BayesPharma: Bayesian methods for pharmacology models

Madeline J. Martin

Elayne Vieira Diaz

P. Walter German

Elyssa B. Margolis

Matthew J. O’Meara

In pharmacology, many experiments seek to measure how a reductive biological system responds to one or more treatments. Here we present BayesPharma, a collection of Bayesian methods to help analyze these experiments. BayesPharma is an package to analyzing pharmacology data using Bayesian methods. It is built around the ecosystem and facilitates a principled Bayesian workflow. BayesPharma can be used to fit and analyze foundational pharmacology models; serve as a pedagogical framework for learning Bayesian methods; and help build and analyze sophisticated pharmacological models.

## Introduction

As pharmacology experiments increase in complexity, it becomes increasingly challenging to analyze them. Although various frameworks exist for fitting pharmacology models using maximum likelihood estimation, such as GraphPad (GraphPad Software, LLC 2023) and the dose-response curves R package(Ritz et al. 2015). It is much less accessible, however, to fit pharmacology models using Bayesian statistics, which is a principled approach quantify model uncertainty before and after data collection. In practice, Bayesian modeling requires defining a functional form, a prior distribution over the model parameters, and a likelihood function that models the data generation process. Then, using Markov chain Monte Carlo (MCMC) or variational inference, the prior and the likelihood are combined to generate samples from the posterior distribution over the parameters. These samples can be used directly to answer scientific questions like “what is the parameter” and “how confident should we be?”

There has been substantial progress in general computational frameworks to facilitate developing and applying Bayesian models. A key example is the package and the ecosystem of supporting tools. provides a language to describe probabilistic models, inference engines, front-end interfaces to many common languages, and a suite of tools to analyze fit models(Carpenter et al. 2017; Bürkner 2017; Vehtari, Gelman, and Gabry 2017a; Gabry and Mahr 2017; Kay 2018; Wickham 2009; Wickham et al. 2019; Team and Others 2013). To facilitate model development, the Bayesian Regression Modeling using Stan (BRMS) package in R implements a formula-based interface for similar to (Bates et al. 2015). BRMS supports defining linear and nonlinear predictors, hierarchical models, and a range of pre-specified or custom response functions. Once specified, these models are translated into the modeling language, where they are compiled, run, and then can be analyzed.

In pharmacology, dose-response modeling through the foundational sigmoidal Hill-equation, Michaelis Menten enzyme kinetics, and multi-drug synergy models are widely used to probe biological systems and develop therapeutics. A limitation of the current tools is that it is currently not straightforward for practitioners to implement and analyze these types of models using Bayesian statistics. To fill this gap, here we describe the R package, a novel Bayesian pharmacology modeling framework that builds on and . After reviewing basic Bayesian modeling concepts in the context of dose-response experiments, we describe the package architecture and demonstrate the utility through several case studies.

### Related work

For general Bayesian modeling theory, there are many excellent textbooks(Gelman et al. 2013; McElreath 2016; Gelman and Hill 2006; Johnson, Ott, and Dogucu 2022), online resources(Betancourt 2023; Posit Software, PBC 2023; Herbert Lee 2023) and prescriptive guidance(Depaoli and Schoot 2017; Kruschke 2021; Gelman et al. 2020). While in theory, Bayesian modeling relies on relatively straightforward statistical principles discussed in the following section, in practice, it is often impossible to fit models analytically. Instead, fitting Bayesian models in practice requires using computational simulations or approximations. While these can be implemented from scratch, it is useful to build on a computational framework that supports model specification, working with probability distributions and samples, and implements algorithms to conduct simulation and variational based inference. Frameworks differ by the language ecosystem they build on, how tightly the components are coupled, and the maturity of the framework including support for diverse models and analyses and practitioner support including documentation and usage guides. Both the framework, on which the package is built, and the framework, implemented in C++ and an R front end with a domain language and inference engine(Plummer 2003), grew out of the BUGS model specification language; , implemented in C++ and python with probabilistic programming API and a range of modules for Bayesian inference(Salvatier, Wiecki, and Fonnesbeck 2016); , implemented in Python, Pytoch(Paszke et al. 2019) and JAX(James Bradbury 2018) and wraps arbitrary python code as a probabilistic model, and an emphasis on deep-learning based variational inference; and Turing, implemented in Julia leveraging the expressive type system to and support specifying probabilistic models and performing inference(Bezanson et al. 2017). Broadly, the increase in maturity in Bayesian modeling frameworks is making Bayesian modeling more accessible to practitioners, where it has seen a steady increase in popularity across the social sciences, econometrics, and biostatistics.

Over the past three decades, Bayesian methods have become increasingly common in clinical pharmacology(Ashby 2006; Grieve 2007; Campbell 2017; Yang and Novick 2019; Lakshminarayanan and Natanegara 2019; Cooner 2019; Lesaffre, Baio, and Boulanger 2020; Faya et al. 2021; Ruberg et al. 2023). For example, the first approved COVID-19 Pfizer/Biontech vaccine used a Bayesian clinical trial design (Polack et al. 2020; Senn 2022). While there has been some application of Bayesian modeling for the analysis of high-throughput screening (Wei et al. 2013; Lock and Dunson 2015; Shterev et al. 2021; Ma and Kummer 2021; Tansey et al. 2021), and dose response modeling (Smith and Marshall 2006; Johnstone et al. 2017; Labelle, Marinier, and Lemieux 2019; Gould 2019; Arezooji 2020; Semenova et al. 2021), the models tend to be bespoke and emphasize complexity. In contrast here we aim to lower the barrier of entry while maintaining flexibility.

## Bayesian Modeling Workflow

Bayesian statistics is a principled strategy to fit models to data. The key idea is that before seeing the data, the researcher defines a prior distribution over possible models indexed by model parameters, then the prior is combined with the data through Bayes theorem to produce the posterior distribution. Bayes theorem can be derived from basic facts about probability distributions. While many have encountered examples of probability distributions, e.g., the normal distribution over all real numbers or binomial distribution for the flips of a biased coin, it s worth thinking about what a probability distribution means from a computational perspective. Roughly they are objects that support two types of operations, (1) it is possible to draw samples from them and (2) given region of outcome space, we can ask what is the probability of drawing a sample in that region? For a one-dimensional probability distribution, when we draw many samples we can form a histogram, and once we have, we can ask how ask e.g. how what fraction had values larger than a given value. For two random variables, we can think of a sample as a scatter plot over the different dimensions.

from factoring the joint probability distribution over parameters and data, into conditional probability distributions two different ways, , and . Setting them equal to each other and dividing through by the evidence, , gives:

Inspecting this equation, we see that while the prior and likelihood distributions are explicitly defined in the model, the evidence must be estimated. One way to estimate it is through marginalization, which involves integrating the likelihood over the entire prior, and is typically intractable to compute explicitly. One way to conceptualize the evidence is simply as the (unique) constant that allows the posterior to be properly normalized and integrate to one. While a range of techniques have been developed to estimate the evidence, two important techniques that sidestep the issues are Markov chain Monte Carlo (MCMC) and variational inference. For MCMC, the key idea is to note that the relative posterior probability of two different parameters vs.  reduces to computing likelihood ratios as the prior and the evidence terms cancel out. Thus, by taking small steps in the parameter space and repeatedly evaluating likelihood ratios, it is possible to (in the long run) sample points from the posterior probability distribution. Intuitively, this is akin to running dynamical simulations using the posterior as the Boltzmann distribution. For variational inference, the key idea is to re-write Bayes formula and optimize a parametric function to estimate a lower-bound for the evidence(Blei, Kucukelbir, and McAuliffe 2017). Both of these techniques benefit from knowing the gradient of the likelihood with respect to the parameters, as this allows defining something akin to momentum when taking random steps in parameter for MCMC(Hoffman, Gelman, and Others 2014) and using gradient descent optimization for variational inference(Kucukelbir et al. 2015). Fortunately, gradients can be computed automatically for function composition by using the chain-rule(Carpenter et al. 2015) and is used in deep-learning packages in what is called back-propagation.

