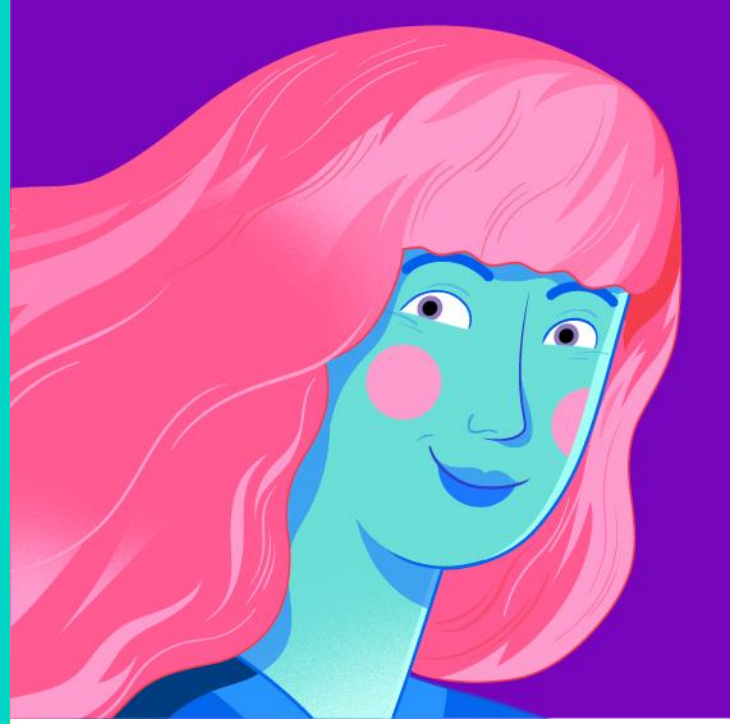
Using computational models to estimate points of departure from in vitro assay data



08/05/2024

Unilever

Learning objectives

- Understand what Next Generation Risk Assessment (NGRA) is, and how different computational models are used in NGRA to analyse data, make predictions and help make safety decisions.
- Introduction to how models are used to estimate points of departure (PODs) from *in vitro* concentration response data.
- Develop an understanding some of the challenges involved in inferring PODs and what approaches can be used to address them.

Unilever

About me

- Degree in Mathematics from the University of Edinburgh
- PhD in Applied Mathematics from the University of Nottingham
- Postdocs in Germany at the University of Freiburg and the University of Heidelberg
- Joined Unilever in 2014, hired as a mathematical modeller
- Science leader in Computational Toxicology

What is Next Generation Risk Assessment?

An exposure-led, hypothesis driven risk assessment approach that incorporates one or more NAMs to ensure that chemical exposures do not cause harm to consumers

Dent *et al* ., (2018) *Comp Tox* 7:20-26

Unilever

Principles of NGRA from ICCR

4 Main overriding principles:

- The overall goal is a human safety risk assessment
- The assessment is exposure led
- The assessment is hypothesis driven
- The assessment is designed to prevent harm

3 Principles describe how a NGRA should be conducted:

- Following an appropriate appraisal of existing information
- Using a tiered and iterative approach
- Using robust and relevant methods and strategies

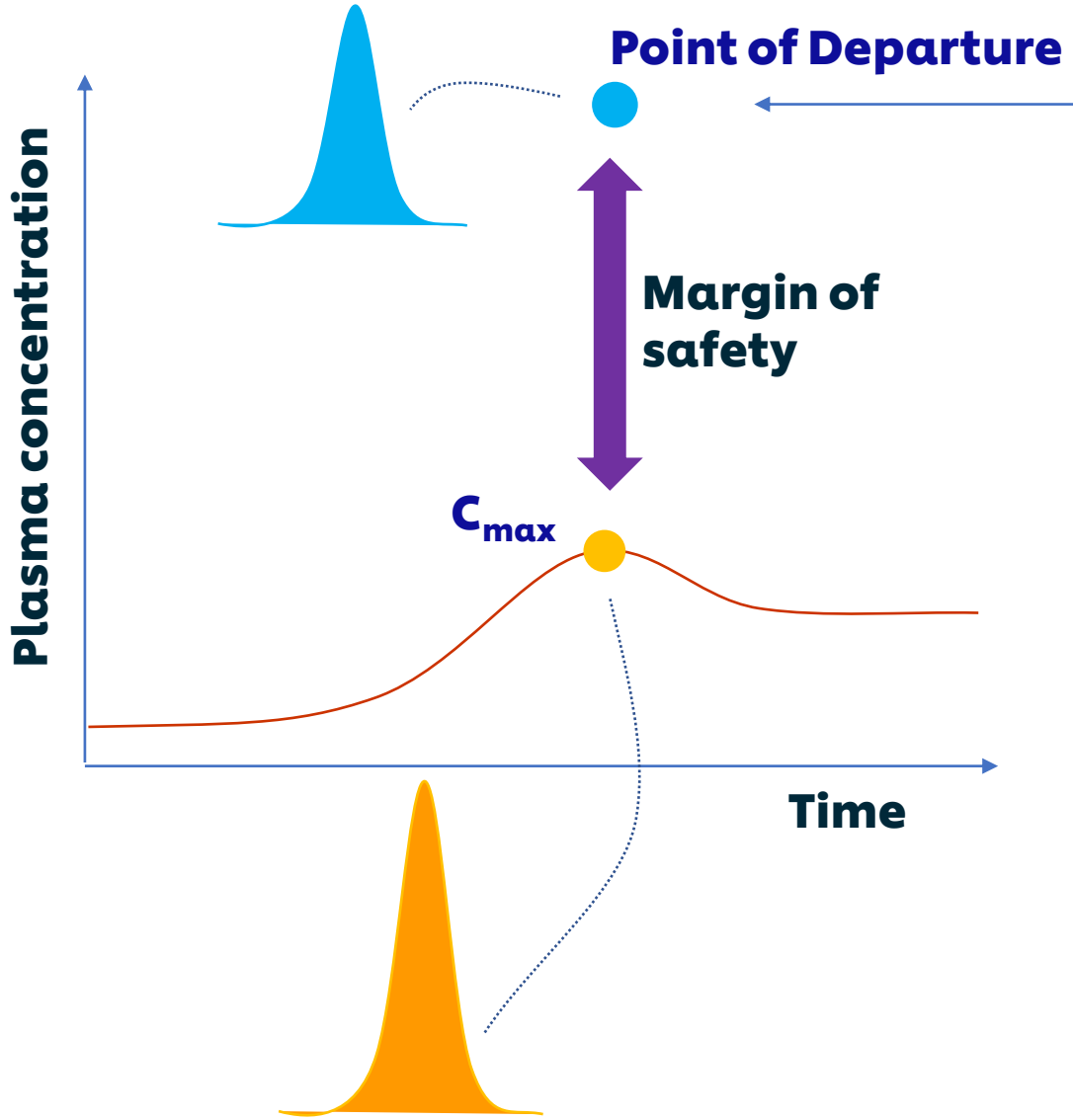
2 Principles for documenting NGRA:

- Sources of uncertainty should be characterized and documented
- The logic of the approach should be transparently documented

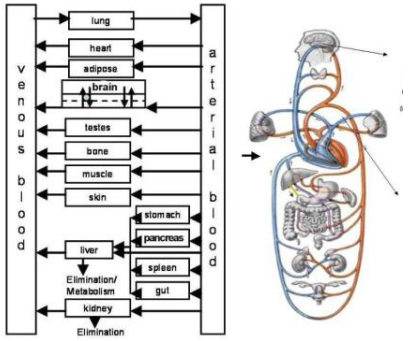
Dent *et al* ., (2018) *Comp Tox* 7:20-26

Unilever

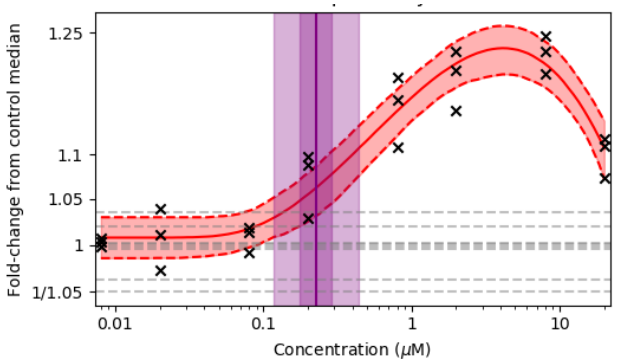
The Margin of Safety Approach



Exposure models (PBK, free/total concentration)

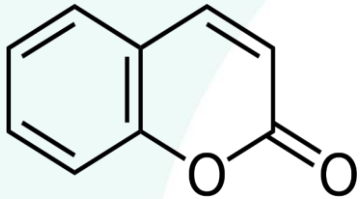


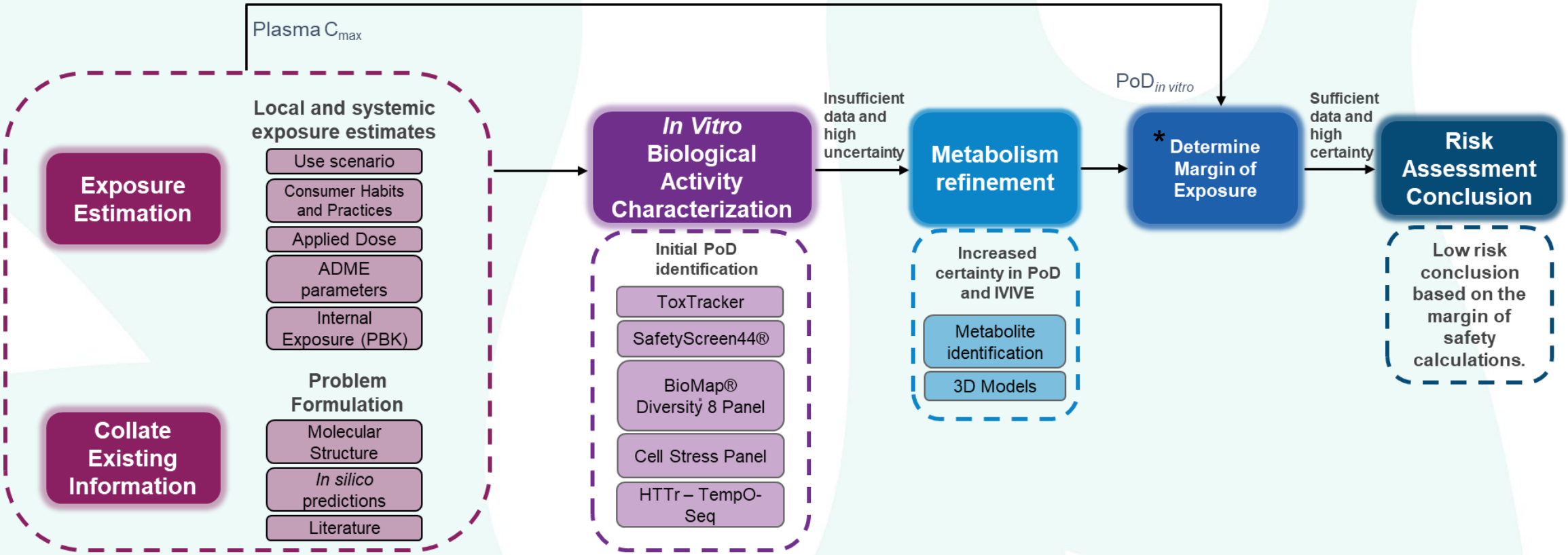
Point of departure derived from *in vitro* concentration-response



