Statistical analysis of whisker stimulation data

Teddy Groves

Table of contents

In order to measure how the vascular responsiveness changed during our experimental protocol, the diameters of different vessel types was recorded before and during whisker stimulation, at baseline, post-hypertension and post-ablation stages. The ratio of the peak compared with the pre-stimulation level for each mouse at each stage, on natural logarithmic scale, also known as the 'log change', was standardised by subtracting the overall mean and dividing by the standard deviation, then treated as a single measurement.

This way of the measurements was chosen to facilitate modelling, as log change is a symmetric and additive measure of relative change (see Tornqvist, Vartia, and Vartia (1985)). Note that when the difference between the two values $v1$ and $v2$ is

far less than 1, the log change $\ln \frac{v2}{v1}$is approximately the same as the classical relative difference $\frac{v2-v1}{v1}$.

Questions

We were interested in two specific questions pertaining to whisker stimulation:

1a) Is there a difference between old and adult mice in the diameter log change for different vessels?

1b) Does sphincter ablation affect diameter log change to a different extent for adult and old mice?

To answer these questions, we fit a Bayesian multilevel regression model.

Missing data

[DESCRIBE MISSING DATA]

We assume that all missing measurements were caused by factors that were unrelated to our main target process (equivalently that the absent measurements were "missing at random"). We therefore did not attempt to model the measurement removal process explicitly.

Implementation

We implemented our model in Stan (Carpenter et al. 2017) and interfaced with data via the Python library cmdstanpy (Stan Development Team 2022). Posterior analysis was facilitated by the library arviz (Kumar et al. 2019). The analysis was orchestrated using the template package bibat (Groves 2023).

All the code implementing the analysis, as well as instructions for reproduction, can be found at https://github/teddygroves/sphincter.

Statistical Model

The final model is shown below:

$$
\begin{aligned}
y_{vtm} &\sim ST(\nu, \hat{y}_{vtm}, \sigma) \qquad (1)\\
\hat{y}_{vtm} &= \mu + \alpha_t^{treatment} + \alpha_v^{vessel} + \alpha_{vt}^{protocol} + \beta_{vt}^{age} \cdot old(m) + \alpha_{vtm}^{mouse}\\
\alpha_t^{treatment} &\sim N(0, \tau^{treatment})\\
\alpha_v^{vessel} &\sim N(0, \tau^{vessel})\\
\alpha_{vt}^{protocol} &\sim N(0, \tau^{protocol})\\
\alpha_{vtm}^{mouse} &\sim N(0, \tau_t^{mouse})\\
\beta_{vt}^{age} &\sim N(0, \tau^{age})\\
\nu &\sim Gamma(2, 0.1)\\
\sigma &\sim HN(0, 0.5)\\
\mu &\sim N(0, 0.7)\\
\tau^{treatment} &\sim HN(0, 0.7)\\
\tau^{vessel} &\sim HN(0, 0.7)\\
\tau^{protocol} &\sim HN(0, 0.7)
\end{aligned}
$$

In equation (1), the term ST indicates the student t distribution, *old* is an indicator function with value $old(m) = 1$ if mouse m is old, and zero otherwise, N indicates the normal distribution, $Gamma$ the gamma distribution and HN the 'half-normal' distribution, i.e. the normal distribution with support only for non-negative numbers.

This model was the end result of fitting a series of Bayesian multilevel models, following the strategy outlined in Gelman et al. (2020). The prior standard deviation 0.7 was chosen because it led to what we judged to be a reasonable allocation of prior probability mass over possible data realisations. The prior for the student t degrees of freedom parameter ν was set following the recommendation in Juárez and Steel (2010).

This model is flexible enough to allow age effects that vary depending on treatment and vessel type, but also allows these parameters to be shrunk towards zero if the data suggests that there is little difference between these categories. Our questions

can be answered by inspecting the parameters β^{age}. If, for a particular vessel type, our model allocates low probability mass to small values of this parameter, this indicates a difference between old and adult mice with respect to this vessel type. In particular, if the values for the sphincter vessel type tend to be far from zero, that would show an age difference for this vessel type.

Results

Figure 1 shows the observed log change measurements with colours illustrating the various categorical information. Note that there is more variation in the baseline log change values than in the values after either treatment.