When using non-standard statistical analysis (such as Bayesian analysis) in fields that are dominated by frequentist statistics (such as Pharmacology), the analysis methods get more scrutiny because they may be unfamiliar to the readers. So, it may be useful to be able to conceptually describe how they compare/contrast with a frequentist analysis, if for no other reason than to keep skeptics at bay. So in case it is useful, here is how I philosophically conceptualize Bayesian analysis relative to frequentist analysis: The scientific method can be conceptualized through building and analyzing models. Consider that we are interested in interacting with a complex system such as a rocket ship, an ecosystem, a person with an illness etc., but we are cautious because doing so may be dangerous, expensive, slow, or unethical. Therefore we may develop a tractable model system to interact with instead (such as a very small rocket ship, zebrafish, or molecular simulation, etc.). If the model system corresponds with the system of interest in relevant ways, then poking and prodding the model system can be used to anticipate how the system of interest will respond when it is in turn poked and prodded. A key step in establishing this correspondence is to collect a small amount of data from the system of interest and use it to design the model system to be representative.

When scientific model systems are constructed by composing conditional probability distributions, they are often called Bayesian models. This is because the key step of going from a nebulous set of possible models to more precise set of models based on observed data uses the celebrated Bayes theorem. While basic Bayesian theory is covered in range of introductory and advanced textbooks, to complement those totally reasonable expositions, here we will give a correct, but informal derivation of Bayesian probability, to helpfully build intuition and facilitate communication with practitioners. Let be a vector of model parameters and let observed data. For concreteness, you can think of as the parameters for specific sigmoid function, and , a set of dose-response measurements. A probability distribution naturally models an event or an occurrence. The event is sometimes called a random variable. Given a region, the chance that the event happens in that region is the probability density over it. Assuming the event definitely happens, then the total density must integrate to unity. A joint distribution over two random variables is just the probability that both events occur. A conditional probability distribution can be thought of as a stochastic function, where evaluation corresponds to setting the values to be conditioned and obtaining the remaining values. Concretely if you have a scatter plot, you could sample points by first sampling along the x-axis and within the narrow range of the chosen x-value, pick a y-value. Or visa versa, first sample along the y-axis and then within the narrow range of the chose y-value, pick the x-value. Then the joint probability distribution over parameters and data, can be written as conditional probability distributions two different ways, , and . Setting them equal to each other and dividing through by gives Bayes Theorem:

With that in mind, let’s look at each term in the equation. The easiest is the , this called the prior must be specified beforehand. We’ll come back to how to choose priors below. The term is called the likelihood, and can be thought of as the data generation process, the “model system” if you will. On the left is the , which is called the posterior. This is a little trickier to understand, but thinking of it as a stochastic function, we can at least inspect the signature: it takes in data, and produces random samples of parameters. Finally the is called the evidence, and is basically there to make the left and the right be equal to each other. As we’ll see below, we don’t actually need to evaluate it, but if we did we could use a trick called marginalization.

How do we sample parameter values from the posterior distribution? One way to conceptualize the problem is to use Boltzmann’s equation to convert the probability distribution to an energy function and temperature. Then to sample, begin at a location in parameter space and simulate the kinematic trajectory as it would move through the parameter space over time. Given enough time, if the parameter can access all parts of the probability distribution, the parameter will equilibrate and give a sample from the probability distribution. While this is guaranteed to happen eventually, in practice a common concern is that for any fixed amount of equilibration, it is unclear if there has been sufficient mixing. We’ll come back to strategies to detect and handle issues with mixing below. A basic kinematic sampling algorithm called Markov chain Monte Carlo (MCMC) says to take a step, and if it has lower energy, accept, if it has higher energy accept with a probability proportional to the difference in probabilities between the two states and the simulation temperature. A key insight is that because only relative probabilities are needed, the global evidence normalization factors cancel and therefore they don’t even need to be computed. More sophisticated algorithms take into account not only the specific energy values but also the energy gradient. Thinking of the object as having inertia, the sampling allows it to overcome certain types of local minima. to take steps that are likely to be accepted, sometimes called Hamiltonian Monte Carlo. A particularly annoying part of MCMC is that in some cases, such as a narrow ravine in the energy landscape, the trajectory will vibrate back and forth without making much forward progress. A clever heuristic that is used by , called No U-turn sampling (NUTs), tries to dampen these types of unproductive moves.

The principled Bayesian workflow Van de Schoot et al. (2020), describes a general protocol building robust Bayesian models and using them to draw inferences. The main steps involve

### BayesPharma package design

We designed the package to support modeling of pharmacology models using the principled Bayesian workflow. As described above, the workflow involves four phases: model specification, model fitting, model evaluation, and interpretation. Here we will describe the interface and how we recommend using it.

To provide data to the package, the user provides R with columns for the response, treatments, and optionally additional covariates such as or . The data is passed to the model function optionally including arguments to customize the formula, prior, initial values, and other arguments to control the model fitting. The package then passes the user input to along with custom code specific for the selected model. Once the model is fit, the resulting object is returned to the user and can be used for analysis. To illustrate

#### Model specification

Here we will describe the components that are needed to specify the model.

**Formula**: The goal of the formula is to describe how the data is generated conditional on the model parameters. Syntactically, model formulas build a , which is similar to the formula specification syntax in base R and other R regression modeling packages. A consists of an equation that declares how the response on the left side is sampled from a parameterized distribution on the right side. For the default linear formulas, the right side specifies mean response with a linear combination of covariates added together with implicitly defined model parameters. For example, the formula

says that is sampled from a distribution with mean where are scalar parameters and is an indicator variable for drug . By default, the sampling distribution is a Gaussian, but other distributions can be specified from the distribution family with a link function using argument. For example, to model count data, which is strictly positive, setting . To model more general sampling equations, can be specified as non-linear by setting , and all model parameters must be explicitly defined. Building on this framework, the package supports a wide range types of regression models including hierarchical models or random effects models and observational models that handle, for example, missing data or measurement error, which are described in detail in . The package extends the formula syntax by defining functions for each model type, such as the sigmoid function to model Hill-equation dose response models. For each model, functions are provided to help build the formula, for example

# This formula...  
demo\_formula <- BayesPharma::sigmoid\_agonist\_formula(  
 predictors = 1 + drug\_id)  
  
# will generate the equivalent formula as this  
demo\_formula\_alt <- brms::brmsformula(  
 response ~ sigmoid(ec50, hill, top, bottom, log\_dose),  
 nl = TRUE,  
 family = brms::student()) +  
 brms::lf(ec50 ~ 1 + drug\_id) +   
 brms::lf(hill ~ 1 + drug\_id) +  
 brms::lf(top ~ 1 + drug\_id) +  
 brms::lf(bottom ~ 1 + drug\_id)

: The data to be passed to the model should be in an R and organized into a “tidy” format. This means that each observation is in a single row and there there are columns for the , and model specific covariates such as for the sigmoid model, and additional experimental covariates that can be used as predictors. See table XXX for the required columns (Treatments and Response) for each of the implemented model types:

To get data in to a tidy form, it is possible to use the range of tools from the R tidyverse libraries including , which can subsets of rows, subsets of columns, compute new columns rowwise with or perform split-apply-combine to operate over subsets of rows, and data in different objects can be merged using SQL-like joins like . If data is organized, e.g., in plate-layout with multiple observations per-row, the package has and , which can be used to transform the shape of objects. When observation measurements and covariates come in in non-standard format, string manipulation is possible through the package. Finally, for exploratory visualization, the package implements the grammar-of-graphics workflow for mapping data in a tidy format to aesthetic attributes of geometric objects in the plot such as the coordinates of points or lines. For a good introduction to data manipulation using the tidyverse, the R for Data Science (2e) book and website (https://r4ds.hadley.nz/) are quite good. Through each of the vigenttes below we will make use of the tidyverse data manipulation.

demo\_data <- tibble::tibble(  
 drug\_id = c("C1", "C2", "C3"),  
 ac50 = c(-8, -7, -6)) |>  
 dplyr::cross\_join(  
 tidyr::expand\_grid(  
 log\_dose = seq(-7, -5, length.out = 10),  
 replica = c(1,2,3))) |>  
 dplyr::mutate(  
 mean\_response = BayesPharma::sigmoid(  
 ac50 = ac50,  
 hill = 1,  
 top = 1,  
 bottom = 0,  
 log\_dose = log\_dose),  
 response = c(  
 stats::rnorm(  
 n = dplyr::n(),  
 mean = mean\_response,  
 sd = ifelse(replica == 1, .2, .8))))

