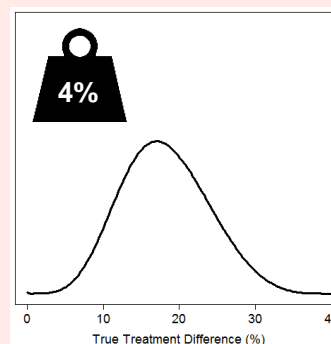
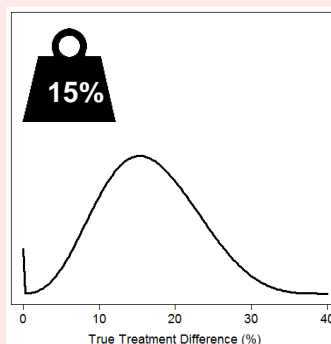
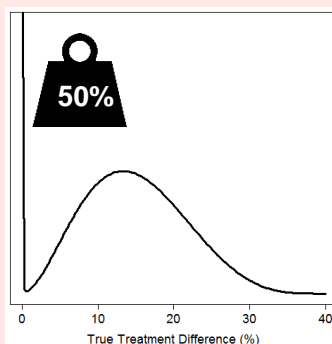
## 2. PoS for early phase asset with novel MOA



### Evolution of PoS: stage-gating the development plan

Timepoint	Start of Phase 2	Early futility (N=50)	PoC interim (N=100)	Start of phase 3
PoS conditional on...		...continue at interim	...continue at interim	...Go in Ph2
PoS	17%	37%	53%	59%*

Prior for treatment difference



where the weights  denotes the probability of a placebo-like compound



**Example: Use of robust mixture priors to extrapolate between population sub-groups**

# Retrospective analysis of adolescent efficacy borrowing from adult efficacy data

---



**Full extrapolation of efficacy from adults to paediatric subjects is proposed for a respiratory asset in partial fulfilment of the PIP/PSP**

**Appropriateness of extrapolation approach:** Similarity of disease presentation and therapeutic approach, common accepted regulatory clinical efficacy endpoints between adults and paediatrics aged 6-17

**Paediatric extrapolation strategy stepwise approach:**

- A positive benefit-risk profile is demonstrated in adults
  - Paediatric PK/PD trial shows that:
    - Adult PK is predictive of paediatric PK
    - Adult PD effect is predictive of paediatric PD effect
  - **Efficacy in adolescents included in two large Phase 3 studies is shown to be consistent with adults**
  - Safety in paediatrics is consistent with the safety profile of the overall population across a number of indications
-

# Retrospective analysis of adolescent efficacy data using informative prior based on adults



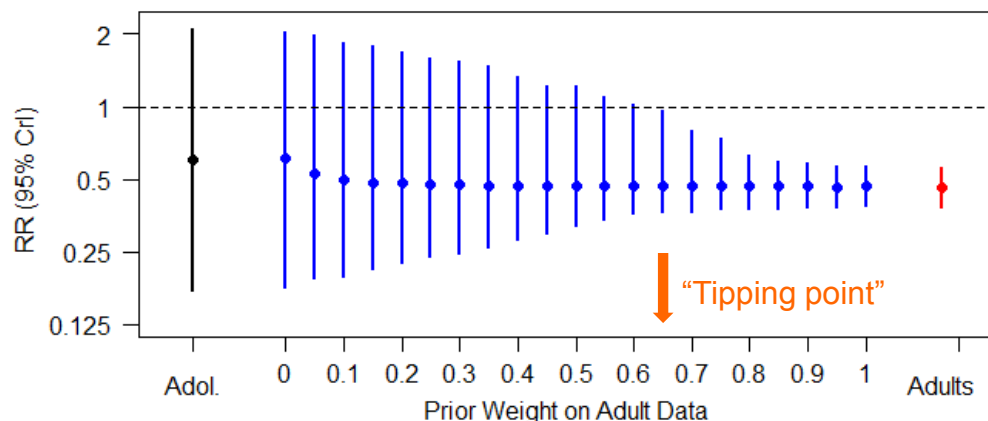
- 34 adolescent subjects were recruited in the two adult Ph III pivotal studies
- Evidence of consistency in efficacy response between adults and adolescents would increase confidence that an extrapolation approach is appropriate