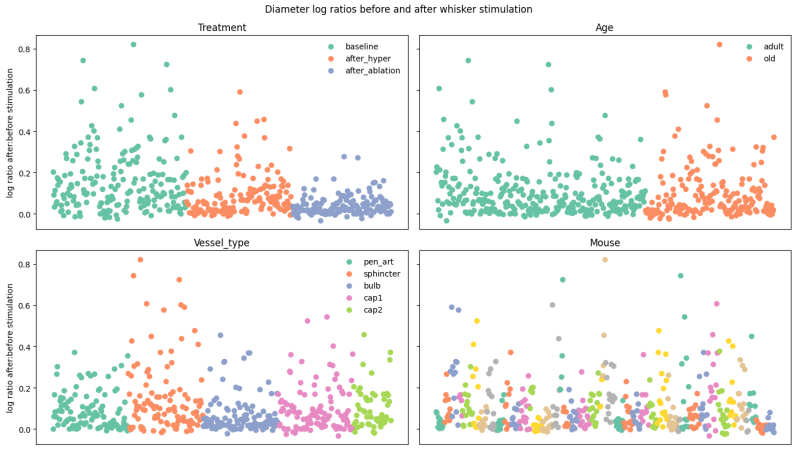

Figure 1: Raw measurements

Figure 2 compares the measurements with our model's posterior predictive distribution. This shows that our model achieved a reasonable fit to the observed data. There is a pattern in the model's bad predictions, in that these tend to be for higher baseline measurements. However, we judged that this pattern was small enough that for our purposes we could disregard it.

Figure 3 answers our questions 1a) and 1b): we found no significant age effects for any vessel or treatment.

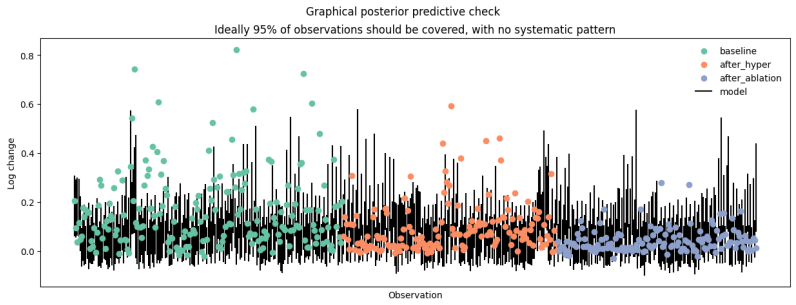

Figure 2: Graphical posterior predictive check

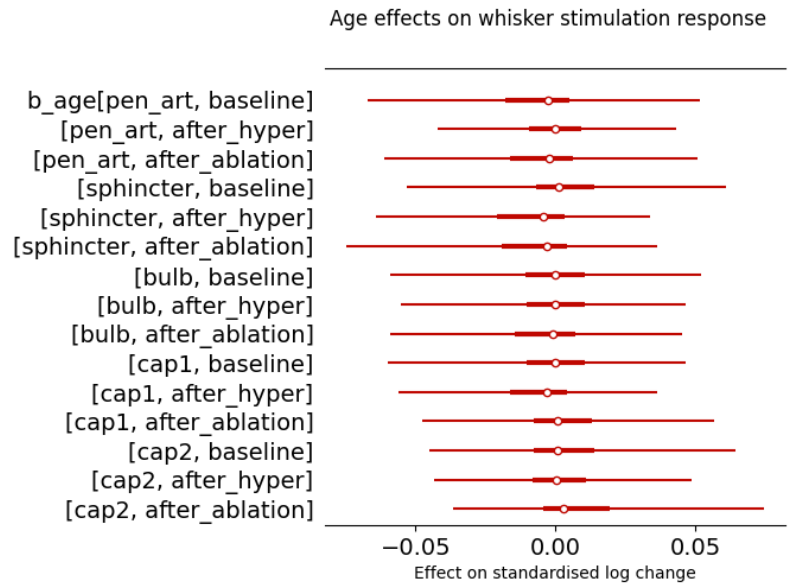

Figure 3: Marginal 2.5%-97.5% posterior intervals for age effects

This does not mean that there were no significant treatment effects, but just that these effects did not depend on age or vessel type. Figure 4 shows the marginal posterior distributions for hypertension and ablation effects relative to the baseline, showing a clear effect in both cases.

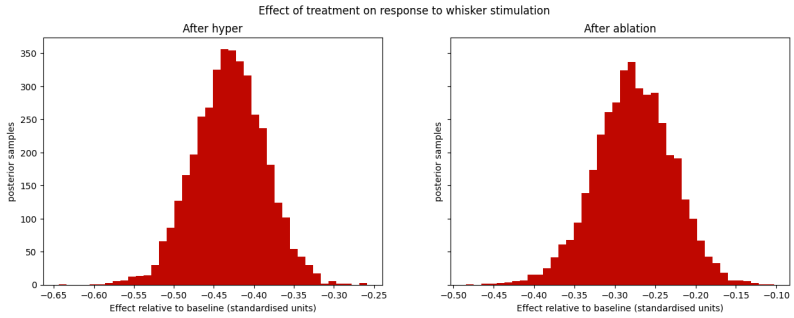

Figure 4: Marginal posterior intervals for treatment effects, relative to the baseline treatment.

These effects were the same for all vessel types, as can be seen from Figure 5, which shows that these were largely irrelevant.

