Bayesian algorithms for analyzing large biomarker panels in support of early phase immuno-oncology drug development

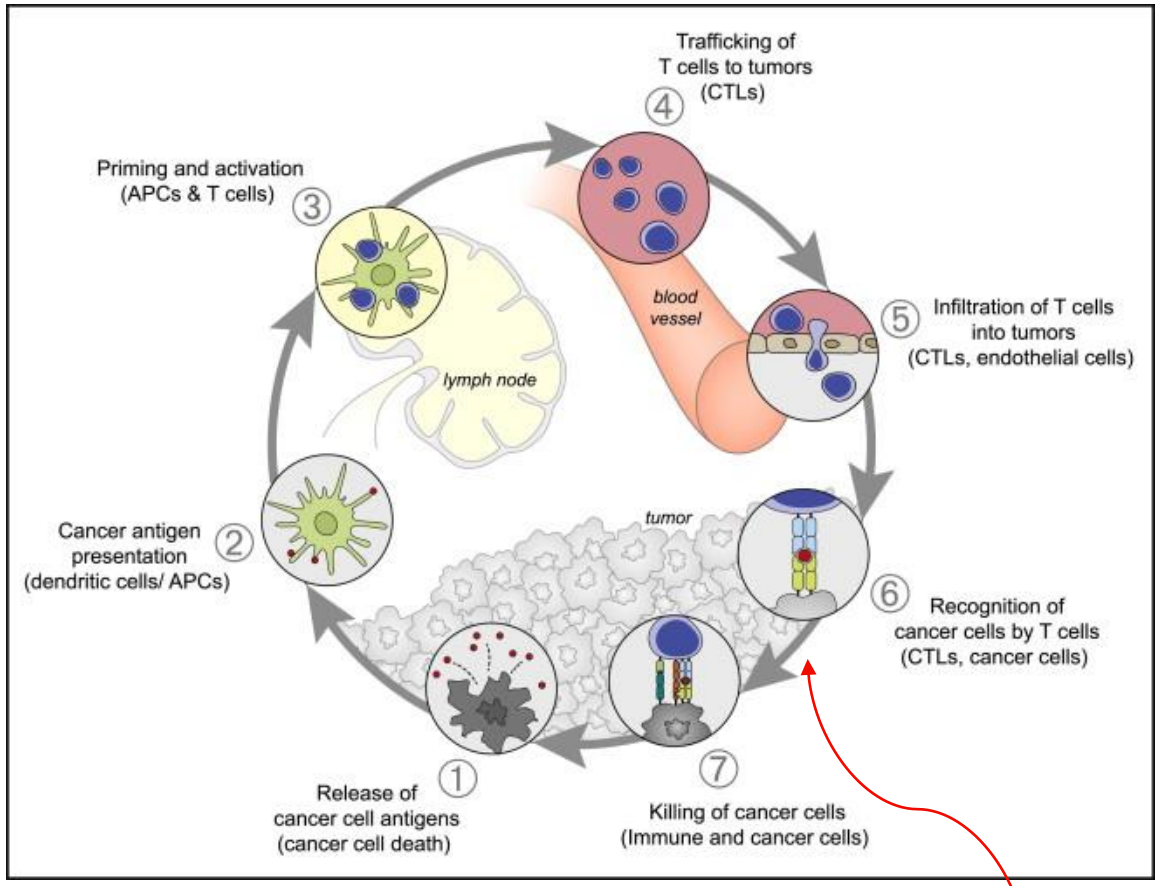
Adarsh Joshi

Oct 13, 2022

BAYES 2022 meeting, Bethesda

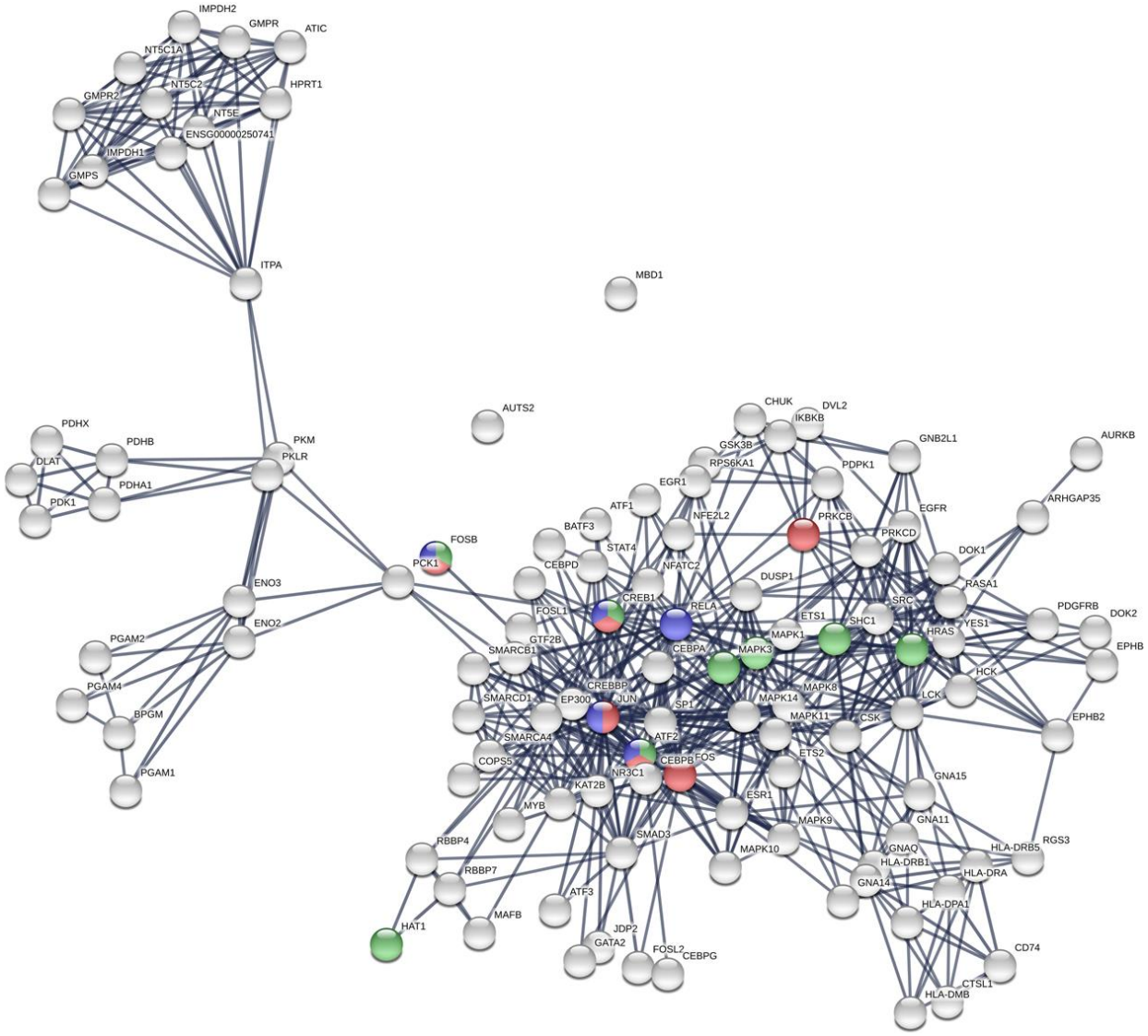
Background

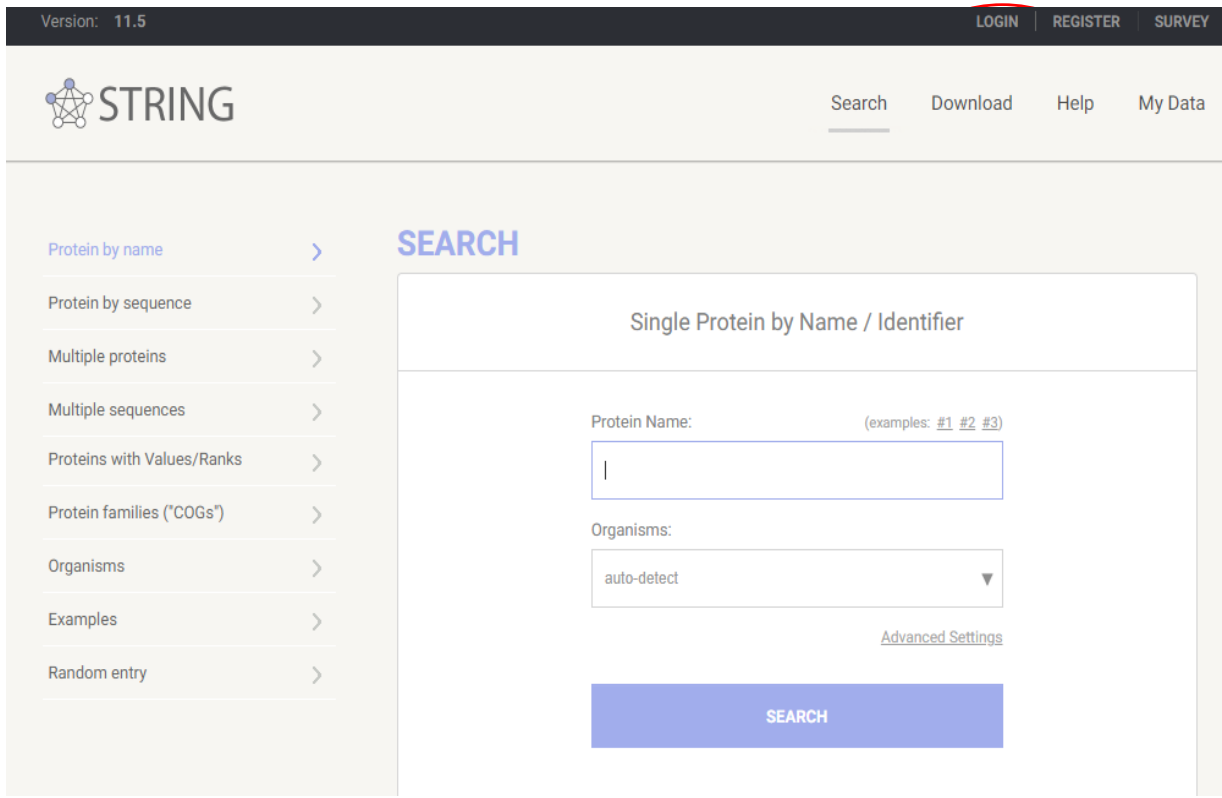
- Success of checkpoint inhibitors have spurred great interest in developing next gen IO therapies that can harness the full potential of the immune system
- Checkpoint inhibitors work only in some tumor types and in subset of patients in those tumor types
- As opposed to TKI therapies, translation from pre-clinical to clinical is extremely poor in IO space
- Collecting rich biological datasets on patients is key to clinical development strategy



