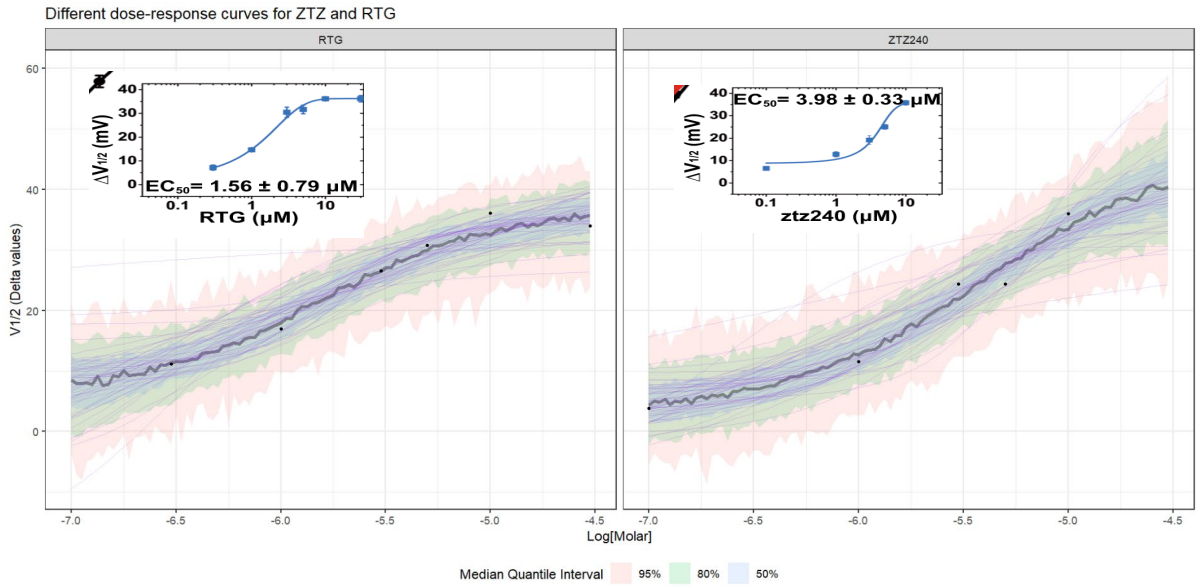
Rotation project

Bayesian approach: Deeper Insights into Uncertainty



Maomlab (computational pharmacology)
6th March-27th April 2023

NILOUFAR POUYAN

PHARMD, PH.D. STUDENT

PROGRAM IN BIOMEDICAL SCIENCES (PIBS)

UNIVERSITY OF MICHIGAN, ANN ARBOR

Cell Research

Article | Published: 03 September 2020

Molecular basis for ligand activation of the human KCNQ2 channel

Xiaoxiao Li, Qiansen Zhang, Peipei Guo, Jie Fu, Lianghe Mei, Dashuai Lv, Jiangqin Wang, Dongwu Lai, Sheng Ye, Huaiyu Yang ✉ & Jiangtao Guo ✉

Cell Research **31**, 52–61 (2021) | Cite this article

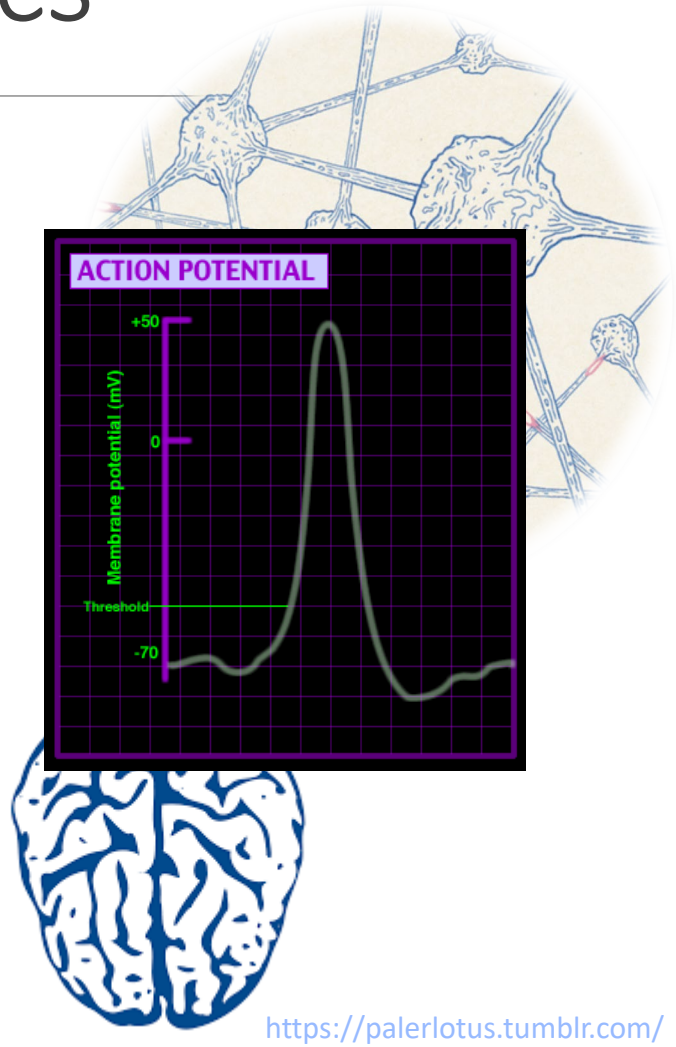
Presentation outline

1. Introduction & Basics ~10 min
   a) Action potential, patch clamp
   b) Li et al. 2021 paper overview
2. Aim & Problem statement ~10min
   a) Bayesian approach (basics) ~5min
3. Analysis workflow ~15 min

Neuronal communication: basics

- Action potential (AP):
  - ✓ The Language of Neuronal Communication
  - ✓ A rapid and transient change in voltage that occurs in response to a stimulus.

- Essential for understanding the function of the nervous system
- Disruptions or abnormalities in action potentials can contribute to various disorders → Drug development



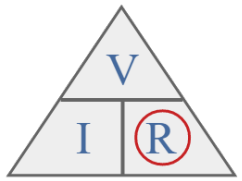
https://palerlotus.tumblr.com/

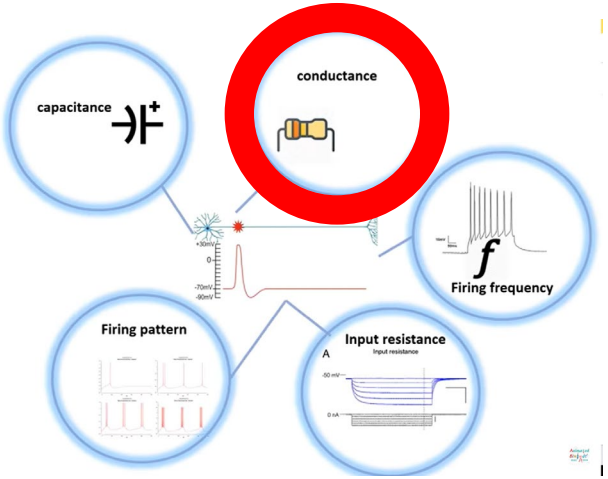
Electrophysiological (ephys) measurement of AP

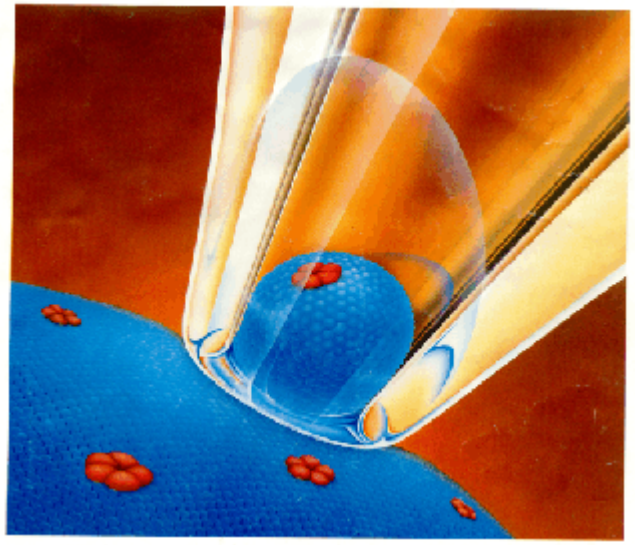
- **Patch clamp recording:** is a technique for studying ion channels in cells, which involves placing a small glass pipette onto the cell membrane to record ion channel activity

- Measuring action potential helps infer Neurons' Biophysiological properties :

I=Current
V=Voltage







https://sites.oxy.edu/linden/Cogsci320s11/patchclamp.htm

- **Clamp type:** which parameter we are controlling (keeping constant)
  - Voltage clamp
  - Current clamp

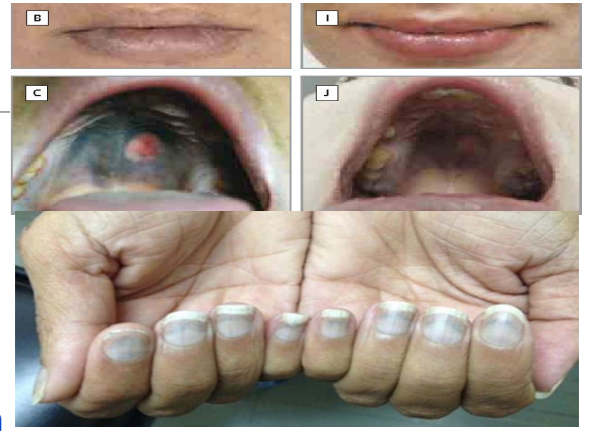
G=I/R

G=I/V

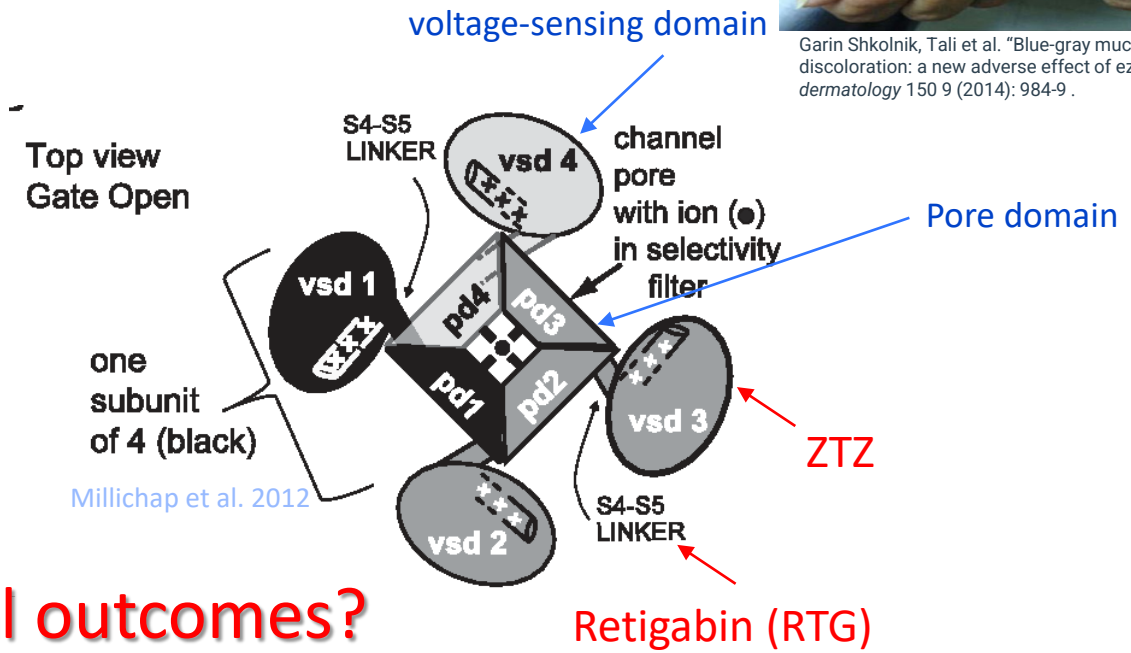
Conductance refers to the ease with which ions can move through the channel pore in response to an applied voltage gradient

Molecular basis for ligand activation of the human KCNQ2 channel. Li et al., 2021

- The voltage-gated potassium channel KCNQ2 is responsible for M-current (potassium current) in neurons and is an important drug target to treat epilepsy, pain and several other diseases related to neuronal hyper-excitability
- Retigabin (RTG)
  - Targets S4-S5 linker
  - Serious ADR → Discontinued
- ZTZ
  - Targets Voltage-sensing domain (Gao et al., 2010)



Garin Shkolnik, Tali et al. "Blue-gray mucocutaneous discoloration: a new adverse effect of ezogabine." *JAMA dermatology* 150 9 (2014): 984-9 .