: A key principle of Bayesian models is that they require specifying priors. For newcomers, understanding how to determine how they should be specified and justified tends to be one of the more challenging parts of the modeling process. From a practical perspective priors can be thought of as just defining a weighted region of parameter space over which to optimize the model to best fit the data. In particular, the more compact and closely aligned the priors are with the data, the easier it is for the model to fit the data. So for setting up and getting started with a new model fit, it is best to give as strong (more constrained) of priors as possible. From a scientific perspective, priors and posterior distributions can be interpreted as capturing the uncertainty in parameters before and after observing the data. So, from this perspective, weaker (less constrained) priors are preferred in order to “let the data speak for itself”. Ultimately however, in Bayesian modeling, it is not possible to completely remove the bias due to the prior. This means that in a complete Bayesian analysis some substantive argument should be made that the inferences from the model are not sensitive to reasonable choices of the prior. In a way, this actually makes choosing the prior less stressful as there is no singular best prior choice, and instead it reflects the scientific questions of the modeling process. For a deeper discussion and practical advice, see: the [prior choice recommendations](https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations) page on the wiki. To facilitate specifying priors for each of the models, the package implements helper functions, for example

# This formula  
demo\_prior <- BayesPharma::sigmoid\_agonist\_prior()  
  
# will generate the equivalent formula as this  
demo\_prior\_alt <- c(  
 brms::prior(prior = normal(-6, 2.5), nlpar = "ec50"),  
 brms::prior(prior = normal(1, 1), nlpar = "hill", lb = 0.01),  
 brms::prior(prior = normal(1, 0.5), nlpar = "top"),  
 brms::prior(prior = normal(0, 0.5), nlpar = "bottom"))

The function takes in code (in this case the [Normal distribution](https://mc-stan.org/docs/2_20/functions-reference/normal-distribution.html), which is defined in the documentation), the defines the non-linear parameter, and optional arguments and give upper or lower bounds. For each prior helper function, individual priors can either be explicitly given to override the defaults, or specified as numeric constants to fix them to a particular value.

: To estimate the posterior distribution using markov chain Monte Carlo sampling, initial parameters values must be given. From these initial parameters, the parameters are simulated in multiple independent chains in such a way that they converge to a sample from the posterior distribution. The initial parameter values should be in a feasible region of parameter space to get the simulation to rapidly mix. For each model, the package provides default initialization values. Typically if a different prior is given then the initial values can be adjusted along with the updated prior.

# This formula  
demo\_init <- BayesPharma::sigmoid\_agonist\_init()  
  
# will generate the equivalent initial values as this  
demo\_init\_alt <- function() {  
 list(  
 b\_ec50 = function(){as.array(-6)},  
 b\_hill = function(){as.array(1)},  
 b\_top = function(){as.array(1)},  
 b\_bottom = function(){as.array(0)})  
}

To summarize the model

| Name | Type | Treatments | Parameters | Response |
| --- | --- | --- | --- | --- |
| Sigmoid | one treatment | log\_dose | top, bottom, AC50, hill | response |
| MuSyC | two treatments | logd1, logd2 | logE[0-3], logC[1,2], h[1,2], logalpha | response |
| tQ | enzyme kinetics | series\_index, time, ET, ST | Kcat, kM | P |

#### Model fitting

Once the components of the model have been specified, to fit it, each model type provides a function to integrate the formula, data, prior, init and additional arguments to build and fit the model.

# This model ...  
demo\_model <- BayesPharma::sigmoid\_model(  
 formula = demo\_formula,  
 data = demo\_data)  
  
# is equivalent to:  
demo\_model\_alt <- brms::brm(  
 formula = demo\_formula,  
 data = demo\_data,  
 prior = demo\_prior,  
 init = demo\_init,  
 control = list(adapt\_delta = 0.99),  
 iter = 8000,  
 stanvars = BayesPharma::sigmoid\_stanvar)  
brms::expose\_functions(demo\_model\_alt, vectorize = TRUE)  
demo\_model\_alt$model\_type <- "sigmoid"

#### Model Evaluation

: To evaluate the model fit involves evaluating the quality of the parameter estimation.

: uses Hamiltonian Markov chain Monte Carlo simulations to sample from the posterior. While simulations are guaranteed to converge to the posterior eventually, for any finite sample, there is a risk that the samples may be biased by the initial values. To assess convergence, a key strategy is to run multiple chains and compare the within chain against the between chain variation in parameter estimates. If the chains have not converged then the inter-chain variation will be high. To quantify (Gelman and Rubin 1992) define ,

To mitigate this bias, there are two key strategies, first simulate independent chains and measure the intra- vs inter-chain correlation.

: A key strategy in developing and checking the quality of a model, is to sample from the prior and posterior distributions and visualize the resulting outcomes. In sampling from the prior, the goal is to make sure that the resulting distributions are consistent with the scientific understanding of the model. For example, if the collected data are counts, but the default Normal distribution family is used, then since the Normal distribution has infinite support, the prior model may generate samples with negative counts, which doesn’t make sense. Seeing these negative counts would motivate using either a family bounded below by zero, or setting a lower bound of zero by setting in the prior specification.

For posterior predictive distributions, the goal is to generate samples from the posterior and qualitatively evaluate if it is consistent with the observed data. For example, if the observed data is multi-modal, but the posterior samples are unimodal this may suggest that a different functional form is needed. Using plots and visualizations can reveal these and other issues with the model that may need to be explicitly modeled. Note that in contrast with typical frequentist modelling, the goal is not to construct the simplest meaningful model in order to have the most power to reject it, the goal is to construct a model that best fits the data in order to give the best interpretation of the observations.

: A general modeling strategy is to begin with simpler models and incrementally increase model complexity to handle nuance as needed. Simpler models are not only easier to interpret, but given a simple model and complex model that explain the data equally, Occam’s razor says that the simplier model should be preferred, in part because it is more likely to generalize to unseen contexts. Evaluating model fit while taking into account model complexity is non-trivial for non-Bayesian models and can be done in ad-hoc ways through the number of parameters, deviance scores etc. However, for Bayesian models, explicitly modeling the model uncertainty gives a principled approach by measuring the marginal likelihood of the data given the model, quantified by expected log-posterior density (ELPD) of held out data. Instead of splitting the data in to train/test split, k-fold or leave-one-out (LOO) cross-validation can be used, where the model is fit to portion of the data and the ELPD is measured on the rest, across multiple partitions. Re-fitting the model for each data-point is computationally intensive, but it can be shown that it can be approximated through Pareto Smoothed Importance Sampling (PSIS) and is implemented in the package (Vehtari, Gelman, and Gabry 2017b). By default the package models compute the criterion. Checking the will summarize the outlier data points (Pareto k-statistics > 0.7). See [Loo interpretation](https://mc-stan.org/loo/articles/online-only/faq.html#elpd_interpretation) for more guidance in how to interpret the summary statistics. Then, given two models for the same data, calling , will give rank the models based their their ELPD. See the case-studies for concrete examples. An alternative strategy to compare models is to use Bayesian Model Averaging (BMA), where multiple models are fit separately and then averaged depending on their ability to explain the data. The models that receive weights greater than zero contribute to explaining the overall data distribution.

: After models have been fit and selected, the posterior distribution can be used to test hypotheses. For example, to test if one parameter is greater than another, the fraction of samples from the posterior in which the condition is true can be interpreted as a p-value.

Together these stages of model fitting and evaluation are the foundation for a principled work, that is outlined in more detail in @(Gelman et al. 2020). Understanding what steps need to be communicated depends on the the context, but in our view the case-studies in the following sections illustrate templates that can be for typical Pharmacology analyses.

# Case Studies

In this section we will consider several models as case studies: the sigmoidal hill model in #{sec:hill}, the MuSyC synergy model in {#sec:MuSyC}, Michaelis-Menten enzyme progress curve in #{sec:michaelis\_menten}. For each we will implement it, and apply it to example data by fitting different models and then we will compare the models based on their fit of the data and inferences that can be made.

