**Assignment 2 - meta-analysis**

Questions to be answered

**1. Simulate data to setup the analysis and gain insight on the structure of the problem.**

* **Simulate one dataset of 100 studies**
  + **(n of participants should follow a normal distribution with mean of 20, sd of 10, but no fewer than 10 participants),**
  + **with a mean effect size of 0.4,**
  + **average deviation by study of .4**
  + **and measurement error of .8.**

**The data you get should have:**

* **one row per study,**
  + **with an effect size, mean and standard error.**

**Build a proper bayesian model to analyze the simulated data.**

* **Then simulate publication bias**
  + **(only some of the studies you simulate are likely to be published, which?),**
* **the effect of publication bias on your estimates**
  + **(re-run the model on published studies, assess the difference),**
* **and discuss what this implies for your model.**

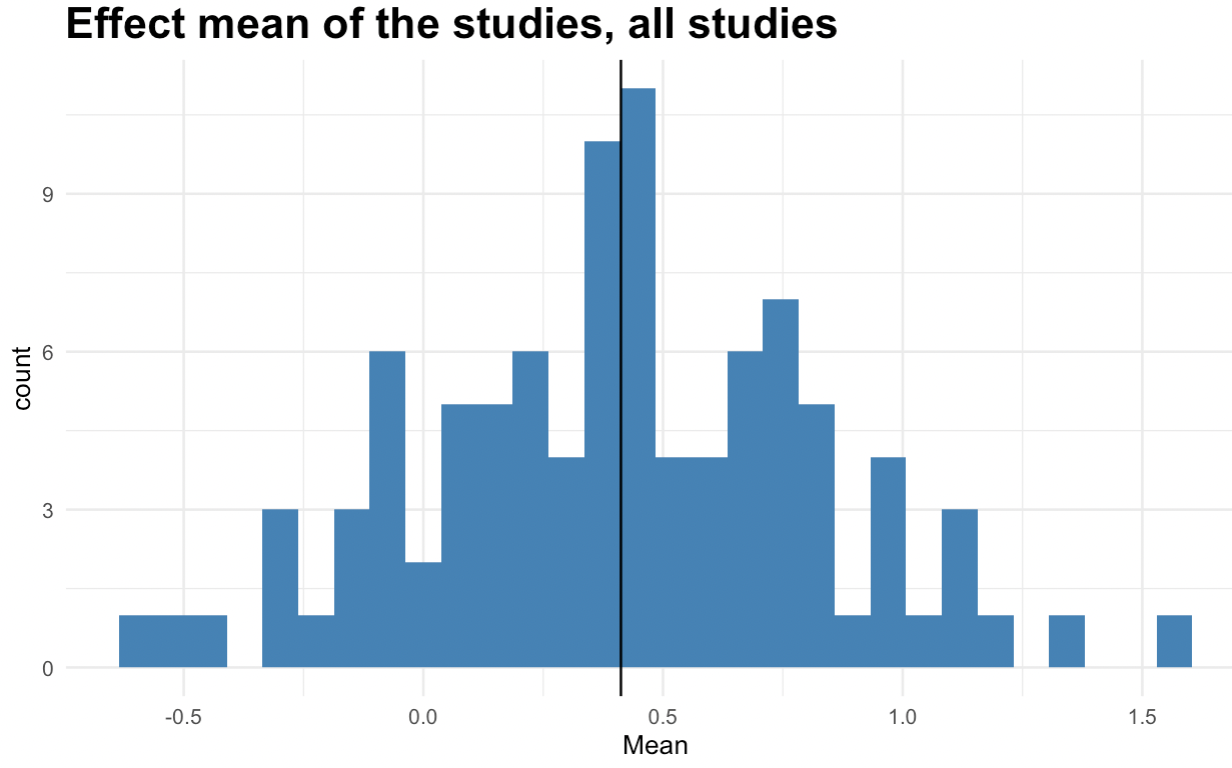
**Part 1**

We simulated a dataset (d) containing 100 studies that included our given parameters with a mean effect size of 0.4, average deviation by study of .4, measurement error of .8 and with the number of participants per study being distributed normally with mean of 20, sd of 10 and a minimum number of 10 participants. Using a for loop we created a Dataframe with 100 rows (one for each study) and 6 columns: effect size, mean and standard error, study, participants and PublishedPos. With the intention of making the simulations more realistic and with a greater variability, we p-hacked the simulations by adding 3 outliers.