Millichap et al. 2012

Differential functional outcomes?

Structure-function characterization:

- Measurements:

  1. Activator-bound structure (cryogenic electron microscopy (cryo-EM) → **high-quality atomic resolution structures** of the KCNQ2 ion channel in complex with ZTZ and RTG.

  2. Ephys measurement to compare the effects
     - ✓ Whole-cell/voltage patch clamp technique
     - ✓ Measured conductance (G):

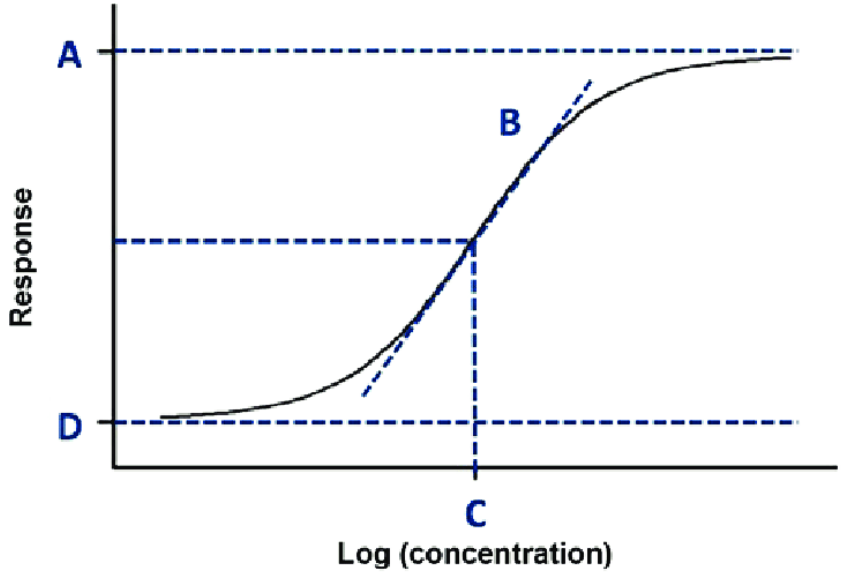
     1. Conductance – voltage relationship (G-V)
     2. The voltage-drug relationship

Sigmoid curve fitting

A. Top ()
B. Slope (hill coeff)
C. Ec50
D. Bottom

1. The **Conductance** – **voltage** relationship (G-V)
2. The **voltage**-**drug** relationship

▪The response of a biological system to increasing doses of a drug or other intervention shows a characteristic **S-shaped curve** (sigmoid shape) → **Dose**-**response** **curve**



Y=Bottom + (Top-Bottom)/(1+10^(LogEC50-X))

Data analysis:

$\Delta V_{1/2} = V_{1/2}$ in control – $V_{1/2}$ in the presence of ztz240

Y=Bottom + (Top-Bottom)/(1+10^(LogEC50-X))

- **G-V curve fit:** Boltzmann equation

G = Gmin + (Gmax - Gmin) / (1 + exp(V - V1/2) / S)

Gmin: Minimum conductance

Gmax: Maximum conductance

S: is the slope factor.

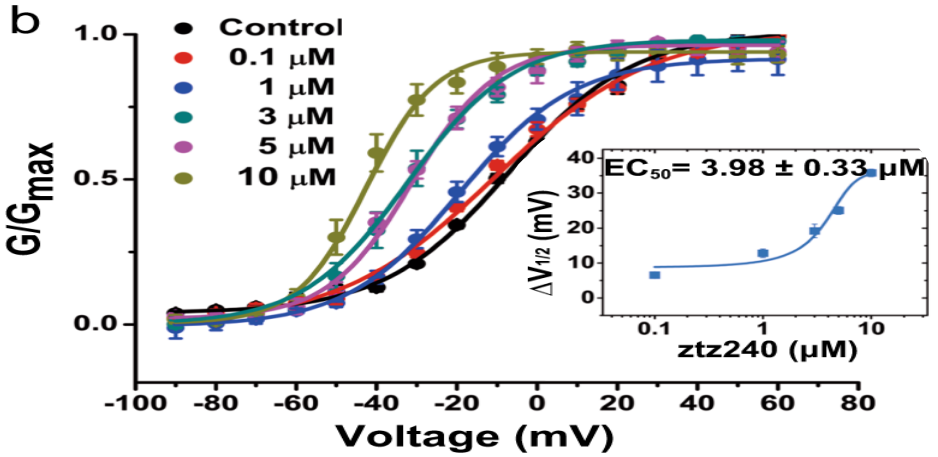
V1/2: Voltage at which 50% of maximum conductance is achieved

- **Dose-response curve:** Hill equation

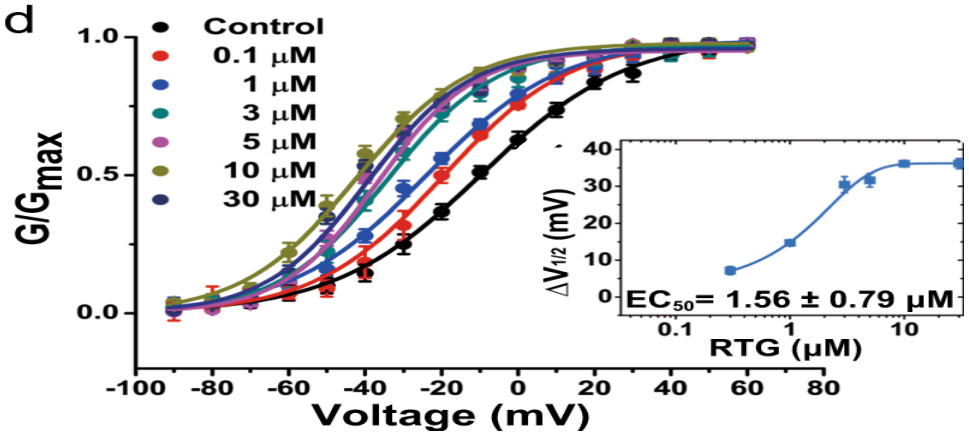
E = Emax / (1 + 10 ^ ((EC50 - C) * P))

Emax: maximum response
C : Drug concentration
PL Hill coefficient
EC50: Drug concentration producing half of the maximum response

ZTZ240



RTG



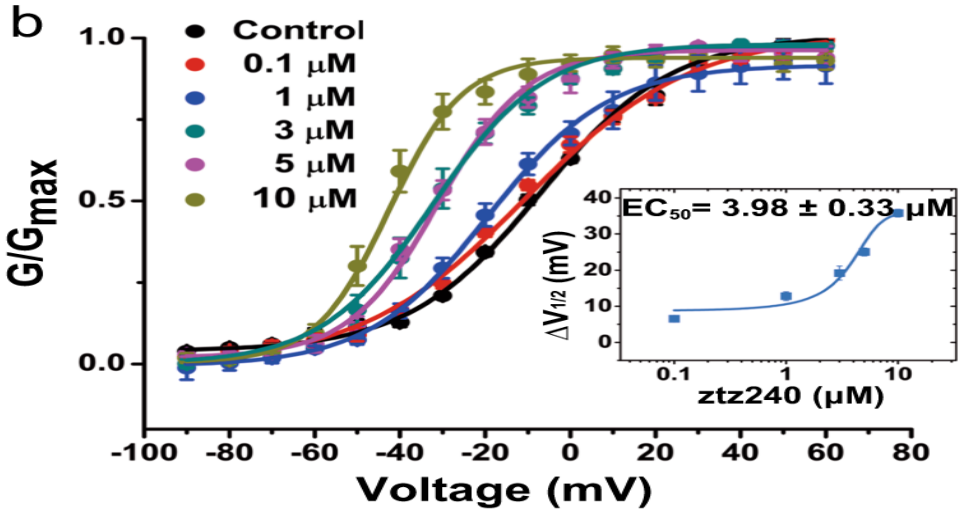
Data Interpretation

- Left shift of G-V → enhances the voltage sensitivity of KCNQ2

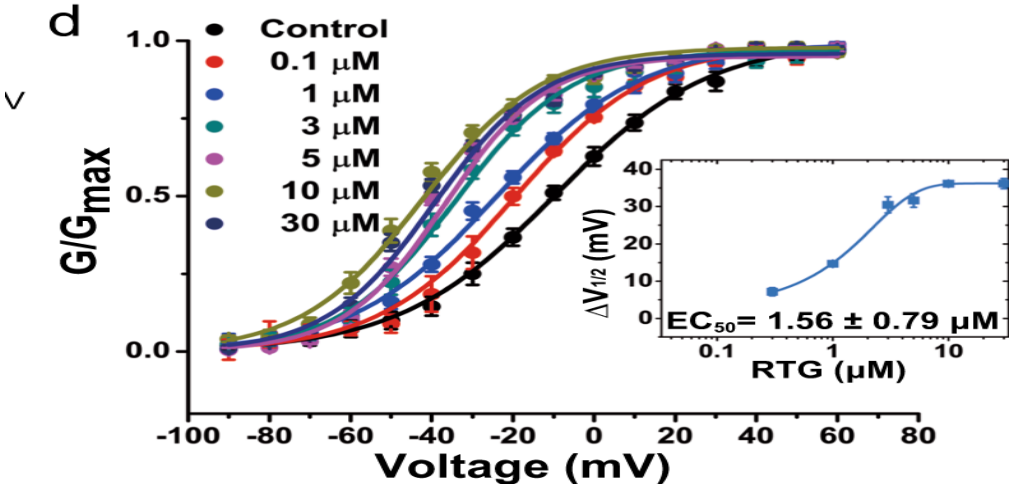
- Different slopes : different V1/2
- Data in conjunction of the structural analysis eventually suggested that these agents RTG and ztz240 may have different activation mechanisms on KCNQ2 gating.

Does means ± SEM offer a reliable and true measure of uncertainty?

ZTZ240



RTG



Parameters are presented by means ± SEM

Measuring uncertainty

✓ **Crucial:** A measure of the confidence& reliability in the results of a measurement or calculation.