	Total N	RR (95% CI)
Adults (≥18yrs)	1093	0.46 (0.38, 0.56)
Adolescents (12-17yrs)	34	0.60 (0.17, 2.10)

We “stress tested” the “similarity” assumption as follows:

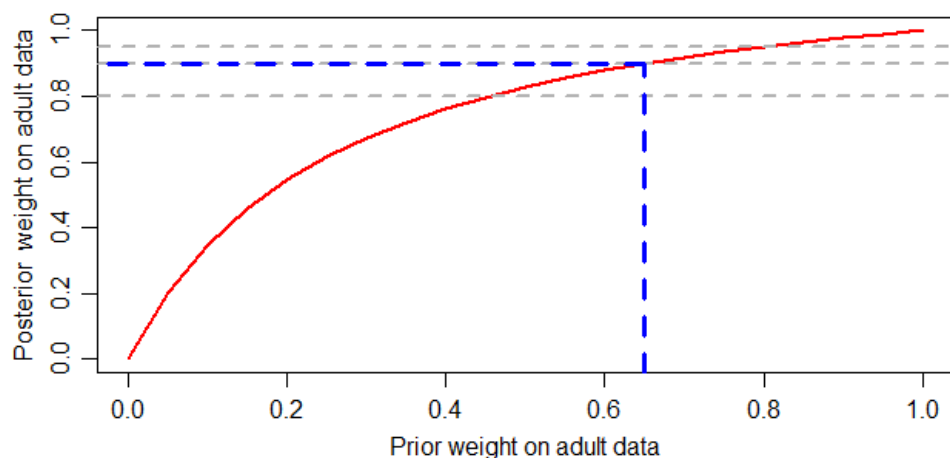
- Bayesian analysis of adolescent subset using adult subset as informative component of robust mixture prior
- Varied prior weight on adult component, to identify minimum weight needed to achieve high (>97.5%) posterior probability of efficacy in adolescents

# Bayesian “tipping point” analysis to stress-test the extrapolation approach



If it is reasonable to assume prior odds of at least 2 to 1 (~65%) in favour of “similarity” assumption

➡ Can conclude there is **substantial evidence\*** of efficacy in adolescents



If it is reasonable to have at least 65% prior confidence in the “similarity” assumption

➡ Evidence from adolescent data increases our confidence in this “similarity” assumption to >90%

Provides quantification of strength of evidence supporting the proposed extrapolation strategy in a situation where adolescent data alone is too limited to infer efficacy

\*High (>97.5%) probability of efficacy utilizing evidence of all reliable sources of efficacy data

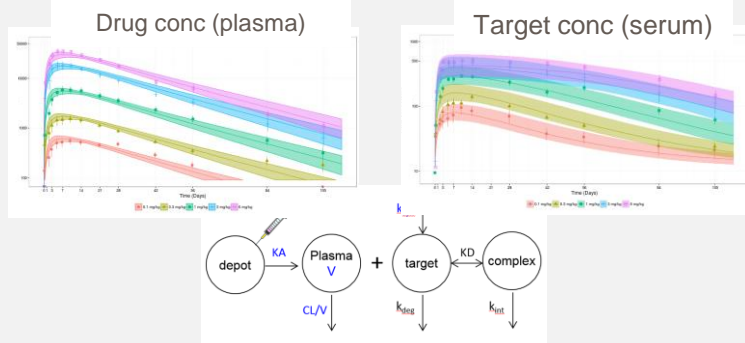
# **Opportunities for using informative priors in clinical pharmacology**