Checkpoint inhibitors

Typical Enrichment analyses





- Standard bioinformatics pipelines are set around comparing two conditions (e.g., DESeq2) or simple linear models (LIMMA)
- Drug development questions can be more broad ranging

Large panels of biomarkers

Questions of interest to IO teams

- Which biological mechanisms are showing a dose effect?
  - What doses are they saturating at?
- Evidence of over-stimulation of the immune system (U-shaped dose effect)?
  - Traditional MTD based dose selection strategy may not pan out
- Evidence of tolerance in repeated dosing?
- Impact of ADA (anti-drug antibodies)?

Statistical considerations

- Small studies with 30-70 patients with mixed tumor types
- Heterogeneity in biological responses
  - Patient populations change over the course of dose escalation
- Batch effects in technologies
- Modeling considerations
  - Heavy tails/skeweness
  - Influential outliers
  - Censoring/missing data

Timelines: need to analyze and provide data summaries with quick turnaround

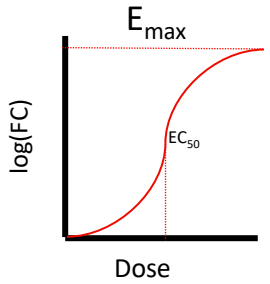
How Bayesian modeling can help

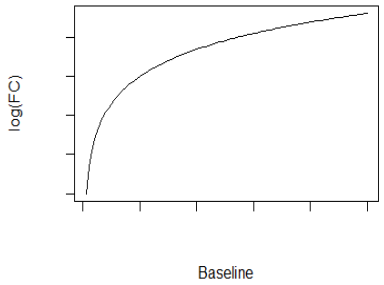
- Ease of modeling
  - Many aspects of modeling can be specified with few lines of code in software
  - MCMC simulation can help get around parameter fitting, degrees of freedom calculation, asymptotic normality
  - Parameter credible intervals have good coverage properties for appropriately chosen priors

- Some things may be known apriori
  - Experience with other drugs in similar class (directionality of covariate effects)
  - Noise level can be established from in-vitro experiments/literature
  - Reported diurnal changes for cytokines: modest ~1.5-2 fold change

Bayesian Emax Model with baseline adjustment

$$\log(FC) = \quad f(dose) + \qquad g(baseline\ covariates) + \qquad noise$$





g() can be a very flexible function modeled with splines

$$Fold\ Change = \quad exp\left(e_0 + \frac{emax * d^h}{ec_{50}^h + d^h}\right) * \exp(g(baseline))$$

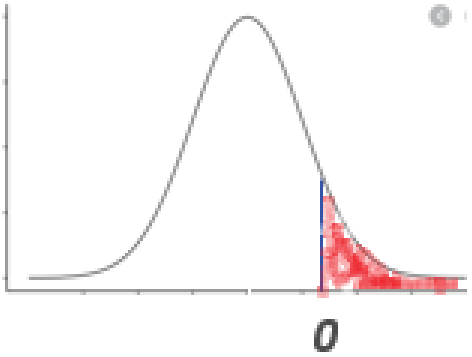
Full prior specification

- le0, lemaxd, emaxb are parameters in the model of log(FC)
- dnorm(0,1.5) means that on the raw scale (FC scale), these parameters can vary from 0.2 – 5.0
- In the dose model, dose is standardized as (dose/ED50)
- The baseline variables are also standardized prior to feeding into the model
- tau1 prior is set so that it can allow a fold change of 0.011-93 in an untreated population

```
#### Below is JAGS code. When dnorm()
## is used to specify a prior, the
## second parameter is precision (1/variance)

## dose model parameters
le0 ~ dnorm(0,1.5)
lemaxd ~ dnorm(0,1.5)


## ed50
ed50d ~ dunif(0,maxd)


### prior for the hill coefficient
lhd ~ dunif(0,1.69)
hd <- exp(lhd)


## baseline effect parameters
for(iter2 in 1:n.adj){
  emaxb[iter2] ~ dnorm(0,1.5)
}


### precision (1/variance) of random noise term
tau1 ~ dgamma(0.5,0.01)

