Appendix for "Exploring the function of the BBSome using clinical data: Meta-analysis of genotype-phenotype associations in Bardet-Biedl Syndrome"

Contents

This document consists of three parts: * Part 1 describes the Bayesian analysis reported in the main manuscript.
* Part2 describes all Bayesian models we tried throughout this project and discusses the reasoning behind our choice of model for the main analysis, in particular why between-study variability is crucial and taking into account complete loss of function (cLOF) useful while age, and sex can be omitted. * Part 3 shows how the conclusions of the paper hold under multiple different models.

The complete source code for the analysis can be found at https://github.com/martinmodrak/bbs-metaanalysis-bayes or Zenodo, DOI: 10.5281/zenodo.2545609

Part 1: Main Bayesian Analysis

The data

First let us examine some of the properties of the data se we are working with - a brief summary follows.

```
## Skim summary statistics
##  n obs: 916
##  n variables: 19
##
## -- Variable type:character -------------------------------
##        variable missing complete   n min max n_unique
##             age     424      492 916   1   7       92
##         case_no       4      912 916   1  20      810
##  mutation_types      72      844 916   3  11        5
##             sex     403      513 916   1   1        2
##
## -- Variable type:factor ----------------------------------
##          variable missing complete   n n_unique
##         age_group     424      492 916        7
##  functional_group       0      916 916        4
##              gene       0      916 916       20
##  loss_of_function       0      916 916        2
##               Sex     403      513 916        2
##            source       0      916 916       85
##
## -- Variable type:numeric ---------------------------------
##                   variable missing complete   n mean n_unique
##                         CI     234      682 916 0.66        2
##                      HEART     692      224 916 0.29        2
##                        LIV     631      285 916 0.3         2
##  loss_of_function_certain       0      916 916 0.51        2
##                        OBE     150      766 916 0.89        2
##                         PD     169      747 916 0.8         2
##                         RD      65      851 916 0.94        2
##                        REN     236      680 916 0.51        2
##                        REP     471      445 916 0.59        2
```

Note in particular, that both age and sex are missing in almost half of the records. Also, the data about individual phenotypes (all the numeric columns) is largely incomplete. Some minor clearing is required to use age, as it is stored as character (a combination of age ranges and ages). For some phenotypes we get values that are not 0 or 1 - those correspond to patients that were monitored in multiple studies, but the phenotype data was inconsistent between studies. In our analysis we treat those patients as exhibiting the phenotype.

`loss_of_function` and `loss_of_function_certain` are different codings of the fact that the protein has certainly lost function (is truncated in both alleles).

For some analyses, we group the genes together according to *functional groups*, those are defined as folows:

functional_group	genes
BBS03	BBS03
BBSome	BBS04,BBS01,BBS07,BBS09,BBS02,BBS05,BBS08,BBS18
Chaperonins	BBS10,BBS06,BBS12
Others	BBS13,BBS14,BBS16,BBS20,BBS21,BBS17,BBS11,BBS19

And here are the counts of individual mutations as observed in the data:

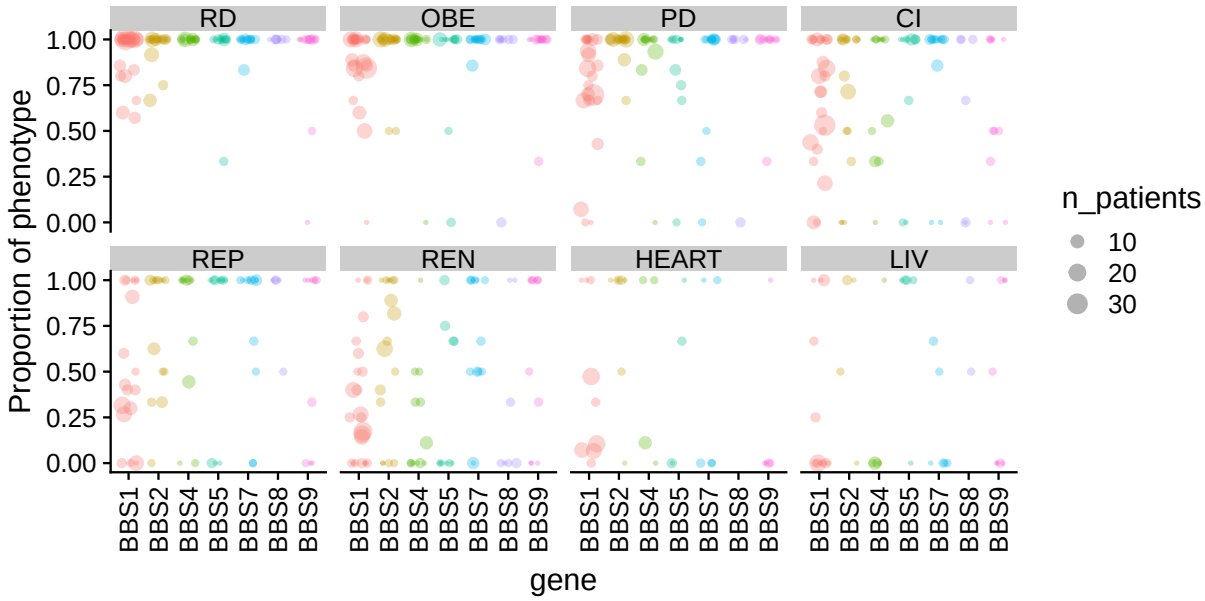
gene	count
BBS01	218
BBS02	85
BBS03	47
BBS04	51
BBS05	34
BBS06	60
BBS07	38
BBS08	19
BBS09	29
BBS10	139
BBS11	1
BBS12	59
BBS13	9
BBS14	58
BBS16	39
BBS17	3
BBS18	1
BBS19	2
BBS20	9
BBS21	15

While we include all of the genes in our computational model, we will mostly show only the most frequent mutations in the results here, those include BBS1 through BBS10 and BBS12.

The phenotypes present in the data are:

```
## [1] "RD"    "OBE"   "PD"    "CI"    "REP"   "REN"   "HEART" "LIV"
```

The data shows considerable between-study (source) variability (showing only the BBSome genes for clarity):



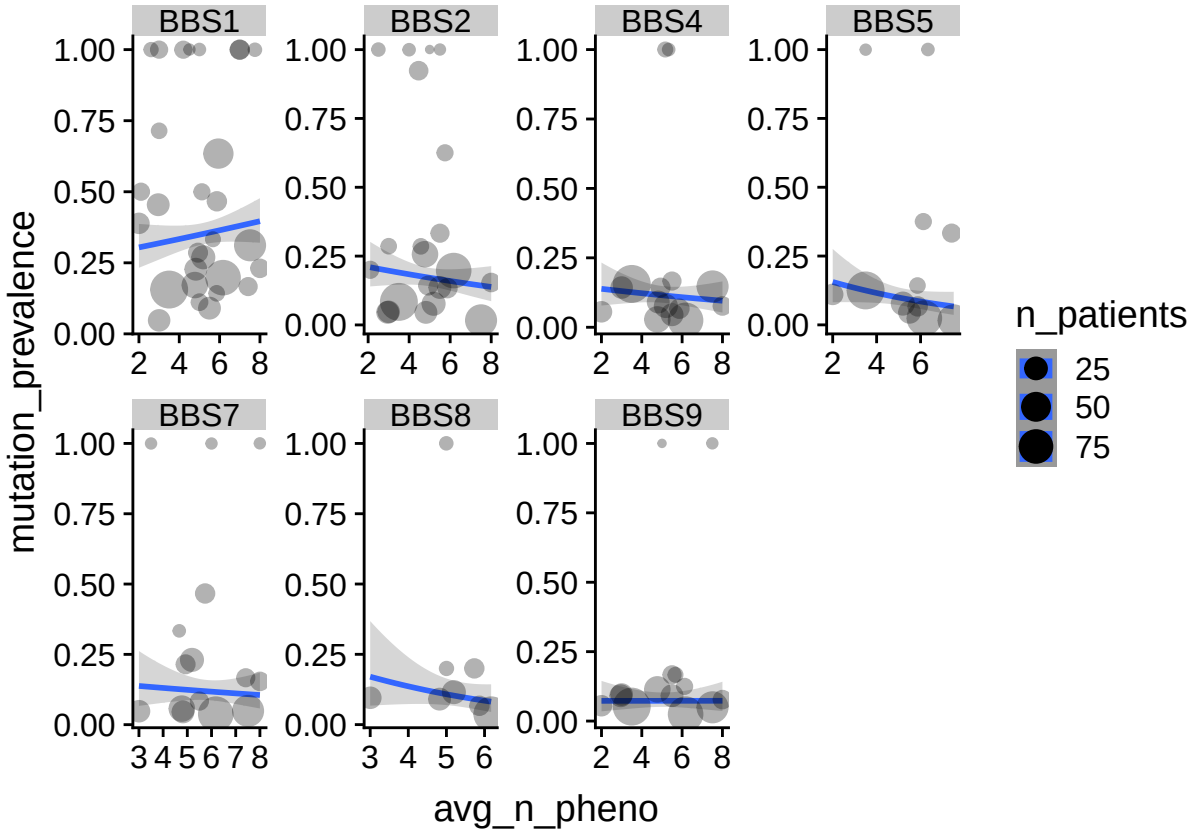
Here, every point represents the proportion of patients with a given mutation that manifested a given

phenotype in one study. The point size represents the number of patients with the mutation in the study. We see that even studies with a relatively large number of patients show very different proportions. We are therefore confident that allowing for between-study heterogeneity is important for analyzing the data correctly. In the attached multiverse analysis, we also attempted to model the studies as homogenous and found it to be a bad fit.

Handling missingness

The most problematic missing data problem is the missingness in phenotype data. There are two distinct sources of missingness: a) a study missing the phenotype value for only some patients or b) a study not reporting the status of the phenotype for any patient. Our analysis assumes the phenotype data to be missing at random, i.e. that the decision to not report a given phenotype in a study and missingness for individual patients is independent of the prevalence of the phenotype in the study population. This is probably not true for missingness at the study level, as investigators are plausibly more likely to report more prevalent phenotypes and more likely to ignore phenotype that was not observed in any patient. Similarly, if data for a specific patient omits a given phenotype (the state of the phenotype is reported as missing data in the original study), it is more likely the phenotype was not present.

However, in our analysis we were unable to find a good way to account for this phenomenon. But since we focus on comparison of individual phenotype prevalence across different mutations within a single study and do not compare phenotypes against each other, this should only be a significant issue if the rate of missingness in phenotype values is correlated with the prevalence of individual mutations present in a study (e.g., if studies with high obesity missingness would also tend to have overabundance of mutations in BBS3). Let's check whether this is the case:



In the figure above, each dot is a single study and shows the average number of phenotypes reported per patient vs. the prevalence of mutations in individual genes. Point size corresponds to the number of patients

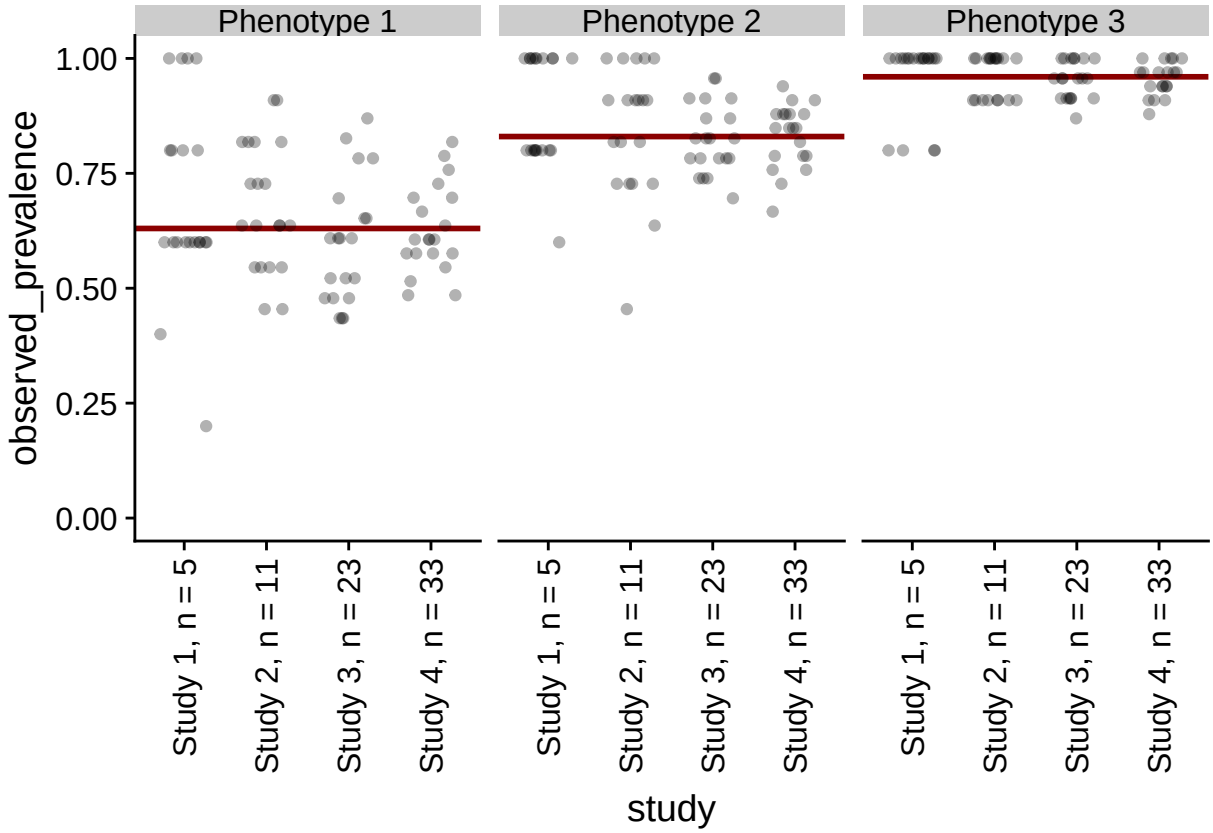
in a given study. Looking at the figure, a strong association of missingness to specific mutations seems implausible, but we can't completely rule out that it biases our results. This is a limitation of our approach and should be taken into account when interpreting our conclusions.

There is large missingness in age and sex data as well. While our main analysis ignores age and sex, the attached multiverse analysis shows that after accounting for between-study differences, age and sex differences are already mostly accounted for. We have also tried to impute age and sex data and show that models including imputed age and/or sex provide almost identical results.

The model - an accessible explanation

As all models, the model we use simplifies and abstracts the medical reality in hope we can arrive at useful conclusions. Our model is a member of generalized linear model family, using logit link and hierarchial terms in a fully Bayesian treatment. Let's unpack this a little, starting with what a logit link does. In the following, we will describe how we handle a single phenotype as the estimation for individual phenotypes is mostly independent.

Our model tries to estimate theoretical *true prevalence* of the phenotype in a population - i.e. the probability that a randomly selected patient from the population will exhibit this phenotype. But all we observe is that each individual either exhibits the phenotype or not. Depending on the number of individuals enrolled in a study, the *observed prevalence* will jump more or less around this true prevalence - let's have a look at an example:



The red line shows the true prevalence of a phenotype (assuming it is the same across multiple studies). Each point shows the observed prevalence of a single possible realization of a study. We see that the observed prevalence can differ substantially from the true value and that the spread decreases with increasing n. In our data we can't however expect large n as the biggest n we have is (Castro-Sanchez et al. 2015) with 33 patients having mutation in BBS01 and in 74% of cases there are less than 5 patients with a given mutation

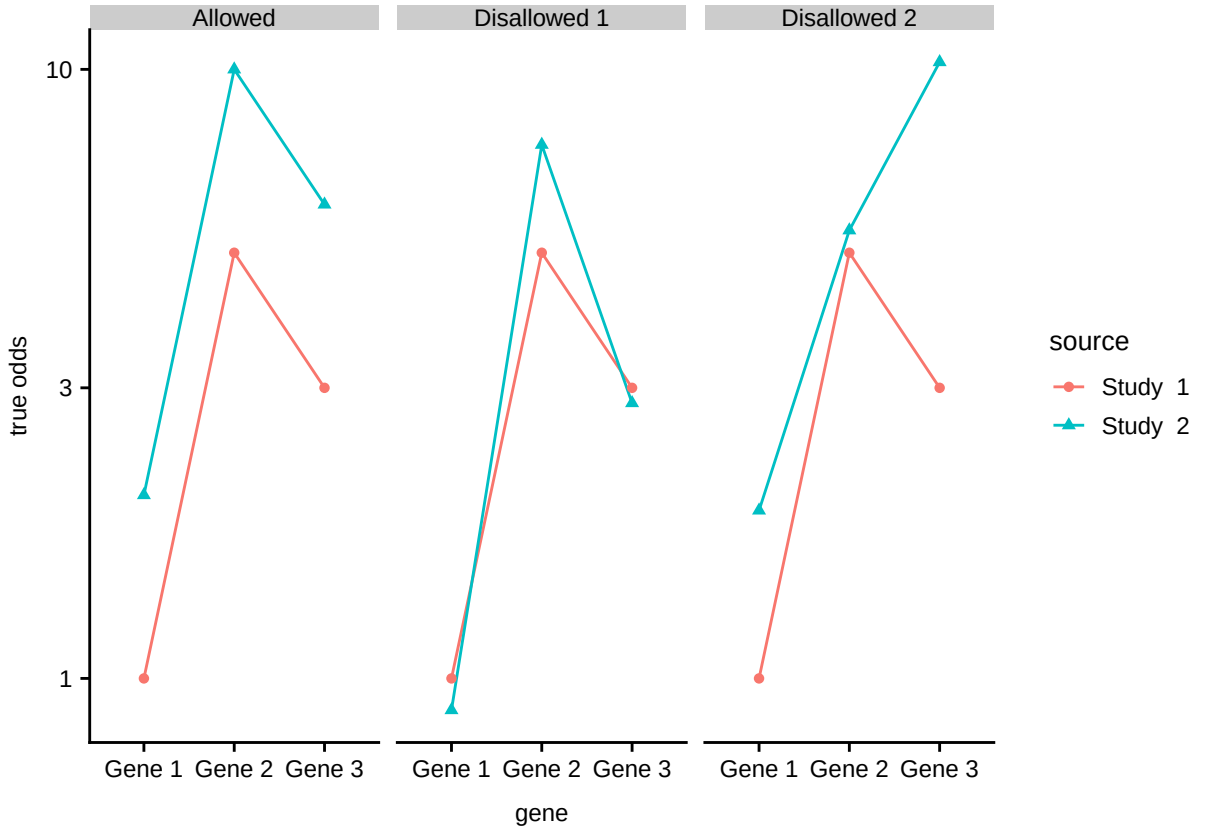
in the same study. Also note that the observed values are clustered at discrete "levels" because e.g., among 5 patients you only can have prevalence of 0.2 or 0.4 and nothing in between. Further, you can see that with high prevalence, there is less variation in the observed prevalence.