Using a tiered approach to conduct risk assessments

0.1% COUMARIN IN FACE CREAM AND BODY LOTION (NEW FRAGRANCE)





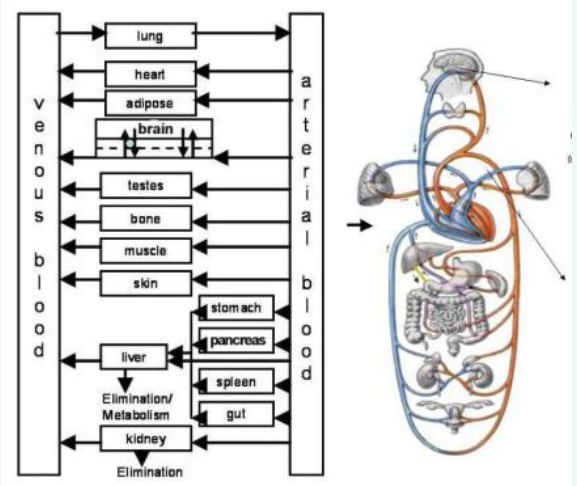
Baltazar *et al.*, (2020) *Tox Sci* Volume 176, Issue 1, 236–252

*or Bioactivity exposure ratio (BER)

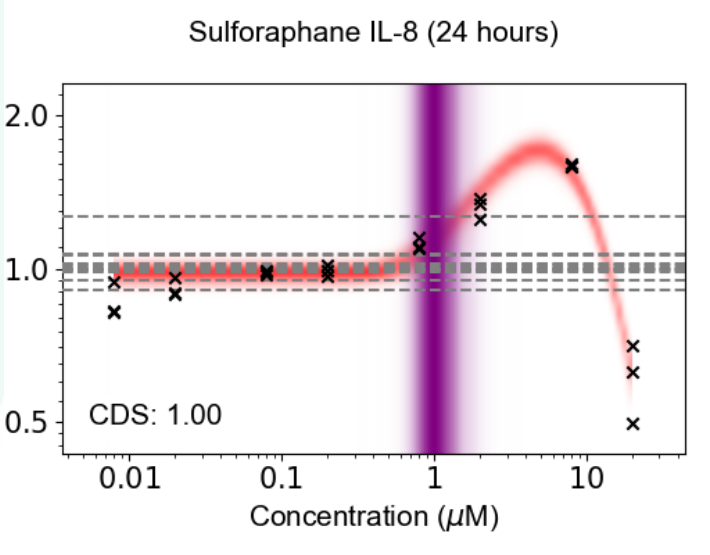
Unilever

Different types of computational approaches used in NGRA

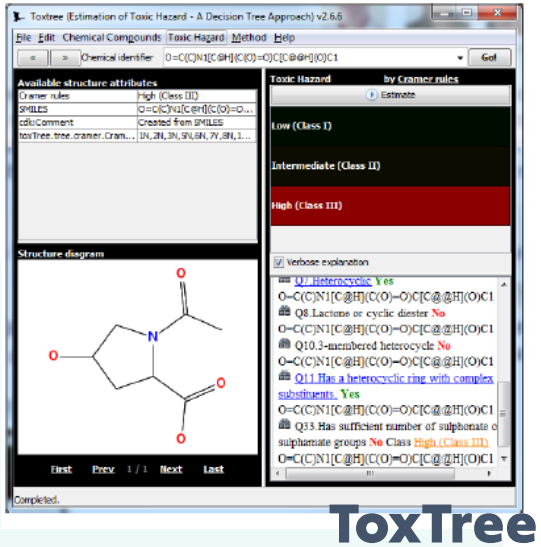
Physiologically-based kinetic (PBK) modelling



Dose response modelling

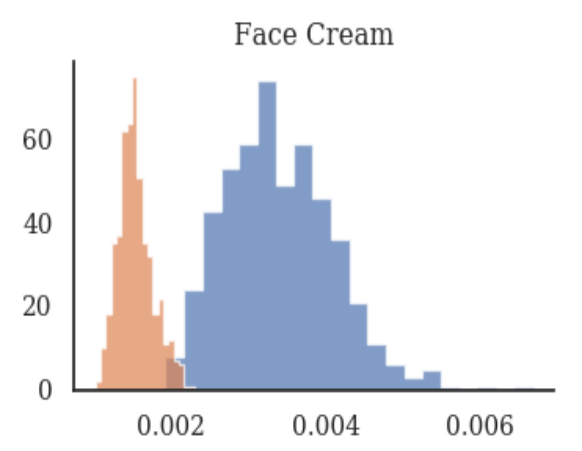


In silico tools

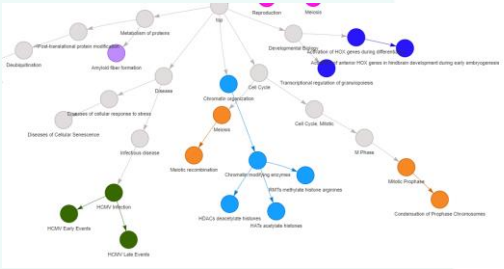


ToxTree

Statistical models of uncertainty and variability

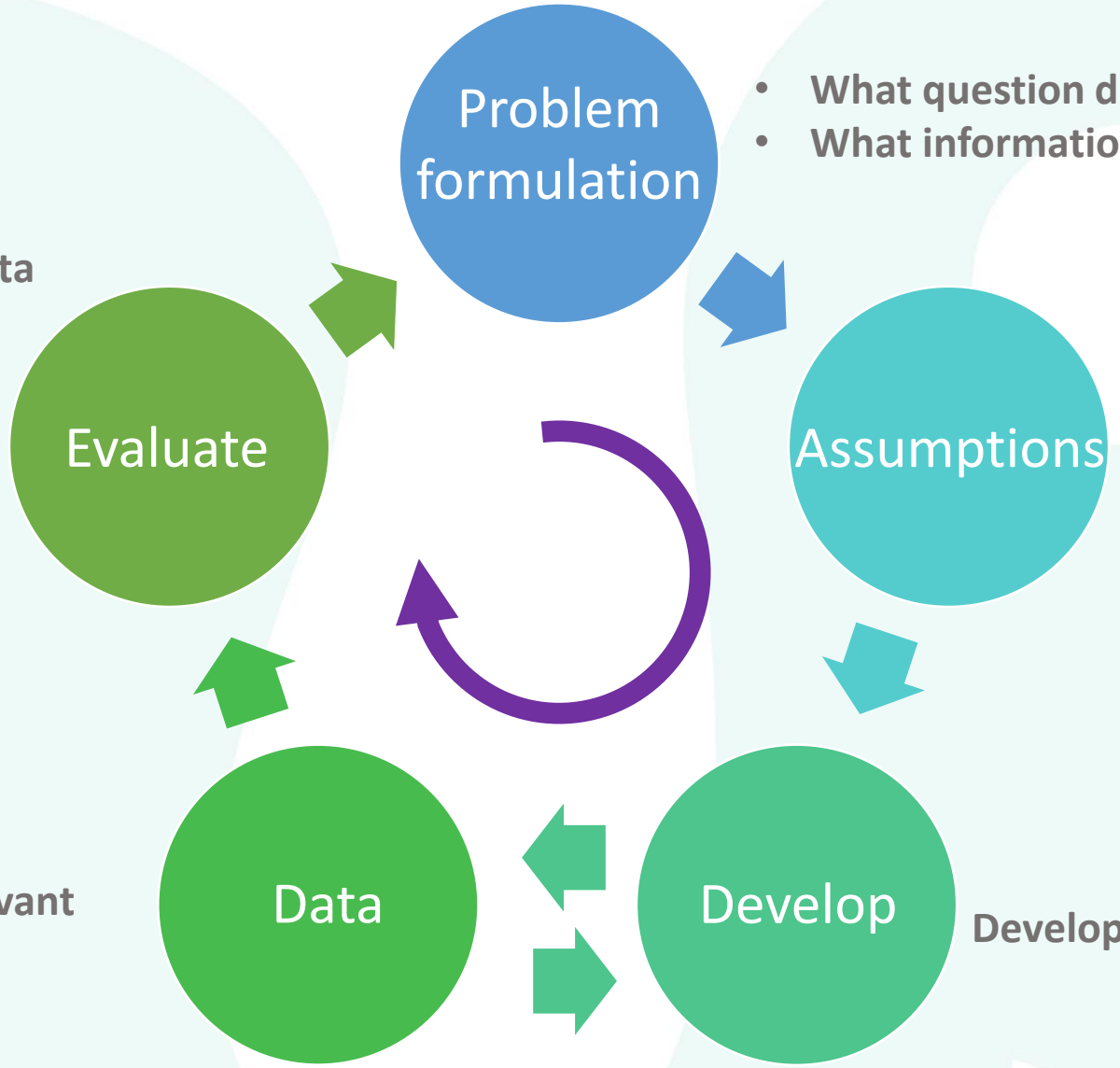


Bioinformatics tools for analysing omics data



Unilever

Principles of model development and the wet-dry cycle



- What question do you want to answer?
- What information do you have available?

Define model assumptions

Develop and implement the model

Generate/curate relevant data

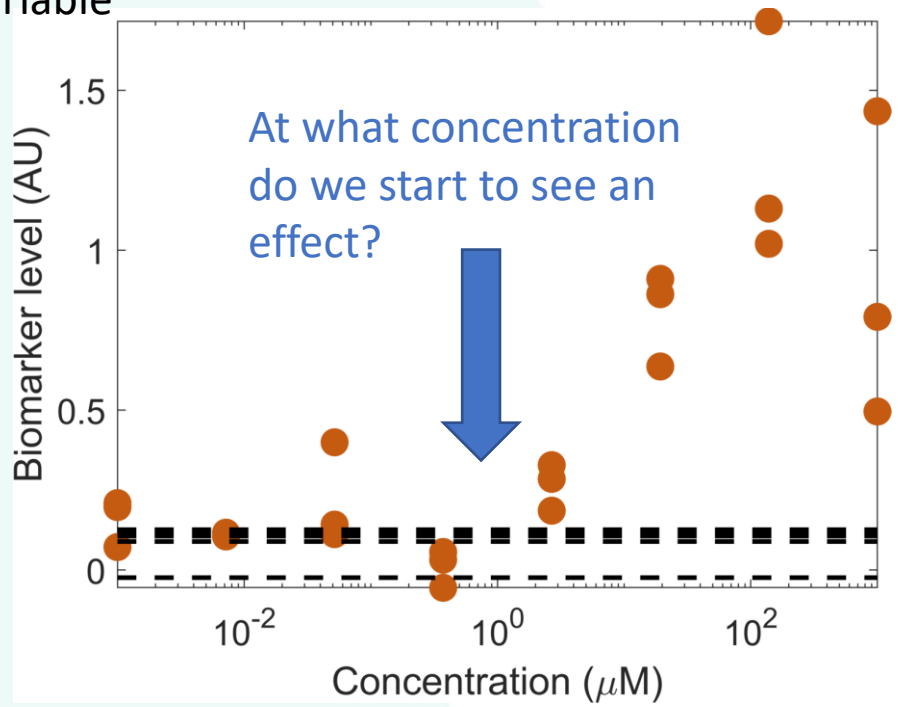
How does the model perform?
Does it describe the data well?