```

Statistics from the Bayesian Emax model

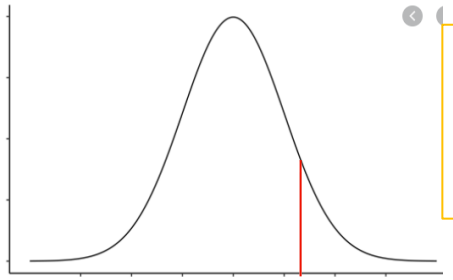
Emax parameter



MCMC approximated distribution

2 x Shaded area = Bayesian equivalent of P-value for dose effect

EC80 (obtained by post-simulation algebra)

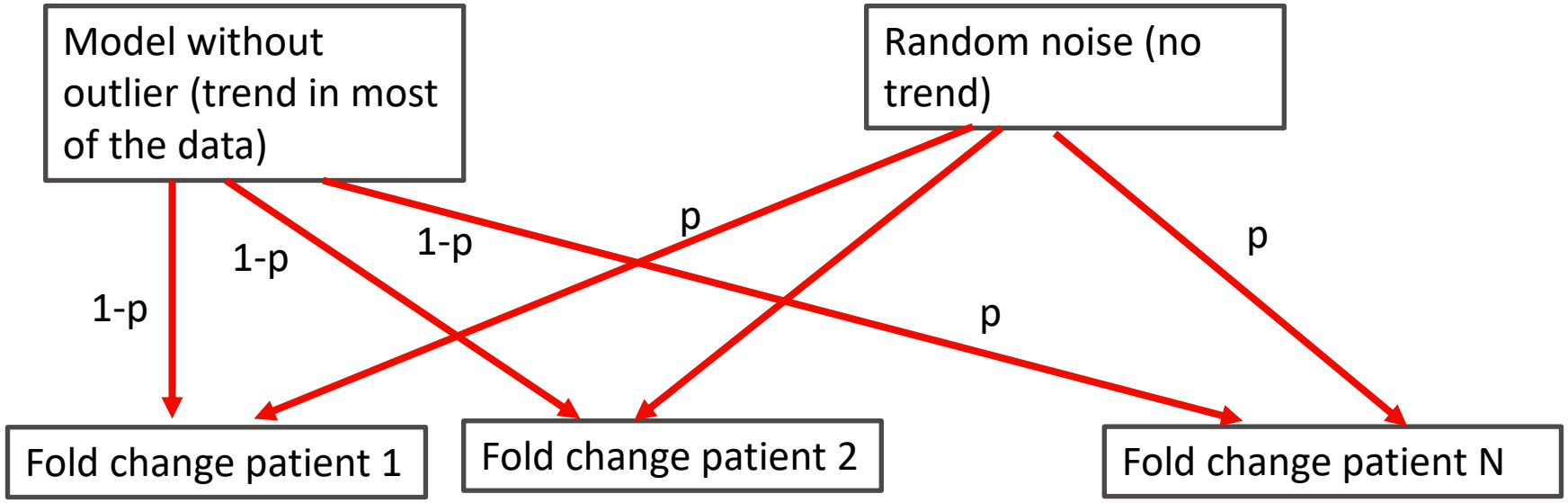


95th percentile

MCMC approximated distribution

Strong evidence of saturation for that marker if 95th percentile is in the range of tested doses

Automated outlier detection and removal via mixture modeling



- Model estimates all the parameters of the overall trend and the probability p simultaneously; p constrained to be less than 2%
- No data point is fully removed from the estimation of the overall trend; a probability of inclusion in the overall trend is assigned to each data point and statistics computed by taking into account those probabilities