## Hill Equation

In this case study we are going to reanalyze the dose response of 4 Kappa Opioid receptor (KOR) antagonists using the package from from a study done by Margolis et al. (-Margolis et al. (2020)). Whole cell electrophysiology in acute rat midbrain slices was used to evaluate pharmacological properties of four novel KOR antagonists: BTRX-335140, BTRX-395750, PF-04455242, and JNJ-67953964

Originally in the paper the dose-response data analysis was done by using the package in R which implements the minimization of negative log likelihood function and reduces to least square estimation for a continuous response. The data was normalized to % baseline then fit to a 4-parameter log-logistic dose response model, setting the top (max response) to 100% and estimating the IC50, its variance, and the bottom (min response).

### Fitting the sigmoid model

Using the package, we can re-fit the sigmoid model with a negative slope, and fixing the top parameter to as the response is normalized to a no-drug baseline.

For the prior, we are going to use a normal distribution because the response values are continuous. First, we will run the analysis with top (max response) parameter prior set to a constant value of 100 because top is normalized to and the default broad prior for the , and parameters. Broad priors represent unbiased uncertainty and provide an opportunity for extreme responses.

The level of informativeness of the prior will affect how much influence the prior has on the model. Here is more [information on prior choice recommendations.](https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations)

kor\_prior <- BayesPharma::sigmoid\_antagonist\_prior(top = 100)  
kor\_prior

prior class coef group resp dpar nlpar lb ub source  
 normal(-6, 2.5) b ic50 <NA> <NA> user  
 normal(-1, 1) b hill <NA> 0.01 user  
 constant(100) b top <NA> <NA> user  
 normal(0, 0.5) b bottom <NA> <NA> user

#### Prior predictive checks

Following the Bayesian workflow, before fitting the model, it is good to check the prior predictive distributions to see if they are compatible with the domain expertise. So, before running the model, we will verify that the prior distributions cover a plausible range of values for each parameter. To do this, we want to sample only from the prior distributions by adding as an argument to the function. We will use the default response distribution of the model ().

kor\_sample\_prior <- BayesPharma::sigmoid\_model(  
 data = kor\_antag |> dplyr::select(substance\_id, log\_dose, response),  
 formula = BayesPharma::sigmoid\_antagonist\_formula(),  
 prior = kor\_prior,  
 init = BayesPharma::sigmoid\_antagonist\_init(),  
 sample\_prior = "only")

And then plot of the prior predictive distributions:

kor\_sample\_prior |>  
 BayesPharma::plot\_density\_distribution()

|  |
| --- |
| KOR antagonists prior distribution |

To sample from the model we will the NUTs Hamiltonian Monte Carlo, and initialize the parameters to the prior means to help with model convergence, using the default values of , , , .

kor\_model <- BayesPharma::sigmoid\_model(  
 data = kor\_antag |> dplyr::select(substance\_id, log\_dose, response),  
 formula = BayesPharma::sigmoid\_antagonist\_formula(  
 predictors = 0 + substance\_id),   
 prior = kor\_prior,  
 init = BayesPharma::sigmoid\_antagonist\_init())

### Analyzing model fit

The generated model summary shows the formula that the expected response is sigmoid function of the log\_dose with four parameters, and a shared Gaussian distribution. Each parameter is dependent on the substance\_id. Since want to fit a separate model for each substance we include a to indicate that there is no common intercept. The consists of data points and the posterior sampling was done in chains each with steps with steps of warm-up. The population effects for each parameter summarize marginal posterior distributions, as well as the effective sample size in the bulk and tail. This gives an indication of the sampling quality, with an ESS of > samples being good for this type of model.

Family: gaussian   
 Links: mu = identity; sigma = identity   
Formula: response ~ sigmoid(ic50, hill, top, bottom, log\_dose)   
 ic50 ~ 0 + substance\_id  
 hill ~ 0 + substance\_id  
 top ~ 0 + substance\_id  
 bottom ~ 0 + substance\_id  
 Data: data (Number of observations: 73)   
 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 Bulk\_ESS Tail\_ESS  
ic50\_substance\_idBTRX\_335140 -8.84 0.20 -9.20 -8.40 1.00 13542 8700  
ic50\_substance\_idBTRX\_395750 -8.24 0.41 -8.92 -7.37 1.00 11435 5487  
ic50\_substance\_idJNJ -9.15 0.32 -9.77 -8.51 1.00 15085 9371  
ic50\_substance\_idPF -6.15 1.06 -7.64 -3.33 1.00 8184 4927  
hill\_substance\_idBTRX\_335140 -1.47 0.61 -2.92 -0.58 1.00 14622 11291  
hill\_substance\_idBTRX\_395750 -0.89 0.49 -2.17 -0.26 1.00 10926 6320  
hill\_substance\_idJNJ -1.01 0.51 -2.38 -0.40 1.00 14262 10996  
hill\_substance\_idPF -0.31 0.25 -0.93 -0.03 1.00 7332 5246  
bottom\_substance\_idBTRX\_335140 -0.00 0.50 -0.99 0.98 1.00 16477 11542  
bottom\_substance\_idBTRX\_395750 0.02 0.50 -0.94 0.99 1.00 19064 12423  
bottom\_substance\_idJNJ -0.01 0.50 -0.99 0.97 1.00 17501 11293  
bottom\_substance\_idPF 0.01 0.50 -0.97 0.98 1.00 18273 12265  
top\_substance\_idBTRX\_335140 100.00 0.00 100.00 100.00 NA NA NA  
top\_substance\_idBTRX\_395750 100.00 0.00 100.00 100.00 NA NA NA  
top\_substance\_idJNJ 100.00 0.00 100.00 100.00 NA NA NA  
top\_substance\_idPF 100.00 0.00 100.00 100.00 NA NA NA  
  
Family Specific Parameters:   
 Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
sigma 32.16 2.88 27.15 38.42 1.00 15048 12016  
  
Draws were sampled using sampling(NUTS). For each parameter, Bulk\_ESS  
and Tail\_ESS are effective sample size measures, and Rhat is the potential  
scale reduction factor on split chains (at convergence, Rhat = 1).

#### Traceplot

The model ran without warning messages meaning there were no parameter value problems or MCMC conflicts. The bulk and tail ESS indicate high resolution and stability. The R-hat for each parameter equals and the shows the chains mixed well indicating the chains converged.

kor\_model |>  
 bayesplot::mcmc\_trace()

|  |
| --- |
|  |

#### Compare prior and posterior marginal distributions

Displayed below is a plot for the prior and posterior distributions of the parameters (prior is pink and posterior is teal). this can be useful for comparing the density distribution of the prior and posterior produced by the model:

BayesPharma::plot\_prior\_posterior\_densities(  
 model = kor\_model,  
 predictors\_col\_name = "substance\_id",  
 half\_max\_label = "ic50",  
 title\_label="")

|  |
| --- |
| KOR antagonists model, compare prior and posterior distributions for each substance |

Displayed below is a plot of the posterior distributions for each parameter with the confidence intervals and mean. This is a useful visual of the model results and to highlight the mode and high density intervals:

BayesPharma::plot\_posterior\_density(  
 kor\_model,   
 predictors\_col\_name = "substance\_id",   
 half\_max\_label = "ic50",  
 title\_label = "")

|  |
| --- |
| KOR Antagonists, posterior distribution for each substance |

Displayed below is a plot of a sample of 100 sigmoid dose-response curves from the posterior distribution (purple) and the median quantile intervals:

BayesPharma::plot\_posterior\_draws(  
 model = kor\_model,  
 title = "")

|  |
| --- |
| KOR antagonists, posterior draws |

### Comparing alternative models

To test the sensitivity of the analysis to the prior, we can re-fit the model with more informative prior:

prior class coef group resp dpar nlpar lb ub source  
 normal(-8.5, 0.5) b ic50 <NA> <NA> user  
 normal(-1, 0.5) b hill <NA> 0.01 user  
 constant(100) b top <NA> <NA> user  
 normal(10, 15) b bottom <NA> <NA> user

Re-fitting the model

## Comparing the Two Models Using LOO-Comparison:

One way to evaluate the quality of a model is for each data-point, re-fit the model with remaining points, and evaluate the log probability of the point in the posterior distribution. Taking the expectation across all points give the Expected Log Pointwise predictive Density (ELPD). Since this is computationally challenging to re-fit the model for each point, if the model fits the data reasonably well, then the ELPD can be approximated using the Pareto smoothed importance sampling (PSIS). Using the LOO, the package, Pareto k value for each data point is computed, where k less than is good, between and is OK, and higher than indicates the data point is not fit by the model well. Evaluating the model for the KOR antagonists, shows that the model fits the data well.