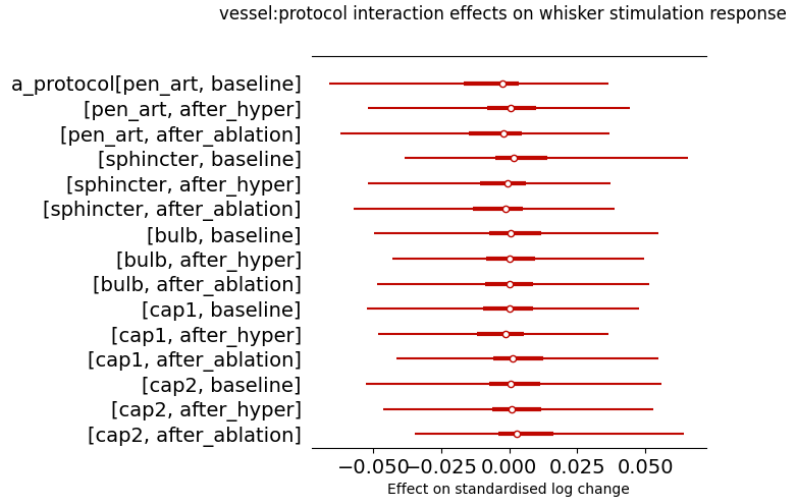

Figure 5: Marginal 2.5%-97.5% posterior intervals for protocol effects

There was also a significant distributional effect of treatment. This is captured by Figure 6, which shows marginal posterior

distributions for the parameters τ^{mouse}. It was higher for the baseline treatment than for either of the interventions, reflecting the pattern noted above in the raw data.

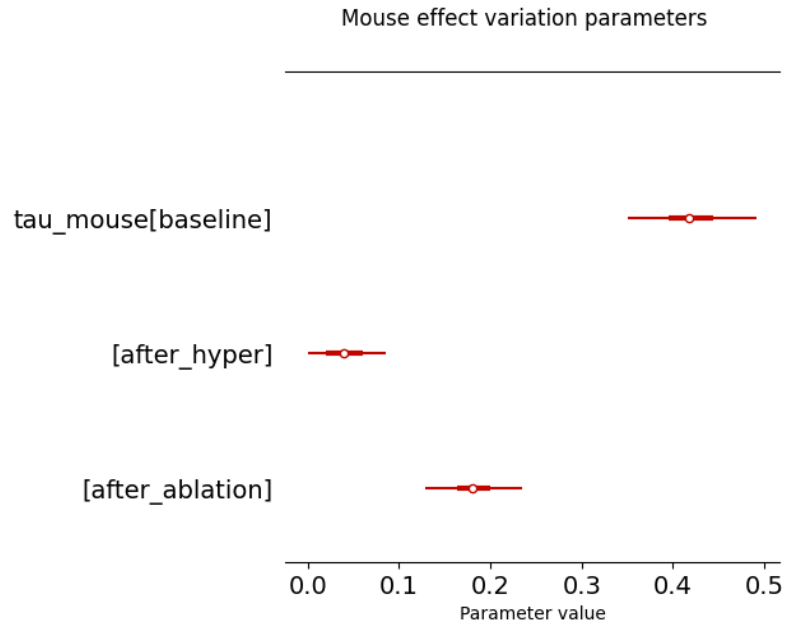

Figure 6: Marginal 2.5%-97.5% posterior intervals for τ^{mouse} parameters.

Pulsatility data

The pulsatility data consisted of fast measurements of diameter and center point for the same mice. These measurements were Fourier-transformed, and the harmonics of the transformed data were interpreted as representing the pulsatility of the measured quantities.

We aimed to answer the following questions:

a) Do diameter or center pulsatility depend on age? If so, is this dependency mediated by the higher average blood pressure of adult mice compared with old mice?

b) Does sphincter ablation affect diameter and/or center pulsatility differently in adult and old mice?

c) How does blood pressure affect diameter and center pulsatility?

d) Do hypertension and sphincter ablation influence diameters, Pd, and Pc differently for different vessels?

To answer these questions we created a pulsatility dataset and fit a series of statistical models to it.

Description of the dataset

We used the first harmonic of each transformed time series as a dependent variable. It might have been preferable to aggregate all the available power harmonics, but this would have complicated our measurement model, and in any case power at the subsequent harmonics was typically negligible compared with the first.

For each measurement the following information was available:

- the identity of the mouse, and that mouse's age (adult or old)
- the vessel type (penetrating artery, bulb and first three capillary orders)
- the treatment (baseline, after hypertension and after ablation)
- the mouse's blood pressure, measured at the femoral artery

The final dataset included 514 joint measurements of diameter and center pulsatility, calculated as described above. These measurements are shown in Figure 7.

Figure 8 shows the relationship between pressure and the measurements in our dataset for both age categories. The light dots show raw measurements and the darker dots show averages within evenly sized bins.