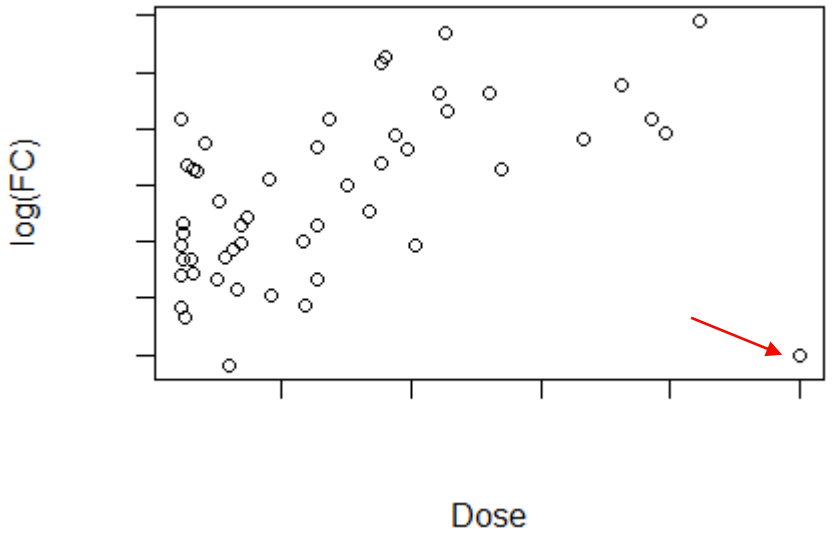
Bayesian non-linear modeling with outlier detection

- Bayesian methodology allows to explicitly model the outlier

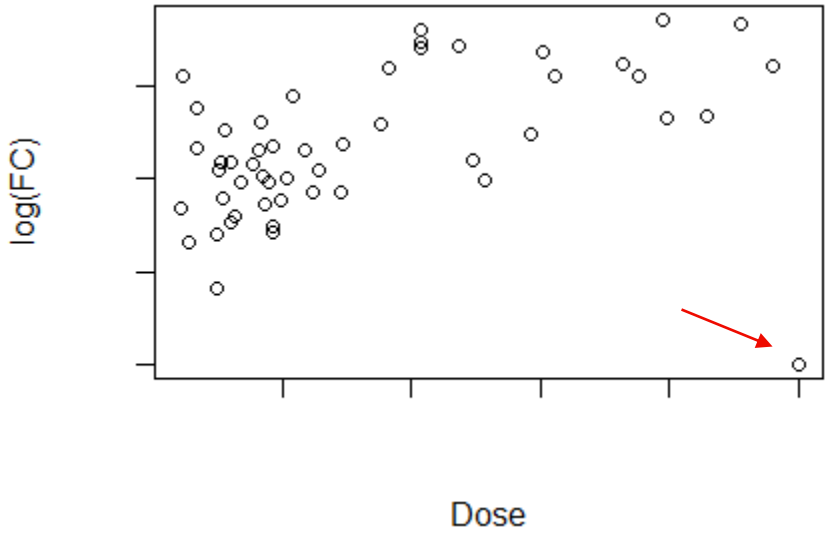
$$\mu_i = f(dose_i) + g(baseline)$$
$$y_i \sim (1-\gamma_i) * N(\mu_i, \sigma) + \gamma_i * N(\lambda, \delta)$$
$$\gamma_i \sim \text{bernoulli(p)}$$
$$p \sim U(0, 0.02)$$

- $f(.)$ could be a complex 4-parameter Emax model
  - Dose and Baseline model to have their own set of parameters
- Bayesian MCMC can be enabled by putting vague priors on the parameters of the Emax models, σ, λ , δ
- Test will be on the Emax coefficient of the dose model
  - This can be obtained directly from MCMC (posterior probability that emax > 0)

Simulated illustration of Bayesian Emax model with outlier detection



Model excludes one outlier with **65%** chance



Model excludes one outlier with **99.7%** chance

Adding random intercept in an Emax model

DoseFinding: Planning and Analyzing Dose Finding Experiments

The DoseFinding package provides functions for the design and analysis of dose-finding experiments (with focus on pharmaceutical Phase II clinical trials). It provides functions for: multiple contrast tests, fitting non-linear dose-response models (using Bayesian and non-Bayesian estimation), calculating optimal designs and an implementation of the MCPMod methodology (Pinheiro et al. (2014) <doi:10.1002/sim.6052>).

Version:	1.0-2
Depends:	lattice, mvtnorm, R (≥ 2.15.0)
Suggests:	numDeriv, Rsolnp, quadprog, parallel, multcomp, ggplot2, knitr, rmarkdown, MASS, testthat
Published:	2021-10-03
Author:	Bjoern Bornkamp [aut, cre], Jose Pinheiro [aut], Frank Bretz [aut], Ludger Sandig [aut]
Maintainer:	Bjoern Bornkamp <bbnkmp at mail.de>

```
data(biom)
## produce first stage fit (using dose as factor)
anMod <- lm(resp~factor(dose)-1, data=biom)
drFit <- coef(anMod)
S <- vcov(anMod)

gsample <- bFitMod(dose, drFit, S, model = "emax",
                   start = c(0, 1, 0.1), nSim = 1000, prior =
```

- bfitMod function has the Bayes implementation
- This uses a two two-stage approach. Can model in covariates (limited) in first step
- Still looking for methods that can add random effects to other parameters