For mathematical convenience, the model does not work with prevalence directly, but with *odds*. Odds are just another way to express prevalence - for example, when the prevalence is 20%, we expect one patients to exhibit the phenotype for each four patients not exhibiting the phenotype, leading to odds $1:4$ or 0.25 and log odds (base 10 here) of roughly -0.6. Unlike prevalence, which is constrained between 0 and 1, odds can be any non-negative number.

Most of the results of the model are reported as comparisons of odds in different populations. For this we use the ratio of the corresponding odds. E.g., when the odds ratio is 2, we expect the first population to have twice as much patients exhibiting the phenotype for each healthy patient than the second population.

Technically the model works with the logarithm of odds. The logit function trasnforms a probability (prevalence) to log odds, hence the logit "link" used in the model. What is the linear part?

Our model assumes that the true odds of a phenotype is a function of four numbers (*coefficients*): the overall odds of the phenotype in the population, a modifier for the gene the patient has damaged and a modifier for the study the patient is enrolled in. One additional modifier is added when the mutation is a certain loss-of-function (cLOF). These four numbers are multiplied to arrive at the final odds for the patient. Assuming only cLOF mutations for simplicity, this means that while the odds of a phenotype are allowed to vary between genes and the overall rate of a phenotype may vary between studies, the odds ratio of different genes is the same across all studies. Let's look at an example:



Above on the leftmost panel we see that the two studies differ in the odds of the phenotype for each gene, but the ratio of odds for Gene 1 to Gene 2 (and 3) is the same in both studies (since the odds are shown on log scale this manifests as a constant gap between the two lines). This type of between study variation is

allowed. On the other two panels, the odds ratio for Gene 1 to Gene 2 (and 3) differs between studies - this type of variation is not allowed by the model.

Another way to describe the allowed case is that, for both studies a mutation in Gene 2 makes a patient five times more likely to exhibit the phenotype than a patient with a mutation in Gene 1, although the base rate of the phenotype may vary between studies.

The cLOF coefficient is held constant for all genes in a given phenotype, meaning that odds for any given phenotype are multiplied by a small number when the mutation is cLOF. Once again this means that relative odds are the same amongs cLOF and other mutations, but absolute odds can be higher (or lower) in cLOF mutations.

This is the "linear" part of the model - we multiply odds, which is the same as adding the logarithm of the odds and addition is a neatly linear thing.

Now the "hierarchical" part. This ties the coefficients in the model in two important ways: i) it assumes small differences in odds across genes and studies are more likely than large differences and updates the estimates accordingly ii) *partial pooling*: the degree to which odds are allowed to "jump around" across both genes and studies is informed by the data, e.g., if the odds are similar for all genes except one, the model will put higher weight on the possibility that the difference in the last mutation is just noise and shrink its estimate towards the average for other genes. On the other hand, if the odds vary wildly across all genes, the model will assume it is more likely this is a true variation and not shrink the estimate much. The amount of shrinkage also depends on the number of observations as estimates where there is a larger number of observations are shrunk less. The variability across studies is pooled in a similar way.

Together those two features result in low risk of overfitting the data, even though we have very little observations for most study - gene - phenotype combinations.

We also allow for a correlation between phenotypes, e.g., that some phenotypes occur frequently together while others rarely manifest in the same patient. Once again the amount of correlation is estimated from data.

Finally, the "Bayesian" part: We follow the Bayesian paradigm, so our estimates of the model coefficients are not a single number, but rather a distribution - some values are more likely than others, but the data are insufficient to let us determine the coefficients with high certainty. Therefore, we never report exact numbers but rather 50% and 95% *credible intervals* of the distribution. Unlike confidence intervals in frequentist analysis, we can directly interpret the 95% credible interval as the interval that contains the true value with 95% probability - assuming our model is correct (which it is not, but we hope it is still a useful abstraction).

The model - mathematical formulation

We use a generalized linear model with logit link and hierarchial terms. Let us dive into the details. For the model, we expand the data into long form, i.e. each row in the dataset corresponds to a combination of patient and reported phenotype (a patient with reported values for 3 phenotypes would correspond to 3 rows in the long form dataset). The model is specified with the following brms formula, using the Bernoulli family with logit link function:

```
## phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source) + phenotype:loss_of_function_certain
```

this can be expressed in mathematical notation as:

$$Y_i \sim Bernoulli(\mu_i)$$
$$\text{logit}(\mu_i) = \alpha + \beta^1_{p_i} + \beta^2_{p_i,g_i} + \beta^3_{p_i,s_i} + c_i\beta^4_{p_i}$$

Where $p_i \in \{1, ..., P\}$ is the index of the phenotype for i-th row, $g_i \in \{1, ..., G\}$ is mutated gene for i-th row and $s_i \in \{1, ..., S\}$ is the index of the source study for i-th row. c_i is 1 when the mutation on the i-th row is cLOF and 0 otherwise. α is the intercept and β^1, β^2 and β^3 model the overall phenotype prevalence,

phenotype prevalance specific to a given mutation and between-study variability in phenotype prevalence respectively. β^4 models the phenotype-specific effect of cLOF.

Note that this is very similar to running a separate regression for each phenotype, with two exceptions: the overall intercept α is explicitly shared between phenotypes and the structure of the priors introduces some information flow between the other coefficients.

The priors we use for the parameters are:

$$
\begin{aligned}
\alpha &\sim N(0, 2) \\
\beta^1 &\sim N(0, \sigma_1) \\
\sigma_1 &\sim N(0, 2) \\
\beta^2 &\sim \mathcal{N}_P(\mathbf{0}, \mathbf{\Sigma}) \\
\mathbf{\Sigma} &= \boldsymbol{\sigma}_2 \bar{\mathbf{\Sigma}} \\
\bar{\mathbf{\Sigma}} &\sim LKJ_P(1) \\
\sigma_{2,p} &\sim N(0, 2) \\
\beta^3_{p,s} &\sim N(0, \sigma_{3,p}) \\
\sigma_{3,p} &\sim N(0, 2) \\
\beta^4_p &\sim N(0, 2)
\end{aligned}
$$

Note that the prior on β^1 is P-dimensional multivariete normal $\mathcal{N}_P$, explicitly modelling the correlation $\bar{\mathbf{\Sigma}}$ between the prevalence of individual phenotypes and per-phenotype variance $\sigma_{2,p}$, while the other priors are univariete normal.

The $N(0, 2)$ priors on the various parameters are mildly skeptical in that they exclude that any of the parameter would explain odds ratio larger than ~ 50. The attached multiverse analysis shows that the results are almost identical when the priors are different.