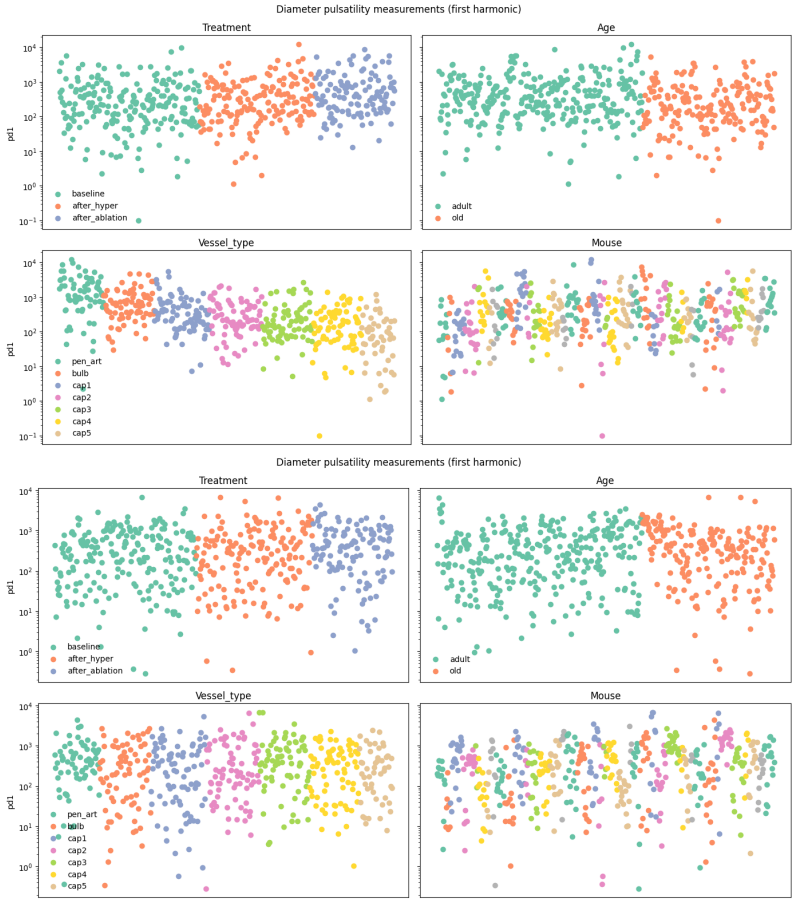

Figure 7: The modelled measurements, shown in order of the coloured categories.

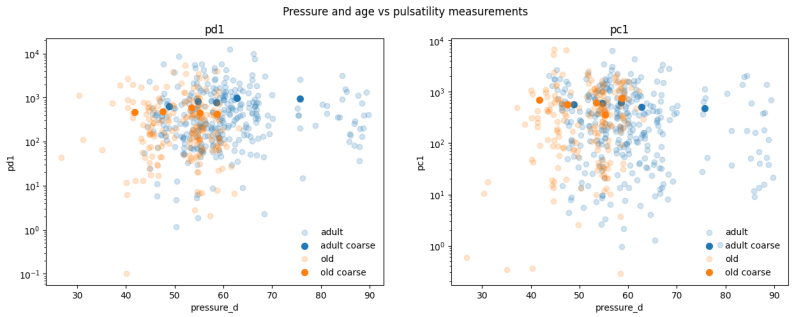

Figure 8: Pulsatility measurements plotted against the corresponding pressure measurements and coloured according to age. Darker dots indicate averages within evenly sized pressure bins.

Missing data

Data from one mouse (index 310321) were excluded after some extreme measurements were observed:

310321 is a mouse where we did not see any whisker response, it reacted to angiotensin II, but the BP increase was abrupted for a short while and then reestablished. Perhaps due to a clot or a bubble in the venous catheter. This resulted in a biphasic and slow BP increase

As with the whisker stimulation data we assumed that all absent measurements were missing at random.

Statistical model

We knew from prior studies that the power harmonics should individually follow exponential distributions [REFERENCE FOR THIS]. This consideration motivated the use of exponential generalised linear models for both the center and diameter pulsatility measurements. In this model, given measurement y and linear predictor η the measurmeent probability density is given by

$$p(y \mid \eta) = Exponential(y, \lambda) = \lambda e^{-\lambda y} \tag{2}$$

$$\ln \frac{1}{\lambda} = \eta \tag{3}$$

The log link function (3) was chosen so that linear changes in the term η induce multiplicative changes in the mean $\frac{1}{\lambda}$ of the measurement distribution, as we believed the effects we wanted to model would be multiplicative.

We compared four different ways of parameterising η based on the information available about a given measurement, corresponding to three hypotheses about the way the data were generated.

The simplest, or "basic", model postulates that the linear predictors η_d^{basic} and η_c^{basic} are the sums of independent parameters corresponding to non-interacting features as follows:

$$\begin{aligned} \eta_{d,n}^{basic} &= \mu_{d,age(mouse(n))} + \alpha_{d,treatment(n)}^{treatment} + \alpha_{d,vessel\ type(n)}^{vessel\ type} \\ \eta_{c,n}^{basic} &= \mu_{c,age(mouse(n))} + \alpha_{c,treatment(n)}^{treatment} + \alpha_{c,vessel\ type(n)}^{vessel\ type} \end{aligned} \tag{4}$$

The basic model was chosen as a plausible baseline against which to compare the other models.