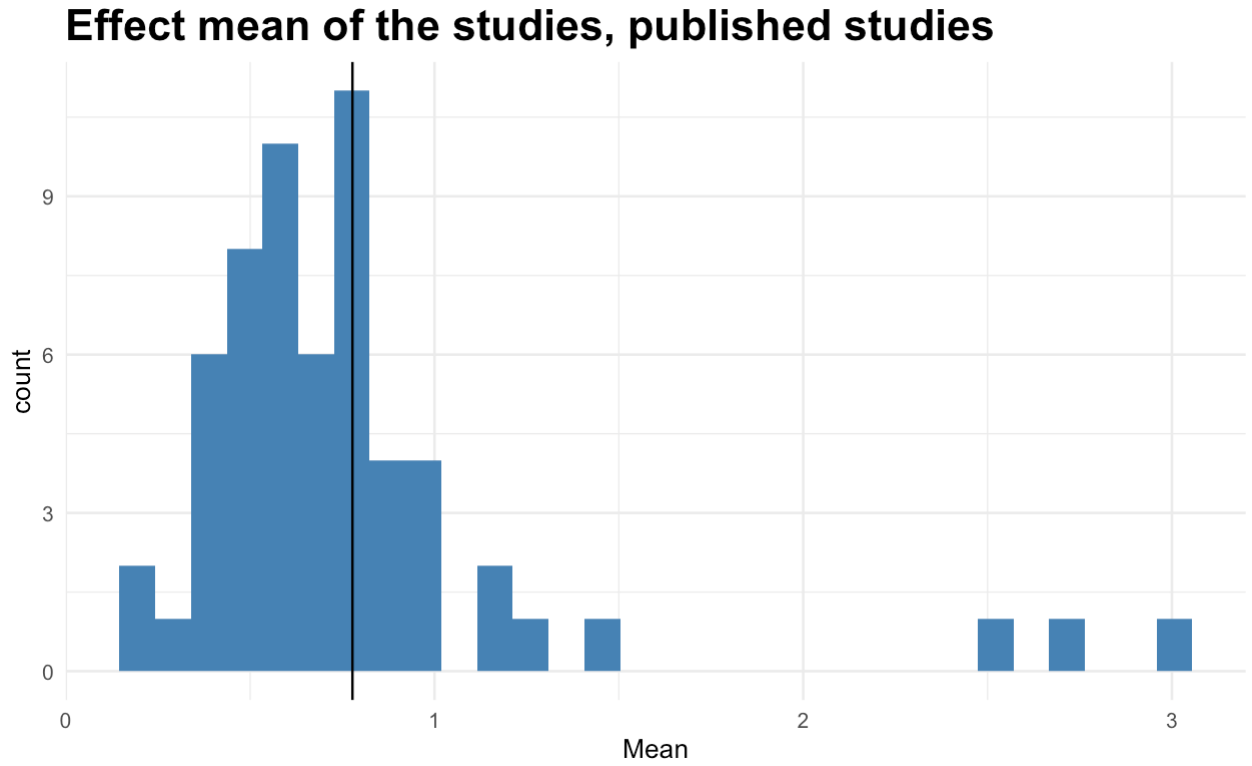
The column PublishedPos simulates a potential publication bias as it contains information on the simulated publication of the study. We know from the literature that a publication is easier to achieve with significant and often positive results, this column marks 90% of the studies as published if they show positive significant effect sizes and only 10% percent if they do not add up to these criterions.

Based on the PublishedPos column, we created a new dataset of the simulated data containing solely the “published” studies to be able to asses the effect of publication bias on our estimates.

The following plot shows the mean effect size for all the 100 studies, with the vertical line annotating the mean of the means at 0.41:

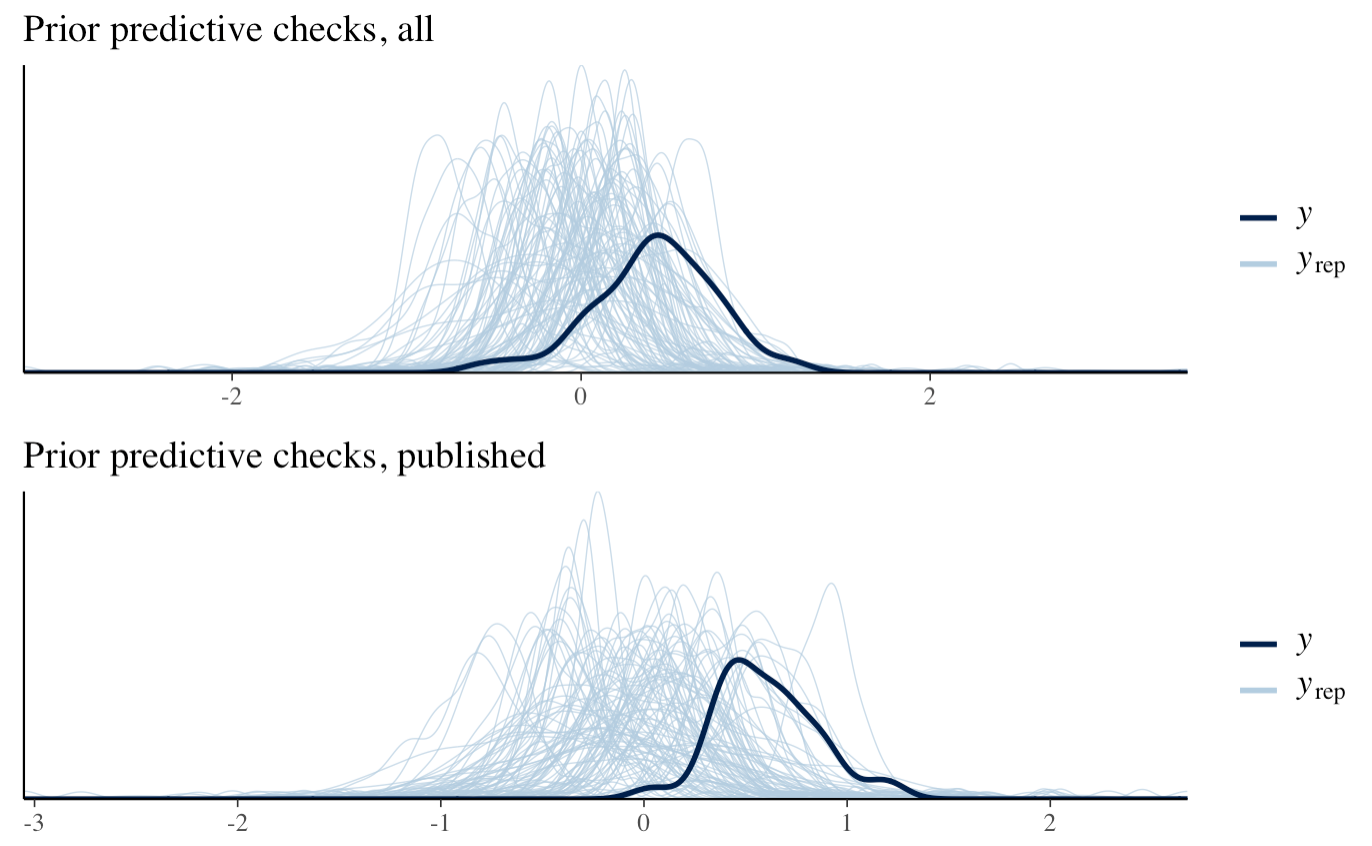


This following plot holds information of all the “published” studies. It shows a mean of 0.78 which is close to twice the size of the previous and with a much more skewed distribution not resembling the perfectly normal distribution from the simulated data in the previous plot:

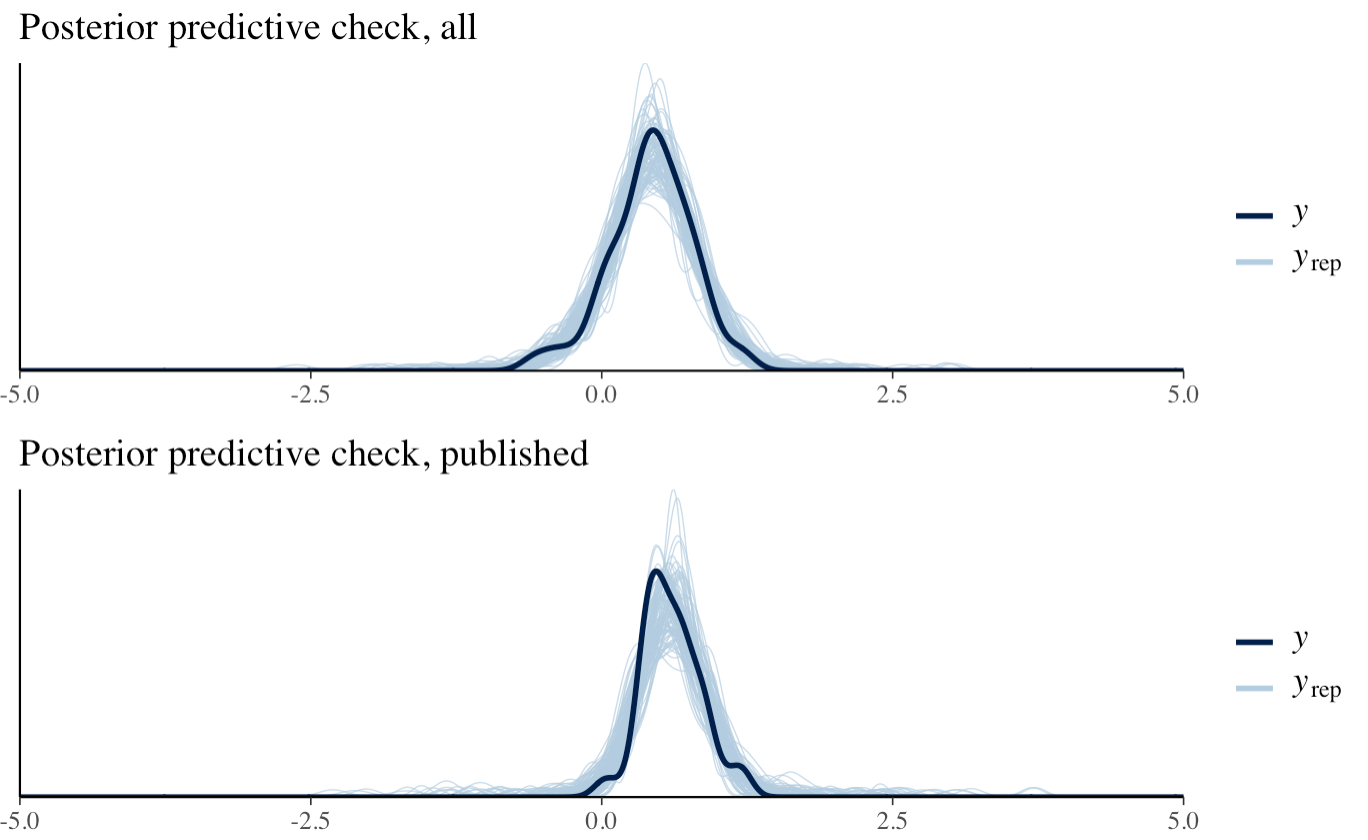
After having simulated the Data we followed a typical Bayesian work flow. Defining the formula: Study\_f <- bf(Study\_effect | se(Standard\_error) ~ 1 + (1 | Study)), and setting some weekly informed priors with mu = 0 a value of 0.3 for the intercept and 0.2 for the standard deviation.

The plot displays the predictions from our prior distributions in blue and the simulated data in black to assess if our priors capture the range of the simulated data.