# 1. Probability of Pharmacological Success

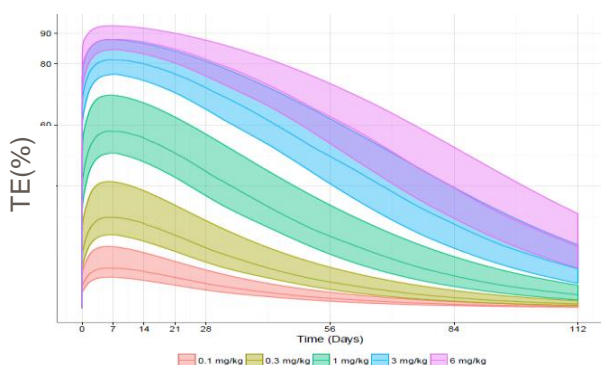


Illustrative example: Proof of pharmacology based on target engagement

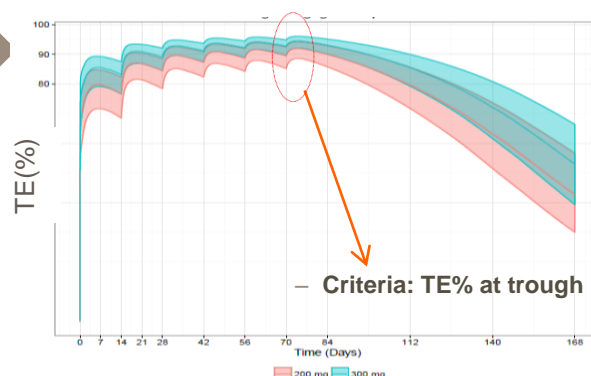
FTIH study: 5 arms, samples for PK and target in plasma (rich) and site of action (sparse)



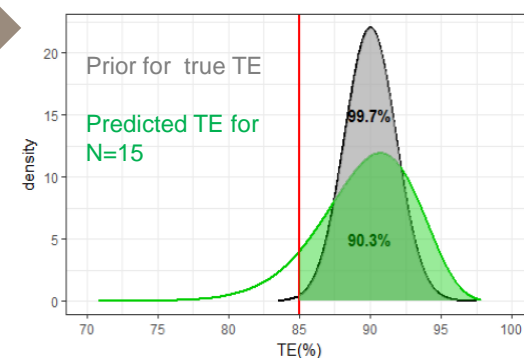
Model-informed analysis to derive distribution of PD effect



Predicted Target Engagement after repeat dosing (future trials)



Prior distribution for %TE and clinical trial simulation to predict observed %TE for given trial design (dose, N)



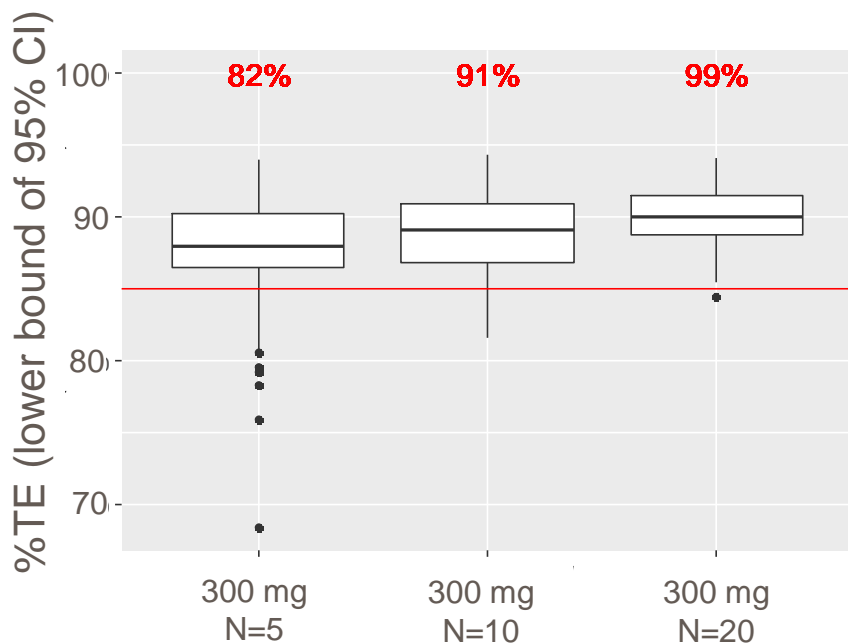
**TE is not an observation but a model derived estimate!**