The "interaction" model was chosen to explore whether there were any important interaction effects that the basic model does not consider. For example, question b) can be interpreted as asking whether there is are any significant effects corresponding to the interaction between age and treatment. Our interaction model calculates the value of the linear predictors $\eta_d^{interaction}$ and $\eta_c^{interaction}$ as depending on age-treatment and age-treatment-vessel type interaction effects as follows:

$$\eta^{interaction}_{d,n} = \mu_{d,age(mouse(n))} + \alpha^{treatment}_{d,treatment(n)} + \alpha^{vessel\ type}_{d,vessel\ type(n)} \tag{5}$$
$$\begin{aligned}
&+ \alpha^{age:treatment}_{d,age(mouse(n)),treatment(n)} \\
&+ \alpha^{age:treatment:vessel\ type}_{d,age(mouse(n)),treatment(n),vessel\ type(n)} \\
\eta^{interaction}_{c,n} &= \mu_{c,age(mouse(n))} + \alpha^{treatment}_{c,treatment(n)} + \alpha^{vessel\ type}_{c,vessel\ type(n} \\
&+ \alpha^{age:treatment}_{c,age(mouse(n)),treatment(n)} \\
&+ \alpha^{age:treatment:vessel\ type}_{c,age(mouse(n)),treatment(n),vessel\ type(n)}
\end{aligned}$$

Next, we constructed a model that adds to the basic model parameters that aim to capture possible effects corresponding to the blood pressure measurements. To compensate for collinearity between age and pressure, our "pressure" model does not use the observed pressure as a predictor, but rather the age-normalised pressure, calculated by subtracting the mean for each age category from the observed pressure measurement. The model for the linear predictors $\eta^{pressure}_d$ and $\eta^{pressure}_c$ is then as follows:

$$\eta^{pressure}_{d,n} = \mu_{d,age(mouse(n))} + \alpha^{treatment}_{d,treatment(n)} + \alpha^{vessel\ type}_{d,vessel\ type(n)} \tag{6}$$
$$\begin{aligned}
&+ \beta^{pressure}_{d,age(mouse(n))} \cdot norm\ pressure_n \\
\eta^{pressure}_{c,n} &= \mu_{c,age(mouse(n))} + \alpha^{treatment}_{c,treatment(n)} + \alpha^{vessel\ type}_{c,vessel\ type(n)} \\
&+ \beta^{pressure}_{c,age(mouse(n))} \cdot norm\ pressure_n
\end{aligned}$$

Finally, we made a model that includes a pressure effect but no age-specific parameters from the pressure model. This is to test whether any age effects are due to the collinearity between age and pressure. The pressure-no-age model's linear predictors $\eta^{pressure-no-age}_d$ and $\eta^{pressure-no-age}_c$ are calculated as shown in equation (7). Note that, unlike in equation (6), the μ and $\beta^{pressure}$ parameters in equation (7) have no age indexes.

$$
\begin{aligned}
\eta_{d,n}^{pressure-no-age} &= \mu_d + \alpha_{d,treatment(n)}^{treatment} + \alpha_{d,vessel\ type(n)}^{vessel\ type} \\
&\quad + \beta_d^{pressure} \cdot norm\ pressure_n \\
\eta_{c,n}^{pressure} &= \mu_c + \alpha_{c,treatment(n)}^{treatment} + \alpha_{c,vessel\ type(n)}^{vessel\ type} \\
&\quad + \beta_c^{pressure} \cdot norm\ pressure_n
\end{aligned} \tag{7}
$$

In all of our models the α parameters were given independent, semi-informative, hierarchical prior distributions to allow for appropriate information sharing. The $\beta^{pressure}$ parameters were given independent, semi-informative, non-hierarchical prior distributions.

Results

We estimated the leave-one-out log predictive density for each model using the method described in Vehtari, Gelman, and Gabry (2017) and implemented in Kumar et al. (2019). The results of the comparison are shown below in Figure 9.

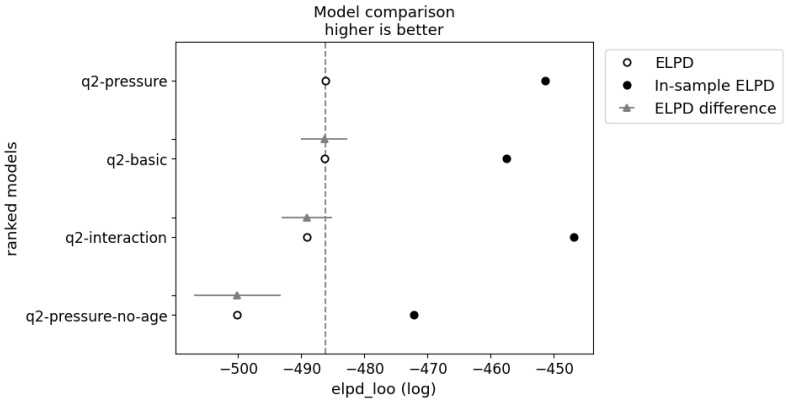

Figure 9: Comparison of estimated leave-one-out log predictive density (ELPD) for our pulsatility models. The main result is that the pressure-no-age and interaction models are clearly worse than the pressure model, as shown by the separation of the relevant grey and dotted lines.