Main results

First, a summary of the model fit as posterior intervals for main model parameters:

```
##  Family: bernoulli
##   Links: mu = logit
## Formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 + phenotype) || source)
##    Data: data (Number of observations: 4680)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Group-Level Effects:
## ~gene (Number of levels: 20)
##                                  Estimate Est.Error l-95% CI u-95% CI
## sd(phenotypeRD)                      0.68      0.50     0.03     1.84
## sd(phenotypeOBE)                     0.36      0.29     0.01     1.07
## sd(phenotypePD)                      1.81      0.58     0.83     3.11
## sd(phenotypeCI)                      0.86      0.29     0.41     1.49
## sd(phenotypeREP)                     0.66      0.29     0.19     1.30
## sd(phenotypeREN)                     1.29      0.38     0.71     2.16
## sd(phenotypeHEART)                   0.82      0.58     0.04     2.18
## sd(phenotypeLIV)                     1.44      0.63     0.32     2.85
## cor(phenotypeRD,phenotypeOBE)        0.04      0.33    -0.59     0.65
## cor(phenotypeRD,phenotypePD)         0.08      0.31    -0.53     0.66
## cor(phenotypeOBE,phenotypePD)        0.11      0.33    -0.55     0.71
```

```
## cor(phenotypeRD,phenotypeCI)        -0.10      0.32    -0.68     0.50
## cor(phenotypeOBE,phenotypeCI)        0.07      0.32    -0.56     0.67
## cor(phenotypePD,phenotypeCI)        -0.14      0.28    -0.64     0.40
## cor(phenotypeRD,phenotypeREP)        0.00      0.31    -0.58     0.61
## cor(phenotypeOBE,phenotypeREP)       0.07      0.32    -0.55     0.66
## cor(phenotypePD,phenotypeREP)       -0.04      0.32    -0.62     0.58
## cor(phenotypeCI,phenotypeREP)        0.22      0.29    -0.41     0.72
## cor(phenotypeRD,phenotypeREN)       -0.12      0.31    -0.68     0.53
## cor(phenotypeOBE,phenotypeREN)       0.04      0.32    -0.59     0.65
## cor(phenotypePD,phenotypeREN)       -0.09      0.26    -0.57     0.41
## cor(phenotypeCI,phenotypeREN)        0.41      0.24    -0.12     0.81
## cor(phenotypeREP,phenotypeREN)       0.31      0.28    -0.27     0.79
## cor(phenotypeRD,phenotypeHEART)     -0.08      0.33    -0.69     0.57
## cor(phenotypeOBE,phenotypeHEART)     0.05      0.33    -0.61     0.66
## cor(phenotypePD,phenotypeHEART)      0.12      0.33    -0.53     0.72
## cor(phenotypeCI,phenotypeHEART)      0.23      0.32    -0.44     0.77
## cor(phenotypeREP,phenotypeHEART)     0.08      0.32    -0.54     0.66
## cor(phenotypeREN,phenotypeHEART)     0.19      0.31    -0.43     0.73
## cor(phenotypeRD,phenotypeLIV)       -0.13      0.32    -0.69     0.53
## cor(phenotypeOBE,phenotypeLIV)      -0.09      0.33    -0.68     0.56
## cor(phenotypePD,phenotypeLIV)        0.18      0.27    -0.38     0.68
## cor(phenotypeCI,phenotypeLIV)       -0.17      0.29    -0.70     0.42
## cor(phenotypeREP,phenotypeLIV)      -0.10      0.31    -0.67     0.52
## cor(phenotypeREN,phenotypeLIV)      -0.06      0.27    -0.56     0.47
## cor(phenotypeHEART,phenotypeLIV)     0.04      0.32    -0.59     0.65
##                                Eff.Sample Rhat
## sd(phenotypeRD)                      1099 1.00
## sd(phenotypeOBE)                     1590 1.00
## sd(phenotypePD)                      1371 1.00
## sd(phenotypeCI)                      2356 1.00
## sd(phenotypeREP)                     2248 1.00
## sd(phenotypeREN)                     2204 1.00
## sd(phenotypeHEART)                   1523 1.00
## sd(phenotypeLIV)                     1585 1.00
## cor(phenotypeRD,phenotypeOBE)        5214 1.00
## cor(phenotypeRD,phenotypePD)          964 1.00
## cor(phenotypeOBE,phenotypePD)         944 1.00
## cor(phenotypeRD,phenotypeCI)         1654 1.00
## cor(phenotypeOBE,phenotypeCI)        1345 1.00
## cor(phenotypePD,phenotypeCI)         3884 1.00
## cor(phenotypeRD,phenotypeREP)        2601 1.00
## cor(phenotypeOBE,phenotypeREP)       2713 1.00
## cor(phenotypePD,phenotypeREP)        3745 1.00
## cor(phenotypeCI,phenotypeREP)        4215 1.00
## cor(phenotypeRD,phenotypeREN)        1193 1.00
## cor(phenotypeOBE,phenotypeREN)       1362 1.00
## cor(phenotypePD,phenotypeREN)        3614 1.00
## cor(phenotypeCI,phenotypeREN)        2503 1.00
## cor(phenotypeREP,phenotypeREN)       2623 1.00
## cor(phenotypeRD,phenotypeHEART)      3283 1.00
## cor(phenotypeOBE,phenotypeHEART)     3403 1.00
## cor(phenotypePD,phenotypeHEART)      4691 1.00
## cor(phenotypeCI,phenotypeHEART)      3791 1.00
## cor(phenotypeREP,phenotypeHEART)     3572 1.00
```

```
## cor(phenotypeREN,phenotypeHEART)       3723 1.00
## cor(phenotypeRD,phenotypeLIV)          1404 1.00
## cor(phenotypeOBE,phenotypeLIV)         1623 1.00
## cor(phenotypePD,phenotypeLIV)          4356 1.00
## cor(phenotypeCI,phenotypeLIV)          3360 1.00
## cor(phenotypeREP,phenotypeLIV)         2563 1.00
## cor(phenotypeREN,phenotypeLIV)         3776 1.00
## cor(phenotypeHEART,phenotypeLIV)       2880 1.00
##
## ~phenotype (Number of levels: 8)
##               Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
## sd(Intercept)     1.74      0.55     0.93     3.09       2089 1.00
##
## ~source (Number of levels: 85)
##                     Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
## sd(phenotypeRD)         2.09      0.47     1.30     3.15       1820 1.00
## sd(phenotypeOBE)        2.32      0.45     1.55     3.30       1338 1.00
## sd(phenotypePD)         1.87      0.37     1.25     2.66       1358 1.00
## sd(phenotypeCI)         1.64      0.32     1.09     2.36       1182 1.00
## sd(phenotypeREP)        2.75      0.60     1.77     4.08       1288 1.00
## sd(phenotypeREN)        1.85      0.36     1.22     2.62       1179 1.00
## sd(phenotypeHEART)      3.27      0.84     1.91     5.18       2291 1.00
## sd(phenotypeLIV)        3.06      0.79     1.77     4.82       2216 1.00
##
## Population-Level Effects:
##                                          Estimate Est.Error l-95% CI
## Intercept                                    0.97      0.63    -0.37
## phenotypeRD:loss_of_function_certain        -0.19      0.47    -1.15
## phenotypeOBE:loss_of_function_certain        1.10      0.37     0.39
## phenotypePD:loss_of_function_certain         0.61      0.33    -0.02
## phenotypeCI:loss_of_function_certain         0.51      0.25     0.04
## phenotypeREP:loss_of_function_certain        0.08      0.30    -0.52
## phenotypeREN:loss_of_function_certain        0.74      0.26     0.24
## phenotypeHEART:loss_of_function_certain     -0.92      0.53    -1.98
## phenotypeLIV:loss_of_function_certain       -0.05      0.52    -1.01
##                                          u-95% CI Eff.Sample Rhat
## Intercept                                    2.14       1495 1.00
## phenotypeRD:loss_of_function_certain         0.74       4395 1.00
## phenotypeOBE:loss_of_function_certain        1.84       4952 1.00
## phenotypePD:loss_of_function_certain         1.27       4821 1.00
## phenotypeCI:loss_of_function_certain         1.02       5081 1.00
## phenotypeREP:loss_of_function_certain        0.67       5300 1.00
## phenotypeREN:loss_of_function_certain        1.24       5169 1.00
## phenotypeHEART:loss_of_function_certain      0.13       5508 1.00
## phenotypeLIV:loss_of_function_certain        1.01       5765 1.00
##
## Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample
## is a crude measure of effective sample size, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

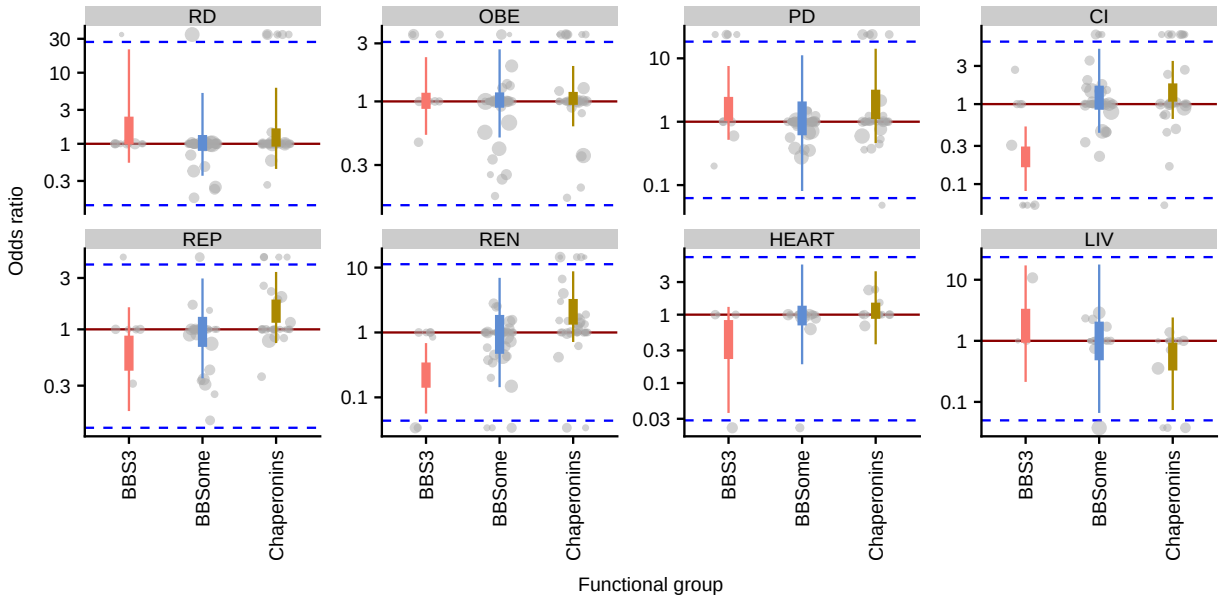
The fitted model parameters themselves are however hard to interpret, as they operate on log odds scale. It is also hard to say how to handle the between-study variability of coefficients. And this variability is substantial - note that the `sd` parameters under `~source` (corresponding to σ_3) admit ranges from ~ 1.1 to ~ 5.1, so the odds of a phenotype, given a mutation can plausibly differ between studies by $1.96 \times \pm 1.1 = \pm 2.16$ to

$1.96 \times \pm 5.09 = \pm 9.99$ *on the log scale* (95% of mass of a normal distribution is within $1.96 \times \sigma$ from the mean).

Instead, we will focus on model predictions. In particular, the results we report can be interpreted as if a new study is drawn at random from the same population of studies as we used (i.e. matching all the inclusion criteria) and we directly observe true odds of all phenotypes for all mutations in this study. That is, the predictions do include between-study variability, our uncertainty about the population of studies, our uncertainty about overall prevalence of the individual phenotypes, our uncertainty about the strength of links between mutations and phenotypes and our uncertainty about correlations between the presence of individual phenotypes. The predictions do NOT include the sampling uncertainty of the hypothetical new study. For example when the hypothetical study has the true odds of a phenotype, given a mutation in BBS12, equal to $1:2$ (0.5), a study on 20 patients can easily observe odds of $3:17$ (0.18) or $11:9$ (1.22) simply due to chance, but we will treat the hypothetical study here as having odds of 0.5 and ignore this additional noise in our results.

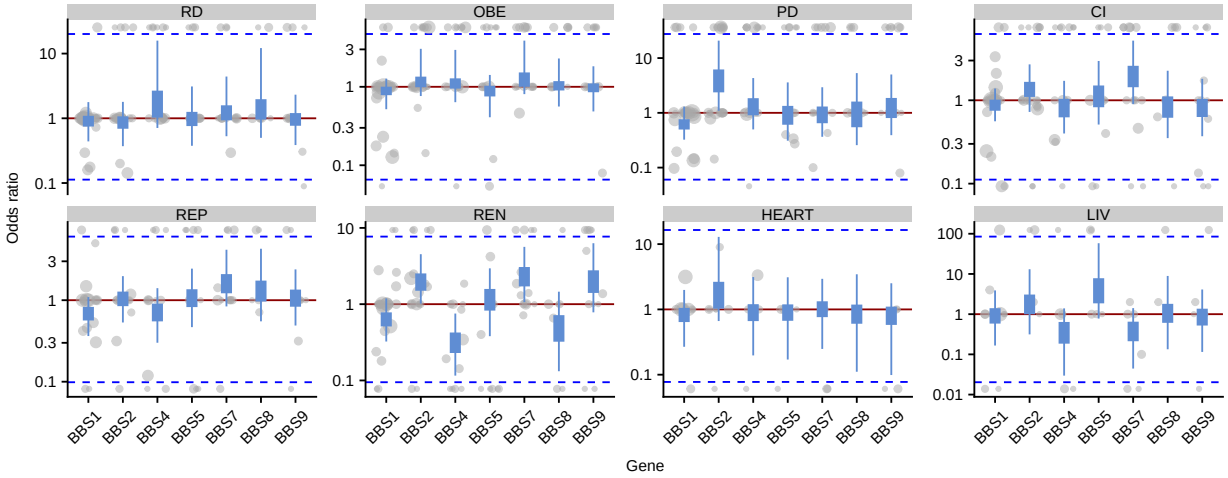
Since those odds can vary wildly between studies, we will focus on various odds ratios (OR) within a single hypothetical study.

Summary for functional groups



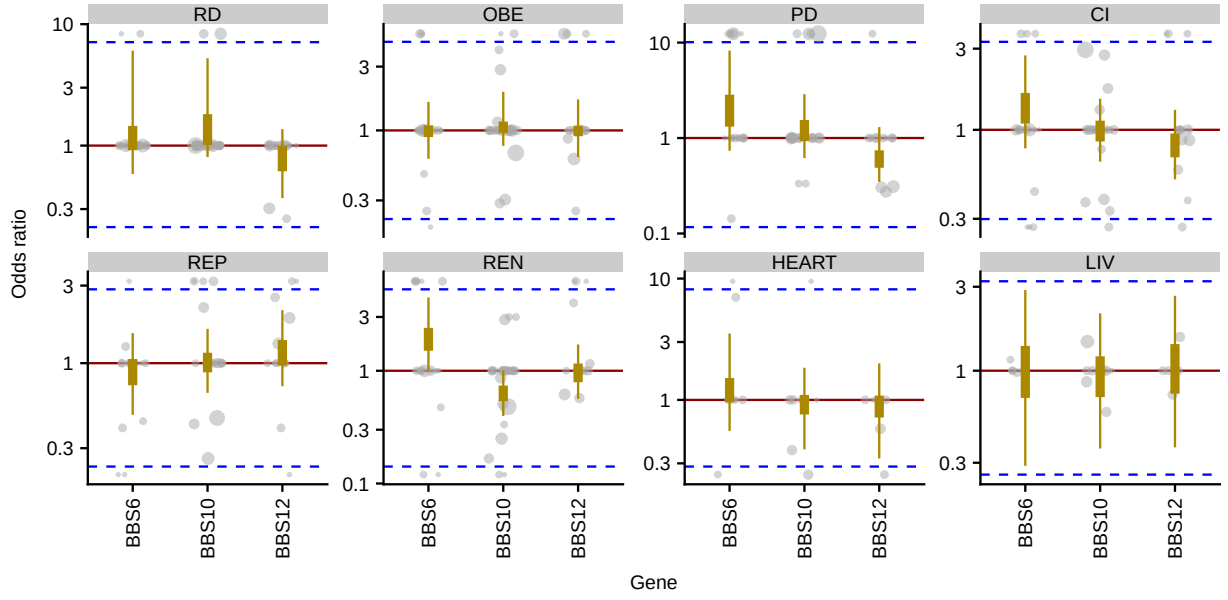
The plot above shows posterior 95% (thin) and 50% (thick) credible intervals for ratio of odds for a phenotype given a random mutation within a functional group to odds for the phenotype given a random mutation across all groups shown. All mutations are assumed to be equally likely - the odds are not weighed by the frequency of the mutations in the dataset. Odds ratios are shown on the log scale and each phenotype has its own scale. Gray dots show the same odds ratio calculated for individual studies included in the meta-analysis. Dots outside the dashed lines correspond to studies where the empirical odds ratio is 0 or infinity. Dot size represents the number of relevant cases in the study.

Summary for BBSome genes



The plot above shows posterior 95% (thin) and 50% (thick) credible intervals for ratio of odds for a phenotype given a mutation in a gene to odds for the phenotype given a random mutation across all genes shown. All mutations are assumed to be equally likely - the odds are not weighed by the frequency of the mutations in the dataset. Odds ratios are shown on the log scale and each phenotype has its own scale. Gray dots show the same odds ratio calculated for individual studies included in the meta-analysis. Dots below the dashed lines correspond to studies where the empirical odds ratio is 0 or infinity. Dot size represents the number of relevant cases in the study.

Summary for Chaperonins



This is the same plot as for the BBSome genes, only showing the members of the chaperonins group.

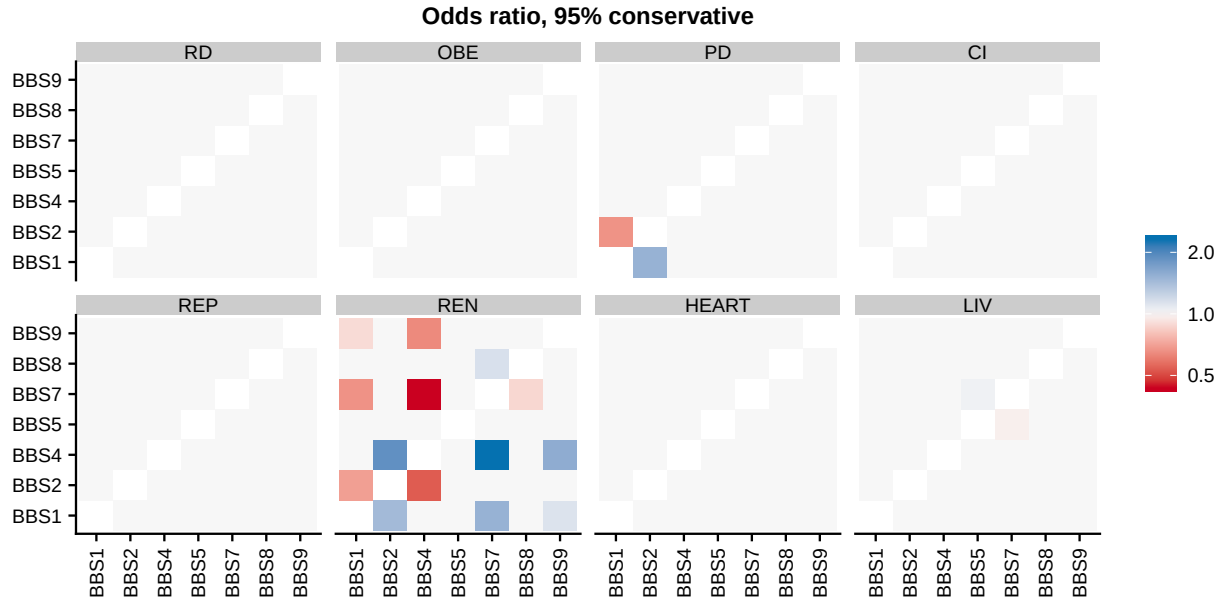
Pairwise comparisons of mutations in BBSome genes

The summary plots above are not well suited to infer pairwise comparisons, as the estimates for the individual genes / functional groups are not independent. In particular, non-overlapping marginal credible intervals in

the summary plot imply that there is a consistent difference, but the converse is not true. If there is a strong positive correlation, there might be a consistent difference even when the above plot would show mostly overlapping marginal posterior intervals.

Pairwise comparisons also have the benefit of better interpretability as we do not need to rely on odds ratio of the phenotype against some average which might not be clinically meaningful. In pairwise comparisons we can directly work with odds ratios for the phenotype given the two mutations in question.

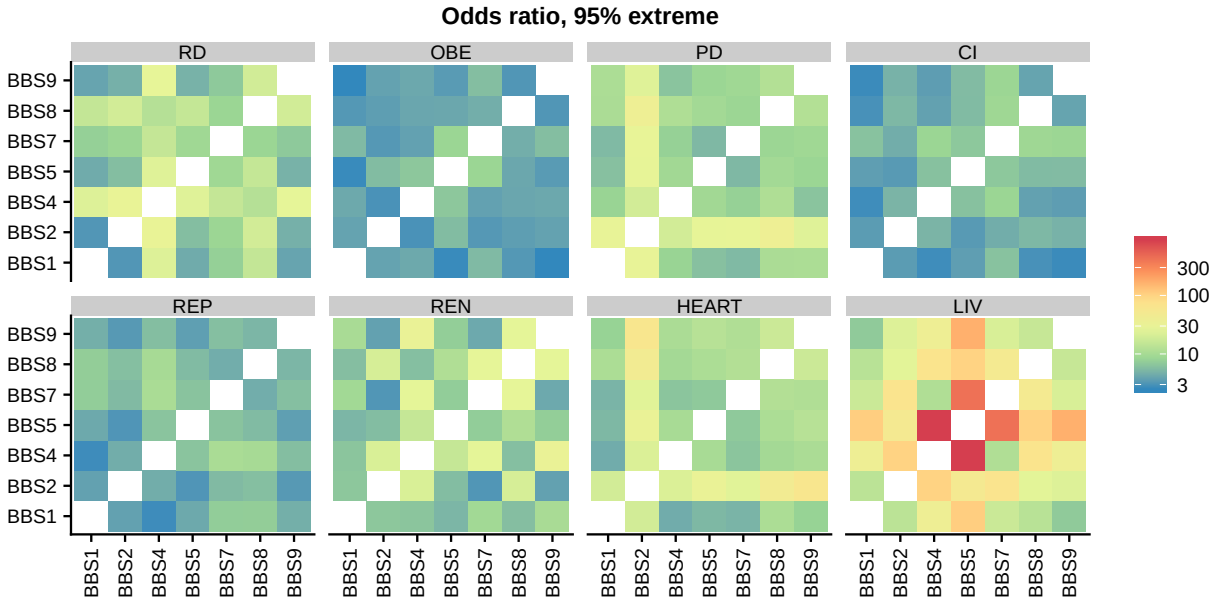
Conservative estimates



The most conservative (closest to one) pairwise odds ratios within 95% posterior credible intervals. The reported odds ratio are for gene on the horizontal axis against the gene on the vertical axis.

This shows pairs where we are fairly certain there is a systematic difference and the minimal magnitude of this difference consistent with the data.

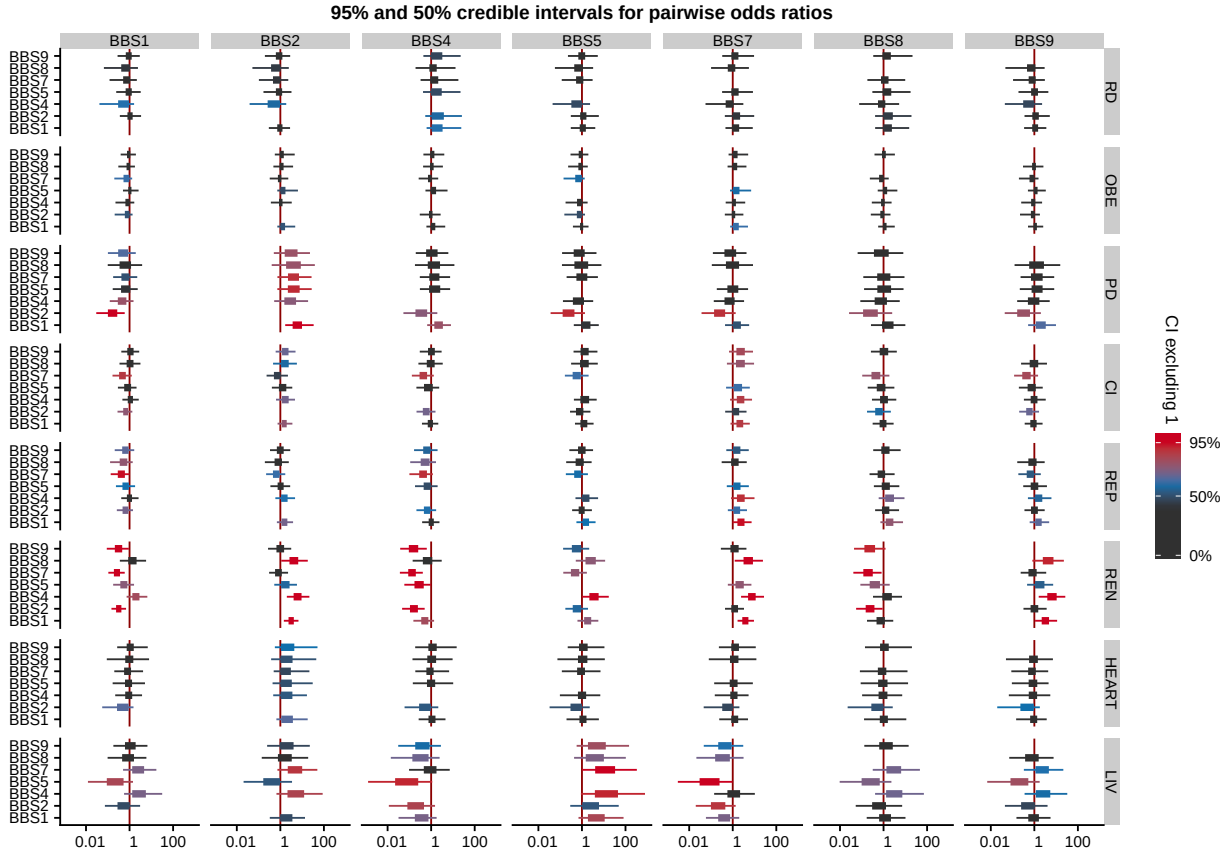
Extreme estimates



The most extreme (furthest from one) pairwise odds ratios within 95% posterior credible intervals. The direction of the effect is not reported as effects in both directions might be similarly plausible - the odds ratio are transformed to be larger than one in all cases.

This shows maximal differences consistent with our model and lets us constrain the differences between mutations for some phenotypes. We see that the data does not let us to put tight constraints on most differences - the tightest we get is OR of 3 - which would still be a very important difference for the clinical prognosis.

Everything at once

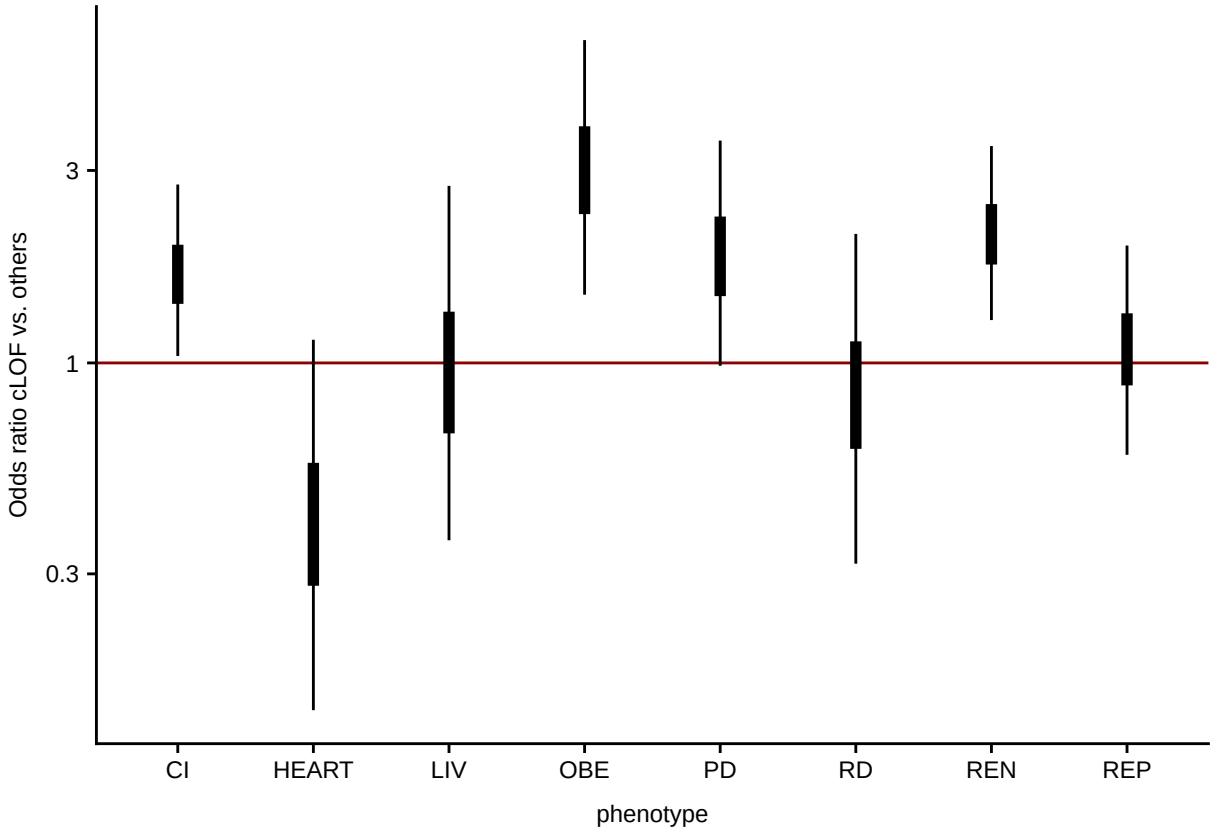


Posterior 95% (thin) and 50% (thick) credible intervals for odds ratio for a phenotype given mutation in the gene on horizontal axis against the gene on the vertical axis. Color indicates the widest central posterior credible interval that does not include one. We deliberately not make any strong cutoff at 95% excluding zero or similar as the tail probabilities have high variance (e.g., where there is less data, one extra positive case could plausibly move this quantity from say 93% to 96%). Odds ratio are shown on the log scale.

This plot integrates the information shown in the plots above and some more.

LOF

What is the effect of certain loss of function? Since LOF does not vary per gene, let us look at the corresponding odds ratio per phenotype.



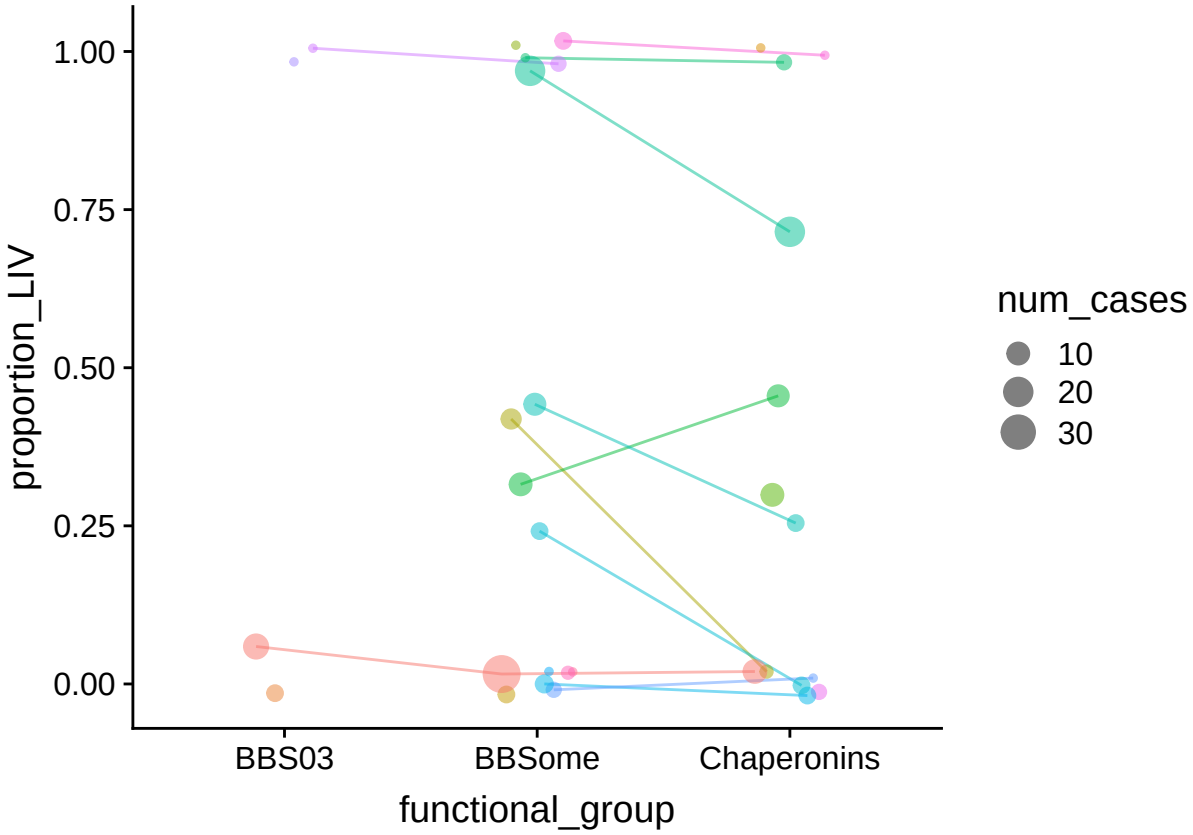
For most phenotypes we (expectedly) see that in cLOF mutations the phenotype is more likely. For LIV, RD and REP, the data is not very conclusive. The surprising part is the high posterior probability assigned to cLOF mutations having less severe phenotype in HEART. This might nevertheless be due to biases in reporting or due to higher lethality of HEART phenotype in cLOF mutations (and thus making the patients not be included in the dataset). It is however likely not a result of low sample size: there are 224 patients with the HEART phenotype reported, spread roughly equally between cLOF and other mutations.

Extra inspection of the LIV phenotype

For the LIV phenotype the frequentist analysis disagrees with the Bayesian analysis presented here. In particular, the frequentist analysis shows BBS3 as least likely to result in LIV and chaperonins as most likely, while the analysis shown here reports exactly opposite trend. First, let's look at aggregate frequencies of LIV phenotype:

```
## # A tibble: 4 x 2
##   functional_group proportion_LIV
##   <fct>                     <dbl>
## 1 BBS03                     0.158
## 2 BBSome                    0.351
## 3 Chaperonins               0.370
## 4 Others                    0.207
```

Indeed this supports the frequentist conclusions (as this is actually what those are based on). However, this aggregate look ignores between-study variability. So let's look what individual studies show:

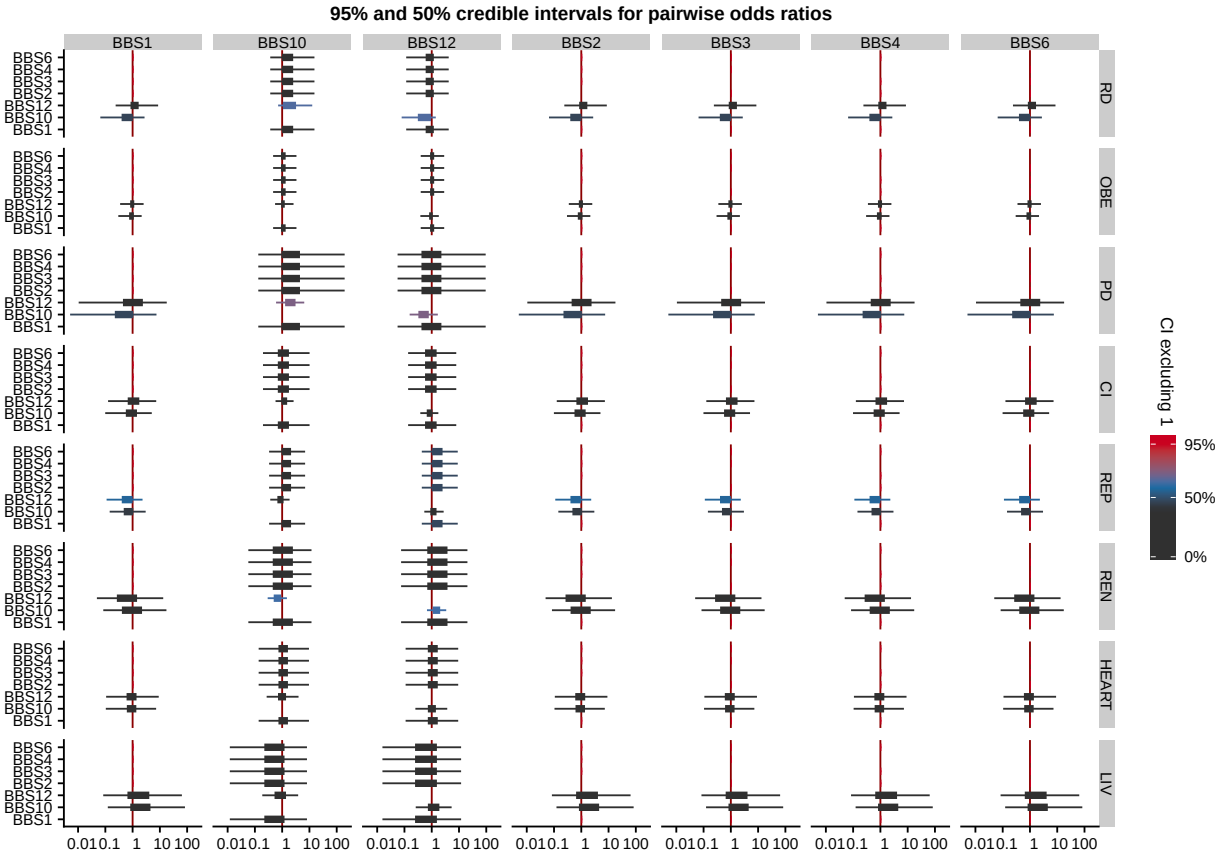


In the plot above, each dot is the proportion of patients with LIV given a mutation in a gene from one of the functional groups in a single study. Lines connect values that are from the same study (note that only one study had mutations from all groups and many had mutations in just one group) - those also have the same color. Size of the points represents the number of patients. We added some jitter to let us differentiate the points - notably all the points near zero and one are actually exactly zero and one.

This plot tells a different story: most individual studies, especially those with larger number of cases show increase in LIV phenotype for BBSome against Chaperonins. There is only one larger study including both BBS3 and BBSome and it has more LIV positive cases for BBS3. The overall opposite trend is driven by a) the only larger studies reporting BBS3 having unusually low proportion of LIV in general and b) studies reporting mutations only for BBS3 or only for BBSome having unusually low proportion of LIV.

It however seems that the evidence in this direction is not very compelling.

For a complete picture, let's look at pairwise differences between the 3 most common BBSome mutations (BBS1, BBS2 and BBS4), BBS3 and the chaperonins (BBS6, BBS10, BBS12).



We see that for the LIV phenotype there is large uncertainty and all 95% intervals include one. We also see that the large differences we have seen in the summary plots for CI and REN do translate into a quite confident estimates of notable pairwise differences between most BBSome genes, BBS3, and chaperonins.

Part 2: Alternative Models & Model Selection

Note: For historical reasons the feature of "certain loss of function" (cLOF) as discussed in the data is called just "lof" in most analysis code. This part will thus use "lof" and "cLOF" interchangeably.

Model descriptions

All models are Bayesian varying intercept logistic regressions using the `brms` package. Appendix 2 (main analysis) includes both accessible and complete mathematical description of the model we chose for the main analysis, which will not be repeated here. Generally, all of the terms in the models are varying intercepts, i.e. the model partially pool the estimates for individual groups (genes, sources, ...) towards population mean to achieve more robust inference.

Base models

Base models are those that work with (a subset of) the original dataset, without any imputation. The models are defined in file `models.R`. They differ in the model formula, subsets of the dataset they use and priors for model coefficients. The syntax for formulas in `brms` is described (in brms manual)[https://rdrr.io/cran/brms/man/brmsformula.html] and will not be explained here. The model may work on filtered dataset - `lof` means filtered for only the mutations with certain LOF, `age`,`sex` and `age_sex` corresponds to filtering for patients with reported age, sex or both. The list of base models follows:

```
## gene_only :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene)
##  data filter:  none
##
## gene_only_filtered_lof :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene)
##  data filter:  lof
##
## gene_lof :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) +
##  phenotype:loss_of_function_certain
##  data filter:  none
##
## gene_lof_per_gene :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype:loss_of_function_certain) || gene) +
##  phenotype:loss_of_function_certain
##  data filter:  none
##
## gene_source :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source)
##  data filter:  none
##
## gene_source_lof :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source) + phenotype:loss_of_function_certain
##  data filter:  none
##
## gene_source_lof_wide :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source) + phenotype:loss_of_function_certain
##  data filter:  none  Note: special priors
```

```
##
## gene_source_filtered_lof :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source)
##  data filter:  lof
##
## gene_source_filtered_lof_wide :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source)
##  data filter:  lof   Note: special priors
##
## gene_source_lof_per_gene :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source) + ((0 + phenotype:loss_of_function_certain) || gene) +
##  phenotype:loss_of_function_certain
##  data filter:  none
##
## gene_source_genecor :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + gene) | phenotype) + ((0 +
##  phenotype) || source)
##  data filter:  none
##
## gene_source_nocor :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) || gene) + ((0 +
##  phenotype) || source)
##  data filter:  none
##
## gene_source_narrow :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source)
##  data filter:  none  Note: special priors
##
## gene_source_very_narrow :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source)
##  data filter:  none  Note: special priors
##
## gene_source_wide :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source)
##  data filter:  none  Note: special priors
##
## gene_source_filtered_sex :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source) + (1 || Sex:phenotype)
##  data filter:  sex
##
## gene_source_filtered_age :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source) + (0 + age_std_for_model || phenotype)
##  data filter:  age
##
## gene_filtered_age_sex :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + (1 ||
```

```
##  Sex:phenotype) + (0 + age_std_for_model || phenotype)
##  data filter:  age_sex
##
## gene_lof_filtered_age_sex :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + (1 ||
##  Sex:phenotype) + (0 + age_std_for_model || phenotype) +
##  phenotype:loss_of_function_certain
##  data filter:  age_sex
##
## gene_lof_per_gene_filtered_age_sex :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + (1 ||
##  Sex:phenotype) + (0 + age_std_for_model || phenotype) + ((0 +
##  phenotype:loss_of_function_certain) || gene) +
##  phenotype:loss_of_function_certain
##  data filter:  age_sex
```

Imputation with mice

Including age or sex in the base models is problematic as it involves tossing out 46% or 44% of data, respectively. This results in wide posterior intervals and weak inferences. To try to ameliorate this we also tested running models on datasets with age and sex imputed, using multiple imputation via the `mice` package. We assume that both age and sex can be related to the functional group of the mutation and to each other. Involving further relations (e.g., individual genes) led to warnings from the `mice` package and we thus didn't use those.

The imputed models differ in the formulas used, but all use the default priors and do not filter the dataset in any way.