Prior predictive checks:

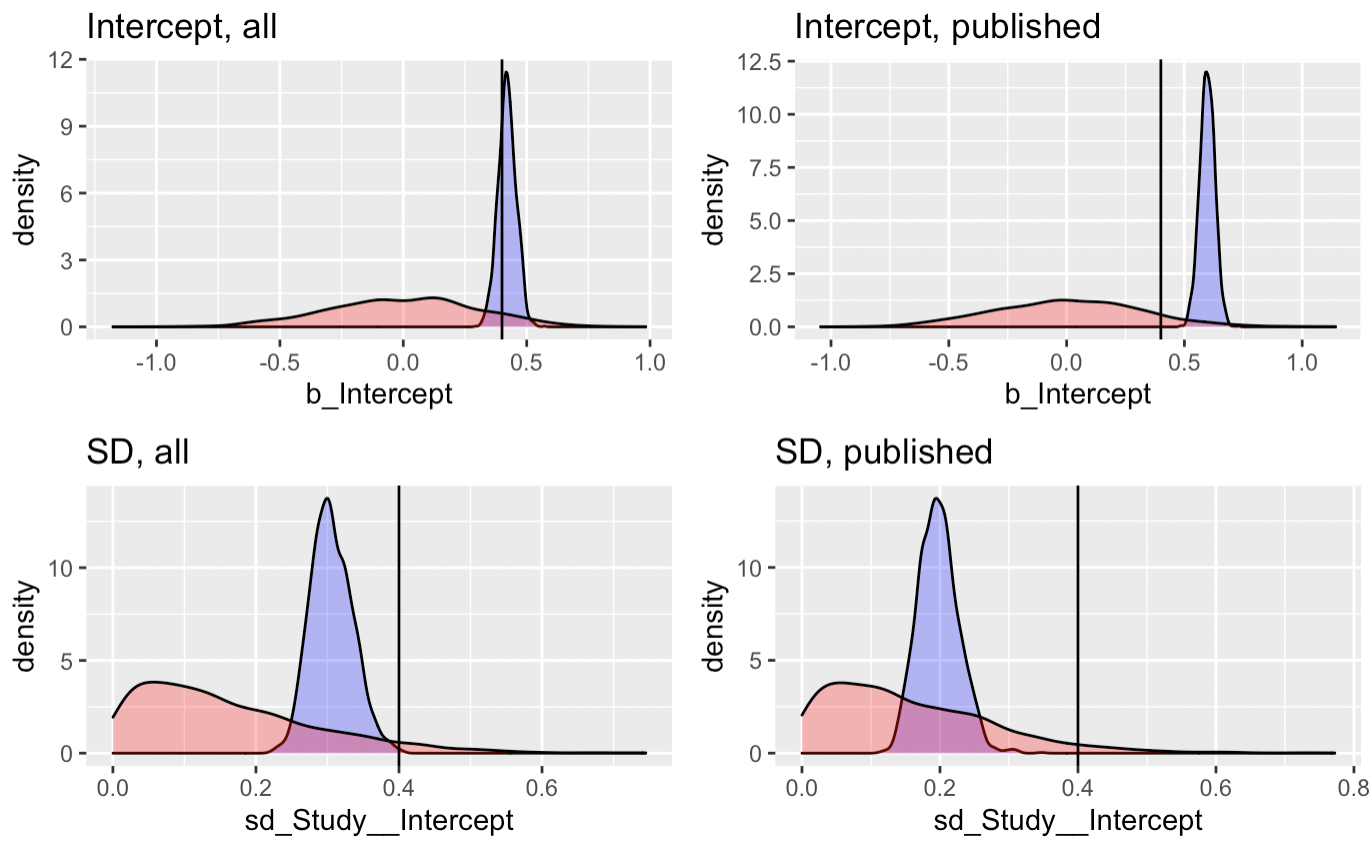


The prior predictive checks look a bit messy but still sensible, so we continue and fit our model to the simulated data.



The Posterior predictive checks look good for both datasets indicating that our model captures the data pretty well.

Some prior-posterior update plots



The prior-posterior update plots show a bit of difference between the two datasets. When looking at the intercepts for both datasets, we see a very narrow distribution that suggest that our posteriors are pretty convincing. The intercept for the published dataset pushes a bit more towards the tail of the posterior, so we could perhaps benefit from making our priors less narrow.

The prior-posterior update plots for the standard deviations look good. Both distributions are nicely placed within the range of the prior and are also rather narrow.

The difference between the two data sets might point towards a publication bias as we see a higher intercept for the published data set. This means that you should be a bit skeptical when looking at the effect sizes of published studies due to a greater possibility of being published when you report a large and, for many occasions, also positive effect size. Potentially, the true effect size is smaller than what is reported in a meta-analysis.