Using models to estimate PODs from concentration-response data

Unilever

Concentration-response data

Continuous variable



Concentration points evenly distributed on log scale (base 10)

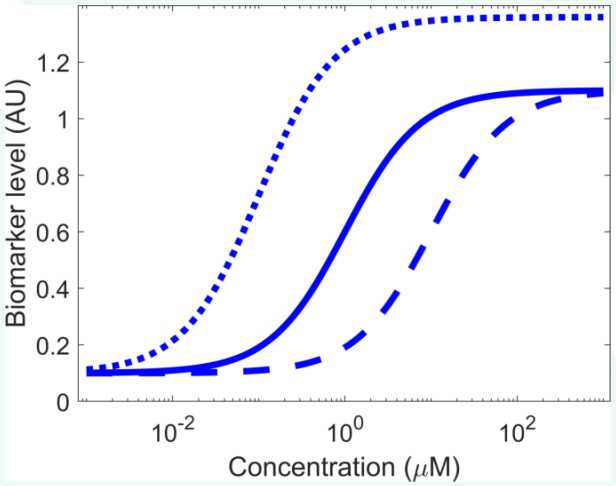
Example data types:

- High content imaging:
  - Fluorescent labelling of specific biomarkers
  - Phenotypic profiling
- Gene expression data:
  - Quantitative reverse transcription PCR (RT-qPCR)
  - Microarray data
  - RNA-seq
- In vitro pharmacological profiling
- Other omics data (e.g., proteomics).

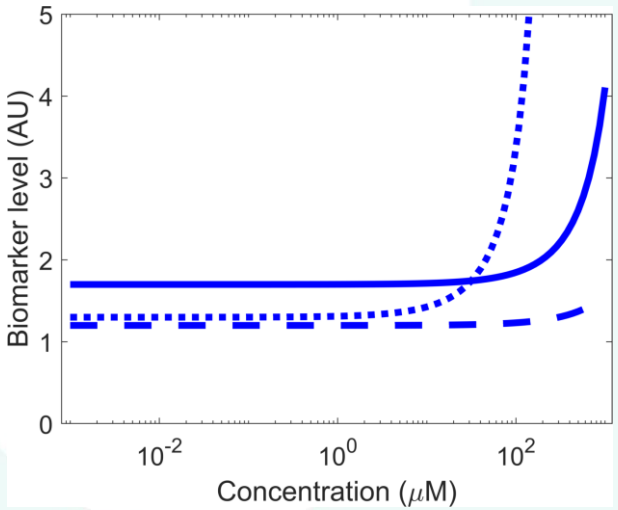
General strategy to estimating PODs from data

- **Problem**: We want to know:
  - Does the chemical have an effect on our biomarker?
  - At what concentration does this effect occur?
- **Typical approach**:
  - Fit one or more models to the data
  - Choose 'best model' based on fit
  - Use the fitted model to estimate quantities of interest – e.g., PODs

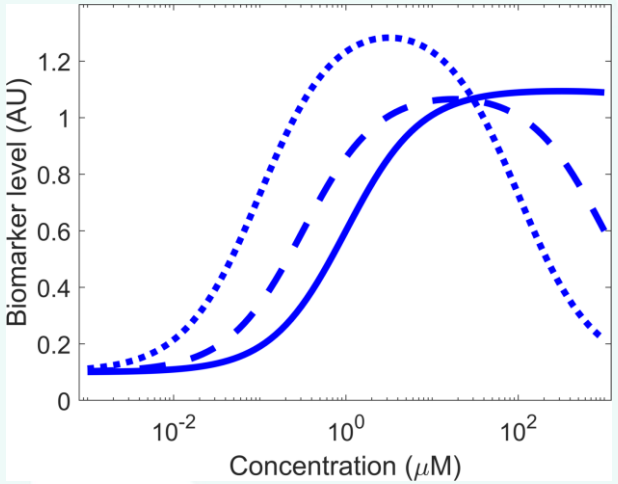
Hill function



Exponential

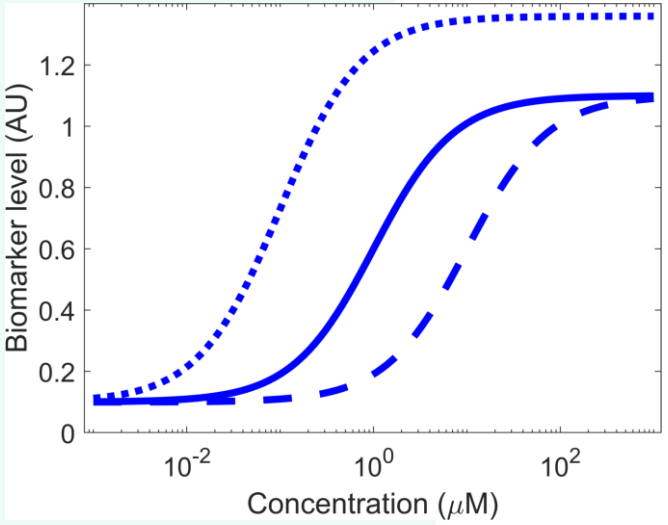


Gain-loss model

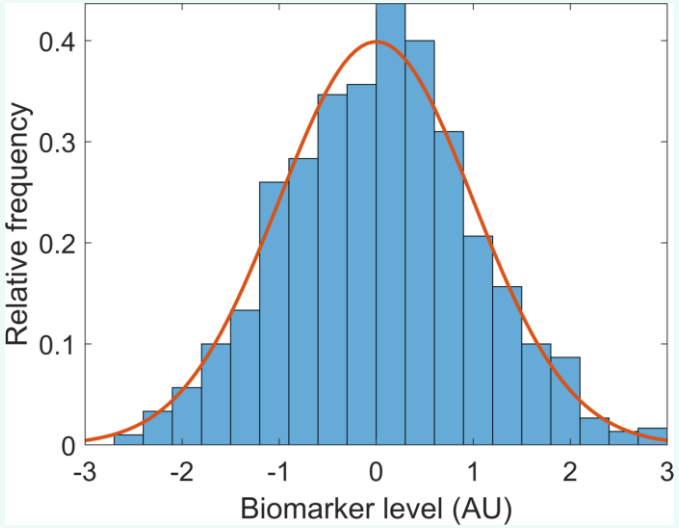


Parametric models

Hill function



Normal distribution

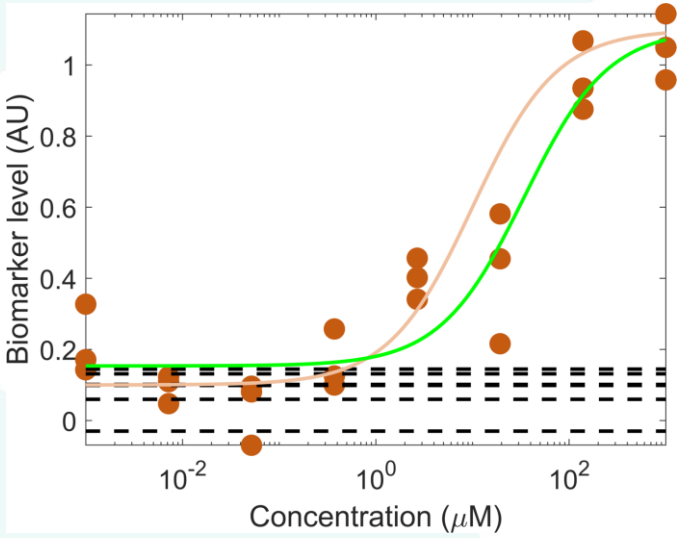


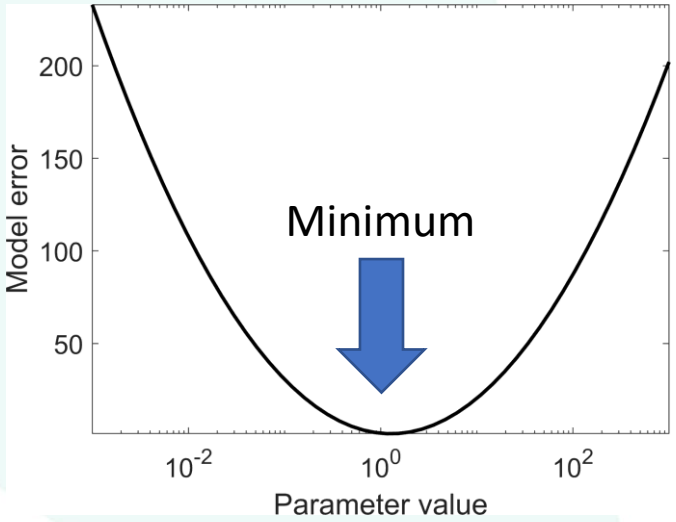
- Main building blocks of the model:
  - Measured data = Mean Response + Observational Noise
  - $y = f(x|C, \theta, V_{max}, h) + \eta$
- Various **parameters** that need to be estimated from the data:
  - $f(x|C, \theta, V_{max}, h) = V_{max} \frac{x^h}{x^h + \theta^h} + C$
  - $\eta \sim N(0, \sigma)$

Parameters:
C, θ, V_{max}, h and σ

Fitting models by maximising the Likelihood function

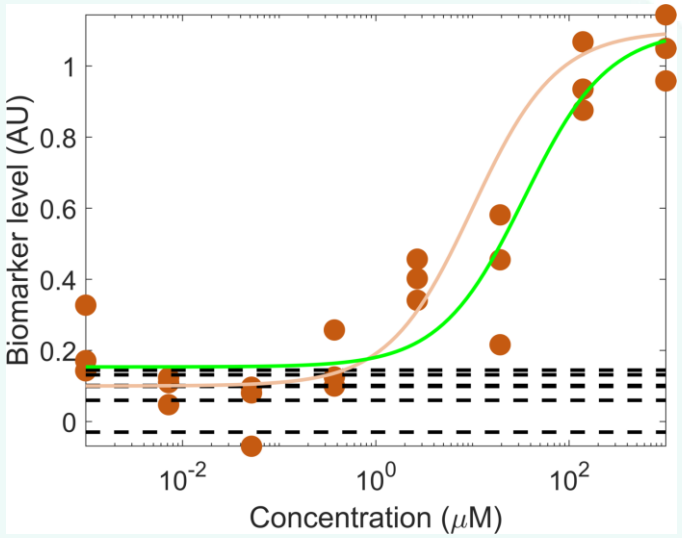
- Formally, the **likelihood** is the probability of the data given a parameter value.
- However the data are fixed, so it should really be thought of as function of the parameters:
  - $\mathcal{L}(\boldsymbol{p}) = P(D|\boldsymbol{p})$
- Often we actually work with the **negative log-likelihood**:
  - $-\log(\mathcal{L}(\boldsymbol{p}))$
- To fit the model to data, we find parameters that maximise the likelihood.
  - This is the same as minimising the **negative log-likelihood.**
- Under certain conditions, this is equal to minimising the sum of squared residuals, i.e., $\sum\left(D_i - f(x_i|C, \theta, V_{max})\right)^2$