We evaluated our models' fit to data using prior and posterior predictive checking, with the results for the pressure model shown in Figure 10.

Inspecting of the interaction model output showed that none of the interaction effect parameters that differed substantially from zero, as can be seen in Figure 11.

From this result, together with the worse estimated out of sample predictive performance as shown in Figure 9, we concluded that there were no important interaction effects, so that we could essentially discard the interaction model.

Figure 12 shows the marginal posterior distributions for other effect parameters in all three models.

Answers to specific questions

To address whether there are important age effects, Figure 13 plots the distribution of age effect differences (adult - old) for each measurement type in the pressure model.

This graph shows that, in this model, the mean parameter for diameter pulsatility in adult mice was higher than for old mice in every single posterior sample - in other words there is a clear trend for older mice to have lower diameter pulsatility. There is a smaller opposite trend for center pulsatility measurements, but it is not clearly separated from zero, indicating that the direction of the effect is not fully settled.

The next question concerns whether the age effects are explained by the generally higher blood pressure of the adult mice. This question mostly answered by the poorer estimated out of sample predictive performance of the pressure-no-age model compared with the other models as shown in Figure 9. This shows that there is information in the age labels beyond what is contained in the pressure measurements. It is nonetheless possible that different pressure explains the difference between old and adult mice, but that the pressure measurements did not reflect the true pressure at the measured vessels. This is plausible since the pressure measurements were taken at a different location.