```
## gene_imputed_age_sex :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + (1 ||
##  Sex:phenotype) + (0 + age_std_for_model || phenotype)
##
## gene_source_imputed_sex :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source) + (1 || Sex:phenotype)
##
## gene_source_imputed_age :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source) + (0 + age_std_for_model || phenotype)
##
## gene_source_imputed_age_sex :
##  formula: phenotype_value ~ (1 || phenotype) + ((0 + phenotype) | gene) + ((0 +
##  phenotype) || source) + (0 + age_std_for_model || phenotype)
```

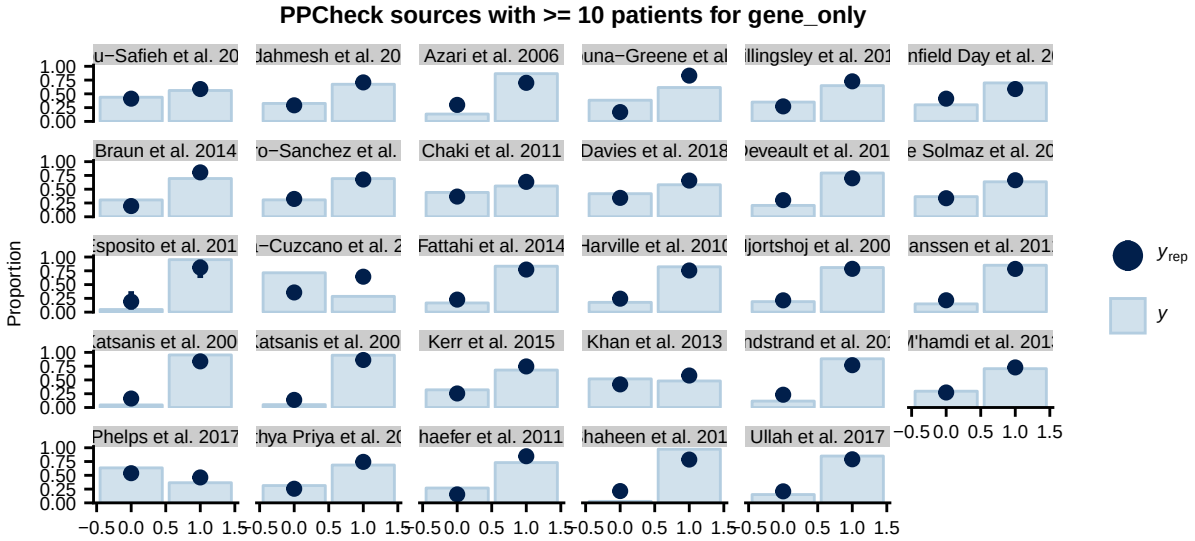
Choosing the model for main analysis

In general we use *posterior predictive checks* (PPCheck) to asses model fit. PPCheck is performed by predicting posterior distribution of possible outcomes implied by the fitted model. This distribution can then be compared to what is actually observed in the data and discrepancies can be noted and used to guide model expansion/selection. In our case, we focus on the prevalance of positive phenotypes across various subdivisions of the data. Of prime interests are subdivisions *not* taken into account by a model - if the model explains groupings that were not included, it is a sign that it works well. If the model consistently misestimates groups it is not aware of, it is an indication that such a group should be involved. We use the `bayesplot` package to perform PPChecks. See (Gabry et al. 2018, Visualisation in Bayesian workflow)[https://arxiv.org/abs/1709.01449] for a more thorough discussion of PPChecks.

In most models, we assume phenotype correlations because the data were selected for containing at least two phenotypes and diagnosis criteria is based on having multiple phenotypes. Both of these processes could have introduced correlations. However, the fitted correlations are not conclusive and do not increase the explanatory power of the model (models with different correlation structures are also included).

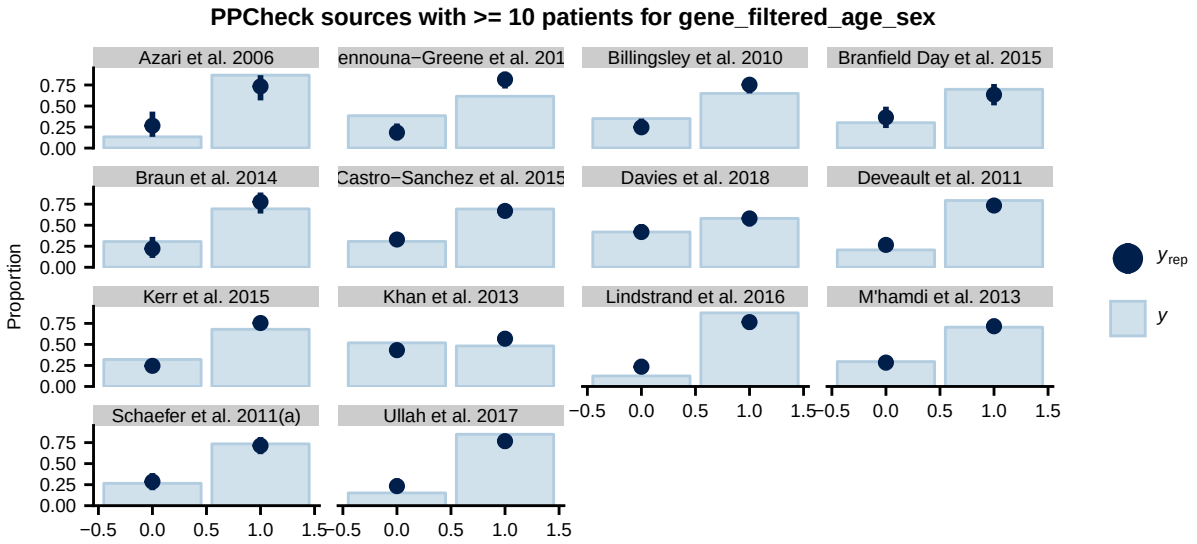
Between-study variability needs to be included

Modeling between-study variability is simply a good practice for any meta-analysis, but PPChecks can convince us that models ignoring it do not fit the data well. Let's start by looking at overall prevalance for studies with at least 10 patients and how the most basic model (`gene_only`, taking only the gene into account) fares:

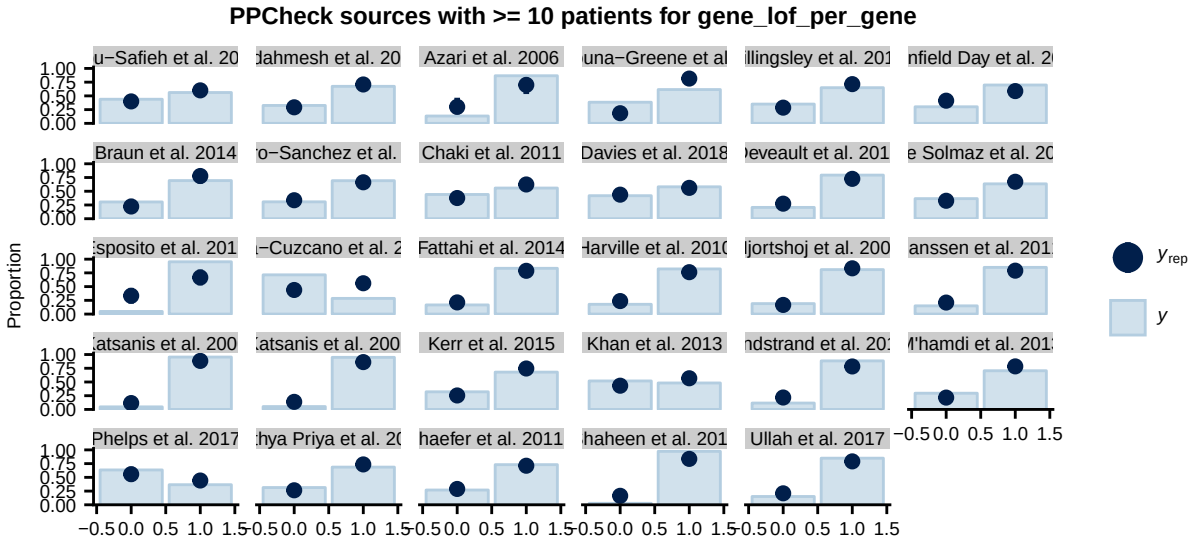


Here, the bars represent actual prevalence in the data, the dots the posterior mean for those subgroups and the lines posterior 95% credible interval. Note that we clump together all phenotypes for simplicity. We see that the model is overly certain in its predictions and the posterior credible intervals frequently miss the actual counts in a study. 11 / 25 sources have predictions not very consistent with data.

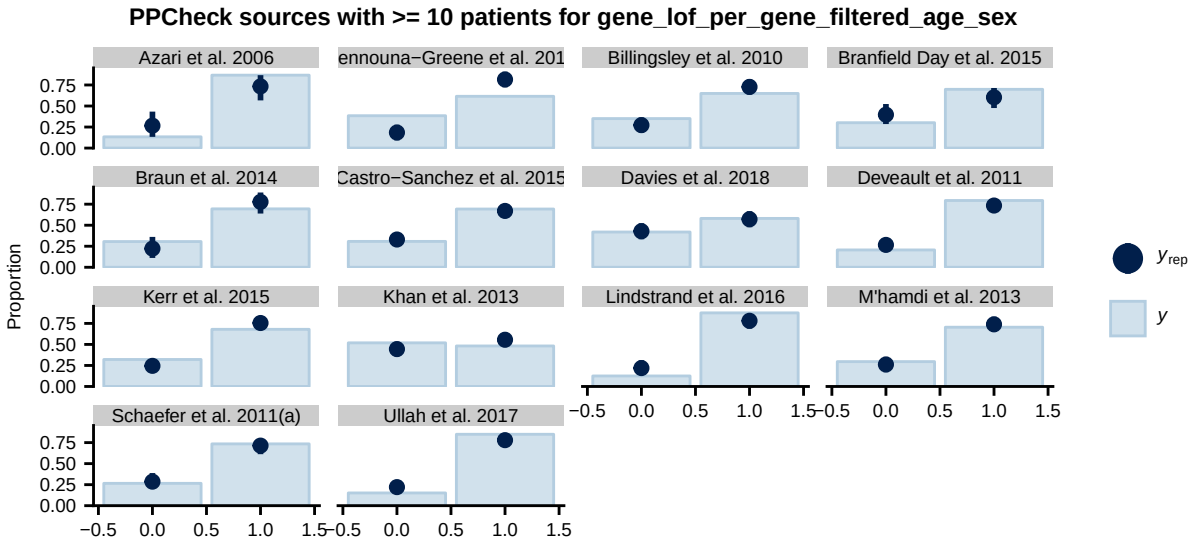
This does not resolve when including cLOF and/or age and sex (here using the model filtered for age, as we don't really care about wide posterior intervals)



6/14 source are noticeably off in gene_filtered_age_sex

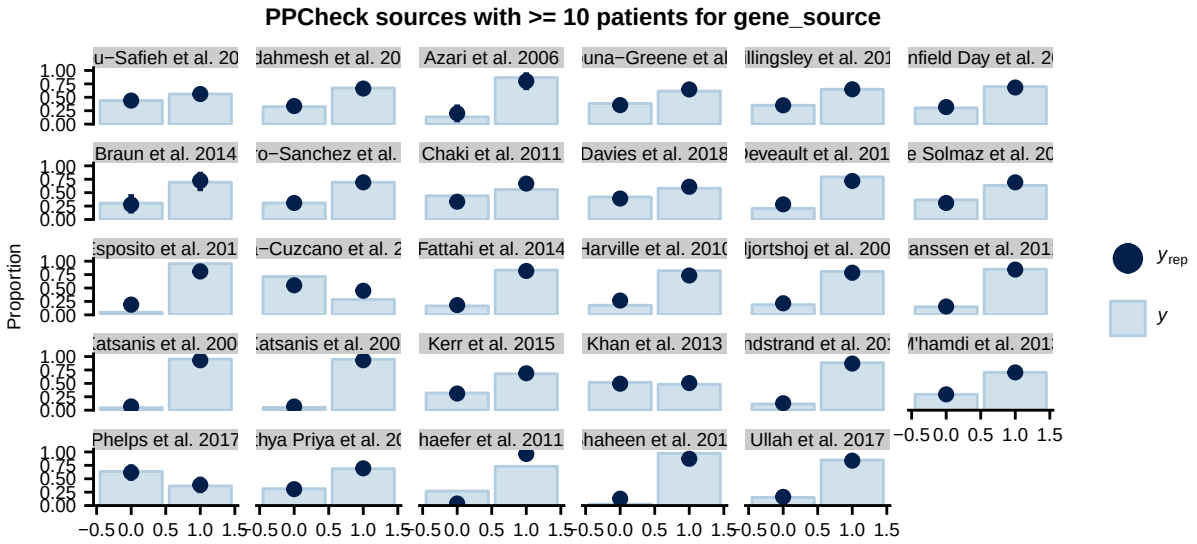


10/25 sources are noticeably off gene_lof_per_gene

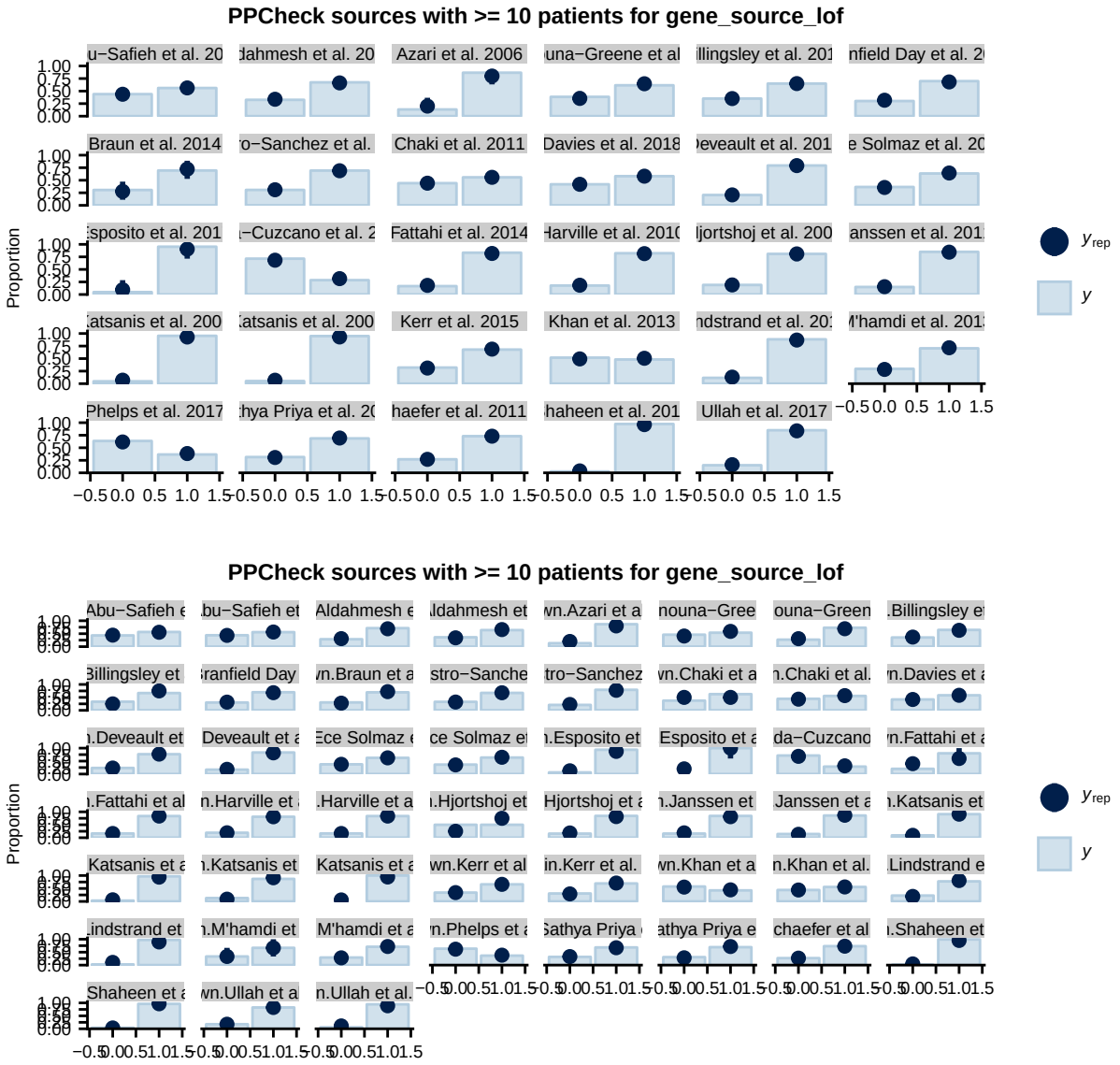


5/14 sources are noticeably off for gene_lof_per_gene_filtered_age_sex

Including source explicitly alleviates a large fraction (but not all) of the problems:



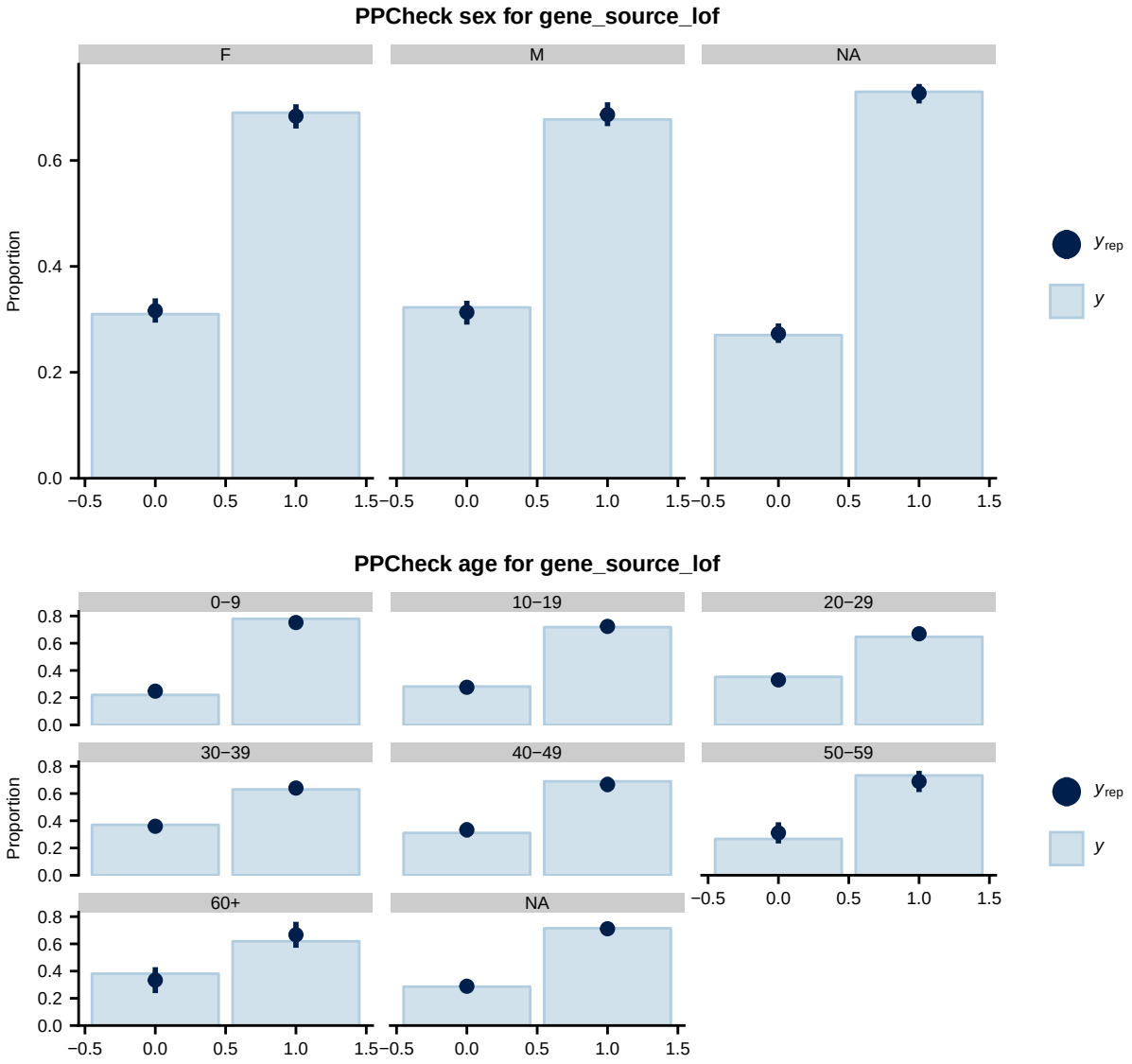
We see 6/25 sources problematic in gene_source. However including cLOF and source makes the fit good across all sources:



It therefore seems that both source and cLOF are important factors.

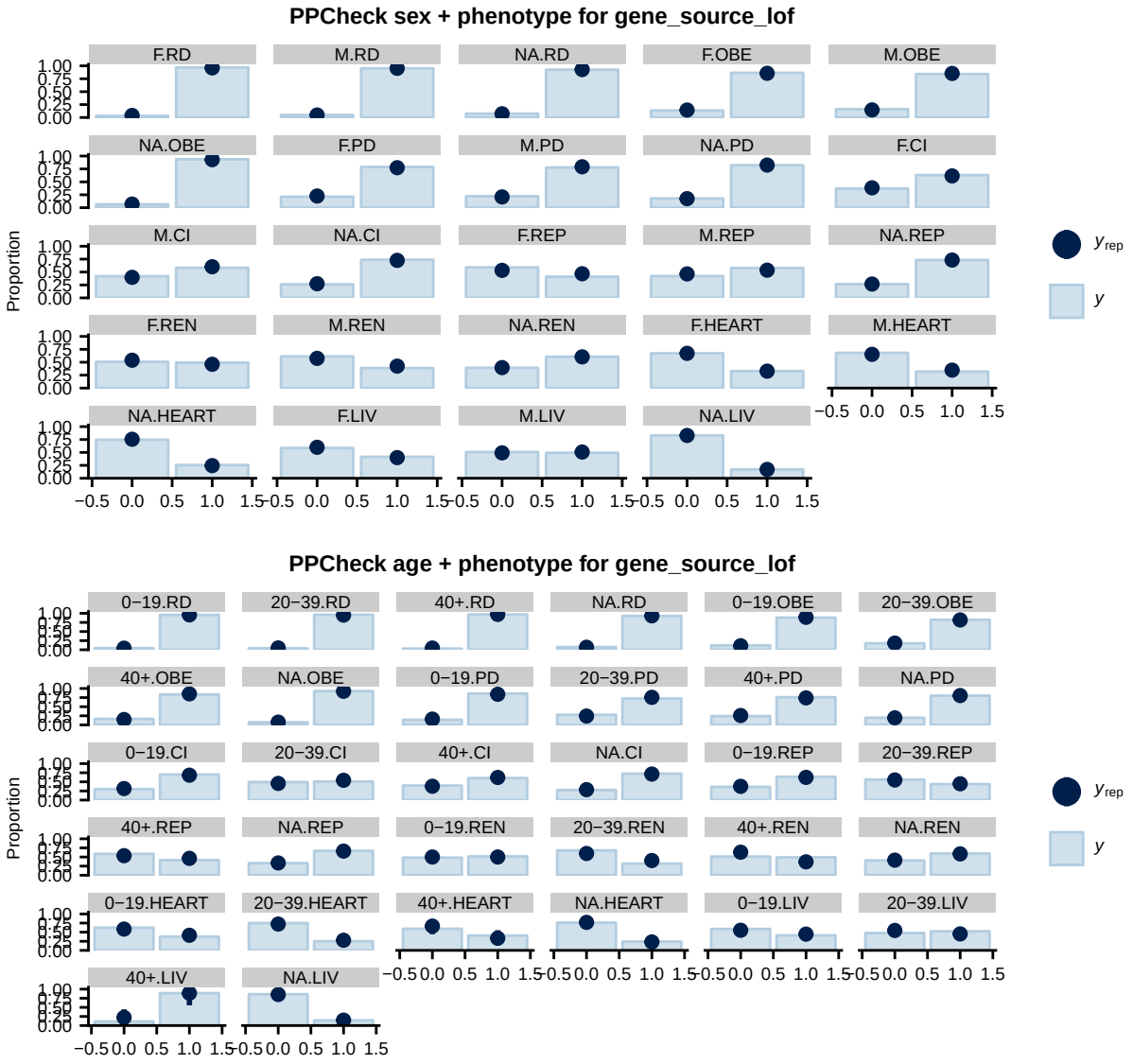
Between-study variability mostly explains age and sex differences

Age and sex differences don't necessarily need to be included, as they are sufficiently well explained by the `gene_source_lof` model. First let's look at overall prevalance by sex and age group:



Note that missing data are treated as a separate age category.

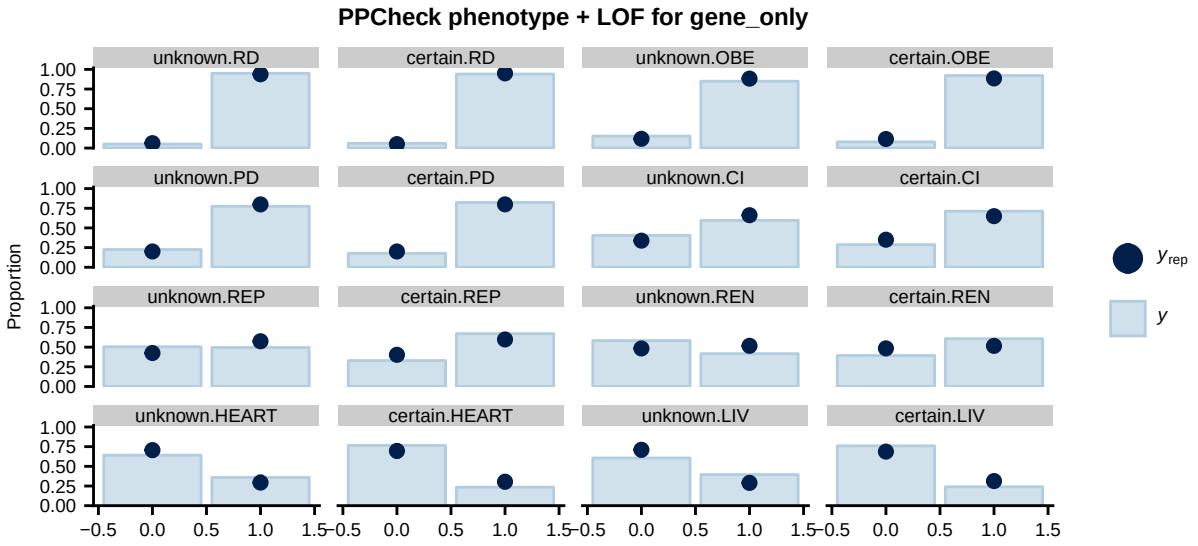
No big problems in the bulk age/sex groups. We can also look at age and sex by individual phenotypes (we collapse some of the age groups to make the age + phenotype plot readable:



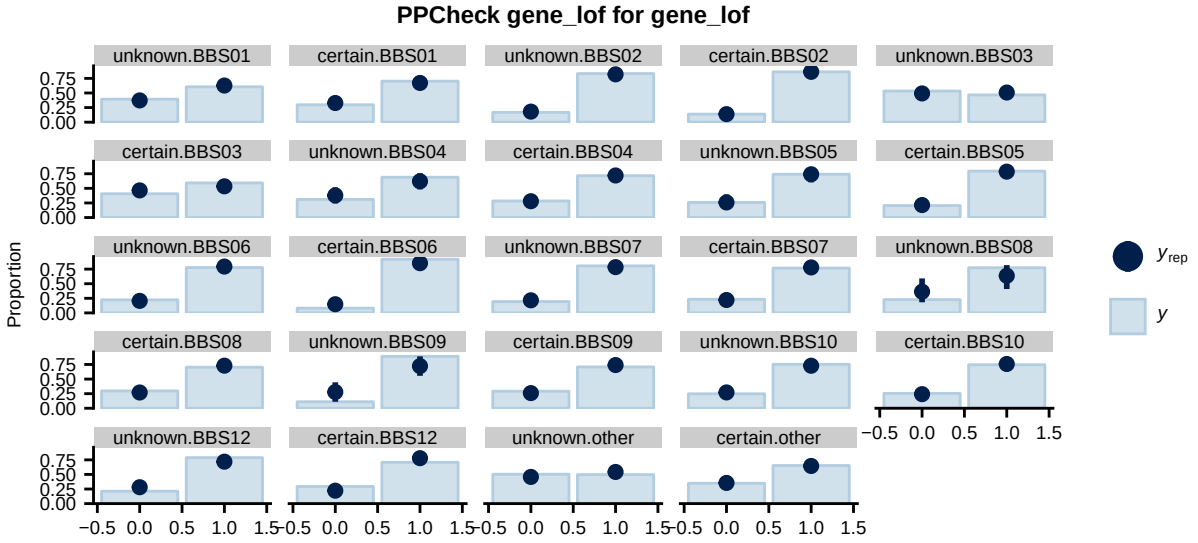
Once again only very minor problems (e.g., age "20-39" for REN phenotype). We think those can be safely ignored.

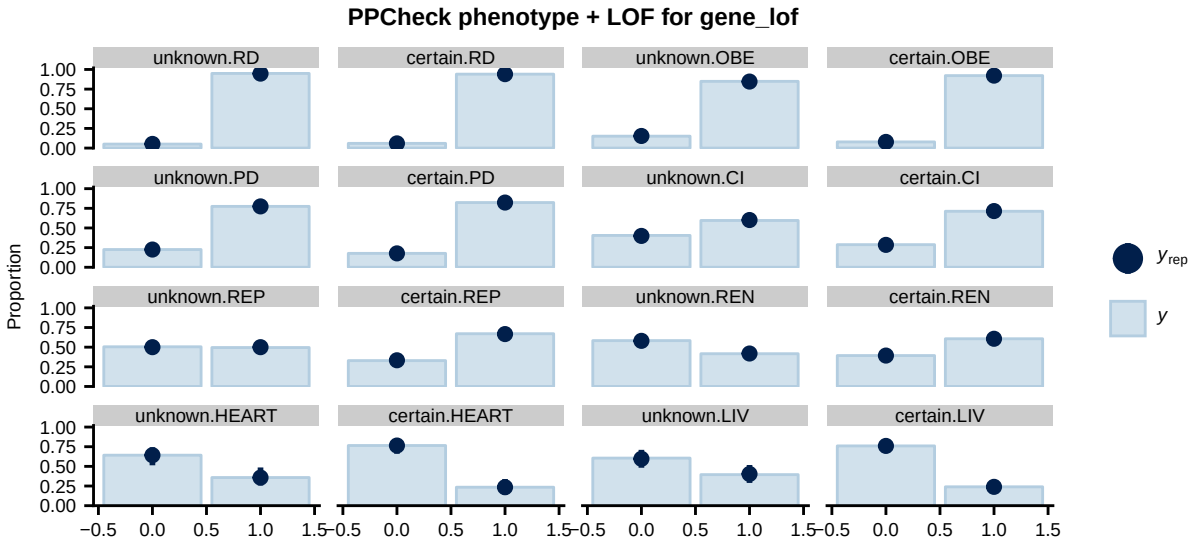
Loss-of-function differences are of a relatively minor importance

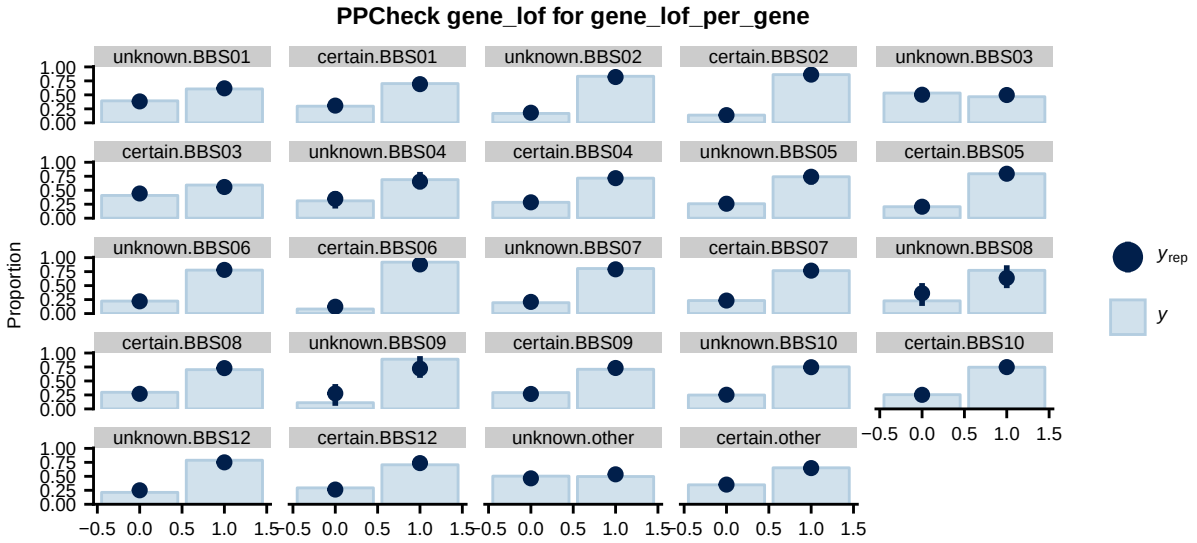
Using the most simple model, we see that cLOF difference are not well modelled for the CI, REP, REN and LIV phenotypes.

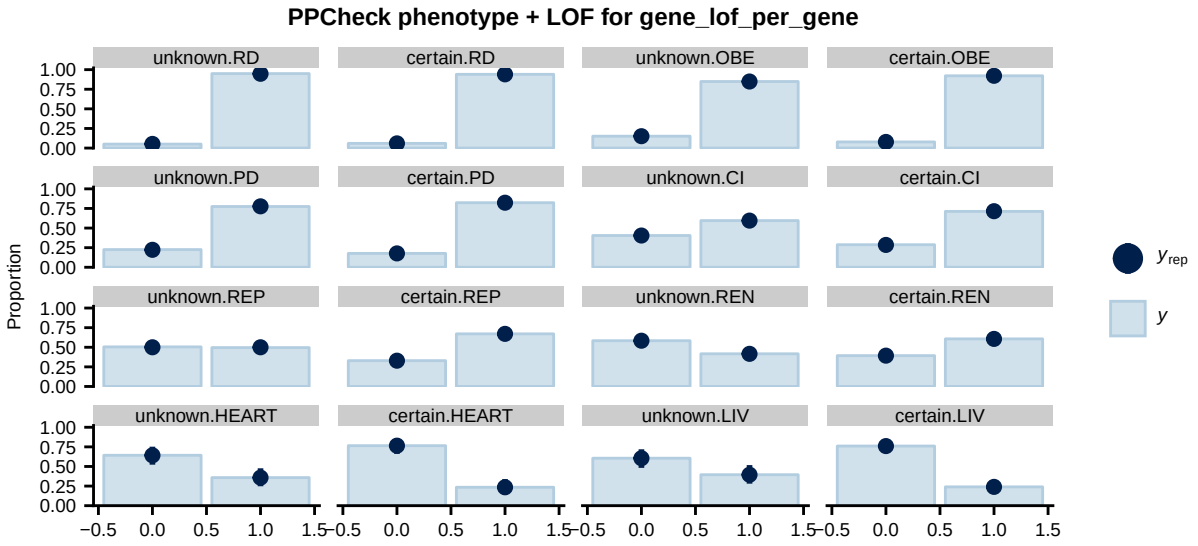


This disappears when adding a per-phenotype cLOF term to the model (i.e. for a given phenotype the effect of cLOF is assumed equal across all mutations):

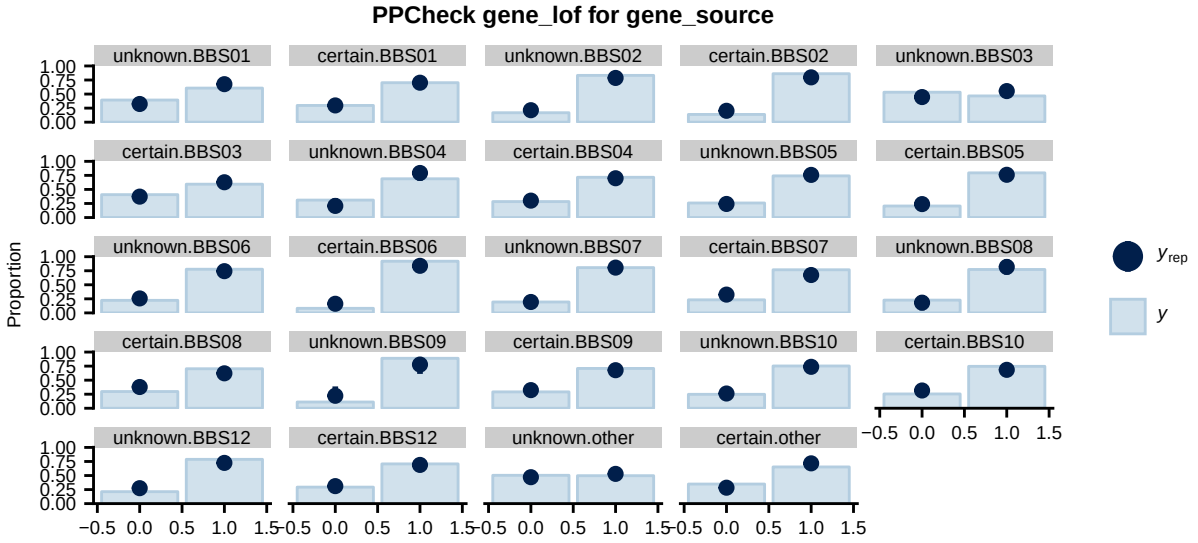


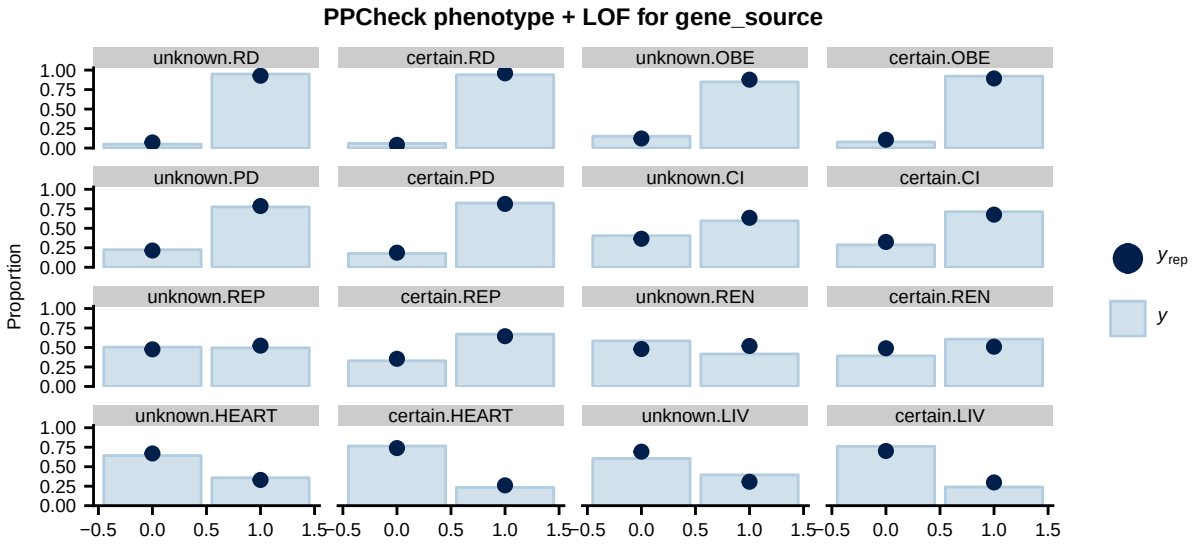

PPCheck phenotype + LOF for gene_lof
unknown.RD
certain.RD
unknown.OBE
certain.OBE
unknown.PD
certain.PD
unknown.CI
certain.CI
unknown.REP
certain.REP
unknown.REN
certain.REN
unknown.HEART
certain.HEART
unknown.LIV
certain.LIV
Proportion
1.00
0.75
0.50
0.25
0.00
−0.5
0.0
0.5
1.0
1.5
y_{rep}
y


PPCheck gene_lof for gene_lof_per_gene
unknown.BBS01
certain.BBS01
unknown.BBS02
certain.BBS02
unknown.BBS03
certain.BBS03
unknown.BBS04
certain.BBS04
unknown.BBS05
certain.BBS05
unknown.BBS06
certain.BBS06
unknown.BBS07
certain.BBS07
unknown.BBS08
certain.BBS08
unknown.BBS09
certain.BBS09
unknown.BBS10
certain.BBS10
unknown.BBS12
certain.BBS12
unknown.other
certain.other
Proportion
1.00
0.75
0.50
0.25
0.00
−0.5
0.0
0.5
1.0
1.5
y_{rep}
y

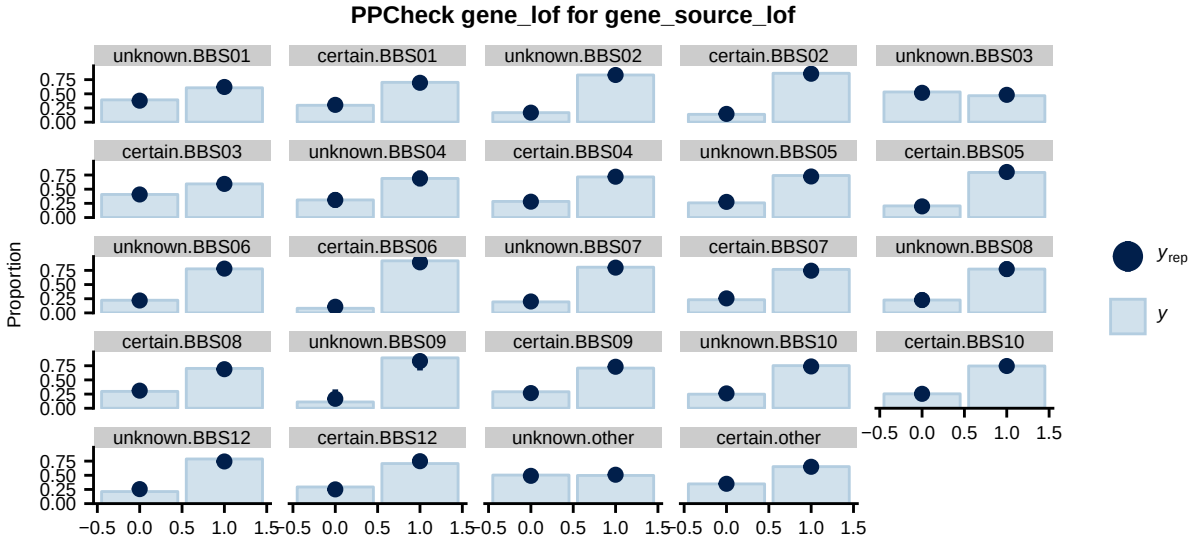


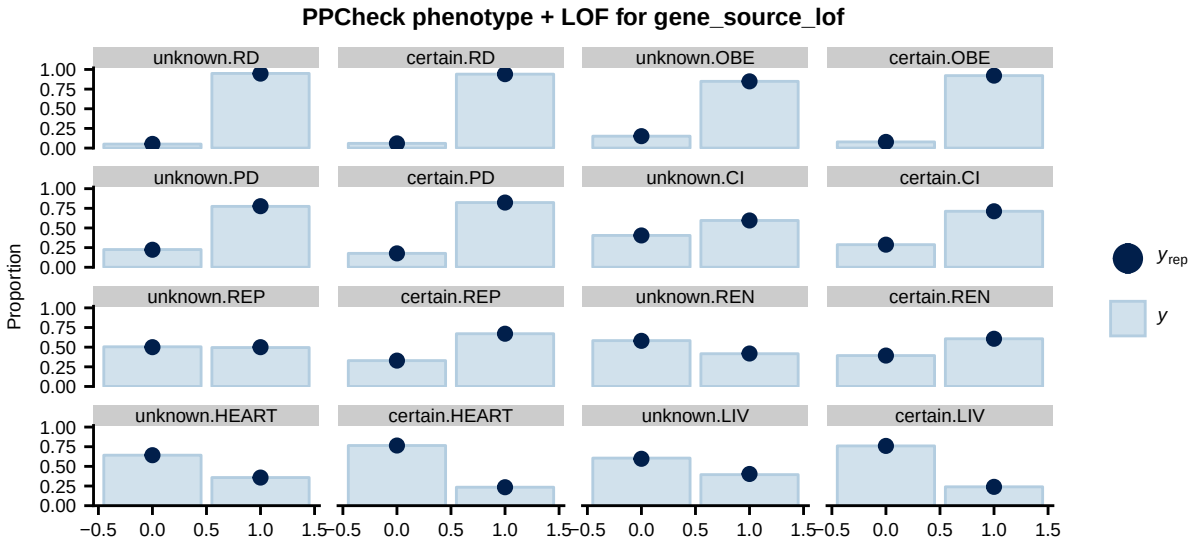
While adding source as a covariate ameliorates the problems with cLOF (it remains problematic only for REN and LIV phenotypes)



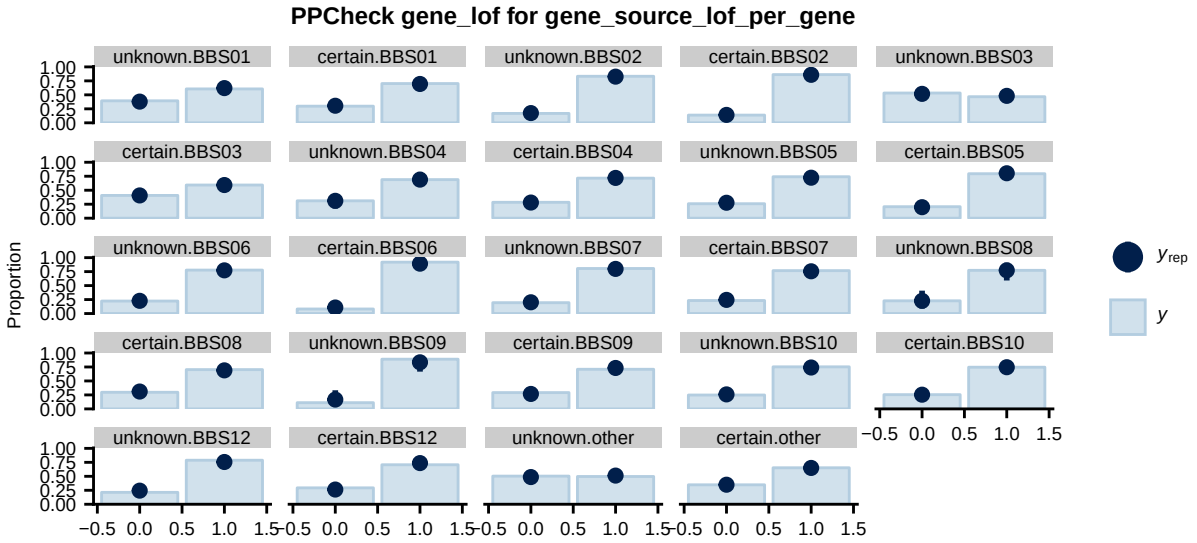


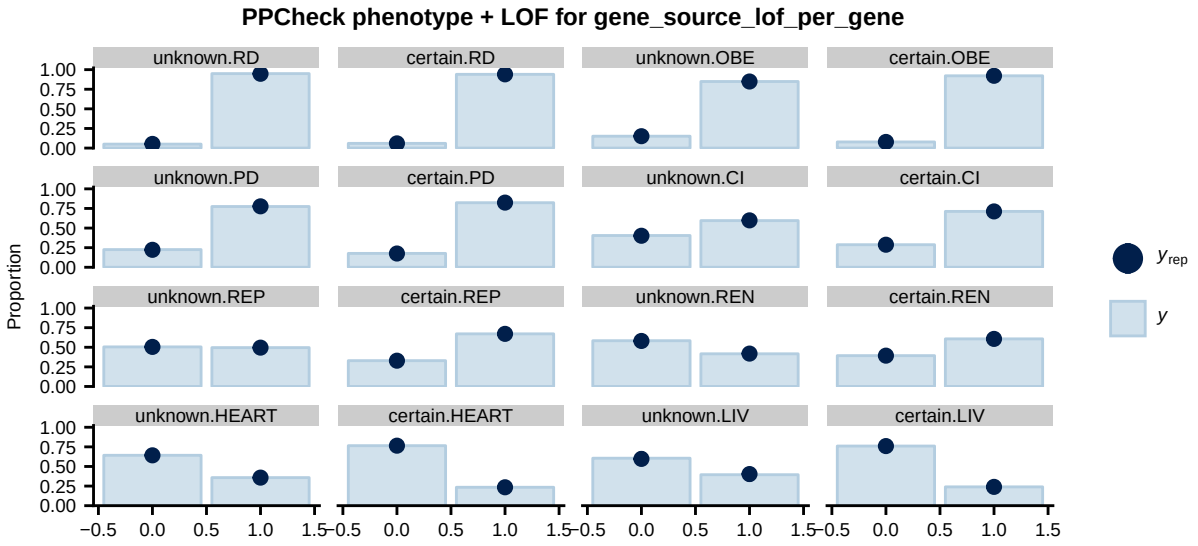
The problem is mitigated when both source and cLOF are included (even when cLOF effect is not allowed to vary with gene)





Having cLOF coefficient differ per gene does not bring noticeable improvements:





Model selection verdict

We have shown that source (between-study variability) has to be included and there is an advantage in including cLOF per phenotype, but not much improvement when including cLOF per phenotype and gene. Since the `gene_source_lof_per_gene` model is too flexible for the limited amount of data (results in very wide posterior intervals, spanning odds ratio up to 10000), we think `gene_source_lof` is a better choice.