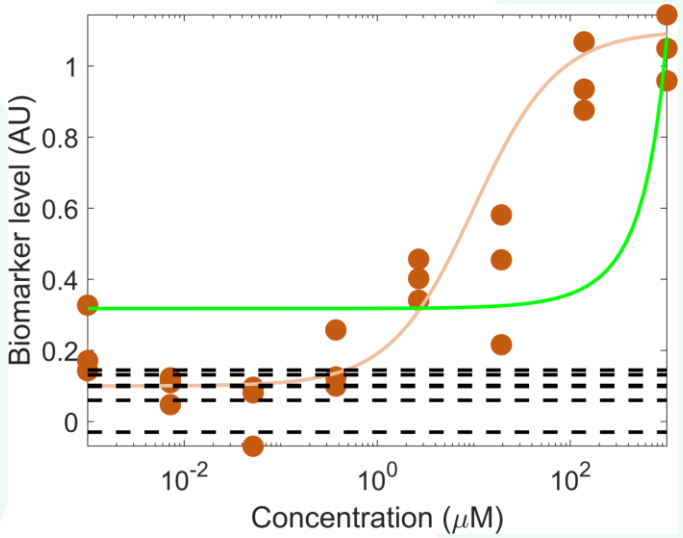


Comparing different model fits

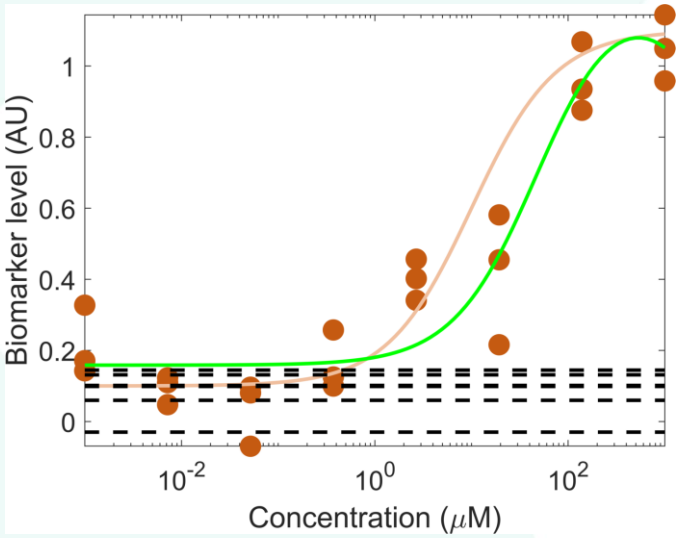
Hill function



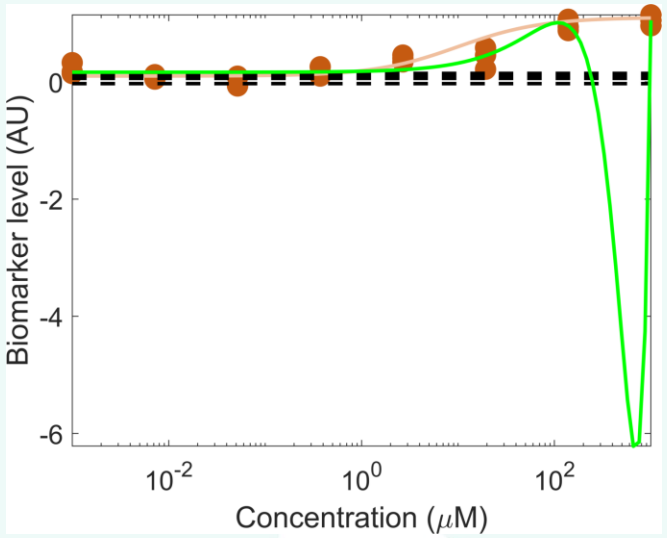
Exponential



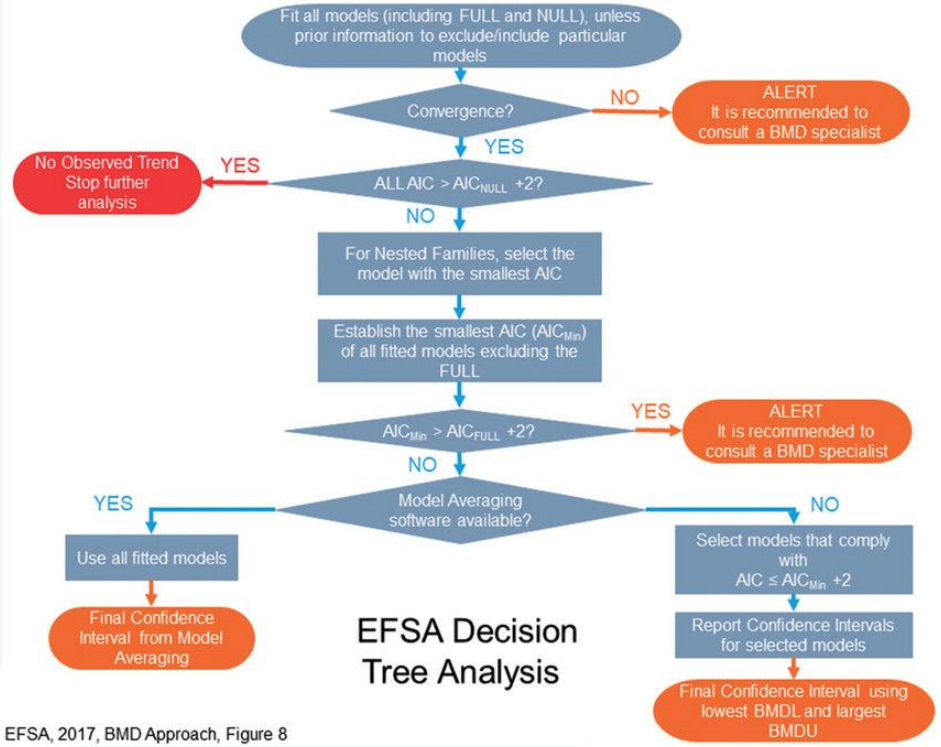
Gain-loss model



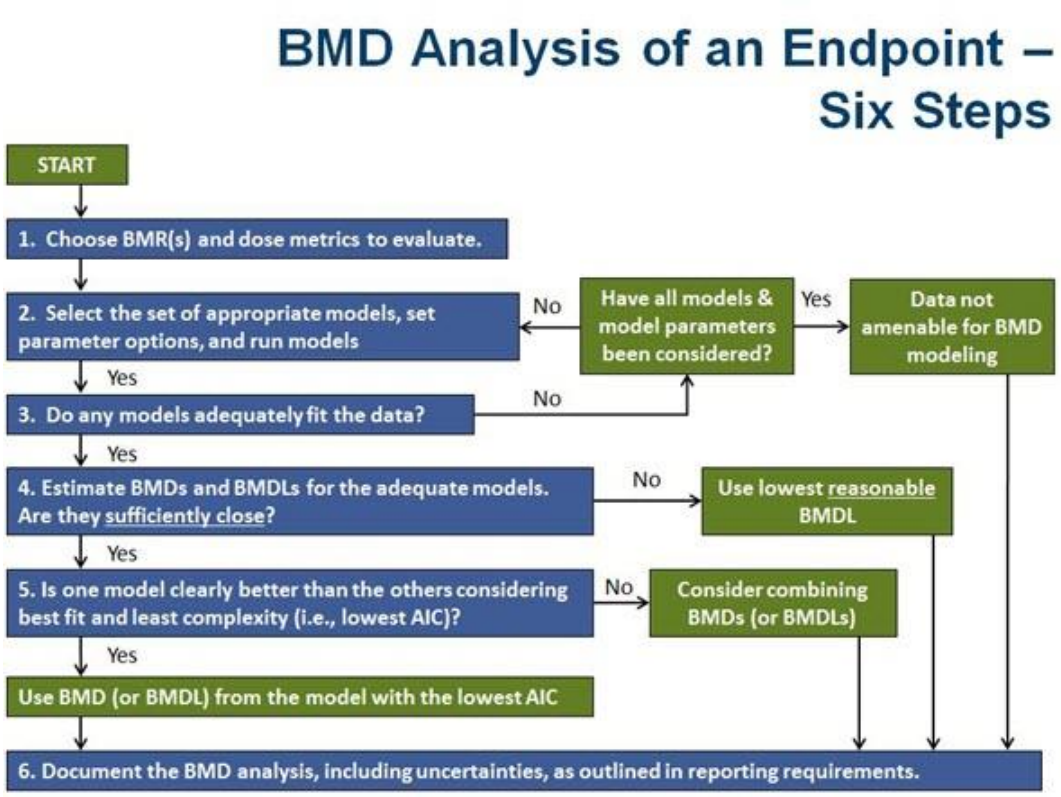
Polynomial



Model selection criteria



EFSA, 2017, BMD Approach, Figure 8



- Different decision trees are used for selecting the ‘best’ model
- Key metric – Akaike Information Criteria (AIC)

Akaike Information Criteria (AIC)

- The AIC or 'Akaike Information Criteria' is a common metric for comparing different models
- A naïve approach would use the Likelihood to select a model – i.e., the model with small error 'wins'
- Generally speaking, the more complex a model (e.g., the more parameters) the more likely it is that it will produce a very small error which is actually overfitting the data.
- The AIC is defined as:

$$\mathbf{AIC} = 2k - \log\left(\mathcal{L}(\boldsymbol{\theta})\right)$$

(where $\mathcal{L}(\theta)$ is the likelihood and k is the number of parameters)

- The preferred model is generally the one with the smallest AIC – it rewards 'good fits' while penalising models that are overly complex (i.e., have a large number of parameters).
- Note it is a *relative* measure used for comparing different models – the AIC says nothing about whether a model fit is good in an absolute sense.
- Another common model selection criteria is the Likelihood Ratio test – which can be used for nested models.
- BMDExpress2, for example, allows users to combined the LR test and AIC to select the best model.