Modeling of censored data

- Data could be censored for multiple reasons
  - BLQ
  - BLPS: ELISA assays can fail an entire plate if too many standards fail

- Easy to specify left censoring in JAGS with a pseudo-data augmentation (*ones* trick)

- Individual observations contribute the following terms to the likelihood
  - $f(x^i)$, if observation is observed with value x^i
  - $F(x^i)$, if left-censored at value x^i
  - $[1-F(x^i)]$, if right-censored at value x^i

- Ones trick: replace censored data with $z^i = 1$ for all censored data indices i

- Then within the likelihood function, just enter the term

$$Z^i \sim \text{dbern}(F(x^i)) \text{ for left-censored}$$
$$Z^i \sim \text{dbern}(1-F(x^i)) \text{ for right-censored}$$

https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-021-04496-8

Testing for Overstimulation (U-shaped effect)

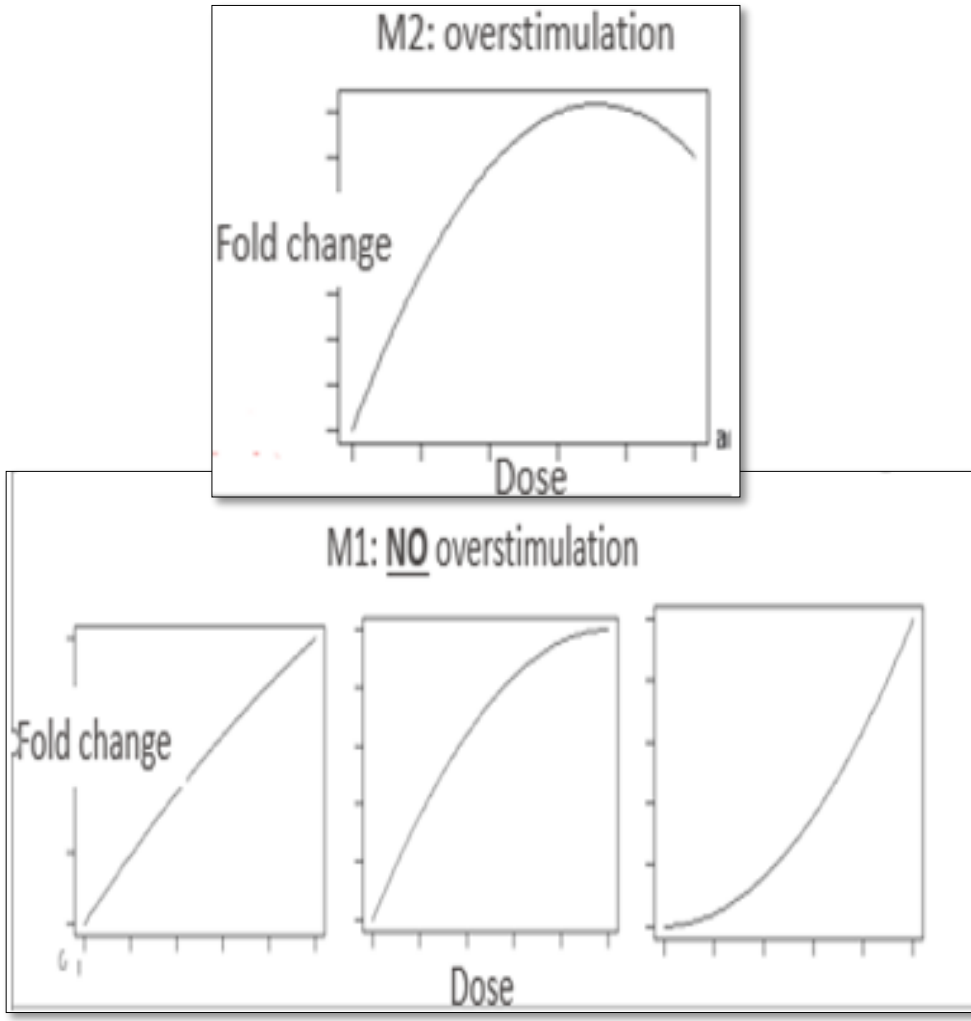
- Cannot just do this in lm() testing for a second order effect
  - You can have a significant second order term but the shape will still be monotonous
  - Need parameter constraints
  - No straightforward way to test this

- Bayesian approach is simple
  - Model the dose response curve as

$$E(y) = -a * dose^2 + b * dose + c$$

  - The peak of this curve occurs at dose=b/2a
  - Just need the posterior probability P(max(tested doses) > b/2a)

- Additional constraints needed: a,b > 0, can be easily incorporated using a log-normal priors for those parameters



Random effects modeling of longitudinal data

- Mis-specification of random effect when it is clearly not gaussian can interfere with the inference on parameters of interest