Part 3: Multiverse Analysis

Note: For historical reasons the feature of "certain loss of function" (cLOF) as discussed in the data is called just "lof" in most analysis code. This part will thus use "lof" and "cLOF" interchangeably.

A wide variety of models was tested. Those are briefly described in the comparison. Their exact formulations can be found in Part 2.

Defining Precise Criteria

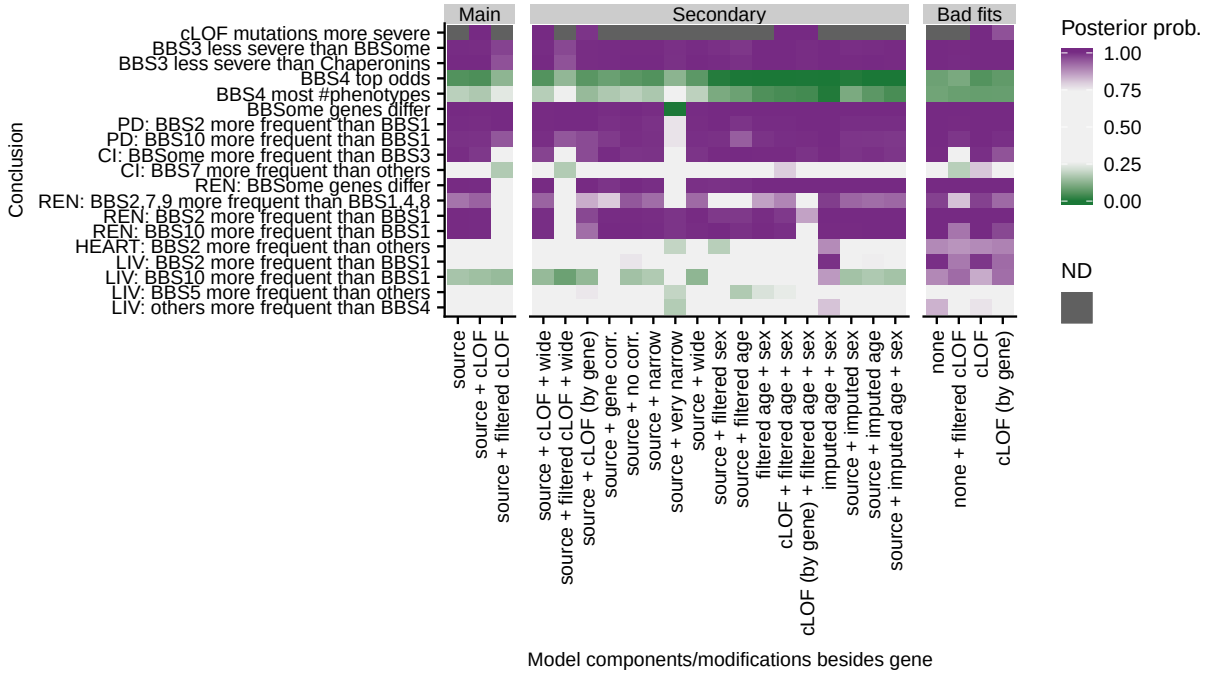
There is some flexibility in defining the exact model configurations consistent with a given conclusion. In general, when discussing whether there is or isn't a difference, the Bayesian model will always say "yes there is a (possibly small) difference" - the posterior probability of a difference of exactly 0 is 0. Instead we choose a threshold of a "clinically relevant effect", which for us is odds ratio outside of (0.5,2) and any effect outside of these bounds is counted as "different".

Similarly, other plain-English statements need to be transformed to an exact predicate that can be considered true or false for a given assignment of numerical values to model coefficients, to let us evaluate its posterior probability. The exact definition of the individual tested statements follows.

- The severity of BBS is worse in patients with LOF mutations than in patients with other mutations.
  - Ignored for models that do not include LOF
  - Measured as posterior probability, that the LOF effect is positive (OR > 1) for at least 5 BBSome genes in 3 phenotypes.
- The data suggest that mutations in BBS3 have lower severity than mutations in different functional groups of genes.
  - Measured as posterior probability, that odds ratio is < 1 for at least 3 phenotypes for at least 1/2 of pairwise gene comparisons.
- The data suggest a difference between the severity of BBS in patients with mutations in different BBSome subunits.
  - Measured as posterior probability, that there is clinically relevant effect for at least 3 phenotypes for at least 5 pairwise comparisons.
- BBS4 phenotype is the most severe of all BBSome-encoding genes.
  - Probability that odds for BBS4 are among the top 3 odds for at least 5 phenotypes
  - Probability that a patient with BBS4 has the highest total number of phenotypes present.
- Mutations in different BBSome subunits predispose to different renal phenotype.
  - Measured as posterior probability, that there is clinically relevant effect for at least 5 pairwise comparisons.
- Differences between small groups of genes - Probability that all pairwise odds ratios are greater/less than 1 (depending on the direction of the comparison)
  - Cognitive impairment is less frequent in BBS3 patients compared to other patients (all canonical BBS genes).
  - Cognitive impairment is more frequent in BBS7 patients compared to other patients with mutations in BBSome-encoding genes.
  - Renal involvement is less frequent in BBS1, BBS4 and BBS8 patients compared to BBS2, BBS7, BBS9 patients.
  - Patients with BBS2 mutations are more likely to have heart anomalies compared to patients with other mutations in BBSome-encoding genes.
  - Patients with BBS5 mutations are more likely to have liver anomalies compared to patients with other mutations in BBSome-encoding genes.
  - Patients with BBS4 mutations are less likely to have liver anomalies compared to patients with other BBSome mutations.
- Individual differences in phenotype between two genes: directly the probability that the OR > 1 for patients with cLOF mutation.
  - Polydactyly is more frequent in BBS2 patients compared to BBS1 patients.
  - Polydactyly is more frequent in BBS10 patients compared to BBS1 patients.

- Renal involvement is less frequent in BBS1 patients compared to BBS2 patients.
- Renal involvement is less frequent in BBS1 patients compared to BBS10 patients.
- Liver involvement is less frequent in BBS1 patients compared to BBS2 patients.
- Liver involvement is less frequent in BBS1 patients compared to BBS10 patients.

Bayesian Comparison

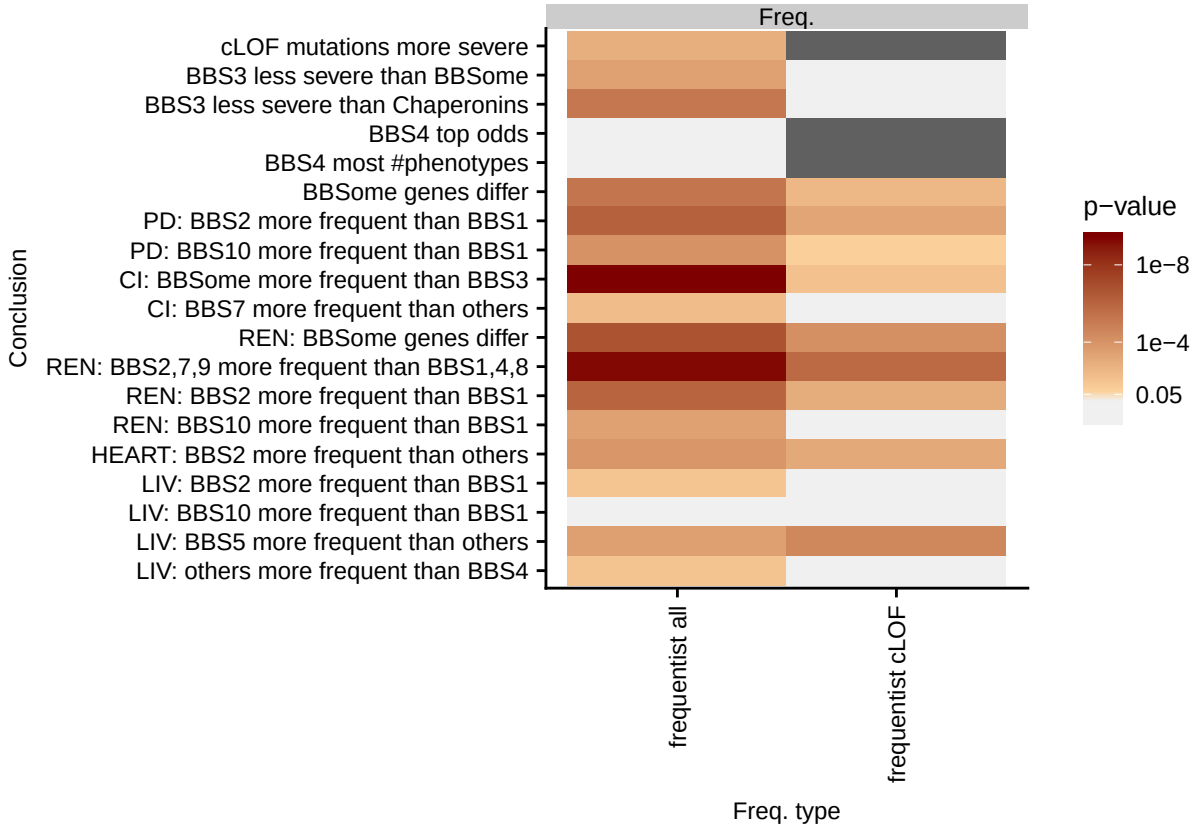


A heatmap of posterior probability of statements characterizing individual conclusions. All Bayesian models include gene as covariate but may also include additional covarietes: source, age, sex and certain loss of function (cLOF) - either as a global covariate or by gene. Since age and sex are not available for all data, we can either fit the model only to patients where those are reported (filtered) or impute missing data (imputed). Instead of using cLOF as a covariate, we can fit the model using only patients with cLOF mutations (filtered cLOF). For most models we include a correlation structure across phenotypes (e.g., that two phenotypes occur frequently together across all genes), but this structure may be absent (no corr.) or replaced with a correlation structure across genes (gene corr. - e.g., that two genes have similar pattern of effects across all phenotypes). We also tried modifying the width of prior distributions (wide, narrow, very narrow). See Appendix 4 for a detailed description of all models and the imputation procedure. Dark grey indicates that the question could not be evaluated for the given model (currently only asking for cLOF differences in models that exclude cLOF). The "Bad fits" category is reserved for models we now fit the data badly, as discussed in Appendix 4.

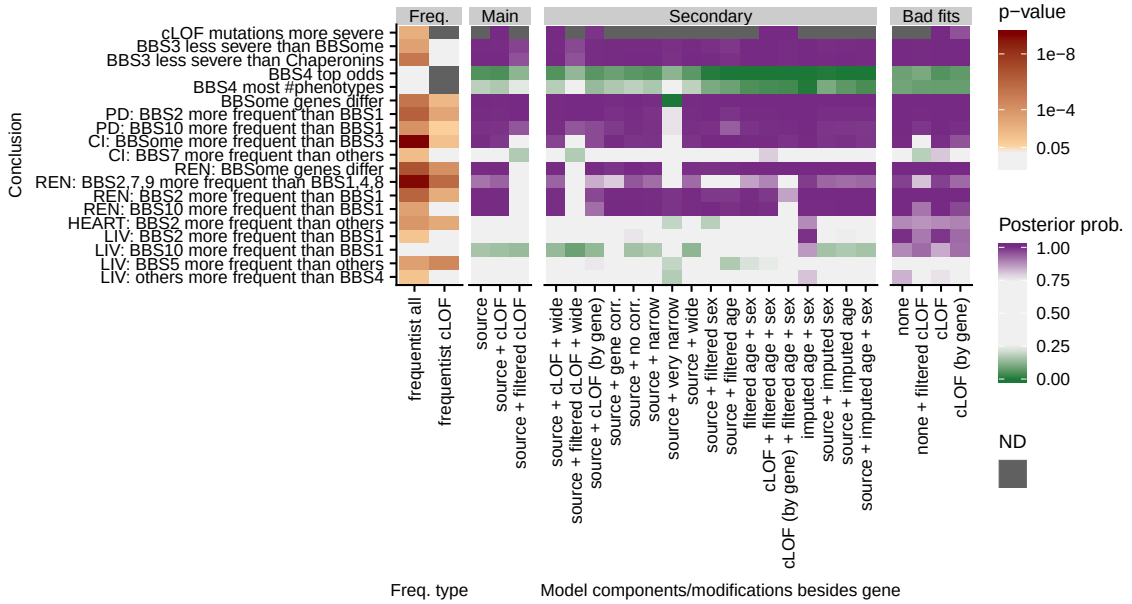
Some patterns to notice:

- When fitting filtered datasets, there is less certainty implying less strong evidence in both directions
- Using very narrow priors on gene coefficients ($N(0, 0.1)$, i.e., that almost all odds ratios should be less than ~ 1.2) unsurprisingly results in little evidence for directional differences between genes.
- Other than noted above, the conclusions are not sensitive to model choice.

Combining with frequentist results



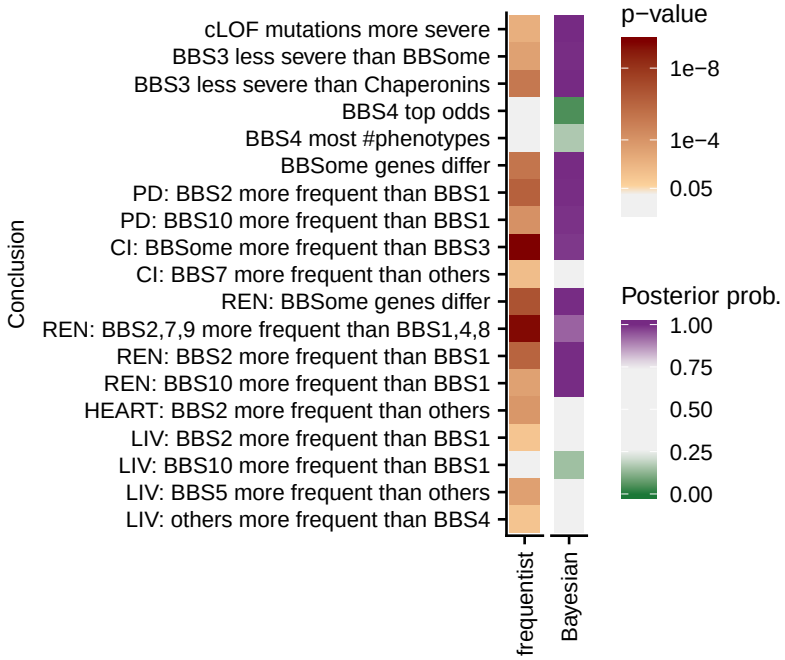
Heatmap of p-values from frequentist analysis for the same conclusions (exact computations described in the main body of the paper).



This is a combined plot for both frequentist and Bayesian analzsis, once again shown mostly consistent results even when frequentist analyses are taken into account. The only big difference is that some of the HEART

and LIV conclusions are not supported by most Bayesian analyses, or only those ignoring between-study variability.

Finally a brief summary of the main analyses as shown in the main paper.



Original computing environment

This report was built from Git revision `fd5c5722f7a7161e0610a44aa0d20138f9511bea` on 21 January, 2019