❖ Traditional approaches to statistical inference

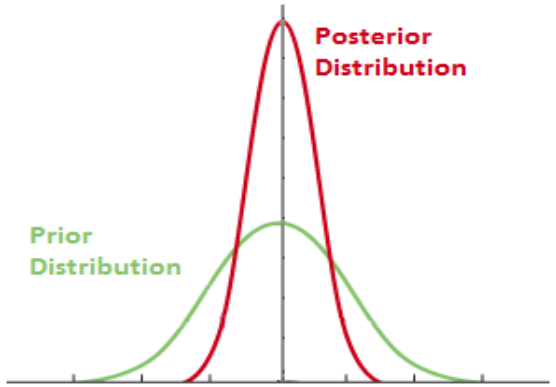
- Often rely on **Fixed:**
  - Point estimates of parameters (EC50)
  - Fixed statement of uncertainty of those parameters (SEM)
- Assume true parameter value is fixed and known, and that any variation in estimates is due to random sampling error:

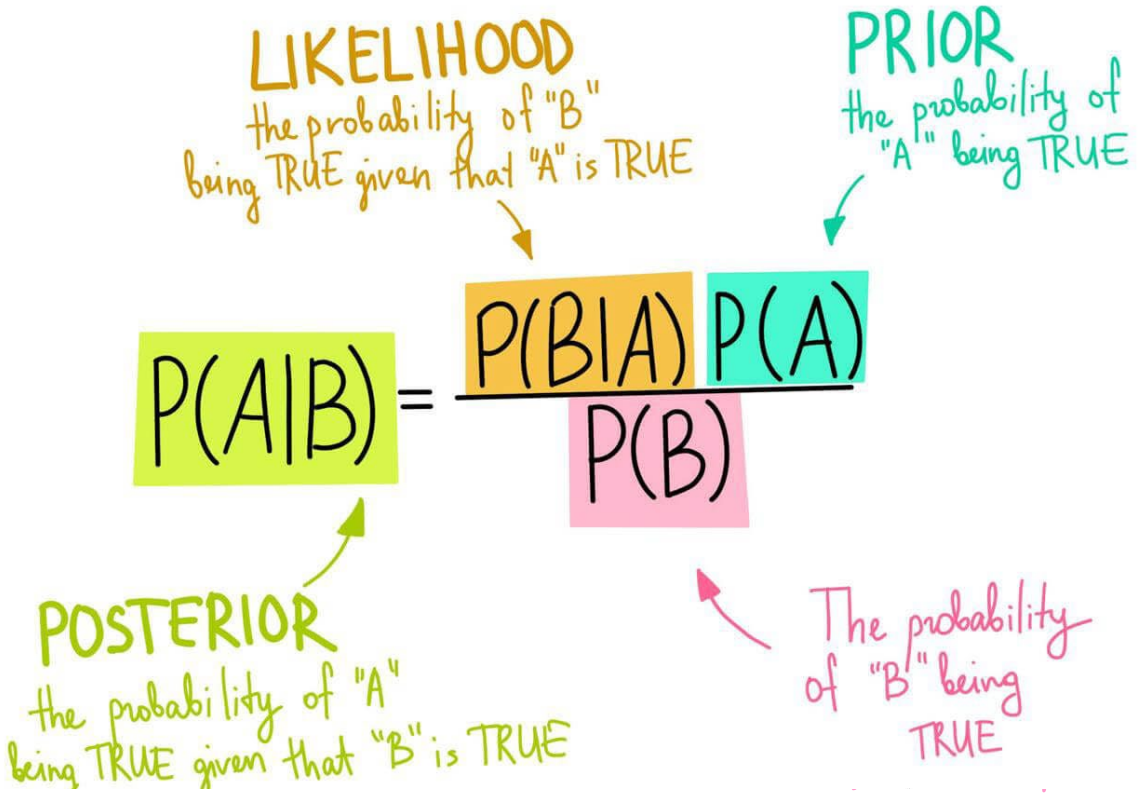
Unrealistic for real-world data (sampling bias/errors) → Assumption violation?

- Incorrect conclusions or decisions based on flawed data → Overconfidence?
- Bayesian approaches: prior knowledge and incorporate uncertainty in the analysis

Bayes theorem: general overview

- Bayes theorem
  - ✓ Instead of point estimates→ Assigns probabilities to events
  - ✓ Updates **our beliefs** (expert opinion /prior knowledge) about the probability of an event using new information

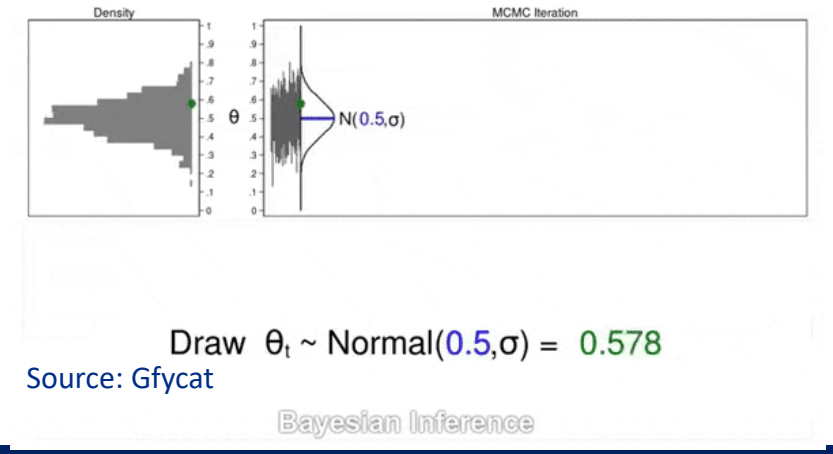




Bayesian approach: define and fit a probabilistic model

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}.$$

- Ingredients
  - ➢ Data
  - ➢ Priors: the information model has before seeing the data
    - ✓ Prior distribution over a set of parameters.
  - ✓ Sampling from the posterior distribution : Markov Chain Monte Carlo (MCMC)sampling



Source: Gfycat

- **Init** is used to specify the initial values of the parameters that the MCMC algorithm starts with.

- Inference from MCMC output
  - ✓ Checking **convergence** (output is stable)
  - ✓ Independent of the initial values (init) and has an adequate number of samples.

General steps:

1. **Model specification:**
   - ✓ Prior selection and check
   - ✓ Init and MCMC features
2. **Model fitting**
3. **Model check**

Data preparation and visualization

	ztz240	X	Voltage	X.2	X.3	X.4	X.5	X.6	X.7	X.8	run1	run2	run3	run4	run5	X.14
1	NA	10?MZTZ240	voltage	conductance: control	NA	NA	NA	NA	NA	NA	Conductance:10 ?M ZTZ240	NA	NA	NA	NA	NA
2	NA		-90	0.040264	0.063194	0.019954	0.04040334	0.01752997	NA	NA	0.022528	0.048057	0.051135	-0.01631143	-0.013289100	NA
3	NA		-80	0.046872	0.068207	0.034528	0.05741659	0.03390494	NA	NA	0.015431	0.003098	0.044276	-0.01135711	0.005603406	NA
4	NA		-70	0.073298	0.104097	0.026942	0.02267043	0.04621841	NA	NA	0.018977	0.056488	0.054565	0.03965878	0.039010960	NA
5	NA		-60	0.057247	0.108639	0.032161	0.03340073	0.04702025	NA	NA	0.053645	0.107997	0.192728	0.05311240	0.071035130	NA
6	NA		-50	0.090069	0.140648	0.058697	0.03158382	0.06090859	NA	NA	0.230904	0.253305	0.539853	0.25732510	0.222396300	NA
7	NA		-40	0.09543	0.179261	0.104180	0.08220429	0.07798038	NA	NA	0.49768	0.520286	0.831337	0.48989620	0.616456300	NA
8	NA		-30	0.195874	0.268817	0.189531	0.16539110	0.14111790	NA	NA	0.628553	0.734518	0.922009	0.71242080	0.873794000	NA
9	NA		-20	0.309978	0.363838	0.357675	0.32487890	0.32928450	NA	NA	0.728599	0.804861	0.922214	0.79550400	0.921583600	NA
10	NA		-10	0.425608	0.524267	0.574629	0.45725310	0.45729180	NA	NA	0.789736	0.864506	1.019162	0.79300110	0.976433200	NA

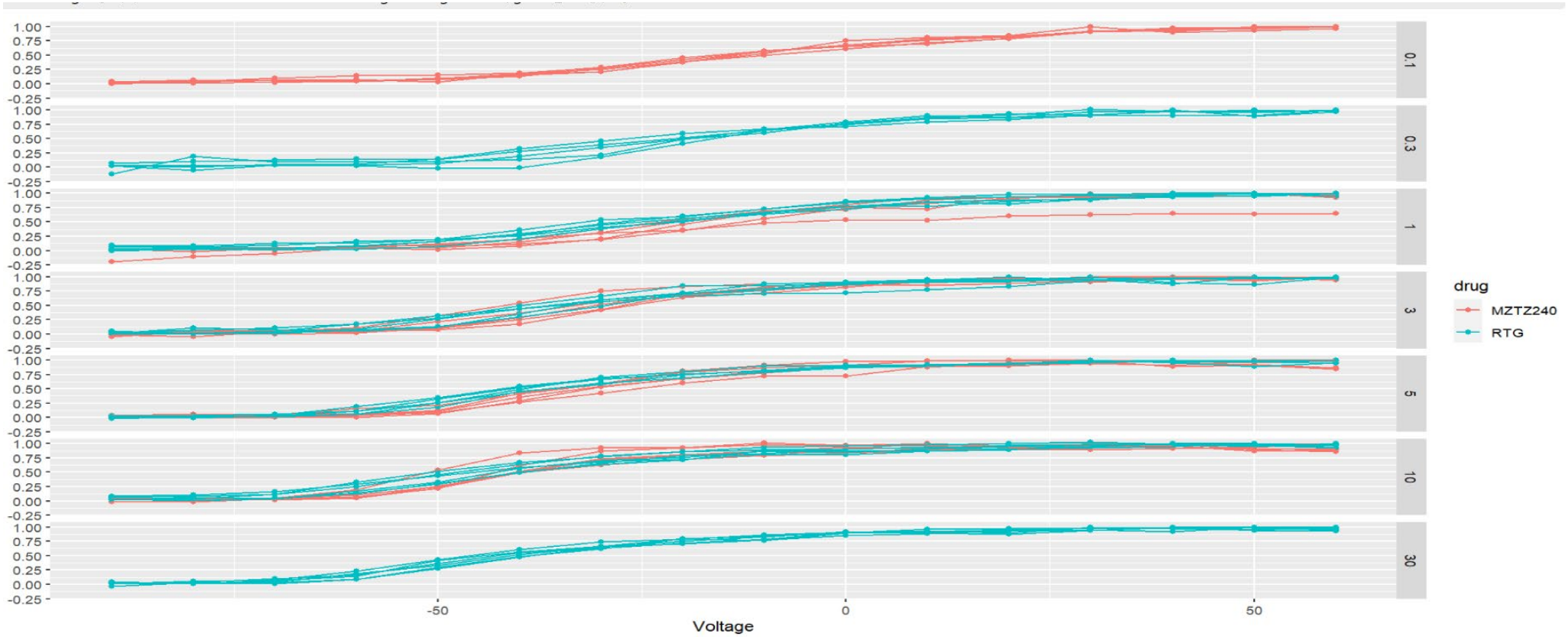
	RTG	Voltage	control	X	X.1	X.2	X.3	X.4	X.5	X.6	X.7	run1	run2	run3	run4	run5	run6
1		-90	0.007921	0.026723	0.078453	0.010713	0.031323	0.063989	NA	NA	NA	0.022152	0.011134	0.012763	0.006849	-0.036020	0.041841
2		-80	0.079208	0.035162	0.109316	0.003964	0.043949	0.065248	NA	NA	NA	0.03481	0.024291	0.011002	0.053788	0.027811	0.011245
3		-70	0.017822	0.021097	0.149069	0.034104	0.045858	0.111638	NA	NA	NA	0.098101	0.040486	0.035293	0.078222	0.005862	0.064340
4		-60	-0.0297	0.039381	0.141880	0.029178	-0.025830	0.076874	NA	NA	NA	0.164557	0.183198	0.083967	0.235844	0.084234	0.143325
5		-50	0.051485	0.063291	0.177707	0.043222	0.045446	0.097680	NA	NA	NA	0.360759	0.330972	0.293093	0.433397	0.270407	0.415935
6		-40	0.065347	0.108298	0.247751	0.051977	0.056101	0.130080	NA	NA	NA	0.553797	0.520243	0.484761	0.603532	0.473152	0.567138
7		-30	0.130693	0.151899	0.317742	0.132778	0.154709	0.253883	NA	NA	NA	0.626582	0.657895	0.624189	0.735312	0.656394	0.666860
8		-20	0.320792	0.284107	0.419695	0.305560	0.257153	0.378905	NA	NA	NA	0.723101	0.711538	0.800325	0.779725	0.755368	0.791877
9		-10	0.40396	0.427567	0.537552	0.508494	0.456400	0.468623	NA	NA	NA	0.783228	0.768219	0.838313	0.862408	0.833279	0.766537
10		0	0.586139	0.603376	0.659796	0.667787	0.630032	0.628538	NA	NA	NA	0.901899	0.850202	0.893969	0.907853	0.920342	0.896732