# 1. Probability of Pharmacological Success

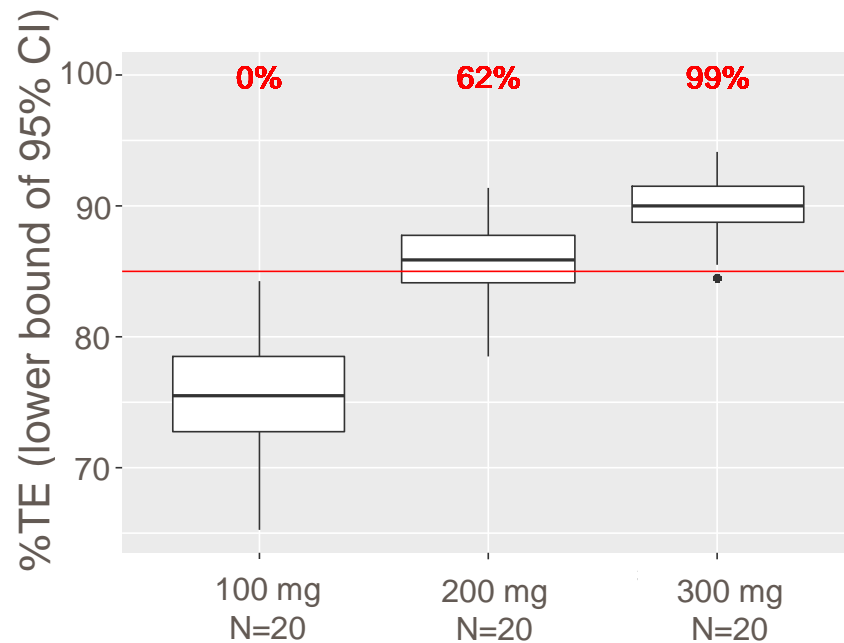


Assurance for planned POM study with different doses and sample sizes

Successful trial: 95% probability  $\%TE_{\text{trough}} > 85\%$  at steady state



Fixed dose / Different sample size



Fixed sample size / Different doses

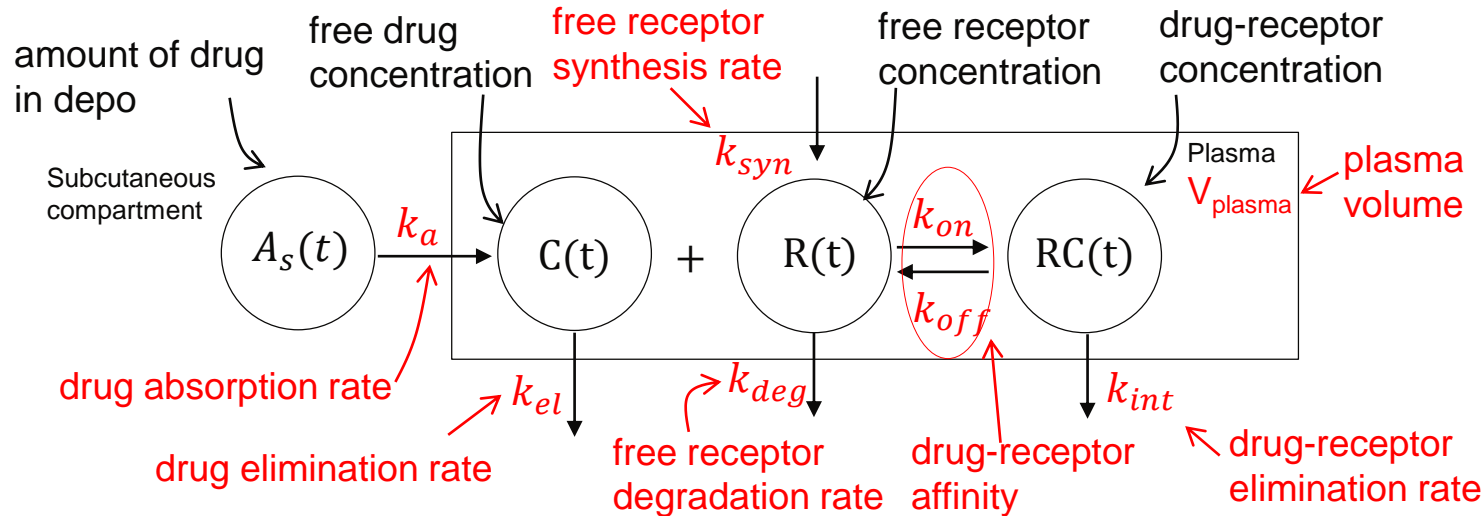
Boxplots show distribution of simulated trial outcomes (lower bound of 95% CI for %TE) under each scenario