```
## R version 3.5.1 (2018-07-02)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 17134)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] mice_3.3.0         lattice_0.20-35    bindrcpp_0.2.2
##  [4] svglite_1.2.1      knitr_1.20         cowplot_0.9.3
##  [7] bayesplot_1.6.0    tidybayes_1.0.1    forcats_0.3.0
## [10] stringr_1.3.1      dplyr_0.7.6        purrr_0.2.5
## [13] readr_1.1.1        tidyr_0.8.1        tibble_1.4.2
## [16] tidyverse_1.2.1    here_0.1           readxl_1.1.0
## [19] skimr_1.0.3        brms_2.5.2         Rcpp_0.12.19
## [22] rstan_2.18.1       StanHeaders_2.18.0 ggplot2_3.1.0
##
## loaded via a namespace (and not attached):
##   [1] minqa_1.2.4              colorspace_1.3-2
##   [3] ggridges_0.5.1           rsconnect_0.8.8
##   [5] rprojroot_1.3-2          ggstance_0.3.1
##   [7] markdown_0.8             base64enc_0.1-3
##   [9] rstudioapi_0.8           svUnit_0.7-12
##  [11] DT_0.4                   fansi_0.4.0
##  [13] mvtnorm_1.0-8            lubridate_1.7.4
##  [15] xml2_1.2.0               codetools_0.2-15
##  [17] splines_3.5.1            bridgesampling_0.5-2
##  [19] shinythemes_1.1.1        jsonlite_1.5
##  [21] nloptr_1.2.0             LaplacesDemon_16.1.1
##  [23] broom_0.5.0              shiny_1.1.0
##  [25] compiler_3.5.1           httr_1.3.1
##  [27] backports_1.1.2          assertthat_0.2.0
##  [29] Matrix_1.2-14            lazyeval_0.2.1
##  [31] cli_1.0.1                later_0.7.5
##  [33] htmltools_0.3.6          tools_3.5.1
##  [35] igraph_1.2.2             coda_0.19-1
##  [37] gtable_0.2.0             glue_1.3.0
##  [39] reshape2_1.4.3           cellranger_1.1.0
##  [41] nlme_3.1-137             crosstalk_1.0.0
##  [43] ps_1.1.0                 lme4_1.1-18-1
##  [45] rvest_0.3.2              mime_0.5
##  [47] miniUI_0.1.1.1           gtools_3.8.1
```