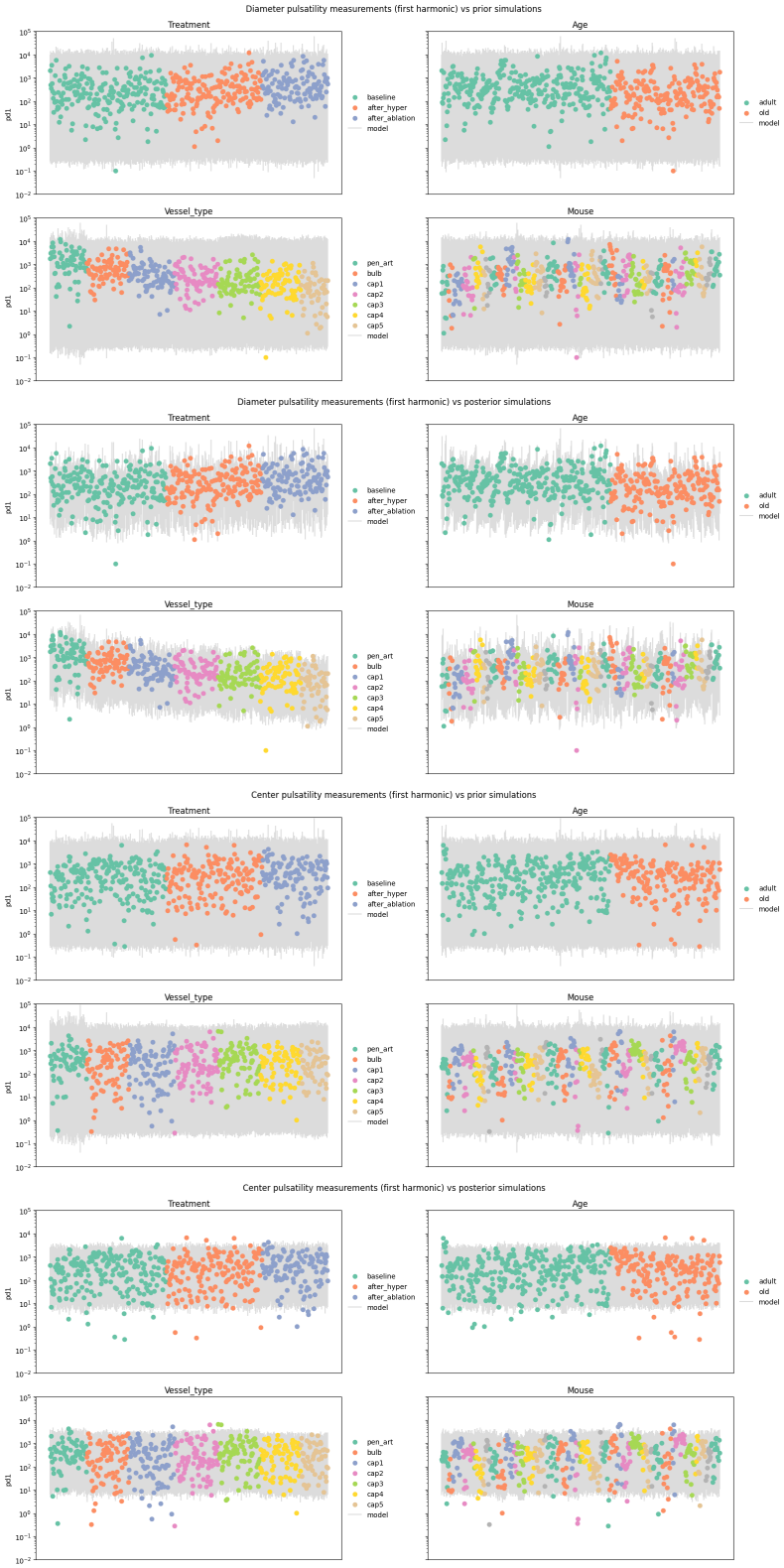

Figure 10: Prior and posterior predictive checks for the pressure model.

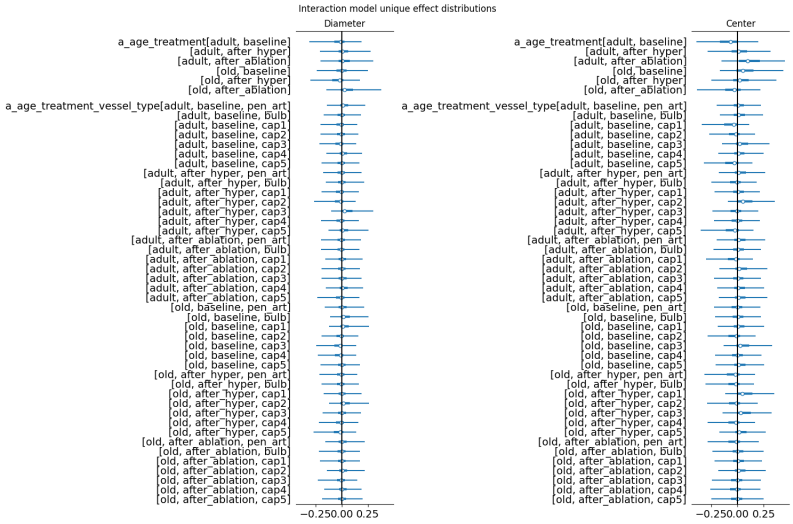

Figure 11: Marginal posterior quantiles for the unique effects in the interaction model.

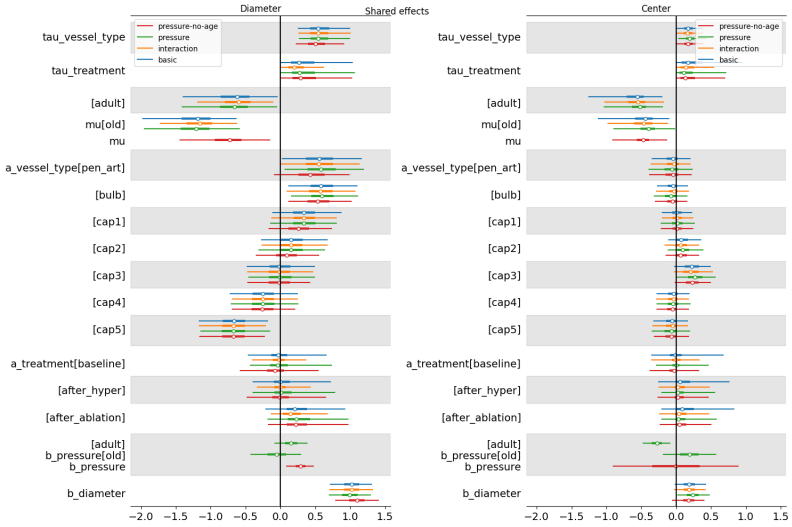

Figure 12: Marginal posterior quantiles for shared model effects.

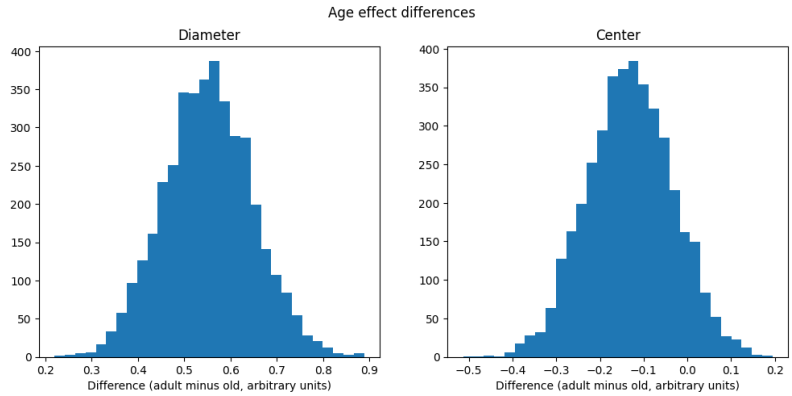

Figure 13: Posterior distribution of age effect differences for each measurement type.