## 2. Informative priors in the analysis of pharmacological models



### An illustrative “semi-physiological” ODE model

#### Model concept



■: variables measured at infusion and/or over study time ‘t’ in patient samples

■: unknown parameters to be estimated from study data

#### Data

$$z_{\text{drug}}(t) = \hat{C}(t) + \widehat{RC}(t) + e_{\text{drug}}(t)$$

$$z_{\text{target}}(t) = \hat{R}(t) + \widehat{RC}(t) + e_{\text{target}}(t)$$

$[e_{\text{drug}}(t), e_{\text{target}}(t)] \sim \text{error probability distribution}$

#### Model ODEs

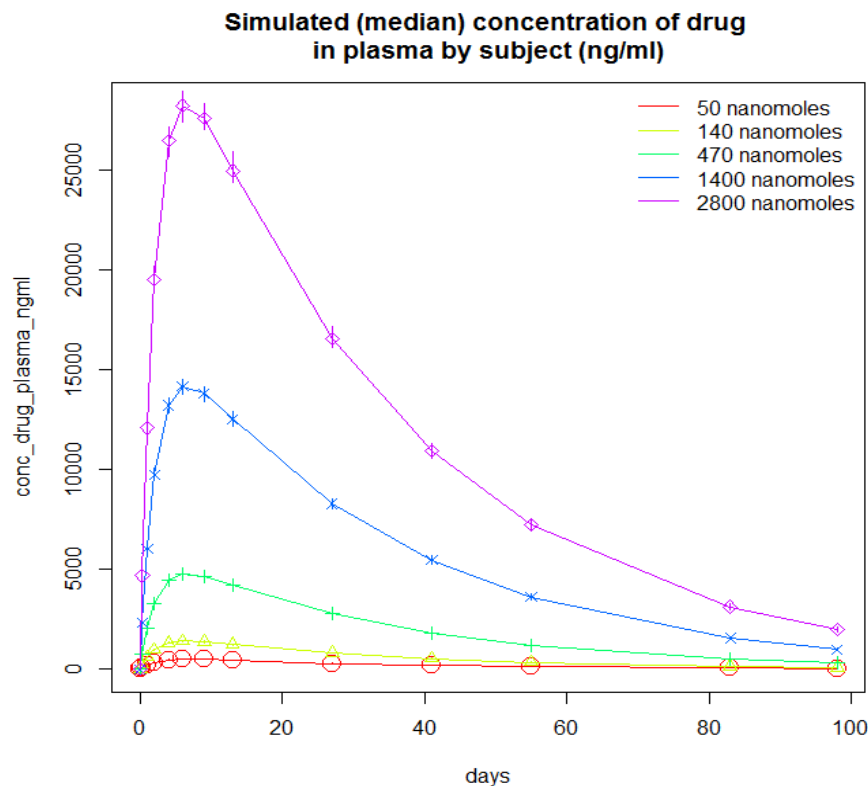
$$\begin{aligned} \frac{dA_s(t)}{dt} &= -k_a \cdot A_s(t) & A_s(0) &= \text{Dose} \\ \frac{dC(t)}{dt} &= \frac{k_a \cdot A_s(t)}{V_{\text{plasma}}} - k_{el} \cdot C(t) - k_{on} \cdot C(t) \cdot R(t) + k_{off} \cdot RC & C(0) &= 0 \\ \frac{dR(t)}{dt} &= k_{syn} - k_{deg} \cdot R(t) - k_{on} \cdot C(t) \cdot R(t) + k_{off} \cdot RC & R(0) &= R_0 = R_{ss} \\ \frac{dRC(t)}{dt} &= k_{on} \cdot C(t) \cdot R(t) - k_{int} \cdot RC(t) - k_{off} \cdot RC(t) & RC(0) &= 0 \end{aligned}$$