Data frame preparation: tidyr &dplyr

Voltage	dose	drug	run number	Acquisition	response	log_dose	drug_dose
50	0.3	RTG	run1	0.904251	0.904251	50	0.3
50	0.3	RTG	run2	0.891848	0.891848	50	0.3
50	0.3	RTG	run3	1.000000	1.000000	50	0.3
50	0.3	RTG	run4	0.955212	0.955212	50	0.3
50	0.3	RTG	run5	0.967095	0.967095	50	0.3
60	0.3	RTG	run1	1.000000	1.000000	60	0.3
60	0.3	RTG	run2	0.966458	0.966458	60	0.3
60	0.3	RTG	run3	0.977570	0.977570	60	0.3
60	0.3	RTG	run4	1.000000	1.000000	60	0.3
60	0.3	RTG	run5	1.000000	1.000000	60	0.3
-90	10	MZTZ240	run1	0.022528000	0.022528000	-90	10
-90	10	MZTZ240	run2	0.048057000	0.048057000	-90	10
-90	10	MZTZ240	run3	0.051135000	0.051135000	-90	10
-90	10	MZTZ240	run4	-0.016311430	-0.016311430	-90	10
-90	10	MZTZ240	run5	-0.013289100	-0.013289100	-90	10
-80	10	MZTZ240	run1	0.015431000	0.015431000	-80	10
-80	10	MZTZ240	run2	0.003098000	0.003098000	-80	10
-80	10	MZTZ240	run3	0.044276000	0.044276000	-80	10
-80	10	MZTZ240	run4	-0.011357110	-0.011357110	-80	10
-80	10	MZTZ240	run5	0.005603406	0.005603406	-80	10

Head and tail of the merged data frame

Preliminary data visualization: ggplot2



Model specification:
BayesPharma::sigmoid_agonist_model[1]

```
function (data, formula = sigmoid_agonist_formula(), prior =
sigmoid_agonist_prior(),
    init = sigmoid_agonist_init(), iter = 8000, control =
list(adapt_delta = 0.99),
    stanvar_function = sigmoid_stanvar, expose_functions = TRUE,
    ...)
```

Prior selection and check

Init and MCMC features

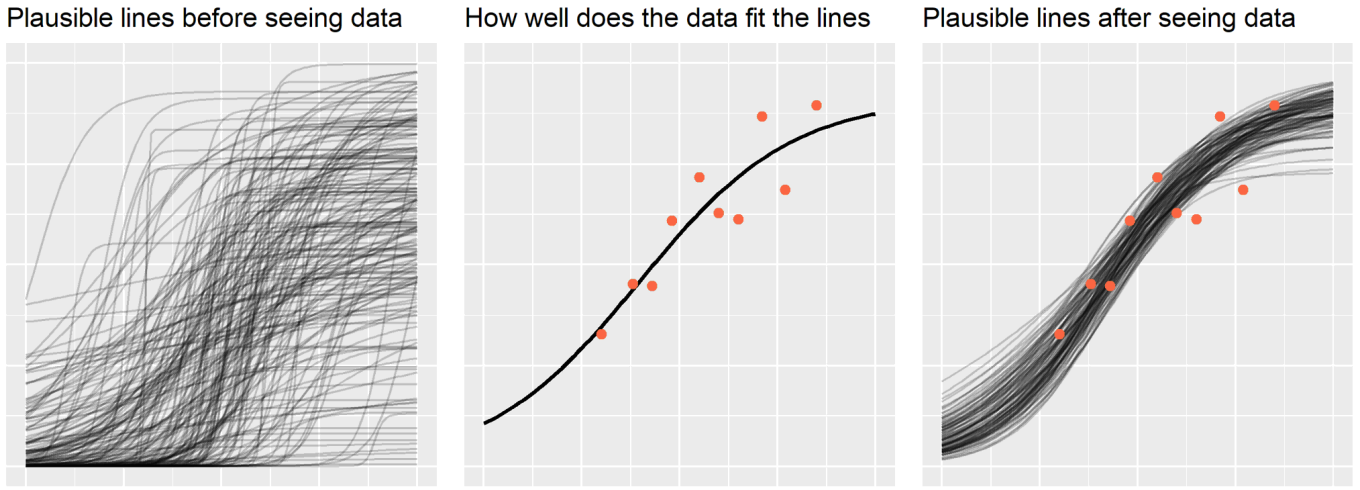
```
Formula: response ~ sigmoid(ec50, hill, top, bottom, log_dose)
```

E = bottom + (top - bottom) / (1 + 10 ^ ((EC50 - X) * Hill))

1. Martin, M., & O'Meara, M. (2023). BayesPharma: Tools for Bayesian Analysis of Non-Linear Pharmacology Models [Manual]. Retrieved from https://api.github.com/repos/maomlab/BayesPharma/pages

Model specification: Hill function parameters

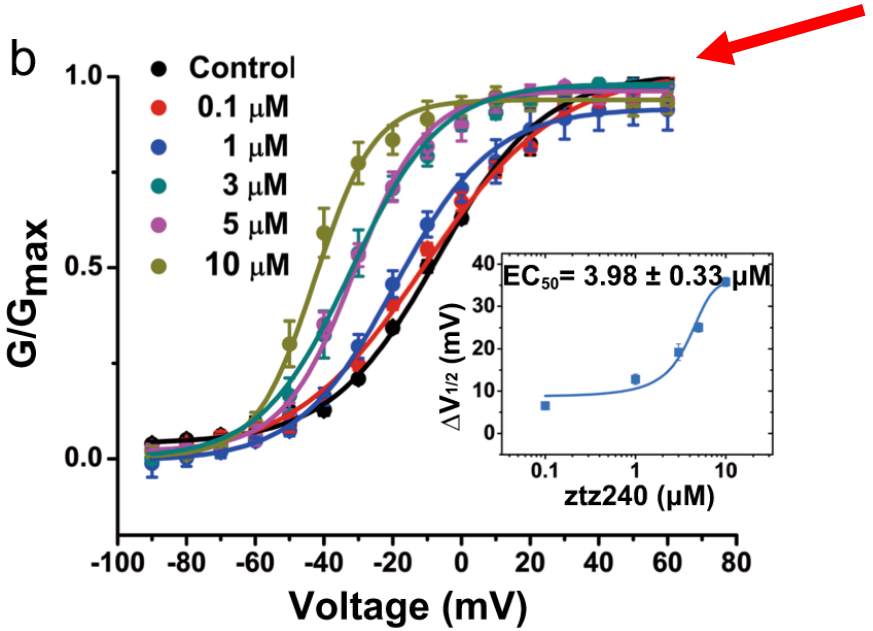
Prior selection



$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}.$$

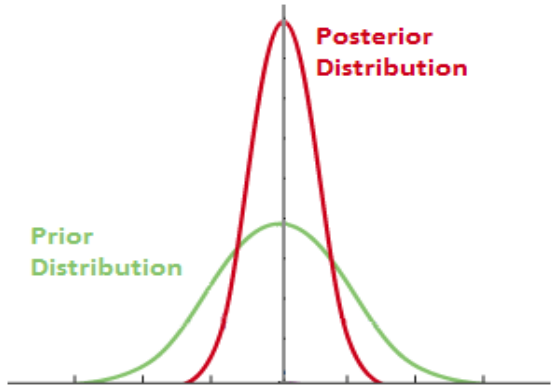
Source: https://www.tjmahr.com/bayes-theorem-in-three-panels/

G-V curve fitting



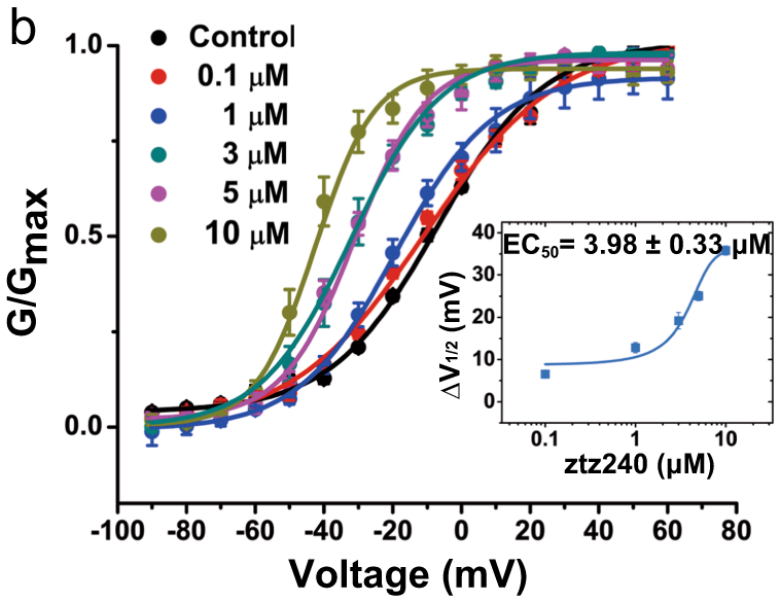
Model specification: Prior selection

1. Prior: A probability distribution that reflects **our uncertainty** about the value of a parameter **before we observe any data**
2. Good priors:
    - ✓ Being **informative enough** to guide the analysis towards a plausible parameter space, while also **not being too restrictive** to allow for unexpected or rare outcomes.
    - ✓ Compatible with domain expertise
    - ✓ Broad priors represent unbiased uncertainty