**2. What is the current evidence for distinctive vocal patterns in schizophrenia?**

**Use the data from Parola et al (2020) -** [**https://www.dropbox.com/s/0l9ur0gaabr80a8/Matrix\_MetaAnalysis\_Diagnosis\_updated290719.xlsx?dl=0**](https://www.dropbox.com/s/0l9ur0gaabr80a8/Matrix_MetaAnalysis_Diagnosis_updated290719.xlsx?dl=0) **focusing on pitch variability (PITCH\_F0SD).**

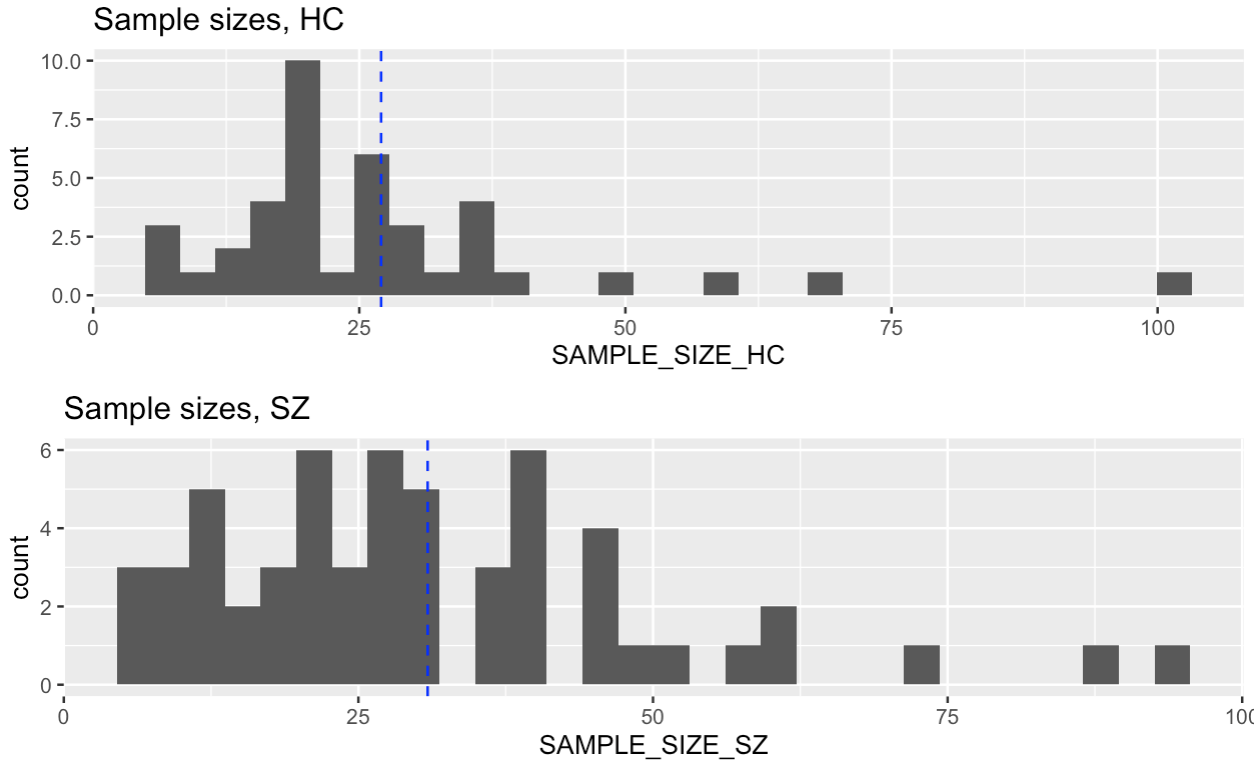
* **Describe the data available (studies, participants).**
* **Using the model from question 1 analyze the data,**
  + **visualize and report the findings:**
    - **population level effect size;**
    - **how well studies reflect it;**
    - **influential studies,**
    - **publication bias**

**Part 2**

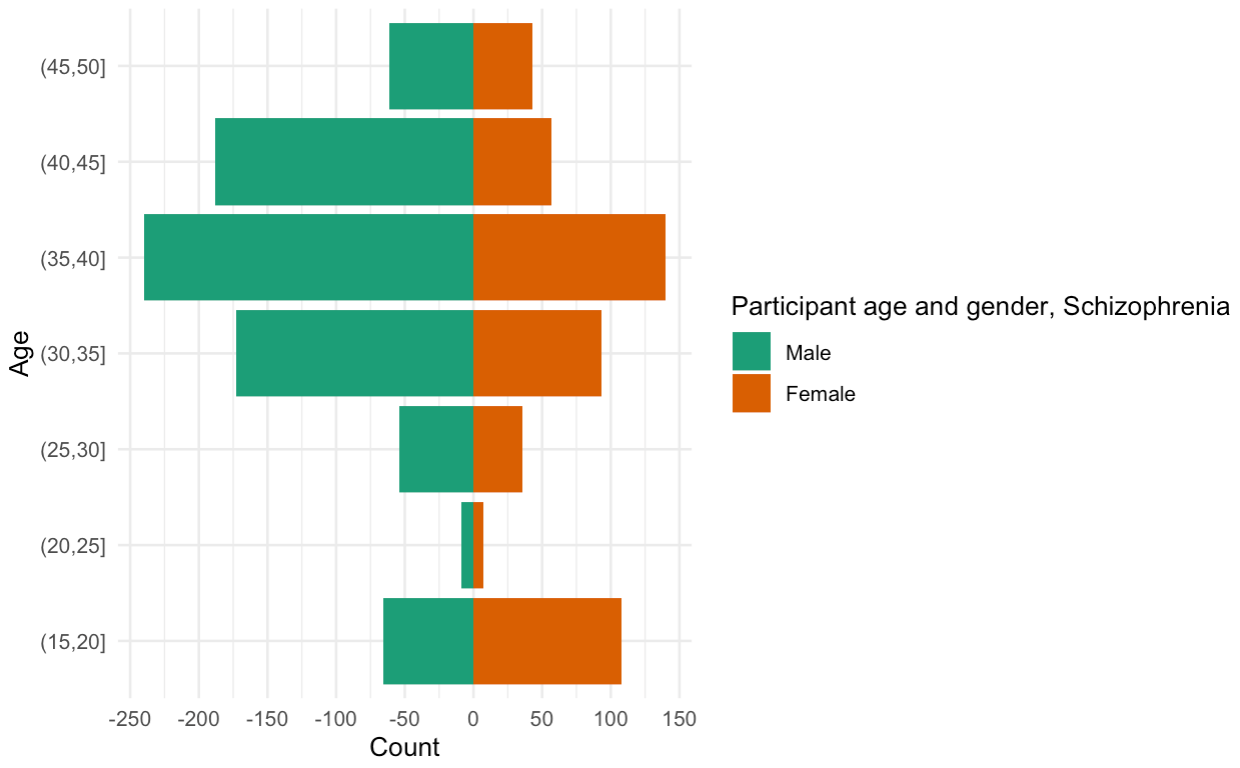
The data from Parola et al (2020) consists of a dataframe with the results and information from 48 studies all investigating vocal patterns in patients with schizophrenia. The studies vary greatly in sample sizes, both for the control group and for the participants with schizophrenia. The schizophrenic sample sizes vary from min 6 to a maximum of 101 participants with a mean of 27. The control group as well has a minimum of 6 but a maximum of 94 participants and a mean of 31.