No problematic observations found. Returning the original 'loo' object.

NULL

Since ELPD is a global measure of model fit, it can be used to compare models. Using from the LOO package, returns the and for each model relative the model with the lowest ELPD. The , the model with more informative prior, is the preferred model, but not significantly.

No problematic observations found. Returning the original 'loo' object.

elpd\_diff se\_diff  
kor\_model2 0.0 0.0   
kor\_model -0.9 1.2

##Analysis Using the drc Package

Here we will analyze the KOR antagonist data using the drc package and compare it to the results from the analysis.

We will fix the top to and fit the , , and .

drc\_models <- kor\_antag |>  
 dplyr::group\_by(substance\_id) |>  
 dplyr::group\_nest() |>  
 dplyr::mutate(  
 model = data |>   
 purrr::map(~drc::drm(  
 response ~ log\_dose,  
 data = .x,  
 fct = drc::L.4(fixed = c(NA, NA, 100, NA),  
 names = c("hill", "bottom", "top", "ic50")))))  
  
drc\_models |>  
 dplyr::mutate(summary = purrr::map(model, broom::tidy, conf.int = TRUE)) |>  
 tidyr::unnest(summary) |>  
 dplyr::arrange(term, substance\_id) |>  
 dplyr::select(-data, -model, -curve)

# A tibble: 12 × 8  
 substance\_id term estimate std.error statistic p.value conf.low conf.high  
 <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
 1 BTRX\_335140 bottom 1.31 19.4 0.0675 9.47e- 1 -40.0 42.6   
 2 BTRX\_395750 bottom 29.5 9.40 3.14 7.85e- 3 9.20 49.8   
 3 JNJ bottom -18.1 26.7 -0.681 5.04e- 1 -73.7 37.4   
 4 PF bottom 39.4 30.8 1.28 2.22e- 1 -27.0 106.   
 5 BTRX\_335140 hill 4.06 9.20 0.441 6.65e- 1 -15.5 23.7   
 6 BTRX\_395750 hill 9.82 164. 0.0600 9.53e- 1 -344. 364.   
 7 JNJ hill 1.17 0.580 2.02 5.69e- 2 -0.0378 2.38  
 8 PF hill 1.13 1.33 0.855 4.08e- 1 -1.73 4.00  
 9 BTRX\_335140 ic50 -8.91 0.308 -28.9 1.42e-14 -9.57 -8.26  
10 BTRX\_395750 ic50 -8.97 0.505 -17.8 1.70e-10 -10.1 -7.88  
11 JNJ ic50 -8.77 0.670 -13.1 2.89e-11 -10.2 -7.37  
12 PF ic50 -7.96 1.27 -6.29 2.78e- 5 -10.7 -5.23

Displayed below is the comparison of results from and for each parameter of the dose-response curve. Here we see that the Bayesian method provides a distribution curve as evidence and has smaller confidence intervals than most of the standard errors provided by the method.

Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
ℹ Please use `linewidth` instead.

|  |
| --- |
| KOR antagonists conditional effects. The blue lines are samples from the exttt{BayesPharma} kor\_model posterior distribution, the orange line is the conditional mean, and the purple line is the conditional mean for the exttt{drc} model fit. |

## MuSyC synergy model

When two different treatments are combined they may interact to cause a response. For end-point assays, if the response is stronger or weaker than what would be expected with an additive model, the treatments are said to be epistatic. For sigmoidal dose-response models, however, the analysis may be more complicated. One drug may not only may shift the maximal response (efficacy) of the other, but it may also shift the effective dose and shape of the response (potency). Historically a range of models have been proposed that capture different aspects of synergy, for example the Bliss independence and Loewe additivity are null-models for no synergistic efficacy or potency, respectively. The R package and the python package can be used to visualize treatment interactions, compute a range of synergy scores, and test if the interactions are significant.

Recently Meyer et al. derived an integrated functional synergistic sigmoidal dose-response, which has the Loewe and Bliss models as special cases. They implemented a Bayesian model-fitting strategy in Matlab, and a maximum likelihood model-fitting into the synergy python package. To make the model more accessible to the pharmacology community, in this section, we briefly review the MuSyC functional form, describe a Bayesian implementation in Stan/BRMS, and illustrate using the model to re-analyze how drugs and voltage may interact to modulate the current through a potassium channel.

### MuSyC Functional Form

The functional form for the MuSyC model gives an equation for the response $\color{brown}{E\_d}$ at doses of $\color{teal}{d\_1}$ and $\color{teal}{d\_2}$ of the two treatments and has free parameters $\color{purple}{C\_1}$, $\color{purple}{C\_2}$, $\color{brown}{E\_0}$, $\color{brown}{E\_1}$, $\color{brown}{E\_2}$, $\color{brown}{E\_3}$, $\color{purple}{h\_1}$, $\color{purple}{h\_2}$, $\color{purple}{\alpha}$:

To interpret these parameters if we set $\color{teal}{d\_2}=0$, then which is the Hill equation, which we modeled above . If we then additionally set $\color{teal}{d\_1}=0$ then $\color{brown}{E\_d}=\color{brown}{E\_0}$, in the limit as ${\color{teal}{d\_1}}\rightarrow \infty$ then ${\color{brown}{E\_d}}\rightarrow {\color{brown}{E\_1}}$, and if ${\color{teal}{d\_1}}=\color{purple}{C\_1}$ then ${\color{brown}{E\_d}} = ({\color{brown}{E\_0}} + {\color{brown}{E\_2}})/2$, which is the half maximal response (either the $\color{brown}{\mbox{IC}\_{50}}$ if treatment is an inhibitor or $\color{brown}{\mbox{EC}\_{50}}$ if treatment is agonist). The slope at ${\color{teal}{d\_1}}={\color{purple}{C\_1}}$ is

The evaluation of the functional form for ${\color{brown}{E\_d}}$ is numerically unstable. To transform using the trick, let

Then

### Implementation and usage of the MuSyC model in Stan/BRMS

## Michaelis-Menten enzyme progress model

## Enzyme Kinetic Modeling

Enzymes are proteins that catalyze chemical reactions. Not only do they facilitate producing virtual all biological matter, but they are crucial for regulating biological processes. In the early 20th century Michaelis and Menten described a foundational kinematic model for enzymes, where the substrate and enzyme reversibly bind, the substrate is converted to the product and then released.

$$
E + S \xrightleftharpoons[k\_b]{k\_f} C \xrightarrow{k\_{cat}} E + P
$$

where the free enzyme (E) reversibly binds to the substrate (S) to form a complex (C) with forward and backward rate constants of kf and kb, which is irreversibly catalyzed into the product (P), with rate constant of kcat, releasing the enzyme to catalyze additional substrate. The total enzyme concentration is defined to be the ET := E + C. The total substrate and product concentration is defined to be ST := S + C + P. The Michaelis constant is the defined to be the kM := (kb + kcat) / kf. The kcat rate constant determines the maximum turn over at saturating substrate concentrations, Vmax := kcat \* ET. The rate constants kcat and kM can be estimated by monitoring the product accumulation over time (enzyme progress curves), by varying the enzyme and substrate concentrations.

By assuming that the enzyme concentration is very low (ET << ST), they derived their celebrated Michaelis-Menten kinetics. Since their work, a number of groups have developed models for enzyme kinetics that make less stringent assumptions. Recently (Choi, Rempala, and Kim 2017), described a Bayesian model for the total QSSA model. To make their model more accessible, we have re-implemented it in the Stan/BRMS framework and made it available through the BayesPharma package.