Model specification: Prior selection

Para	Default values : **BayesPharma:: sigmoid_agonist_prior**	**New values** `BayesPharma::sigmoid_agonist_prior`
ec50	normal(-6, 2.5)	normal(-20, 40)
hill	(1, 1)	-
Top	(1, 0.5)	1
Bottom	(0, 0.5)	-



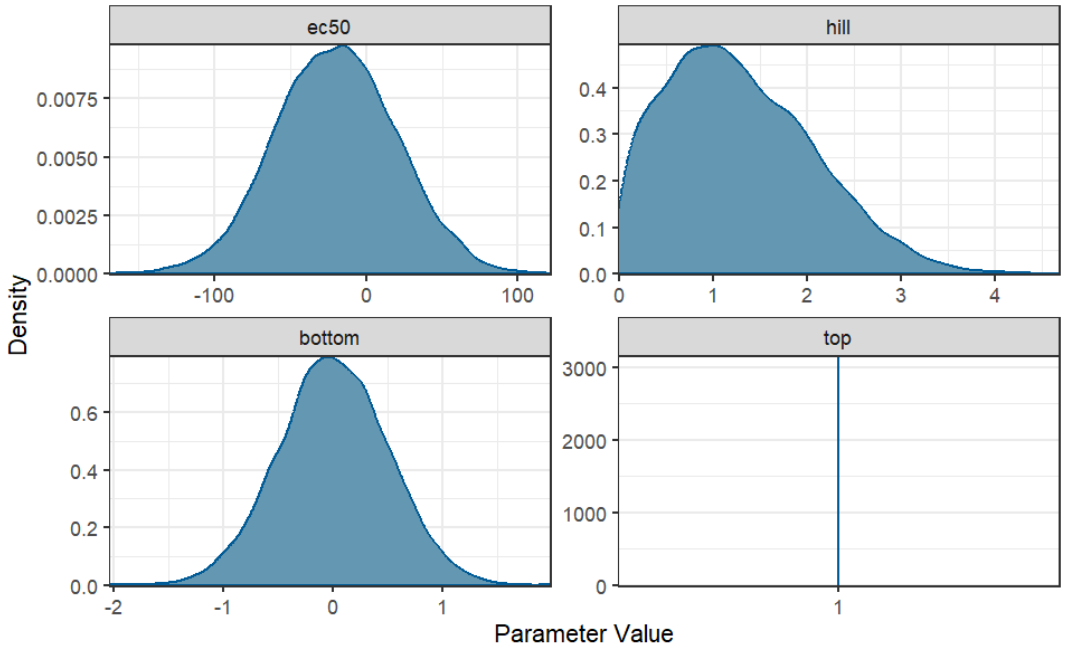
Model specification: Prior predictive checks plausible range?

```
kor_sample_prior <- BayesPharma::sigmoid_agonist_model(
data = DF |> dplyr::select(drug, log_dose, response),
prior = BayesPharma::sigmoid_agonist_prior(top = 1,
ec50 = brms::prior(normal(-20, 40), nlpar = "ec50")),
init=BayesPharma::sigmoid_agonist_init(ec50 = -20),
sample_prior = "only")
```

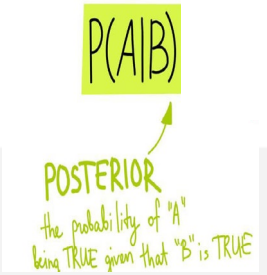
Parameter	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	2.80	4.08	0.09	10.37	1.00	8535	5715

Residuals

MCMC convergence



Model fitting




```{r Model2_with_DF}
Model2<-BayesPharma::sigmoid_agonist_model (
  data=DF,
  #fitting the model for each dose and drug
  formula=BayesPharma::sigmoid_agonist_formula(predictors= 0+ dose:drug),
  #we need to override the default prior for the EC50 to encompass the corresponding values of Voltage instead of drug dose
  prior=BayesPharma::sigmoid_agonist_prior(
    ec50 = brms::prior(normal(-20, 40), nlpar = "ec50")),
  # we have to change the initial if we look at it is still at the last default point of -9
  # we need to switch it to the new one: -20
  init = BayesPharma::sigmoid_agonist_init(ec50 = -20)
)
```

```
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: response ~ sigmoid(ec50, hill, top, bottom, log_dose)
         ec50 ~ 0 + dose:drug
         hill ~ 0 + dose:drug
         top ~ 0 + dose:drug
         bottom ~ 0 + dose:drug
   Data: data (Number of observations: 912)
  Draws: 4 chains, each with iter = 8000; warmup = 4000; thin = 1;
         total post-warmup draws = 16000

Population-Level Effects:
                           Estimate Est.Error l-95% CI u-95% CI Rhat
ec50_dose0.1:drugMZTZ240     -11.63      1.47   -14.48    -8.70 1.00
```

ec50 is allowed to vary based on the interaction between **dose** and **drug**, with no intercept. This is equivalent to fitting a separate **ec50** value for each unique combination of **dose** and **drug**.

Model check

Population-Level Effects:	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
ec50_dose0.1:drugMZTZ240	-11.63	1.47	-14.48	-8.70	1.00	18543	12079
ec50_dose0.3:drugMZTZ240	-19.70	39.57	-98.13	58.93	1.00	28936	11092
ec50_dose1:drugMZTZ240	-19.31	1.25	-21.78	-16.87	1.00	18669	12626
ec50_dose3:drugMZTZ240	-32.17	1.08	-34.29	-30.10	1.00	16983	12784
ec50_dose5:drugMZTZ240	-32.14	0.98	-34.07	-30.23	1.00	19445	13085
ec50_dose10:drugMZTZ240	-43.78	0.83	-45.44	-42.16	1.00	19474	12857
ec50_dose30:drugMZTZ240	-20.08	40.03	-98.33	57.85	1.00	29106	11358
ec50_dose0.1:drugRTG	-19.97	39.91	-98.64	58.86	1.00	28799	11674
ec50_dose0.3:drugRTG	-18.91	1.21	-21.29	-16.49	1.00	19811	13130

Parameter	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	0.06	0.00	0.06	0.06	1.00	27296	11444

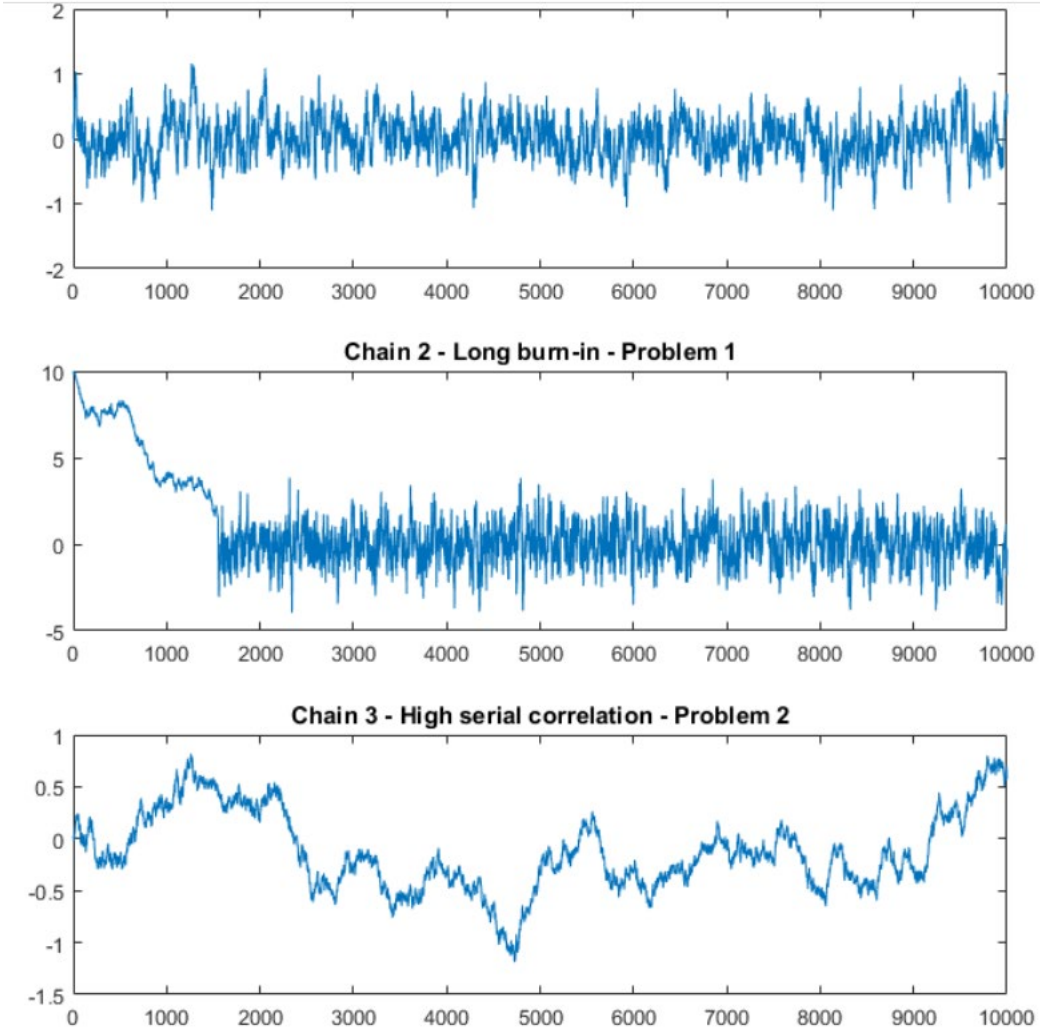
Residuals

MCMC convergence

Model check: Trace plots

- The x-axis: iteration number
- The y-axis: parameter value.
- Each line in the plot : the parameter values for a single chain.
- Problems:
- lack of convergence?
- Autocorrelation? (lack of independence)

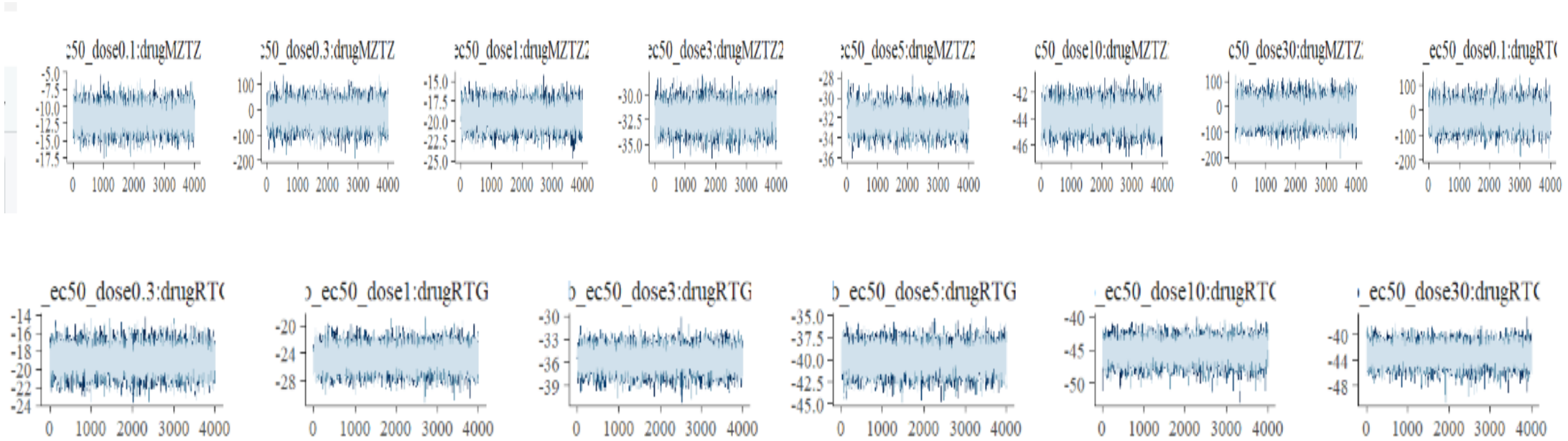
Fat, hairy caterpillar :)



https://www.statlect.com/fundamentals-of-statistics/Markov-Chain-Monte-Carlo-diagnostics