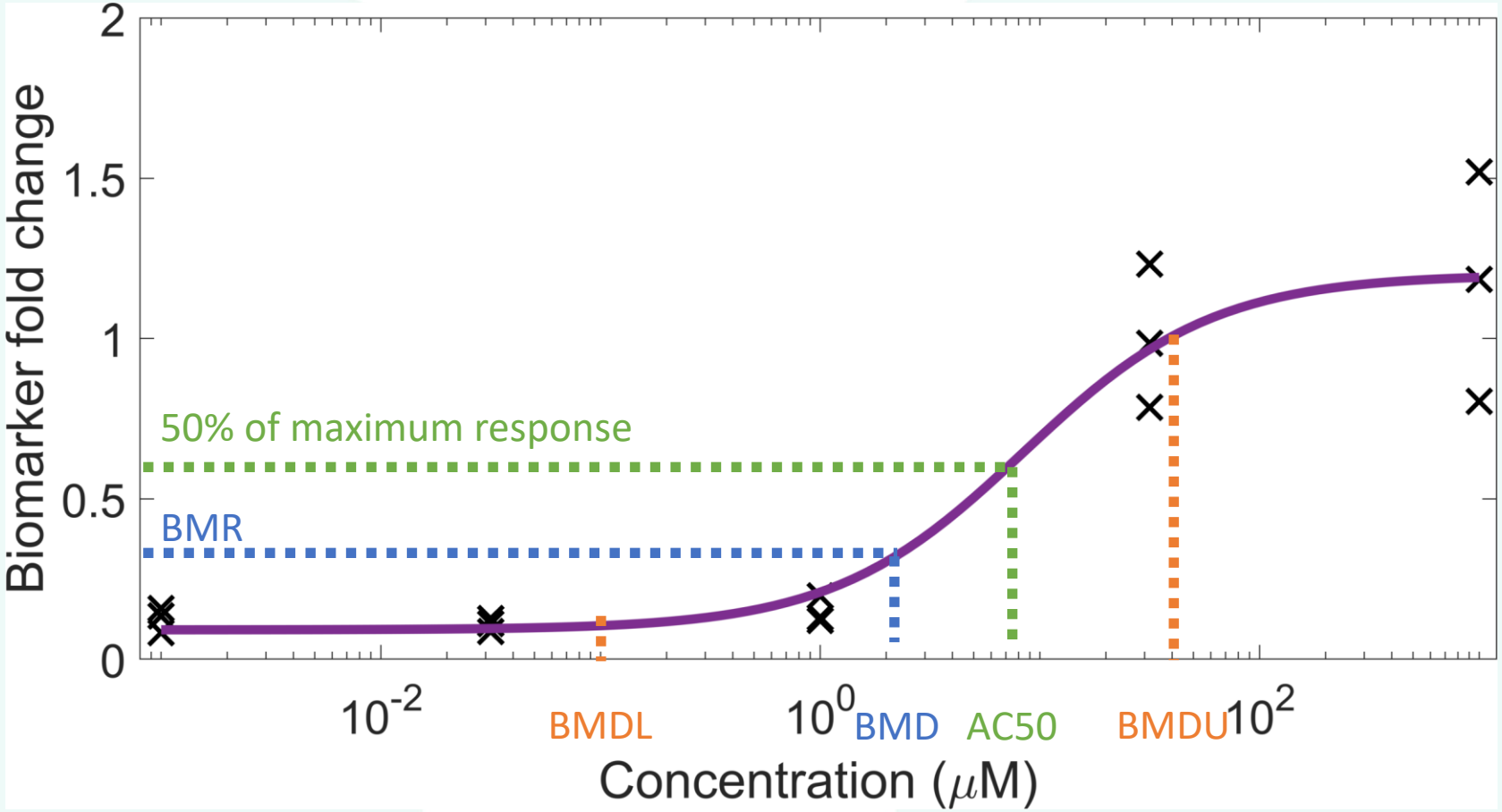
Unilever

Akaike Information Criteria (AIC) Example

Model	$-\log(\mathcal{L}(\theta))$	Number of parameters (p)	AIC
Gain-loss	17.2	4	25.2
Hill function	17.5	3	23.5
Polynomial	17.7	4	25.7
Linear	77.6	2	81.6
Exponential	86.5	3	92.5

Unilever

Estimating the POD using the 'best model' fit

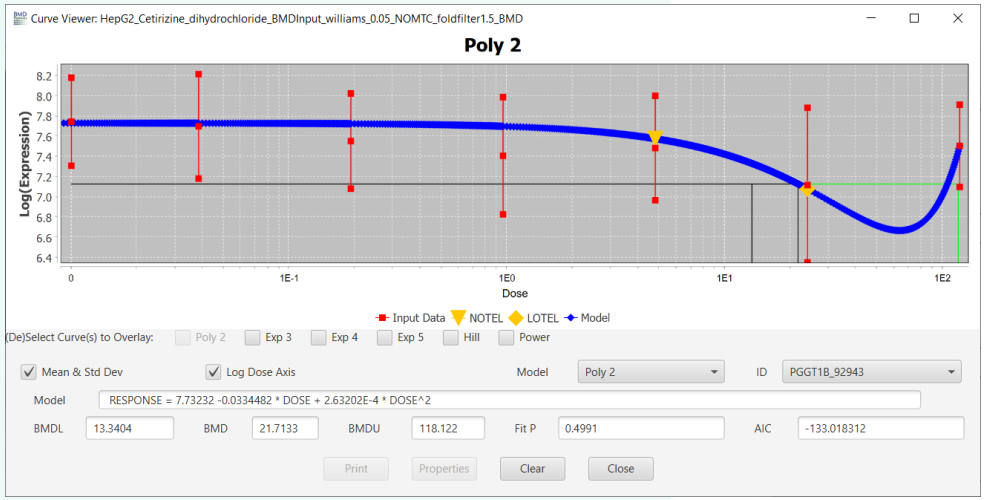


There are several ways to define the BMR/BMD, but generally use:

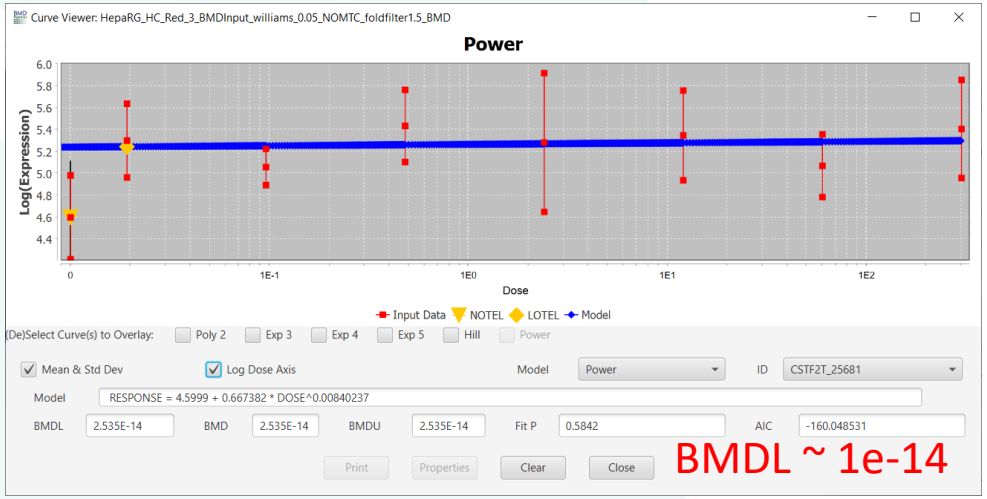
- $\mathrm{BMR} = \mu_{CONTROL} \pm \sigma_{CONTROL}\mathrm{BMRf}$
  (where the BMRf is a multiple of standard deviation of the control)

Challenges of using parametric models

Overfitting?

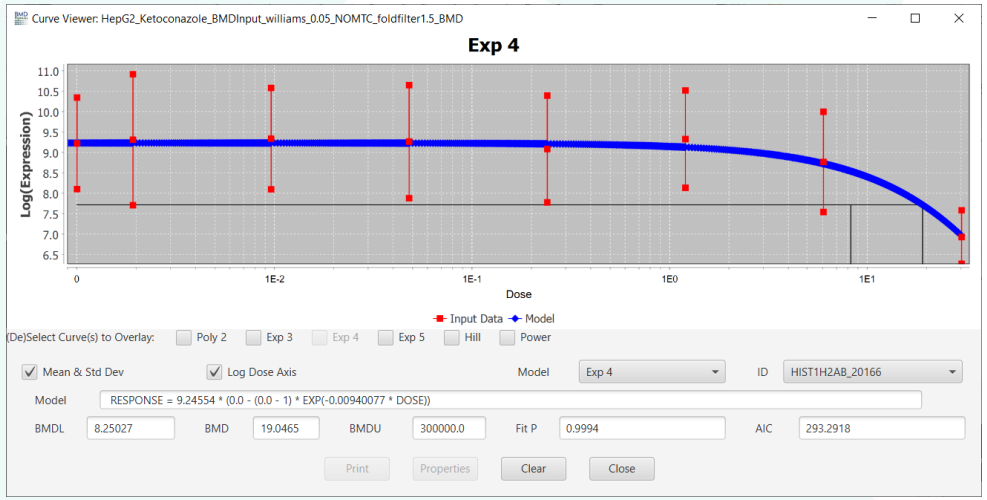


Overestimating the BMDL?

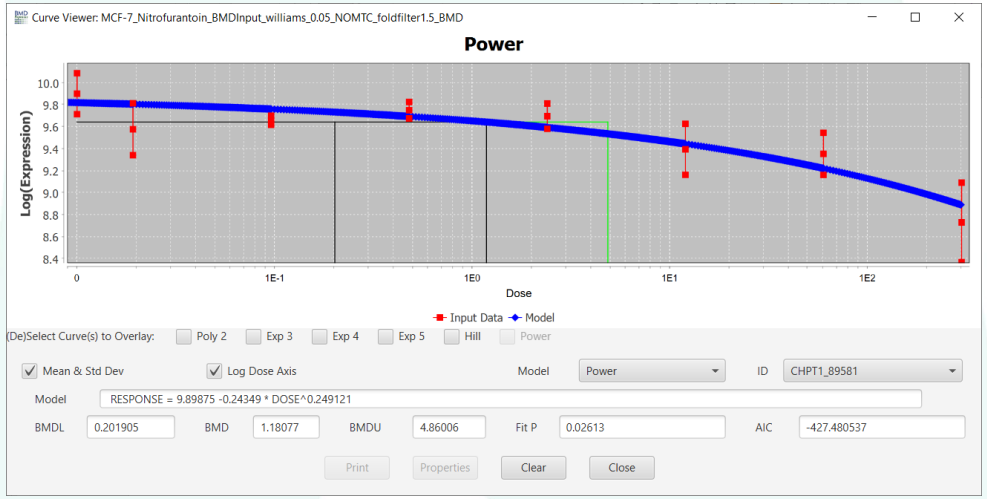


BMDL ~ 1e-14

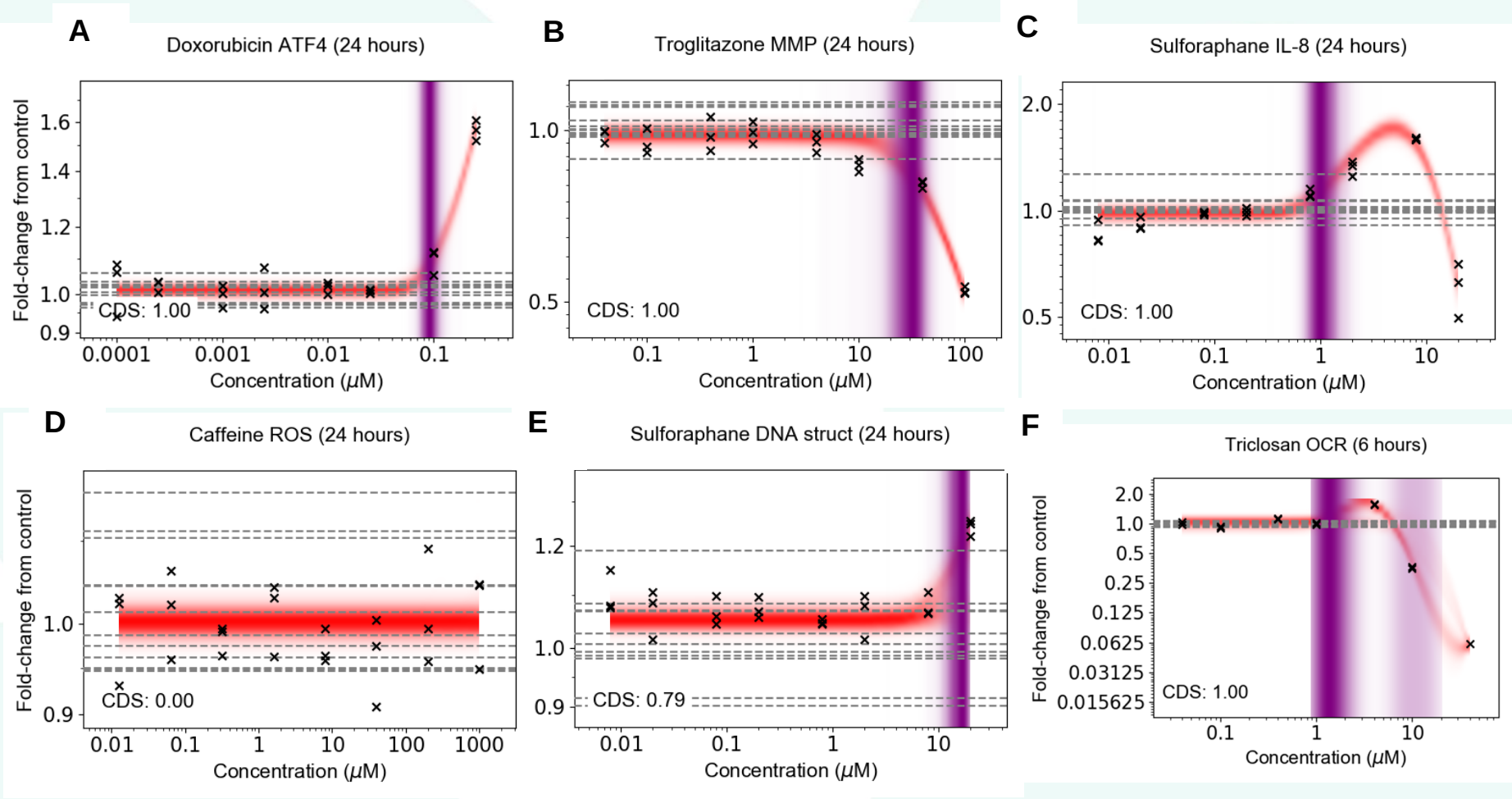
Fit looks good but does not pass BMDL/BMDU criteria