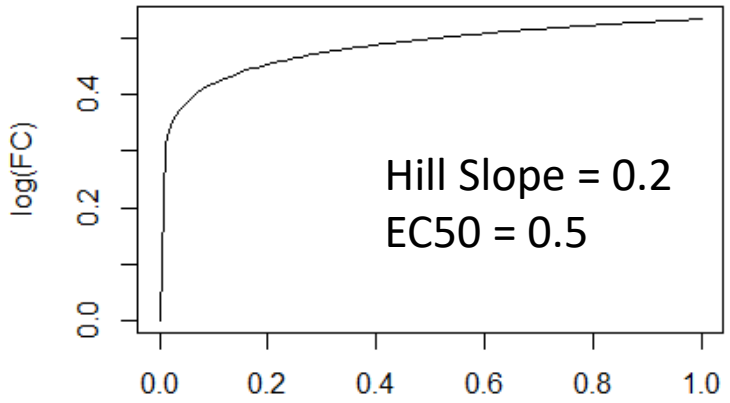
- Two approaches under we are investigating
  - Model random effects with a non-central T
  - Model random effects with a Dirichlet Process

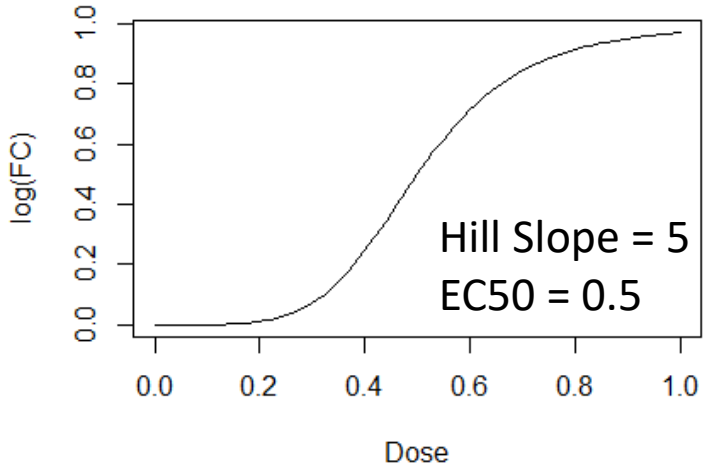
- Non-central T model is a straightforward extension of Gaussian model. Just needs 1 more prior for the non-centrality parameter (later slide)

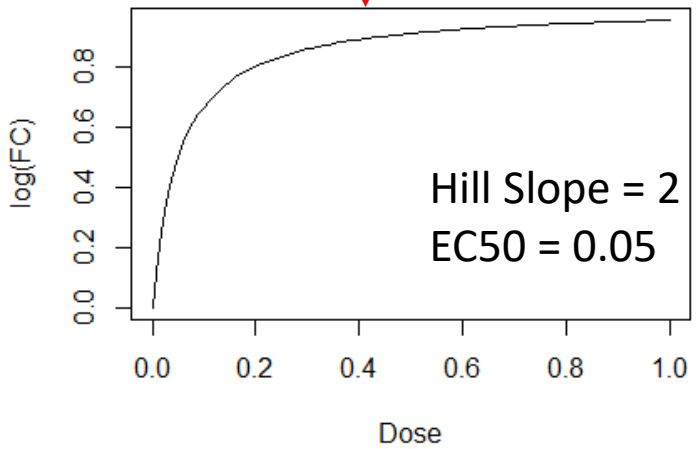
Note on prior selection

- Bayesian theory says priors will get “overwritten” in large data
- However, in small datasets, prior choice can have a considerable effect
- Simulate data from the prior to see if the prior is preferring certain types of shapes over others (see next slide)
- Simulations under realistic parameter assumptions to understand operating characteristics of the prior