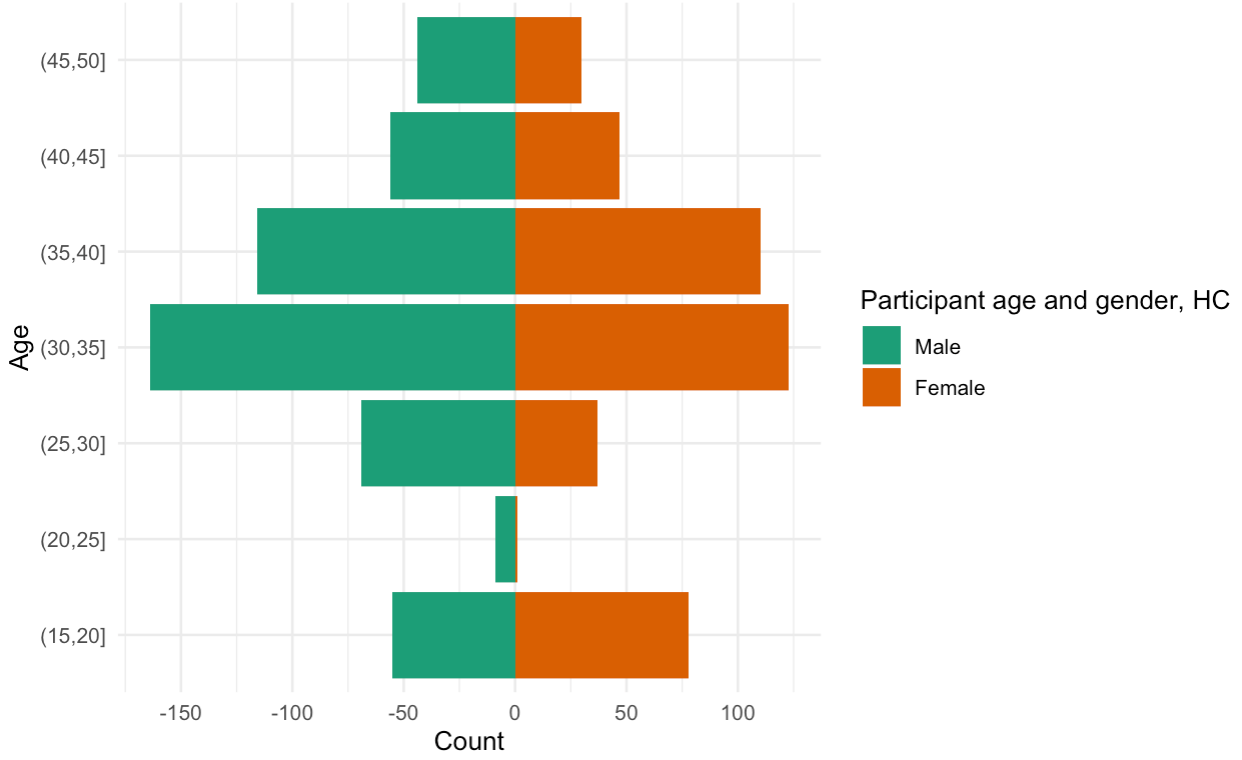
|  |  |  |  |
| --- | --- | --- | --- |
| Group | Sample size min | Samlpe size max | Sample size mean |
| Contol | 6 | 94 | 31 |
| Schizophrenia | 6 | 101 | 27 |

The distribution is visualized in the following histograms with the vertical line marking the mean:



The distribution of men and women and their age for the two groups:



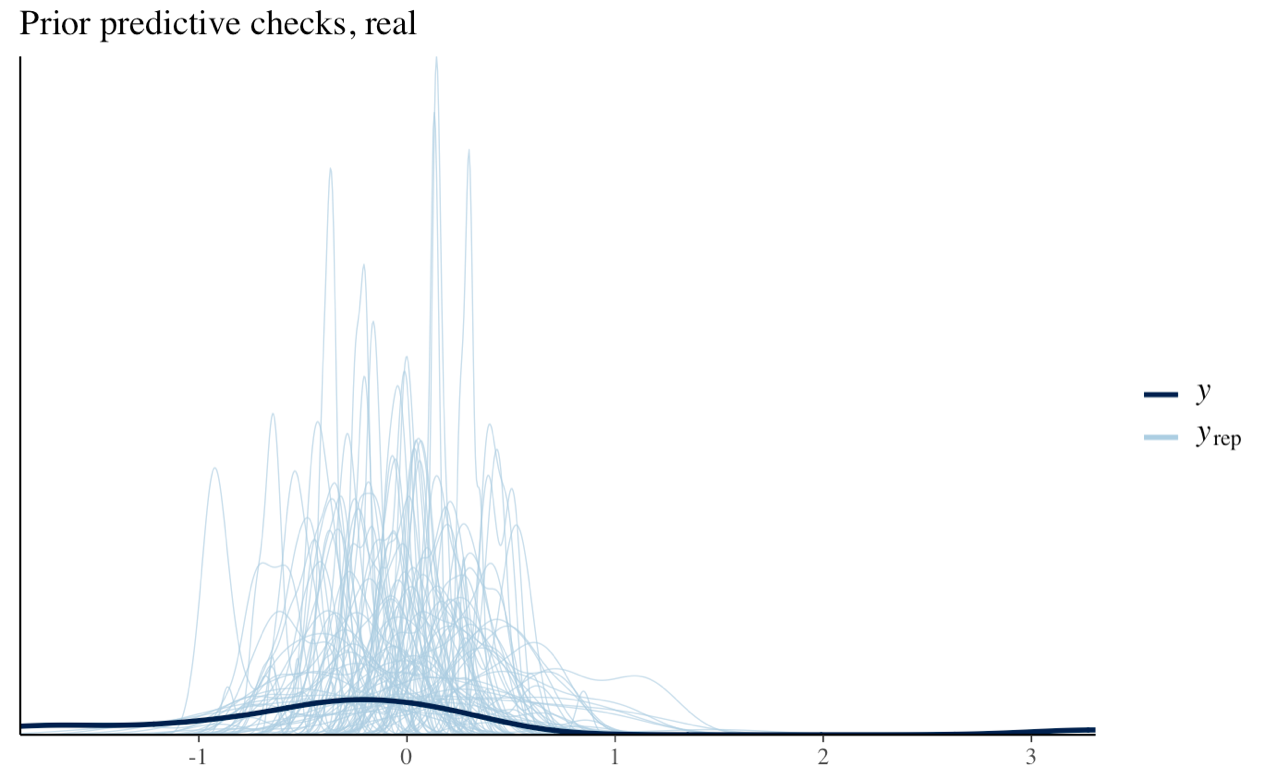


It seems there is an okay balance between the genders for both groups with an preponderance of people in their 30’s.

For the analysis, we used the Metafor package that has a function, escalc, for calculating effect sizes. This is to ensure that the effect sizes from the different studies are compatible, so they can be compared across the studies. The function added two new columns to our dataframe that calculate effect size as cohens d (D) and sampling variances (SE), renamed as "StudyEffect" = yi and "ObservedSigma" = vi.

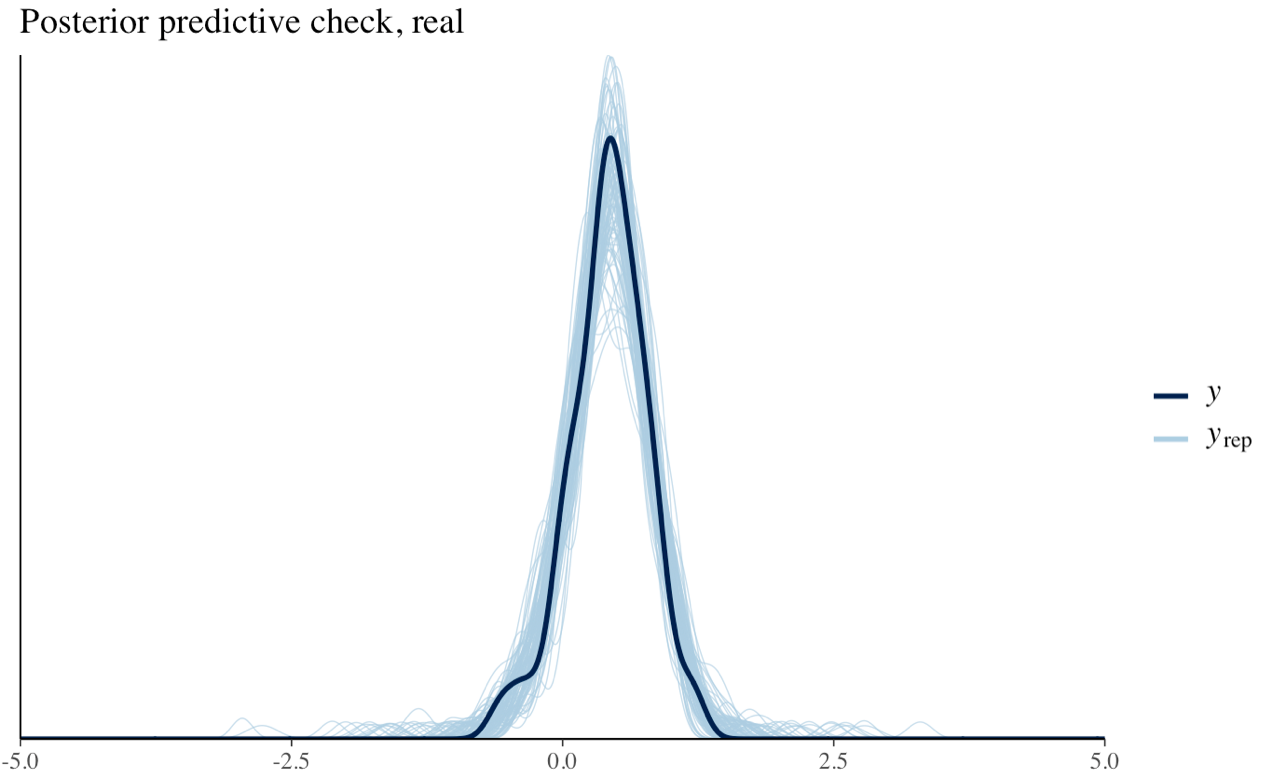
With the new columns we made a dataframe that only focuses on the pitch variability: PITCH\_F0SD, for the two groups across studies. The new dataframe contains the variables: Article, SAMPLE\_SIZE\_SZ, SAMPLE\_SIZE\_HC, PITCH\_F0SD\_SZ\_M, PITCH\_F0SD\_HC\_M, PITCH\_F0SD\_SZ\_SD, PITCH\_F0SD\_HC\_SD, yi, vi.