Fit looks good but does not pass p-value criteria



BIFROST: using non-parametric Bayesian inference to estimate PODs



Hatherell *et al.*, 2020, Identifying and characterizing stress pathways of concern for consumer safety in next generation risk assessment, Tox. Sci. https://doi.org/10.1093/toxsci/kfaa054

Bayesian statistics – what and why

Frequentist probability

- What people are normally taught in school
- Basis for **p-values** and **hypothesis testing**
- Probability reflects the relative frequency at which an event occurs over many repeated trials.
- Only really relevant when dealing with **well-defined random experiments**
- Can't use it to talk about the probability of a 'parameter taking a certain value' or a 'hypothesis being true'.

Bayesian probability:

- Probability reflects the **plausibility** or **belief** in some event being true.
- Provides framework for updating plausibility based on available data.
- For example, can talk about the **probability of a hypothesis being true**, or a parameter taking on a certain value.
- Key terms: credible interval, priors, posterior

Thomas Bayes, 1701-1761

Bayesian statistics – what and why

Bayesian interpretation of probability

- Probability quantifies the plausibility of some event.
- **Bayes' theorem**:

Likelihood $\mathcal{L}(\theta)$

Prior

Posterior

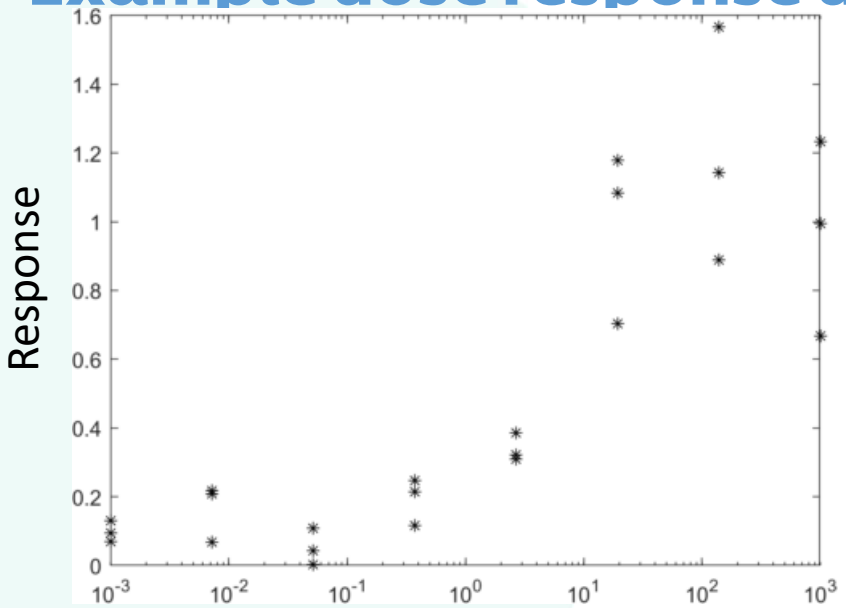
$$P(X|D) = \frac{P(D|X)P(X)}{P(D)}$$

- Here, D is the data and X is a random variable
- E.g., X – V_{max} parameter, D – experimental observations
- The key things are the likelihood, the prior and the posterior:
  - **Posterior**: probability that V_{max} takes a certain value
  - **Likelihood**: probability of the data, given V_{max}
  - **Prior**: probability reflecting initial assumptions V_{max}

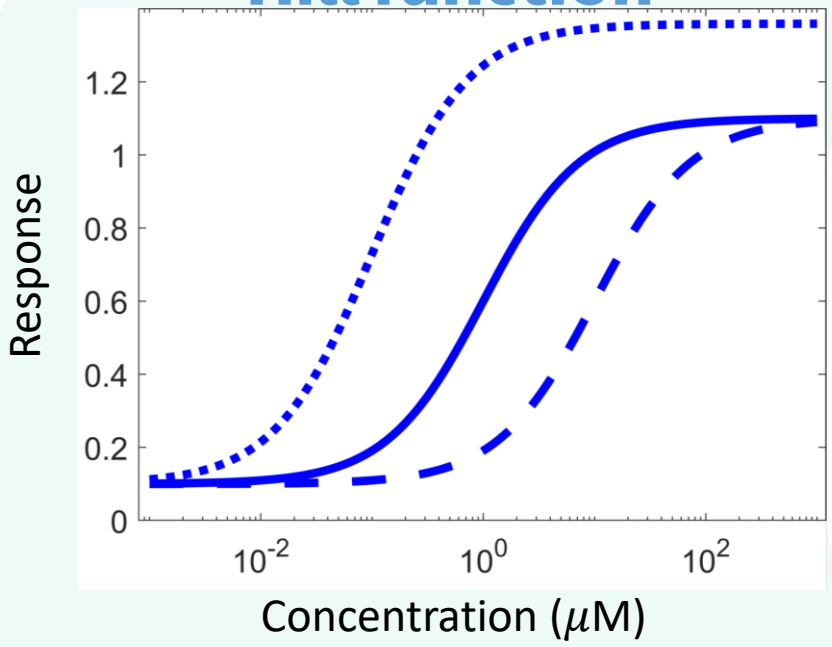
Unilever

Example of using Bayesian inference

Example dose response data



Hill function



Develop

- Hill equation:

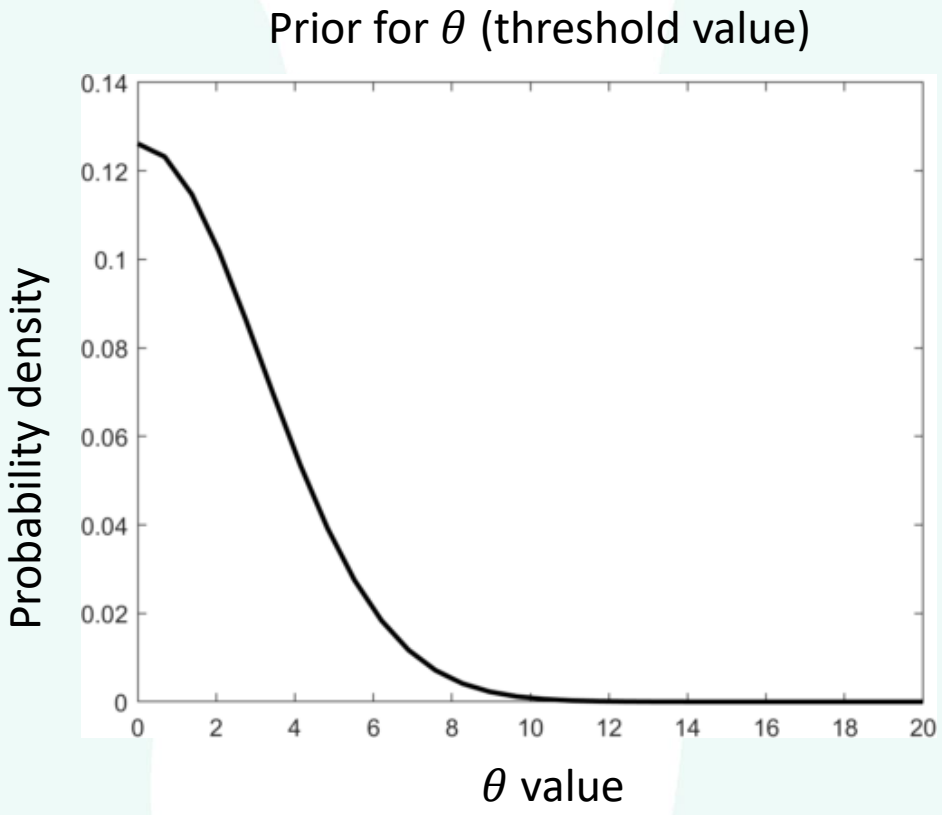
$$f(x|C,\theta,V_{max}) = V_{max}\,\frac{x}{x+\theta} + C$$

- (full Hill equation has exponent on x and θ to obtain sharper curves)

Example of a prior

Develop

- Have parameters θ, C, V_{max} and σ – need to be learned from the data



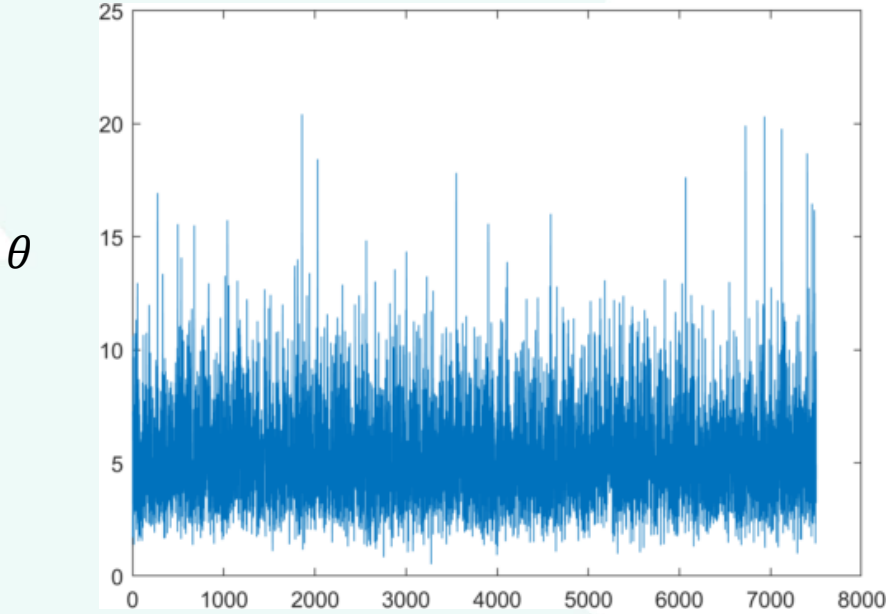
Data

- Typically you only have the measured values that you are fitting to, but you could incorporate prior knowledge (e.g. biologically plausible values) into the prior.