Emax model: specifying prior distribution

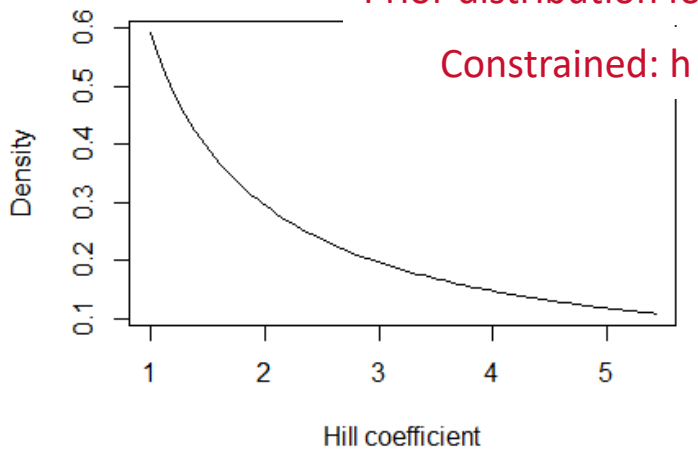




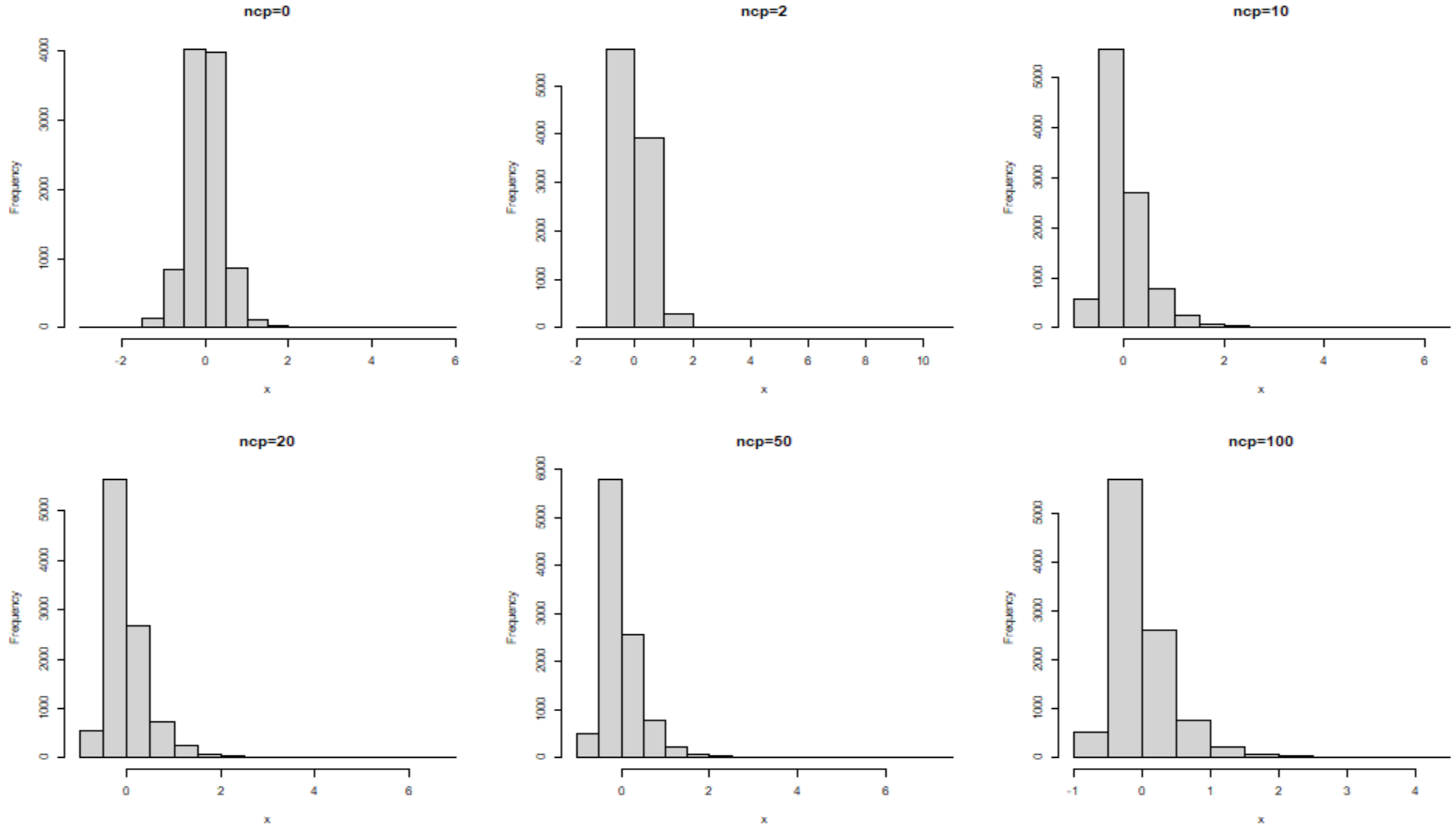


Prior distribution for Hill slope

Constrained: h > 1



Non-central t distribution (scaled/mean-centered)



Shapes are the same for non-centrality parameter > 20; implications for prior setting

Summary

- Biomarkers are becoming of increasing importance in IO development. New technologies have opened unprecedented opportunities to study patient biology from gene expression, protein and cellular level simultaneously

- Rshiny or other visualization tools are of limited value. There is a need to resource teams with statisticians who can bring quantitative decision making to the teams

- Many modeling challenges. There is no time to review diagnostic plots to assess model fit, confirm model assumptions etc. Need smart algorithms that can pick out the most meaningful signals and also provide statistical inference

- Naturally lend opportunities for Bayesian modeling. Availability of Bayesian modeling software can enable wide use of these models without needing to spend significant resources on software development

BACKUP