## Outline for Vignette

Next we will formally define the problem and formulate the model as the solution to an ordinary differential equation. To illustrate, we will consider a a toy system where we assuming the kcat and kM are known and simulate a sequence of measurements using deSolve. We will then implement the ODE in Stan/BRMS using stanvars and show how the parameters of the toy system can be estimated. Since it is common to vary the enzyme and substrate concentrations in order to better estimate the kinematic parameters, we will show how we can improve the Stan/BRMS model to allow multiple observations, each with an arbitrary number of measurements. Then finally, we will consider a real enzyme kinetics data set and use the Stan/BRMS model to estimate the kinematic parameters. We will compare estimated parameters with those fit using standard approaches.

## Problem Statement

Implement the total QSSA model in stan/BRMS, a refinement of the classical Michaelis-Menten enzyme kinetics ordinary differential equation described in (Choi, et al., 2017, DOI: 10.1038/s41598-017-17072-z). From their equation 2:

Observed data:  
 M = number of measurements # The product concentration Pt is measured  
 t[M] = time # at M time points t  
 Pt[M] = product #   
 ST = substrate total conc. # Substrate and enzyme concentrations are  
 ET = enzyme total conc. # assumed to be given for each observation  
  
 Model parameters:  
 kcat # catalytic constant  
 kM # Michaelis constant  
  
 ODE formulation:  
 dPdt = kcat \* ( # Change in product concentration at time t  
 ET + kM + ST - Pt +   
 -sqrt((ET + kM + ST - Pt)^2 - 2\* ET \* (ST - Pt))) / 2  
 initial condition:  
 P := 0 # There is zero product at time 0

In (Choi, Rempala, and Kim 2017) they prove, that the tQ model is valid when

where K = kb/kf is the dissociation constant.

## Simulate one observation

Using the deSolve package we can simulate data following the total QSSA model. Measurements are made with random Gaussian noise with mean 0 and variance of 0.5. To visualize, the true enzyme progress curve is shown in blue, and the enzyme progress curve fit to the noisy measurements with a smooth loess spline is shown in orange. While the smooth fits well, we cannot estimate the parameters for the curve from it.

tQ\_model\_generate <- function(  
 time, kcat, kM, ET, ST) {  
 ode\_tQ <- function(time, Pt, theta) {  
 list(c(theta[1] \* (  
 ET + theta[2] + ST - Pt -  
 sqrt(  
 (ET + theta[2] + ST - Pt)^2 -  
 4 \* ET \* (ST - Pt))) / 2))  
 }  
 deSolve::ode(  
 y = 0, times = time, func = ode\_tQ,  
 parms = c(kcat, kM))  
}  
  
data\_single <- tQ\_model\_generate(  
 time = seq(0.00, 3, by=.05),  
 kcat = 3, kM = 5, ET = 10, ST = 10) |>  
 as.data.frame() |>  
 dplyr::rename(P\_true = 2) |>  
 dplyr::mutate(  
 P = rnorm(dplyr::n(), P\_true, 0.5), # add some observational noise  
 ST = 10, ET = 10)  
  
head(data\_single)

time P\_true P ST ET  
1 0.00 0.0000000 -0.1952277 10 10  
2 0.05 0.7311578 0.3337647 10 10  
3 0.10 1.4243598 1.8453044 10 10  
4 0.15 2.0794197 2.1755056 10 10  
5 0.20 2.6964485 3.2259417 10 10  
6 0.25 3.2758537 4.1491231 10 10

|  |
| --- |
| Simulated enzyme progress curve with parameters kcat=3, kM=5, and total substrace and enzyme concentrations to 10 nM, simulated over 3 seconds |

## Fitting a single ODE observation in BRMS

To implement in BRMS, we can use the stanvars to define custom functions. The key idea is call the ODE solver, in this case the backward differentiation formula (bdf) used to solve stiff ODEs, passing a function ode\_tQ that returns dP/dt, the change in product at time t. The ode\_tQ function depends on the product at time t as the state vector, the kinematic parameters to be estimated kcat and kM and the user-provided data of the enzyme and substrate concentrations ET and ST. To call ode\_dbf we pass in the initial product concentration and time (both equal to zero), measured time-points, parameters and user defied data. Finally we, extract the vector of sampled vector of product concentrations which we return.

stanvars\_tQ\_ode <- brms::stanvar(scode = paste("  
vector tQ\_ode(  
 real time,  
 vector state,  
 vector params,  
 data real ET,  
 data real ST) {  
   
 real Pt = state[1]; // product at time t  
 real kcat = params[1];  
 real kM = params[2];  
 vector[1] dPdt;  
 dPdt[1] = kcat \* (  
 ET + kM + ST - Pt  
 -sqrt((ET + kM + ST - Pt)^2 - 4 \* ET \* (ST - Pt))) / 2;  
 return(dPdt);  
}  
", sep = "\n"),  
block = "functions")

stanvars\_tQ\_single <- brms::stanvar(scode = paste("  
vector tQ\_single(  
 data vector time,  
 vector vkcat,  
 vector vkM,  
 data vector vET,  
 data vector vST) {  
   
 vector[2] params;  
 params[1] = vkcat[1];  
 params[2] = vkM[1];  
 vector[1] initial\_state;  
 initial\_state[1] = 0.0;  
 real initial\_time = 0.0;  
 int M = size(time);  
  
 vector[1] P\_ode[M] = ode\_bdf( // Function signature:  
 tQ\_ode, // function ode  
 initial\_state, // vector initial\_state  
 initial\_time, // real initial\_time  
 to\_array\_1d(time), // array[] real time  
 params, // vector params  
 vET[1], // ...  
 vST[1]); // ...  
   
 vector[M] P; // Need to return a vector not array  
 for(i in 1:M) P[i] = P\_ode[i,1];  
 return(P);  
}  
", sep = "\n"),  
block = "functions")

To use this function, we define kcat and kM as parameters and that we wish to sample P ~ tQ(...). Since all the data points define a single observation, we set loop = FALSE. We use gamma priors for kcat and kM with the shape parameter alpha=4 and the rate parameter beta=1. The prior mean is alpha/beta = 4/1 = 4 and the variance is alpha/beta^2 = 4/1 = 4. We also bound the parameters from below by 0. We initialize each chain at the prior mean and use cmdstanr version 2.29.2 as the backend, and use the default warmup of 1000

model\_single <- brms::brm(  
 formula = brms::brmsformula(  
 P ~ tQ\_single(time, kcat, kM, ET, ST),  
 kcat + kM ~ 1,  
 nl = TRUE,  
 loop=FALSE),  
 data = data\_single |> dplyr::filter(time > 0),  
 prior = c(  
 brms::prior(prior = gamma(4, 1), lb = 0, nlpar = "kcat"),  
 brms::prior(prior = gamma(4, 1), lb = 0, nlpar = "kM")),  
 init = function() list(kcat = 4, kM = 4),  
 stanvars = c(  
 stanvars\_tQ\_ode,  
 stanvars\_tQ\_single))

Fitting the model takes ~15 seconds, with Rhat = 1 and effective sample size for the bulk and tail greater than 1400 for both parameters. The estimates and 95% confidence intervals are good.

Family: gaussian   
 Links: mu = identity; sigma = identity   
Formula: P ~ tQ\_single(time, kcat, kM, ET, ST)   
 kcat ~ 1  
 kM ~ 1  
 Data: dplyr::filter(data\_single, time > 0) (Number of observations: 60)   
 Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;  
 total post-warmup draws = 4000  
  
Population-Level Effects:   
 Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
kcat\_Intercept 2.97 0.47 2.20 4.03 1.00 1232 1536  
kM\_Intercept 4.32 2.09 1.17 9.20 1.00 1201 1387  
  
Family Specific Parameters:   
 Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
sigma 0.52 0.05 0.44 0.62 1.00 1931 1578  
  
Draws were sampled using sampling(NUTS). For each parameter, Bulk\_ESS  
and Tail\_ESS are effective sample size measures, and Rhat is the potential  
scale reduction factor on split chains (at convergence, Rhat = 1).