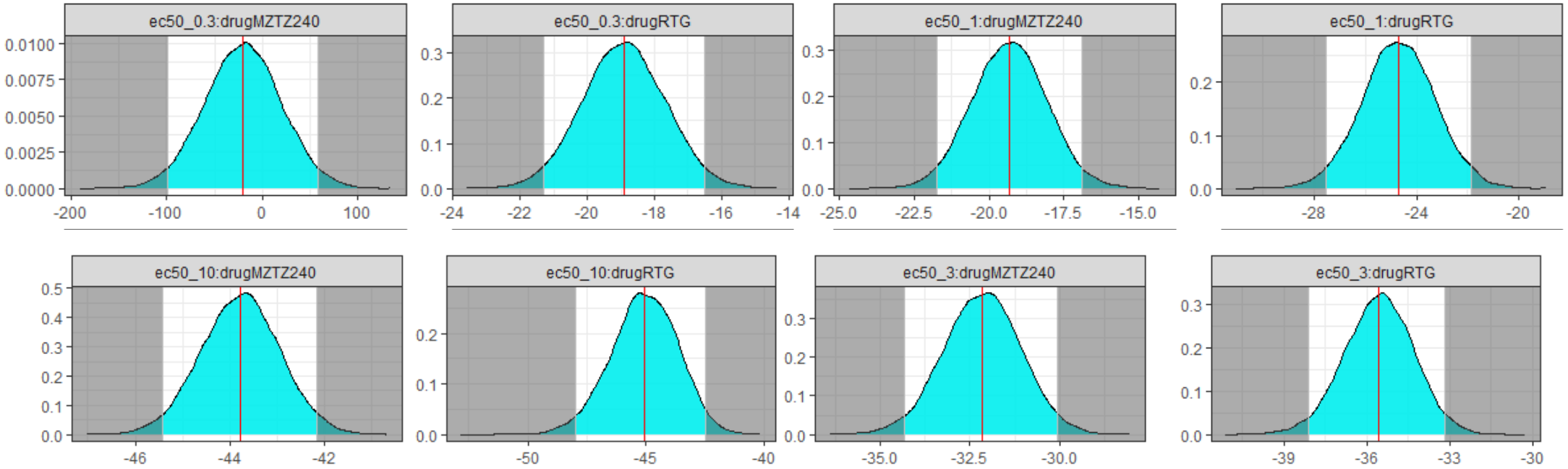
Trace plots:

`mcmc_trace`



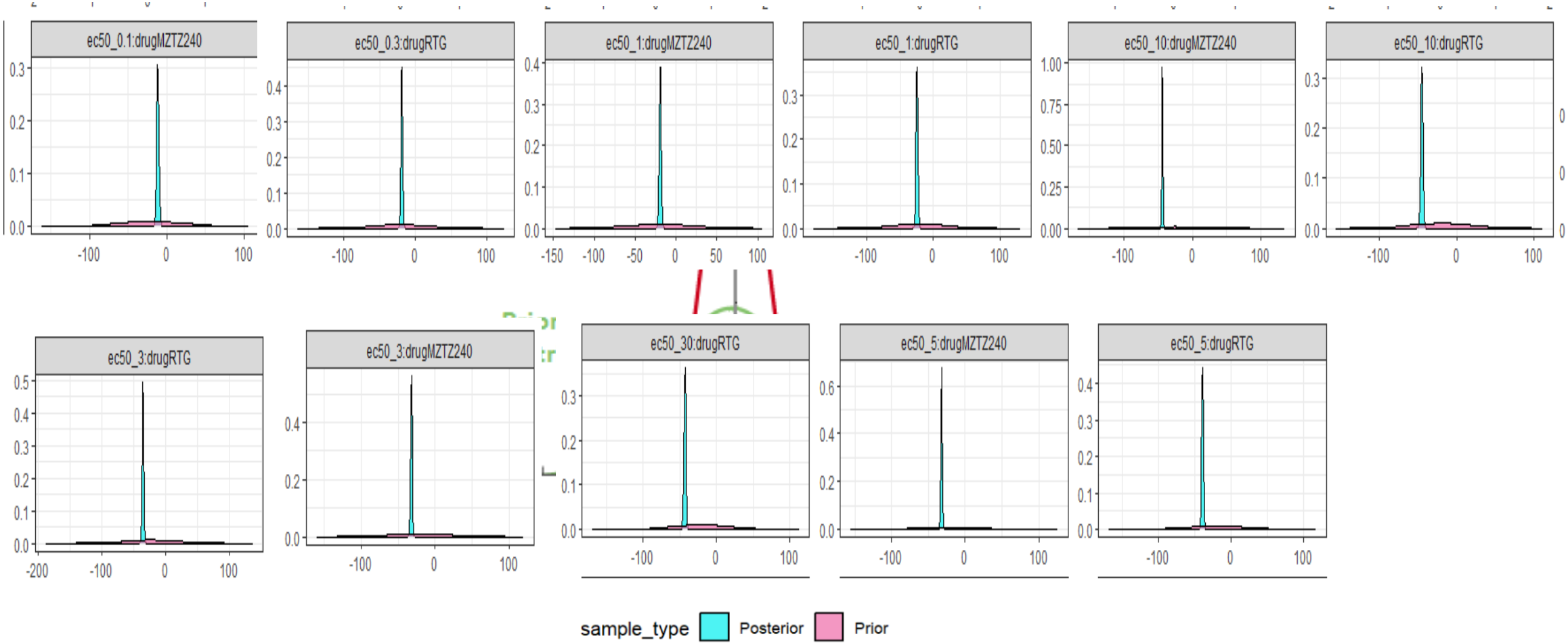
Model check: Posterior density plots

`posterior_densities_plot`

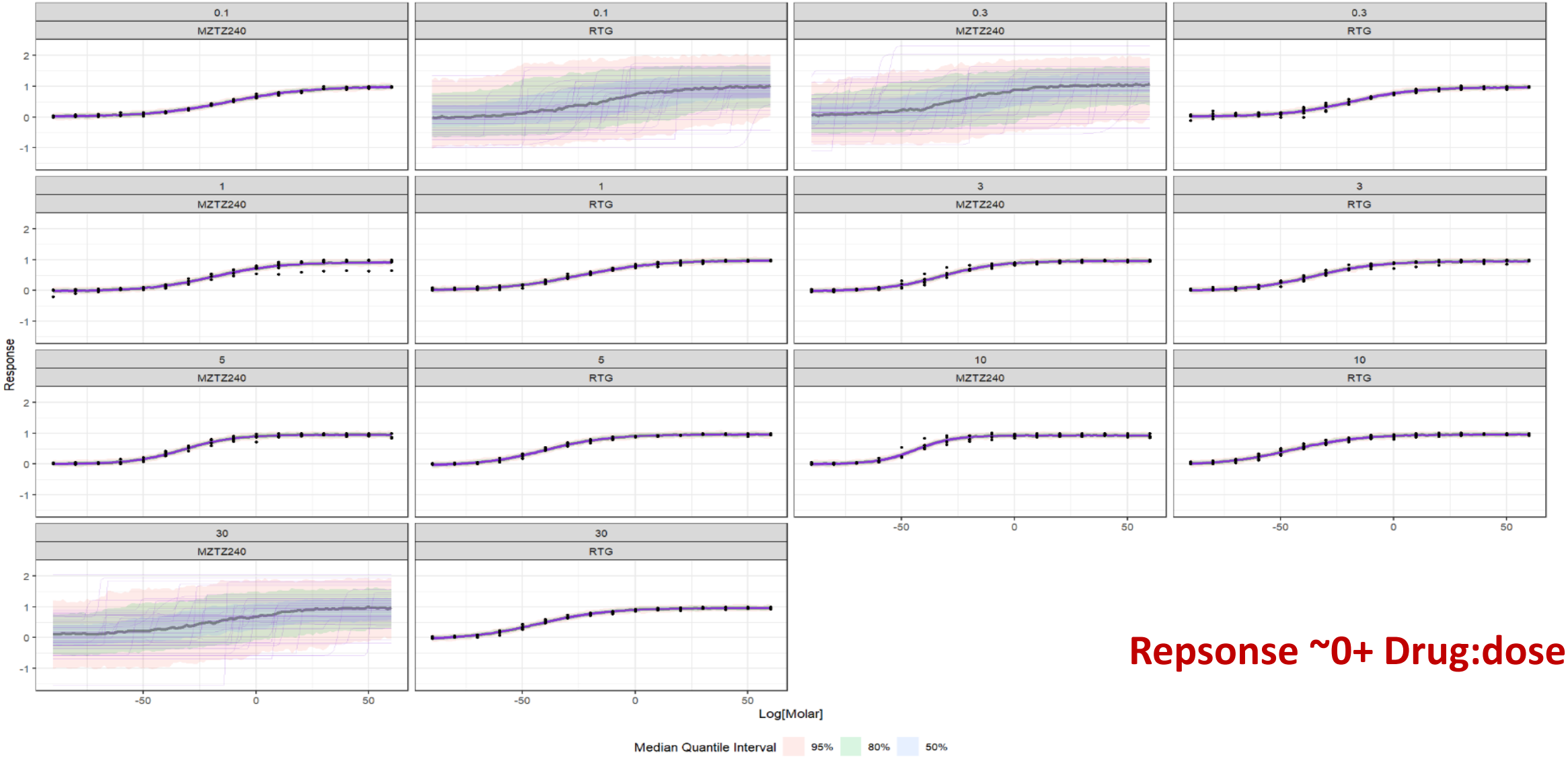


Model check: Prior-posterior plots

`Prior_posterior_densities_plot`

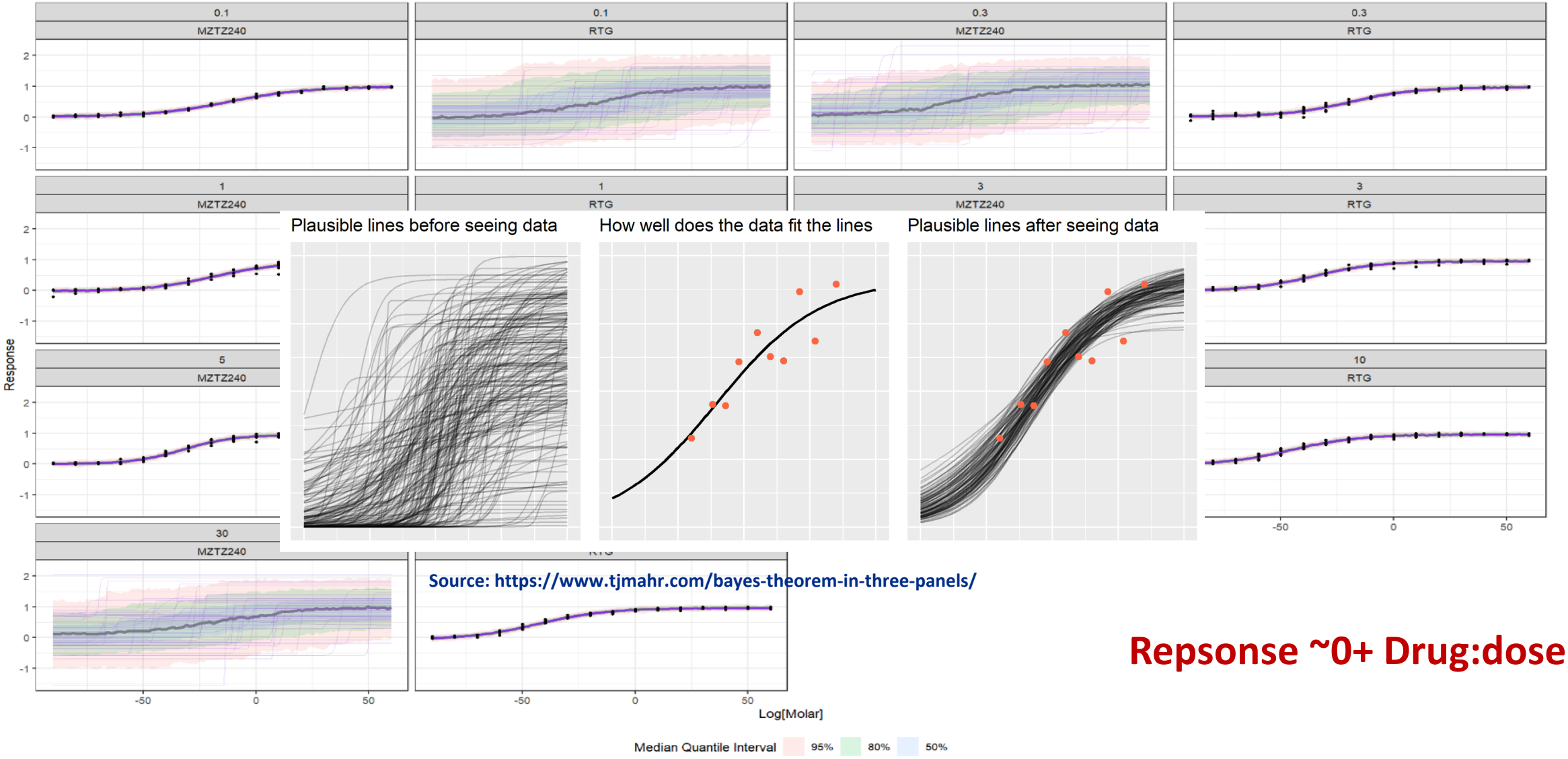


Model check: Posterior draw



Repsonse ~0+ Drug:dose

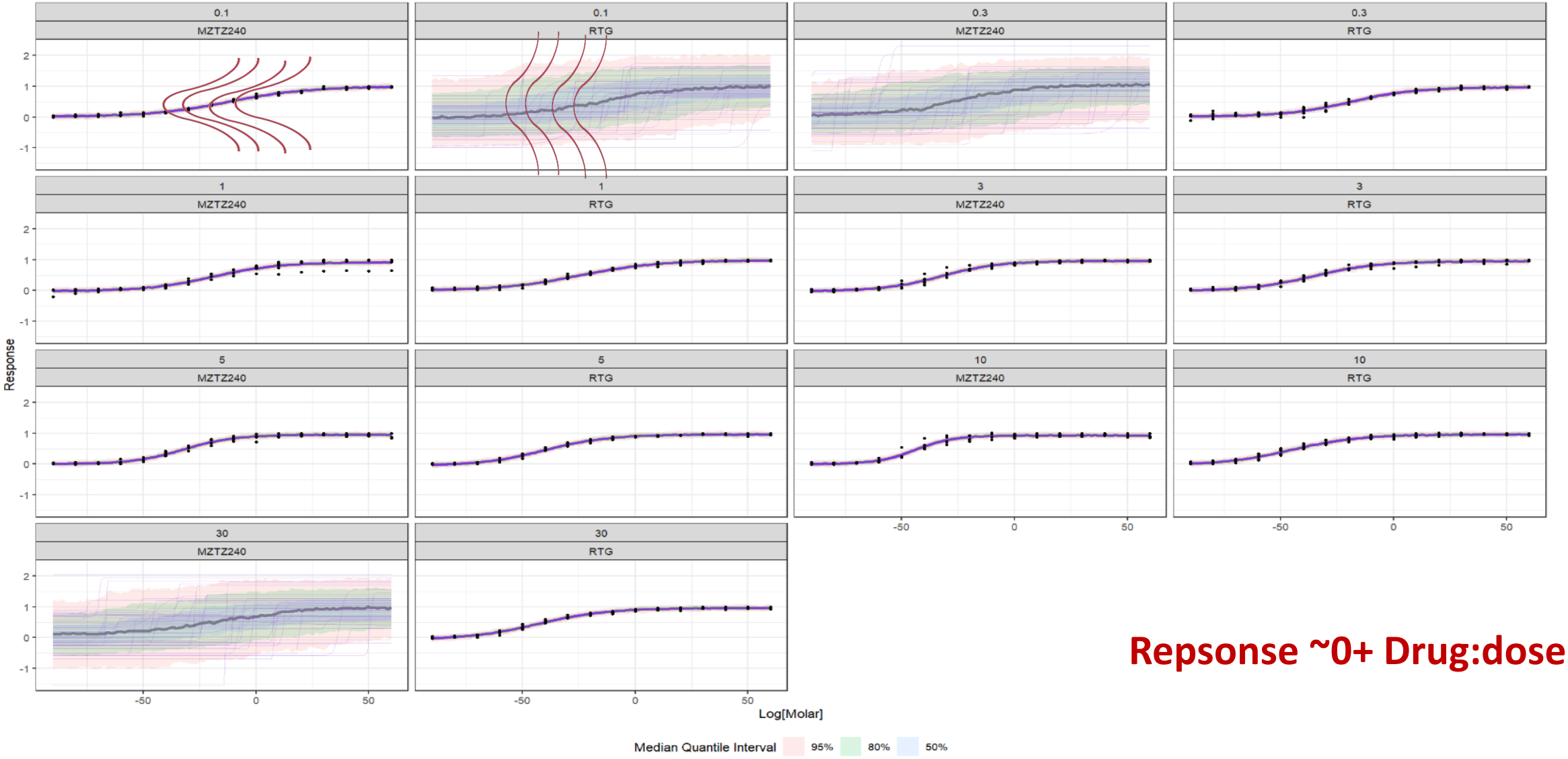
Model check: Posterior draw



Source: https://www.tjmahr.com/bayes-theorem-in-three-panels/

Repsonse ~0+ Drug:dose

Model check: Posterior draw

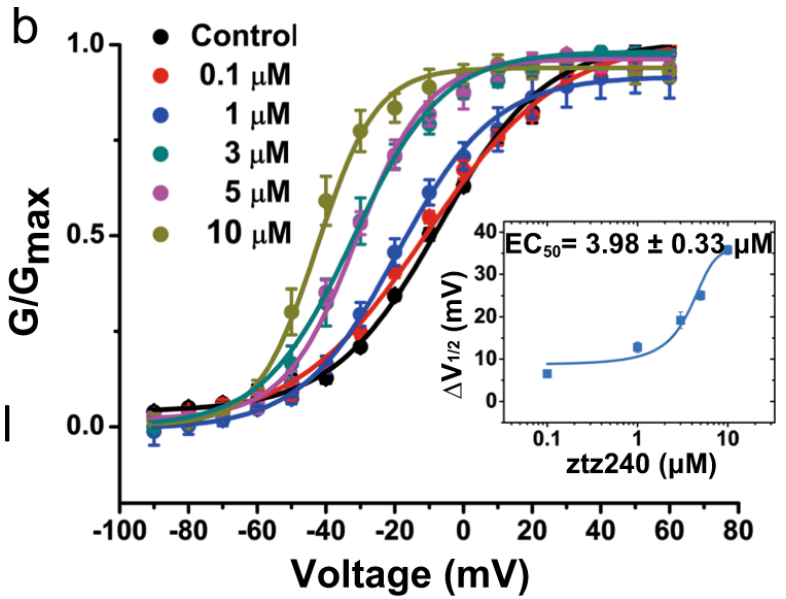


Repsonse ~0+ Drug:dose

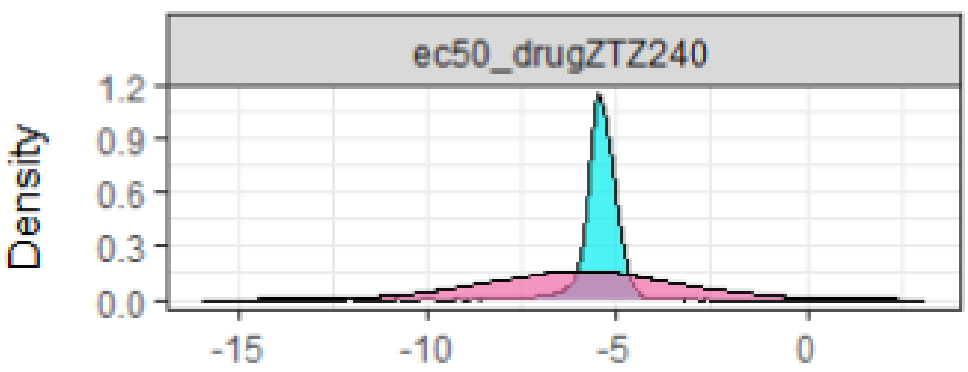
Dose-response curve fit: deltav1/2 vs. dose

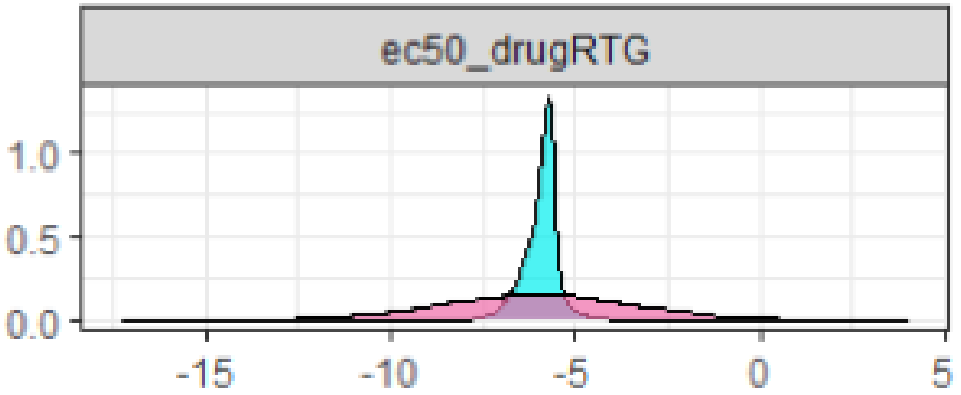
Steps:

1. Extract the ec50 (v1/2) from our conductance/voltage model
2. Fit another sigmoid antagonist model for each dose
3. The paper used delta V1/2, so did we
4. Repeat all the steps