Figure 14 shows the difference in $\beta^{pressure}$ parameters for old and adult mice in the pressure model. This shows noticeable but not fully resolved effects for both measurement types. Specifically, the model thinks that the pressure effect on diameter pulsatility is somewhat more positive for adult mice, while the pressure effect on center pulsatility is somewhat more negative for adult mice. The most interesting thing to take away from this graph is that there is not a clear resolution away from zero in either case. This shows that the effect of normalised pressure is not very pronounced.

Overall, this analysis suggests that higher pressure tends to somewhat increase diameter pulsatility and somewhat decrease center pulsatility, though the results are not definitive. This is partly due to the collinearity between age and pressure that makes the effects of the two difficult to distinguish, and perhaps also due to the limitations of the pressure measurements.

To illustrate the effect of treatments, and specifically sphincter ablation relative to hypertension, Figure 15 shows the difference between the effect for each treatment and the baseline treatment effect. There is a clear effect of ablation to increase diameter pulsatility, a smaller tendency of hypertension to do the same, and no clear effects of either treatment on center pulsatility.

Carpenter, Bob, Andrew Gelman, Matthew D. Hoffman, Daniel

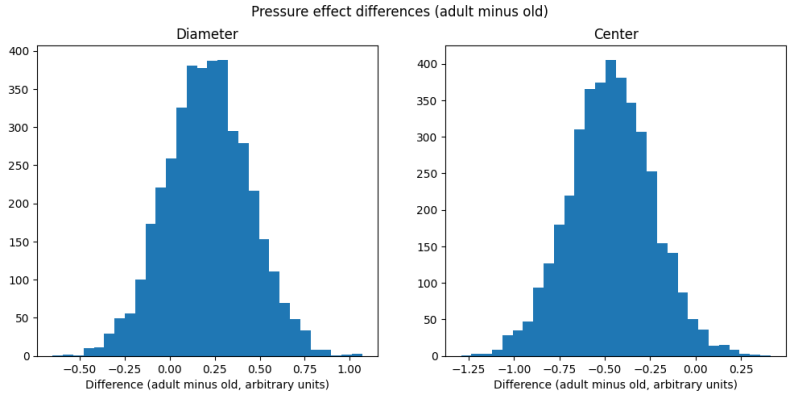

Figure 14: Posterior distribution of pressure effect differences for each measurement type.

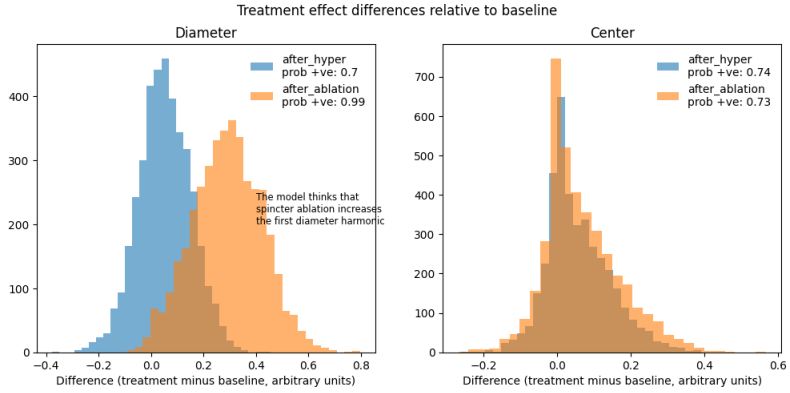

Figure 15: Posterior distribution of treatment effect differences for each measurement type.

Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76 (1): 1–32. https://doi.org/10.18637/jss.v076.i01.

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." *arXiv:2011.01808 [Stat]*, November. https://arxiv.org/abs/2011.01808.

Groves, Teddy. 2023. "Bibat: Batteries-Included Bayesian Analysis Template." https://doi.org/10.5281/zenodo.7775328.

Juárez, Miguel A., and Mark F. J. Steel. 2010. "Model-Based Clustering of Non-Gaussian Panel Data Based on Skew-t Distributions." *Journal of Business & Economic Statistics* 28 (1): 52–66. https://doi.org/10.1198/jbes.2009.07145.

Kumar, Ravin, Colin Carroll, Ari Hartikainen, and Osvaldo Martin. 2019. "ArviZ a Unified Library for Exploratory Analysis of Bayesian Models in Python." *Journal of Open Source Software* 4 (33): 1143. https://doi.org/10.21105/joss.01143.

Stan Development Team. 2022. "CmdStanPy." https://github.com/stan-dev/cmdstanpy.

Tornqvist, Leo, Pentti Vartia, and Yrjo O. Vartia. 1985. "How Should Relative Changes Be Measured?" *The American Statistician* 39 (1): 43–46. https://doi.org/10.2307/2683905.

Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2017. "Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC." *Statistics and Computing* 27 (5): 1413–32. https://doi.org/10.1007/s11222-016-9696-4.