To visualize the posterior distribution vs. the prior distribution, we first sample from the prior, using the same brms::brm call with sample\_prior = "only" the argument.

model\_single\_prior <- model\_single |>  
 stats::update(  
 sample\_prior = "only",  
 iter = 2000)

And to plot, we use tidybayes to gather the draws and ggplot2 to map them to curves, with the prior as the orange curve, posterior as the blue curve, and the true parameter marked as a vertical line.

|  |
| --- |
|  |

Next, we plot the prior and posterior samples as a scatter plot. Note that the high correlation of the kcat and kM parameters in the posterior. This is expected, and typically better estimates require varying the enzyme and substrate concentrations.

|  |
| --- |
|  |

## Fitting multiple observations

Next we will extend the BRMS model to allow fitting common kcat, kM concentrations based on multiple replicas, or varying substrate/enzyme concentrations using BRMS. To demonstrate, we varying the enzyme and substrate concentrations, to better fit the kinematic parameters.

data\_multiple <- tidyr::expand\_grid(  
 kcat = 3,  
 kM = 5,  
 ET = c(3, 10, 30),  
 ST = c(3, 10, 30)) |>  
 dplyr::mutate(observation\_index = dplyr::row\_number()) |>  
 dplyr::rowwise() |>  
 dplyr::do({  
 data <- .  
 time <- seq(0.05, 3, by=.05)  
 data <- data.frame(data,  
 time = time,  
 P = tQ\_model\_generate(  
 time = time,  
 kcat = data$kcat,  
 kM = data$kM,  
 ET = data$ET,  
 ST = data$ST)[,2])  
 }) |>  
 dplyr::mutate(P = rnorm(dplyr::n(), P, 0.5))

|  |
| --- |
|  |

Next we will implement tQ\_multiple as a brms::stanvar object.

stanvars\_tQ\_multiple <- brms::stanvar(scode = paste("  
vector tQ\_multiple(  
 array[] int replica,  
 data vector time,  
 vector vkcat,  
 vector vkM,  
 data vector vET,  
 data vector vST) {  
  
 int N = size(time);  
 vector[N] P;  
 int begin = 1;  
 int current\_replica = replica[1];  
 for (i in 1:N){  
 if(current\_replica != replica[i]){  
 P[begin:i-1] = tQ\_single(  
 time[begin:i-1],  
 vkcat,  
 vkM,  
 vET[begin:i-1],  
 vST[begin:i-1]);  
 begin = i;  
 current\_replica = replica[i];  
 }  
 }  
 P[begin:N] = tQ\_single(time[begin:N], vkcat, vkM, vET[begin:N], vST[begin:N]);  
 return(P);  
}", sep = "\n"),  
block = "functions")

Then we will use it to fit multiple measurements for the same enzyme

model\_multiple <- brms::brm(  
 formula = brms::brmsformula(  
 P ~ tQ\_multiple(observation\_index, time, kcat, kM, ET, ST),  
 kcat + kM ~ 1,  
 nl = TRUE,  
 loop=FALSE),  
 data = data\_multiple,  
 prior = c(  
 brms::prior(prior = gamma(4, 1), lb = 0, nlpar = "kcat"),  
 brms::prior(prior = gamma(4, 1), lb = 0, nlpar = "kM")),  
 init = function() list(kcat = 4, kM = 4),  
 stanvars = c(  
 stanvars\_tQ\_ode,  
 stanvars\_tQ\_single,  
 stanvars\_tQ\_multiple))  
  
model\_multiple

To assess the model fit, we will re-fit the model just sampling from the prior

model\_multiple\_prior <- model\_multiple |>  
 stats::update(  
 sample\_prior = "only",  
 iter = 2000)

|  |
| --- |
|  |

Next we will sample enzyme progress curves from the posterior

|  |
| --- |
|  |

## Over dispersed negative binomial response

The response of an assay results from a measurement of the experimental system. Often the measurements are normalized so that the response for negative control is 1 (e.g. diseased) and the positive control is 0 (e.g. healthy). However when the robustness of the measurement depends on the measured value, this normalization can make it difficult to combine different measurements. An alternative approach is to model the measurements directly, to take into account the uncertainty associated with the response. To illustrate, if the measurement is the relative number of cells having a phenotype, then five out of ten cells and five thousand out of ten thousand cells will have the same response of 0.5, but the former will a less reliable measurement.

Arezooji, Dorsa Mohammadi. 2020. “A Markov Chain Monte-Carlo Approach to Dose-Response Optimization Using Probabilistic Programming (RStan),” November. <https://arxiv.org/abs/2011.15034>.

Ashby, Deborah. 2006. “Bayesian Statistics in Medicine: A 25 Year Review.” *Stat. Med.* 25 (21): 3589–3631. <https://doi.org/10.1002/sim.2672>.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. “Fitting Linear Mixed-Effects Models Using Lme4.” *J. Stat. Softw.* 67 (October): 1–48. <https://doi.org/10.18637/jss.v067.i01>.

Betancourt, Michael. 2023. “Michael Betancourt: Writing.” <https://betanalpha.github.io/writing/>.

Bezanson, Jeff, Alan Edelman, Stefan Karpinski, and Viral B Shah. 2017. “Julia: A Fresh Approach to Numerical Computing.” *SIAM Rev.* 59 (1): 65–98. <https://doi.org/10.1137/141000671>.

Blei, David M, Alp Kucukelbir, and Jon D McAuliffe. 2017. “Variational Inference: A Review for Statisticians.” *J. Am. Stat. Assoc.* 112 (518): 859–77. <https://doi.org/10.1080/01621459.2017.1285773>.

Bürkner, Paul-Christian. 2017. “Brms: An R Package for Bayesian Multilevel Models Using Stan.” *J. Stat. Softw.* 80 (1): 1–28. <https://doi.org/10.18637/jss.v080.i01>.

Campbell, Gregory. 2017. “Bayesian Methods in Clinical Trials with Applications to Medical Devices.” *Communications for Statistical Applications and Methods* 24 (6): 561–81.

Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. “Stan: A Probabilistic Programming Language.” *J. Stat. Softw.* 76 (January). <https://doi.org/10.18637/jss.v076.i01>.

Carpenter, Bob, Matthew D Hoffman, Marcus Brubaker, Daniel Lee, Peter Li, and Michael Betancourt. 2015. “The Stan Math Library: Reverse-Mode Automatic Differentiation in c++,” September. <https://arxiv.org/abs/1509.07164>.

Choi, Boseung, Grzegorz A Rempala, and Jae Kyoung Kim. 2017. “Beyond the Michaelis-Menten Equation: Accurate and Efficient Estimation of Enzyme Kinetic Parameters.” *Sci. Rep.* 7 (1): 17018. <https://doi.org/10.1038/s41598-017-17072-z>.

Cooner, F. 2019. “Considerations and Bayesian Applications in Pharmaceutical Development for Rare Diseases.” *Bayesian Applications in Pharmaceutical Development*. <https://doi.org/10.1201/9781315099798-17/considerations-bayesian-applications-pharmaceutical-development-rare-diseases-freda-cooner>.

Depaoli, Sarah, and Rens van de Schoot. 2017. “Improving Transparency and Replication in Bayesian Statistics: The WAMBS-Checklist.” *Psychol. Methods* 22 (2): 240–61. <https://doi.org/10.1037/met0000065>.

Faya, Paul, Perceval Sondag, Steven Novick, Dwaine Banton, John W Seaman, Jr, James D Stamey, and Bruno Boulanger. 2021. “The Current State of Bayesian Methods in Nonclinical Pharmaceutical Statistics: Survey Results and Recommendations from the DIA / ASA‐BIOP Nonclinical Bayesian Working Group.” *Pharmaceutical Statistics*. <https://doi.org/10.1002/pst.2072>.

Gabry, Jonah, and Tristan Mahr. 2017. “Bayesplot: Plotting for Bayesian Models.” *R Package Version* 1 (0).

Gelman, Andrew, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. 2013. *Bayesian Data Analysis*. Chapman; Hall/CRC.

Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gelman, Andrew, and Donald B Rubin. 1992. “Inference from Iterative Simulation Using Multiple Sequences.” *Stat. Sci.* 7 (4): 457–72.

Gelman, Andrew, Aki Vehtari, Daniel Simpson, Charles C Margossian, Bob Carpenter, Yuling Yao, Lauren Kennedy, Jonah Gabry, Paul-Christian Bürkner, and Martin Modrák. 2020. “Bayesian Workflow,” November. <https://arxiv.org/abs/2011.01808>.

Gould, A Lawrence. 2019. “BMA-Mod: A Bayesian Model Averaging Strategy for Determining Dose-Response Relationships in the Presence of Model Uncertainty.” *Biom. J.* 61 (5): 1141–59. <https://doi.org/10.1002/bimj.201700211>.