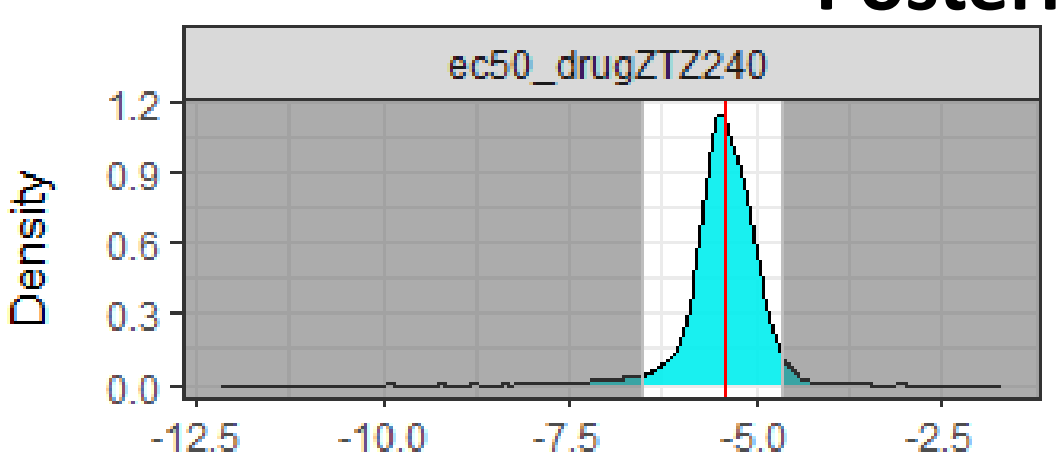


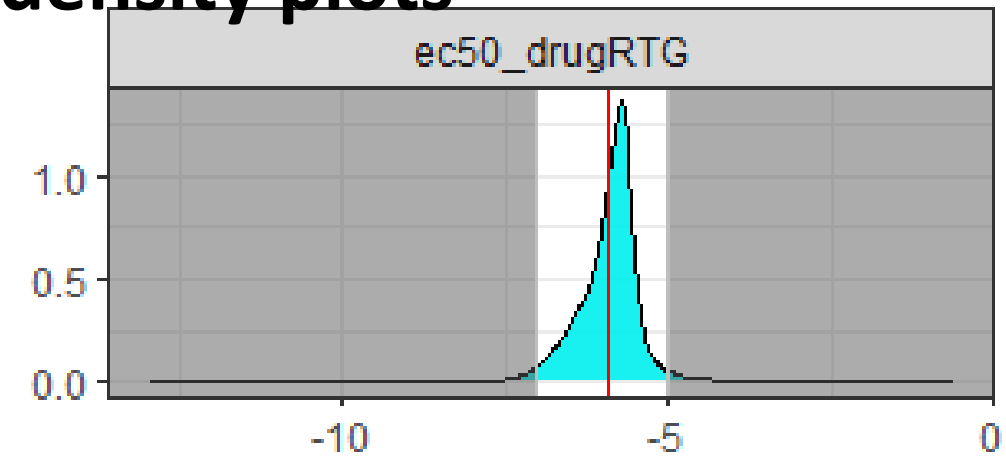
Prior/posterior plots





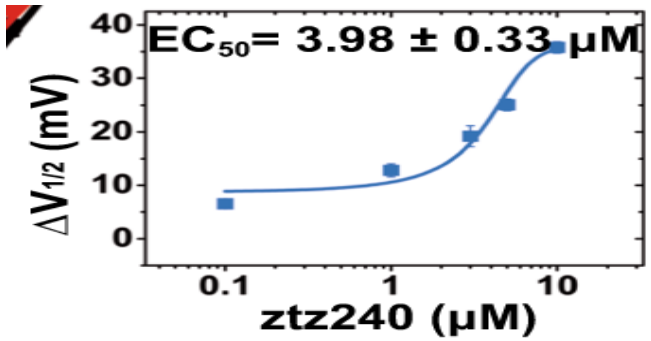
Posterior density plots

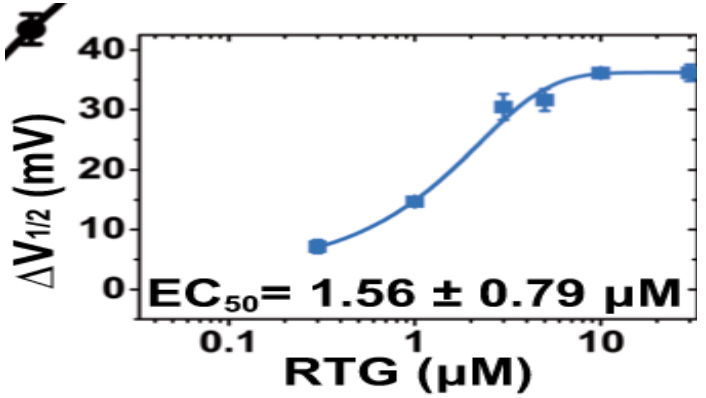




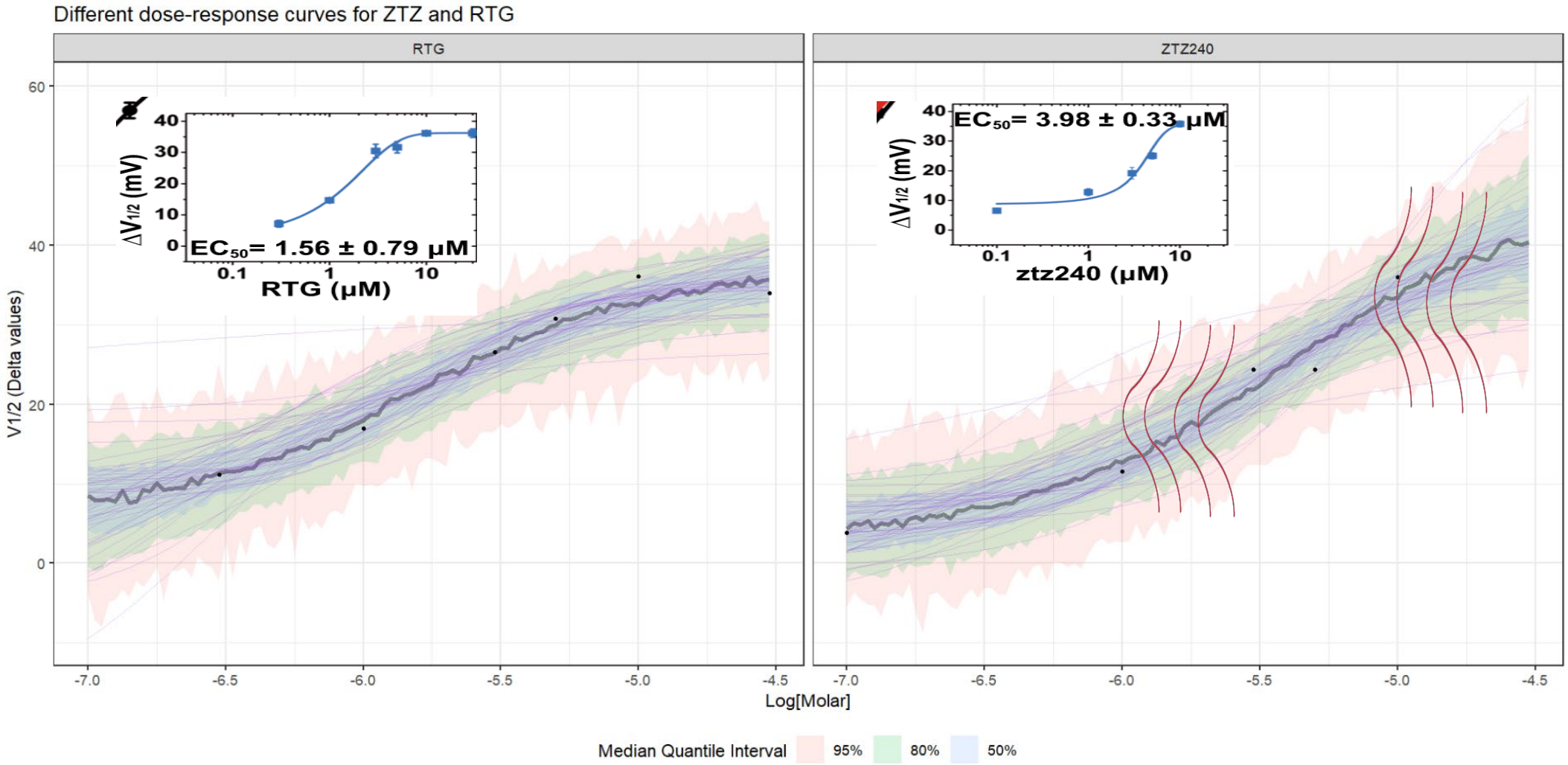
Para	Bayes	Paper
Ec50	3.72	3.98
l-95% CI	3.41	3.65
U-95% CI	25.59	4.31

Para	Bayes	Paper
Ec50	1.2	1.56
l-95% CI	1.11	0.77
U-95% CI	10.53	2.35





Posterior draws



Conclusion

- The Bayesian and the traditional model for both voltage-dependence and dose-response curves give comparable estimates

- Bayesian method suggests a wider uncertainty of the main estimate (Ec50)

Resources

Statistical Rethinking^2

Richard McElreath

@rmcelreath 32.1K subscribers 110 videos

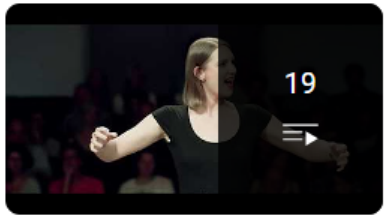
Lectures, mainly for Bayesian statistics, but also professional scientific tal... >

Subscribe

HOME VIDEOS PLAYLISTS COMMUNITY CHANNELS ABOUT

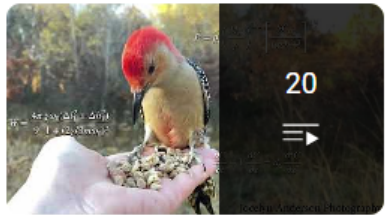
Created playlists

Sort by



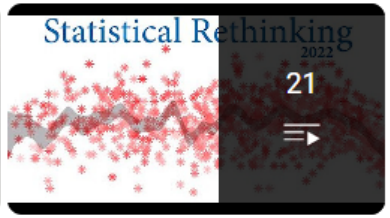
Statistical Rethinking 2023 Music

View full playlist



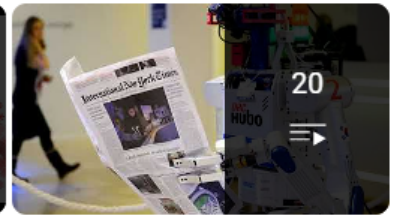
Statistical Rethinking 2023

View full playlist



Statistical Rethinking 2022

View full playlist



Statistical Rethinking Winter 2019

View full playlist

Statistical Rethinking Fall 2017

View full playlist

Statistical Rethinking Winter 2015

View full playlist

Acknowledgements

"I'd like to give a special shoutout to ChatGPT, my trusty sailor in the murky waters of data analysis. Thanks for guiding me through the uncertain tides of Bayesian statistics, Markov Chain Monte Carlo, and posterior probabilities. Without you, I would have been more lost than a frequentist sailor trying to navigate a sea full of priors!"

Thanks for your attention