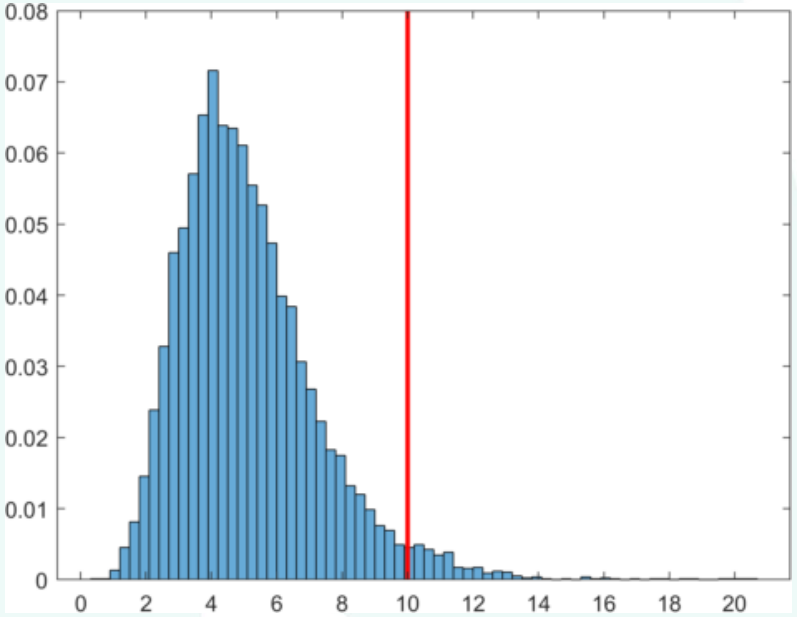
Learning parameters from the data

- One things that's important to know about Bayesian statistics is that for most problems, it is impossible to get an exact solution to the posterior.
- Resort to using methods like **Markov Chain Monte Carlo (MCMC)** to take random samples from the distribution.

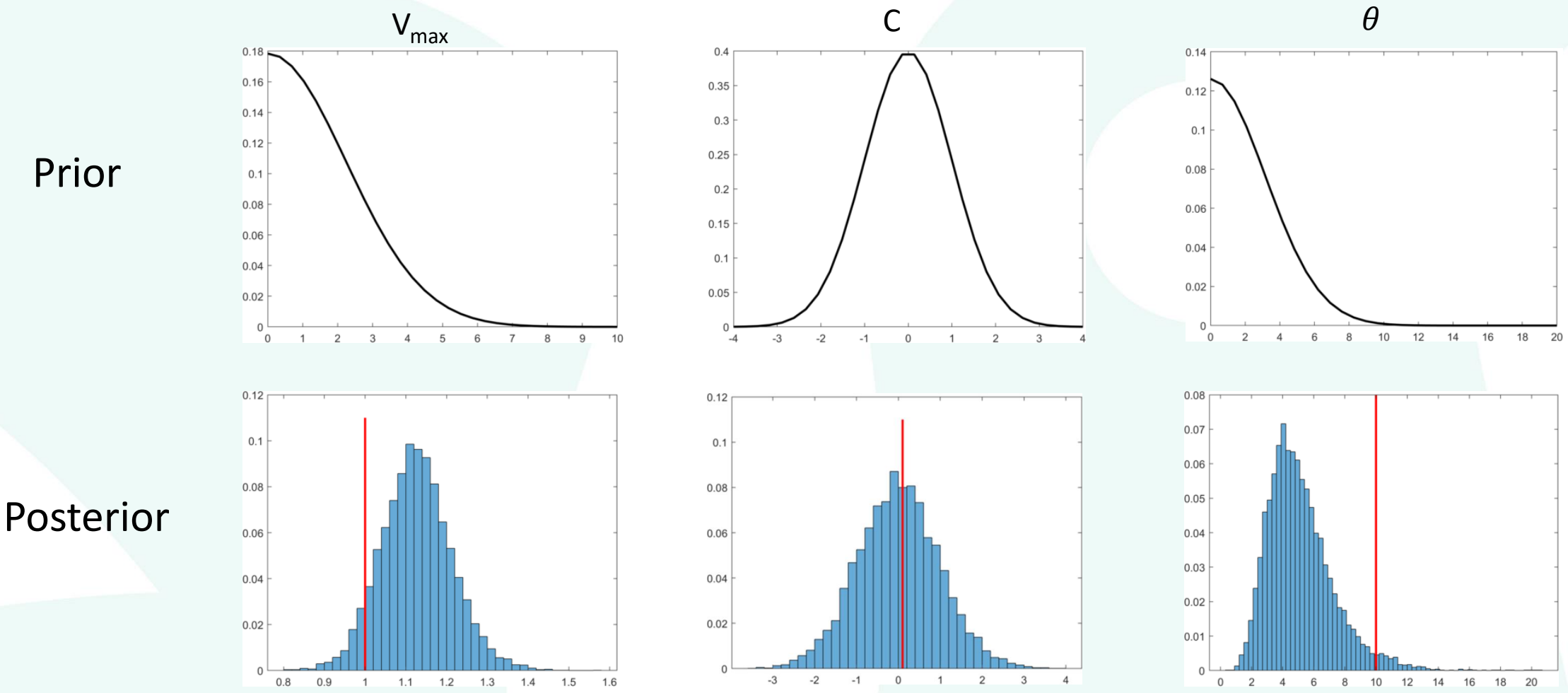
Random samples of θ (from posterior)



Histogram of θ

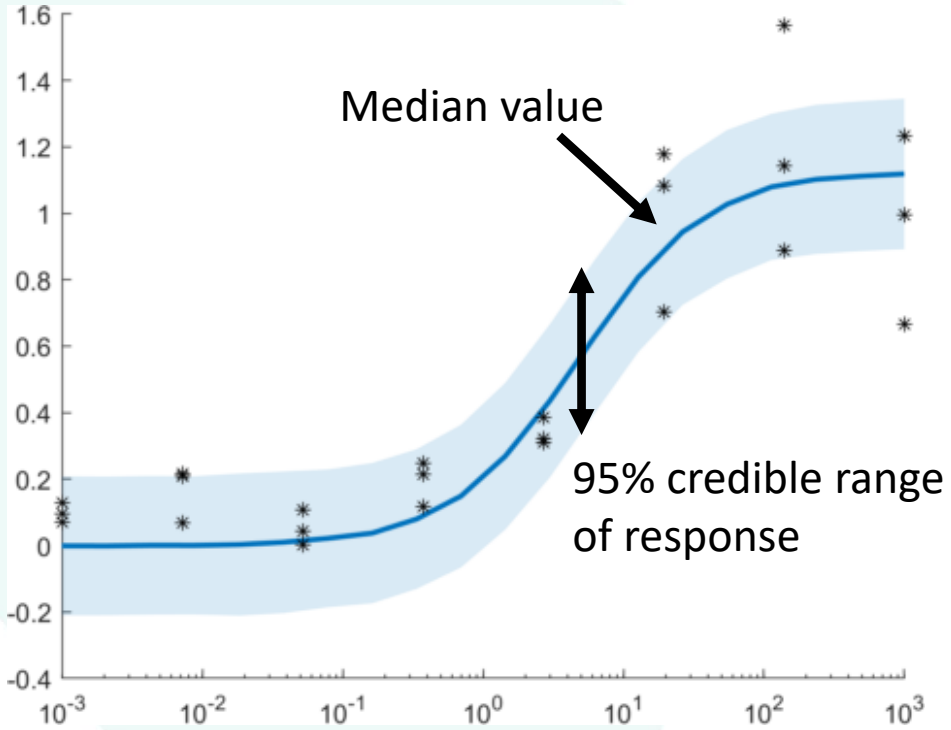


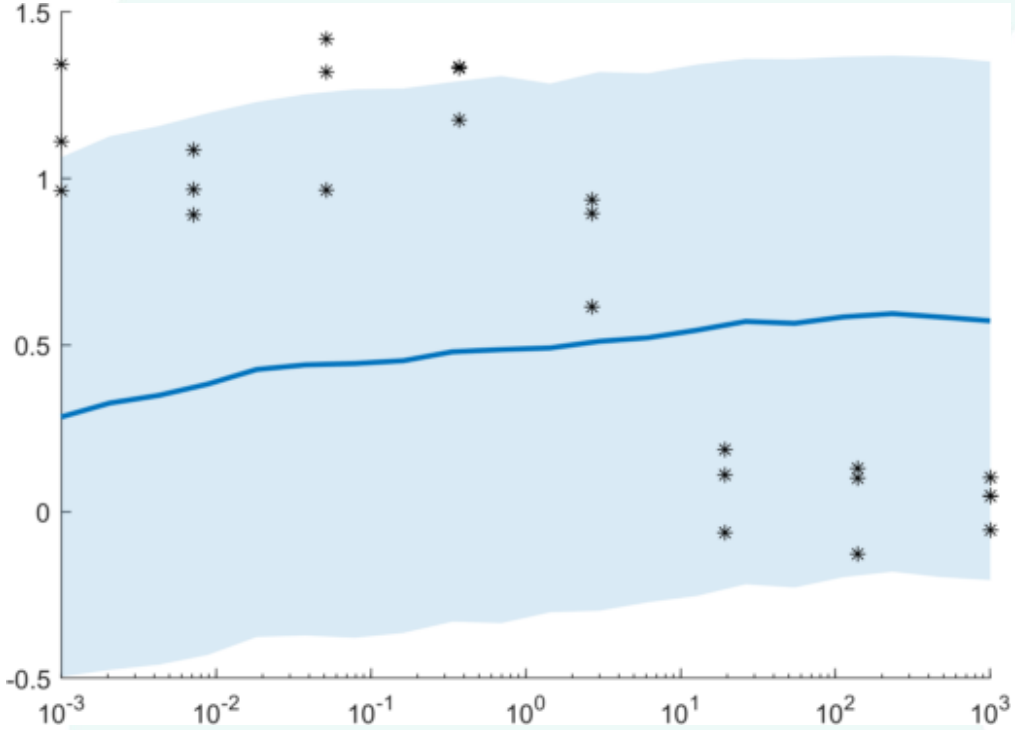
Learning parameters from the data: prior vs posterior



Red horizontal line indicates the ‘true’ value

Evaluating the dose response model





- Bayesian models can be evaluated by comparing the predictive distributions to the training data
- As with the frequentist approach, because you're using a parametric approach you have to fit multiple models to the data and decide which one is best

Examples of Bayesian dose response tools

Pyfit2

Wellcome Open Research

Wellcome Open Research 2017, 1:6 Last updated: 15 MAR 2017

Check for updates

SOFTWARE TOOL ARTICLE

REVISED **Hierarchical Bayesian inference for ion channel screening dose-response data** [version 2; referees: 2 approved]

Ross H Johnstone[1], Rémi Bardenet[2], David J Gavaghan[1], Gary R Mirams[1,3]