GraphPad Software, LLC. 2023. *GraphPad Prism*.

Grieve, Andrew P. 2007. “25 Years of Bayesian Methods in the Pharmaceutical Industry: A Personal, Statistical Bummel.” *Pharm. Stat.* 6 (4): 261–81. <https://doi.org/10.1002/pst.315>.

Herbert Lee. 2023. “Coursera: Bayesian Statistics Specialization.” <https://www.coursera.org/specializations/bayesian-statistics>.

Hoffman, Matthew D, Andrew Gelman, and Others. 2014. “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo.” *J. Mach. Learn. Res.* 15 (1): 1593–623.

James Bradbury. 2018. “JAX: Composable Transformations of Python+NumPy Programs.”

Johnson, Alicia A, Miles Q Ott, and Mine Dogucu. 2022. *Bayes Rules!: An Introduction to Applied Bayesian Modeling*. CRC Press.

Johnstone, Ross H, David J Gavaghan, Ross H Johnstone, Rémi Bardenet, David J Gavaghan, and Gary R Mirams. 2017. “Hierarchical Bayesian Inference for Ion Channel Screening Dose-Response Data [Version 2; Peer Review: 2 Approved].” *Wellcome Open Research* 1.

Kay, M. 2018. “Tidybayes: Tidy Data and Geoms for Bayesian Models.” *R Package Version* 1 (3).

Kruschke, John K. 2021. “Bayesian Analysis Reporting Guidelines.” *Nat Hum Behav* 5 (10): 1282–91. <https://doi.org/10.1038/s41562-021-01177-7>.

Kucukelbir, Alp, Rajesh Ranganath, Andrew Gelman, and David Blei. 2015. “Automatic Variational Inference in Stan.” In *Advances in Neural Information Processing Systems 28*, edited by C Cortes, N D Lawrence, D D Lee, M Sugiyama, and R Garnett, 568–76. Curran Associates, Inc.

Labelle, Caroline, Anne Marinier, and Sébastien Lemieux. 2019. “Enhancing the Drug Discovery Process: Bayesian Inference for the Analysis and Comparison of Dose–Response Experiments.” *Bioinformatics* 35 (14): i464–73. <https://doi.org/10.1093/bioinformatics/btz335>.

Lakshminarayanan, Mani, and Fanni Natanegara. 2019. *Bayesian Applications in Pharmaceutical Development*. CRC Press.

Lesaffre, Emmanuel, Gianluca Baio, and Bruno Boulanger. 2020. *Bayesian Methods in Pharmaceutical Research*. CRC Press. <https://doi.org/10.1201/9781315180212>.

Lock, Eric F, and David B Dunson. 2015. “Shared Kernel Bayesian Screening.” *Biometrika* 102 (4): 829–42. <https://doi.org/10.1093/biomet/asv032>.

Ma, Eric J, and Arkadij Kummer. 2021. “Principled Decision-Making Workflow with Hierarchical Bayesian Models of High-Throughput Dose-Response Measurements.” *Entropy* 23 (6). <https://doi.org/10.3390/e23060727>.

Margolis, Elyssa B, Tanya L Wallace, Lori Jean Van Orden, and William J Martin. 2020. “Differential Effects of Novel Kappa Opioid Receptor Antagonists on Dopamine Neurons Using Acute Brain Slice Electrophysiology.” *PLoS One* 15 (12): e0232864. <https://doi.org/10.1371/journal.pone.0232864>.

McElreath, Richard. 2016. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. CRC Press/Taylor & Francis Group.

Paszke, Adam, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, et al. 2019. “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” December. <https://arxiv.org/abs/1912.01703>.

Plummer, Martyn. 2003. “JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling.” In *3rd International Workshop on Distributed Statistical Computing*. 124th Series.

Polack, Fernando P, Stephen J Thomas, Nicholas Kitchin, Judith Absalon, Alejandra Gurtman, Stephen Lockhart, John L Perez, et al. 2020. “Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine.” *N. Engl. J. Med.* 383 (27): 2603–15. <https://doi.org/10.1056/NEJMoa2034577>.

Posit Software, PBC. 2023. “Bookdown.org Tag: Bayesian.” <https://bookdown.org/home/tags/bayesian/>.

Ritz, Christian, Florent Baty, Jens C Streibig, and Daniel Gerhard. 2015. “Dose-Response Analysis Using R.” *PLoS One* 10 (12): e0146021. <https://doi.org/10.1371/journal.pone.0146021>.

Ruberg, Stephen J, Francois Beckers, Rob Hemmings, Peter Honig, Telba Irony, Lisa LaVange, Grazyna Lieberman, James Mayne, and Richard Moscicki. 2023. “Application of Bayesian Approaches in Drug Development: Starting a Virtuous Cycle.” *Nat. Rev. Drug Discov.*, February, 1–16. <https://doi.org/10.1038/s41573-023-00638-0>.

Salvatier, John, Thomas V Wiecki, and Christopher Fonnesbeck. 2016. “Probabilistic Programming in Python Using PyMC3.” *PeerJ Comput. Sci.* 2 (April): e55. <https://doi.org/10.7717/peerj-cs.55>.

Semenova, Elizaveta, Maria Luisa Guerriero, Bairu Zhang, Andreas Hock, Philip Hopcroft, Ganesh Kadamur, Avid M Afzal, and Stanley E Lazic. 2021. “Flexible Fitting of PROTAC Concentration–Response Curves with Changepoint Gaussian Processes.” *SLAS DISCOVERY: Advancing the Science of Drug Discovery* 26 (9): 1212–24. <https://doi.org/10.1177/24725552211028142>.

Senn, Stephen. 2022. “The Design and Analysis of Vaccine Trials for COVID-19 for the Purpose of Estimating Efficacy.” *Pharm. Stat.* 21 (4): 790–807. <https://doi.org/10.1002/pst.2226>.

Shterev, Ivo D, David B Dunson, Cliburn Chan, and Gregory D Sempowski. 2021. “BHTSpack: Bayesian Multi-Plate High-Throughput Screening of Compounds.”

Smith, Michael K, and Scott Marshall. 2006. “A Bayesian Design and Analysis for Dose-Response Using Informative Prior Information.” *J. Biopharm. Stat.* 16 (5): 695–709. <https://doi.org/10.1080/10543400600860535>.

Tansey, Wesley, Kathy Li, Haoran Zhang, Scott W Linderman, Raul Rabadan, David M Blei, and Chris H Wiggins. 2021. “Dose–Response Modeling in High-Throughput Cancer Drug Screenings: An End-to-End Approach.” *Biostatistics* 23 (2): 643–65. <https://doi.org/10.1093/biostatistics/kxaa047>.

Team, R Core, and Others. 2013. “R: A Language and Environment for Statistical Computing.” Vienna, Austria.

Van de Schoot, Rens, Duco Veen, Laurent Smeets, Sonja D Winter, and Sarah Depaoli. 2020. “A Tutorial on Using the WAMBS Checklist to Avoid the Misuse of Bayesian Statistics.” *Small Sample Size Solutions: A Guide for Applied Researchers and Practitioners; van de Schoot, R. , Miocevic, M. , Eds*, 30–49.

Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2017a. “Erratum to: Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC.” *Stat. Comput.* 27 (5): 1433–33. <https://doi.org/10.1007/s11222-016-9709-3>.

———. 2017b. “Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC.” *Stat. Comput.* 27 (5): 1413–32. <https://doi.org/10.1007/s11222-016-9696-4>.

Wei, Xin, Lin Gao, Xiaolei Zhang, Hong Qian, Karen Rowan, David Mark, Zhengwei Peng, and Kuo-Sen Huang. 2013. “Introducing Bayesian Thinking to High-Throughput Screening for False-Negative Rate Estimation.” *J. Biomol. Screen.* 18 (9): 1121–31. <https://doi.org/10.1177/1087057113491495>.

Wickham, Hadley. 2009. *Ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the Tidyverse.” *J. Open Source Softw.* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.

Yang, Harry, and Steven Novick. 2019. *Bayesian Analysis with R for Drug Development: Concepts, Algorithms, and Case Studies*. CRC Press.