```
##  [49] pan_1.6                    MASS_7.3-50
##  [51] zoo_1.8-4                  scales_1.0.0
##  [53] colourpicker_1.0           hms_0.4.2
##  [55] promises_1.0.1             Brobdingnag_1.2-6
##  [57] parallel_3.5.1             inline_0.3.15
##  [59] RColorBrewer_1.1-2         shinystan_2.5.0
##  [61] yaml_2.2.0                 gridExtra_2.3
##  [63] gdtools_0.1.7              loo_2.0.0
##  [65] rpart_4.1-13               stringi_1.2.4
##  [67] highr_0.7                  dygraphs_1.1.1.6
##  [69] pkgbuild_1.0.1             rlang_0.2.2
##  [71] pkgconfig_2.0.2            matrixStats_0.54.0
##  [73] evaluate_0.11              bindr_0.1.1
##  [75] labeling_0.3               rstantools_1.5.1
##  [77] htmlwidgets_1.3            processx_3.2.0
##  [79] tidyselect_0.2.4           plyr_1.8.4
##  [81] magrittr_1.5               R6_2.3.0
##  [83] mitml_0.3-6                pillar_1.3.0
##  [85] haven_1.1.2                withr_2.1.2
##  [87] xts_0.11-1                 nnet_7.3-12
##  [89] survival_2.42-3            abind_1.4-5
##  [91] modelr_0.1.2               crayon_1.3.4
##  [93] arrayhelpers_1.0-20160527 jomo_2.6-4
##  [95] utf8_1.1.4                 rmarkdown_1.10
##  [97] grid_3.5.1                 callr_3.0.0
##  [99] threejs_0.3.1              digest_0.6.17
## [101] xtable_1.8-3               httpuv_1.4.5
## [103] stats4_3.5.1               munsell_0.5.0
## [105] shinyjs_1.0
```