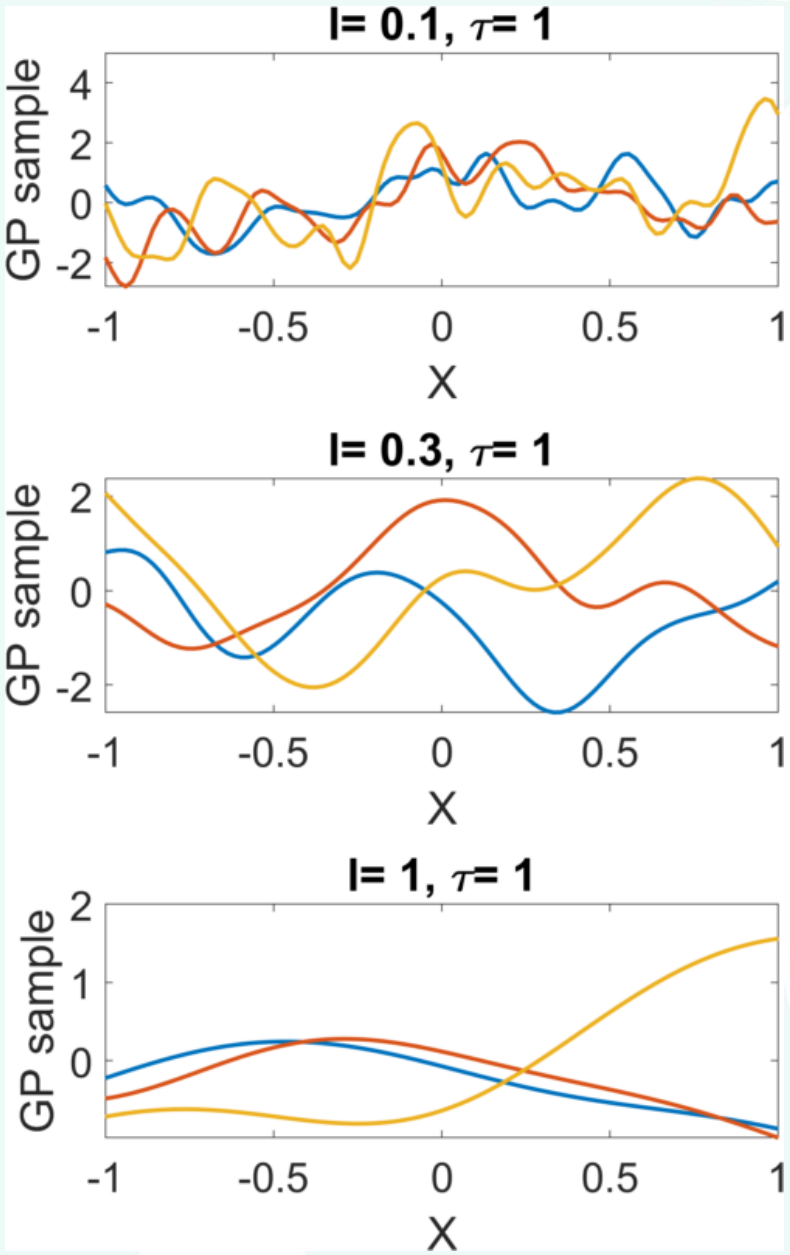
[1]Computational Biology, Department of Computer Science, University of Oxford, Oxford, UK
[2]CNRS & CRIStAL, Université de Lille, Lille, France
[3]Centre for Mathematical Medicine & Biology, School of Mathematical Sciences, University of Nottingham, Nottingham, UK

Pyfit2

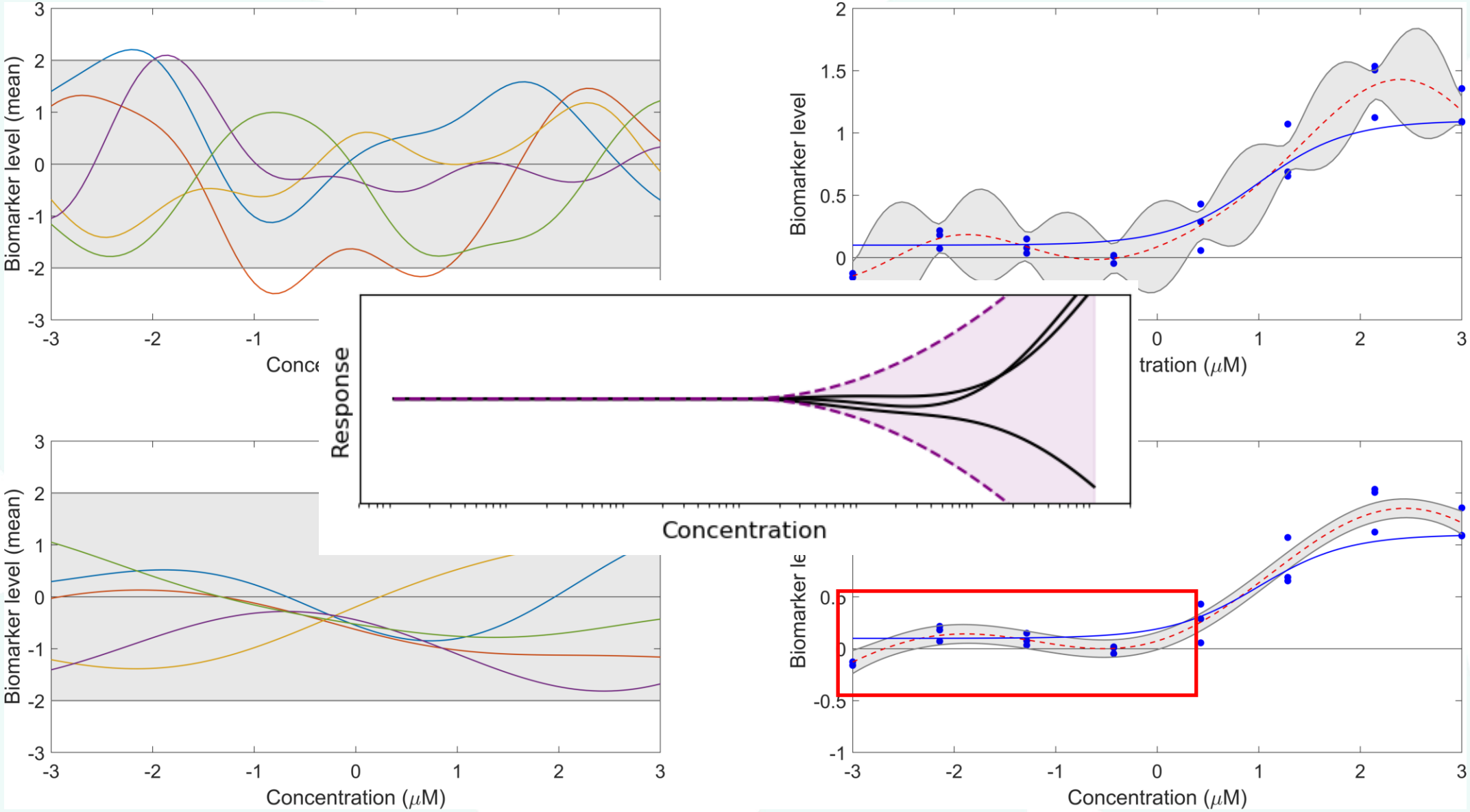
Unilever

Non-parametric approaches

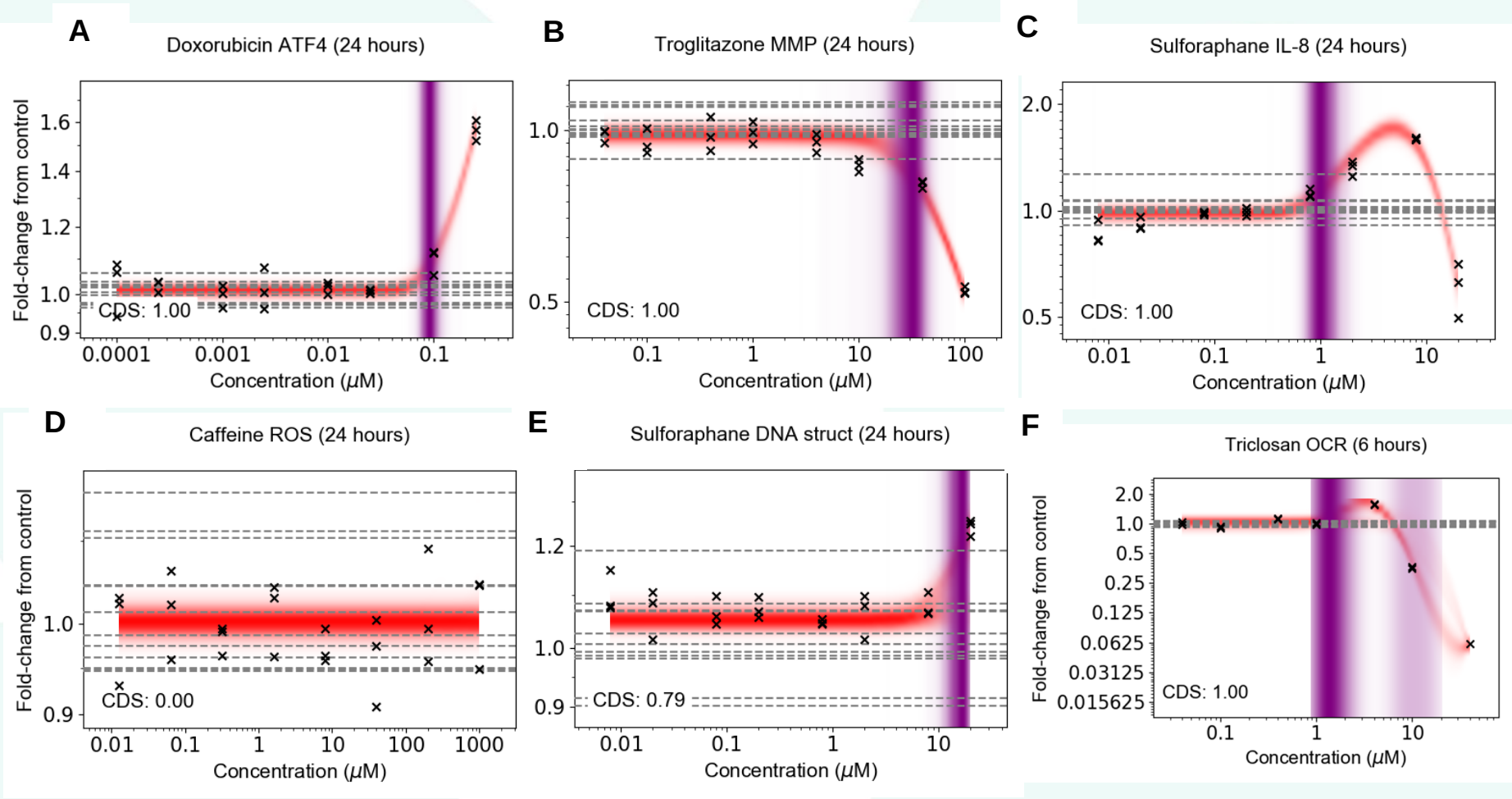
- So far, our approach of:
  1. Fit multiple candidate models, each of which produce a limited range of shapes.
  2. Choose the best model – use this to estimate the POD.
- An alternative approach is to use one model that is very flexible.
- Non-parametric approaches provide a way to do this.
- Gaussian processes (GPs) are an example of this – these allow you to describe different shapes in a probabilistic manner.



Example of using Gaussian Processes to fit data



BIFROST: using non-parametric Bayesian inference to estimate PODs



Hatherell *et al.*, 2020, Identifying and characterizing stress pathways of concern for consumer safety in next generation risk assessment, Tox. Sci. https://doi.org/10.1093/toxsci/kfaa054

Discussion

- Various different models are used in NGRA to help analyse data.
- Two key elements are using PBK models to estimate exposure and concentration-response models to estimate PODs.
- Typically, multiple parametric models are used to fit the data, from which the ‘best model’ can be used to estimate a POD.
- An alterative is to use non-parametric methods, like Gaussian processes.
- While these may be more robust, they can be more computationally complex and there is further go with their acceptance from a regulatory perspective.

Unilever