## 2. Informative priors in the analysis of pharmacological models



### Data simulation from the ODE model



- The ODE model was implemented in Winbugs and R for data simulation and for inference
- Plot shows the average and 95%CIs of simulated\* drug concentrations in plasma for five subjects exposed to different doses over 13 time points

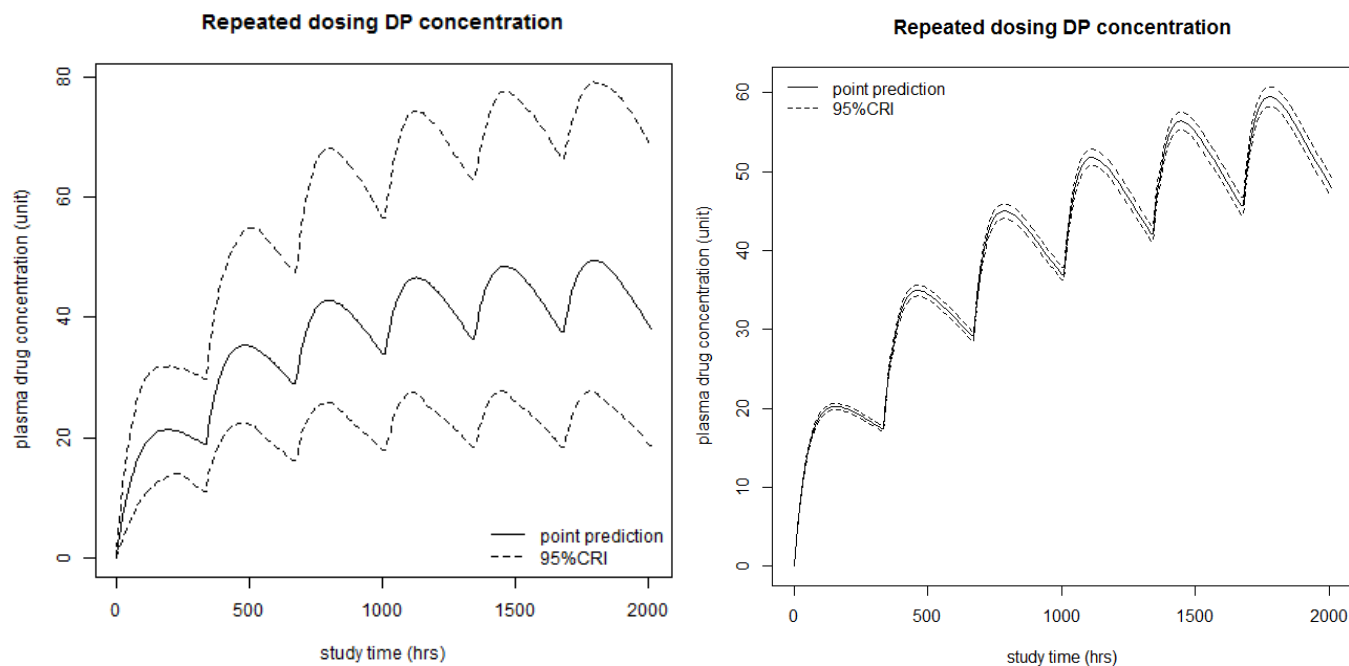
\*Strongly informative priors were used for all model parameters for this (illustrative) simulation

## 2. Informative priors in the analysis of pharmacological models



### Estimation of ODE parameters

- Plots show the average and 95%CI plasma drug concentration over 5 injections predicted from simulated single dose data using **low (left) or high (right) prior precision**.



- Goal is to include real prior information on model parameters (where available), to improve estimation and prediction when simulating from and fitting ODE models to sparse, noisy real data

**Thank you for listening**