For the meta-analysis, we used the same model as for the simulated data: Study\_f <- bf(StudyEffect | se(ObservedSigma) ~ 1 + (1 | Article)) and same priors. Then we followed a normal Bayesian workflow:

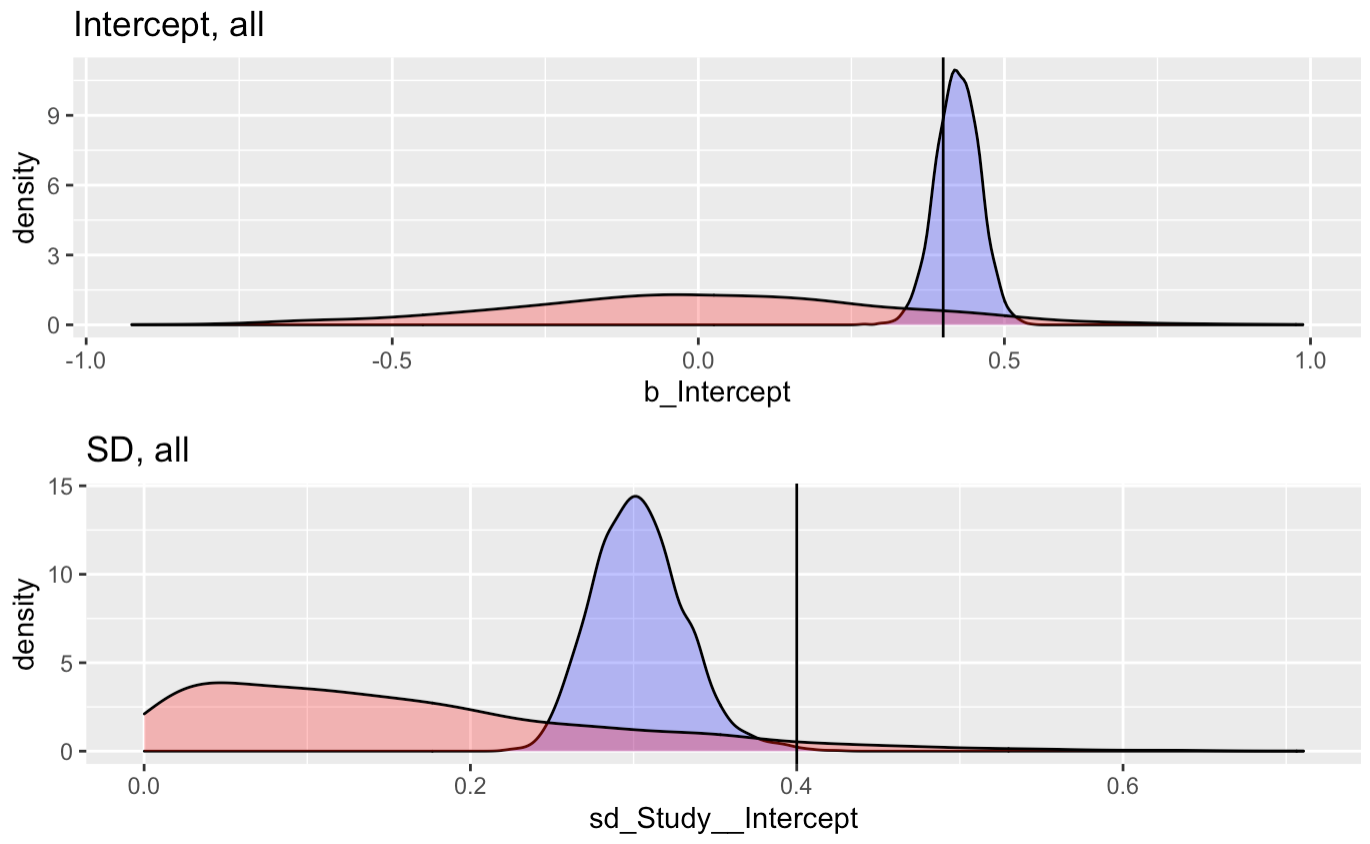


Looks okay. So, we fit our model and see that it captures our data:

Posterior predictive check.



Prior-posterior update-check:



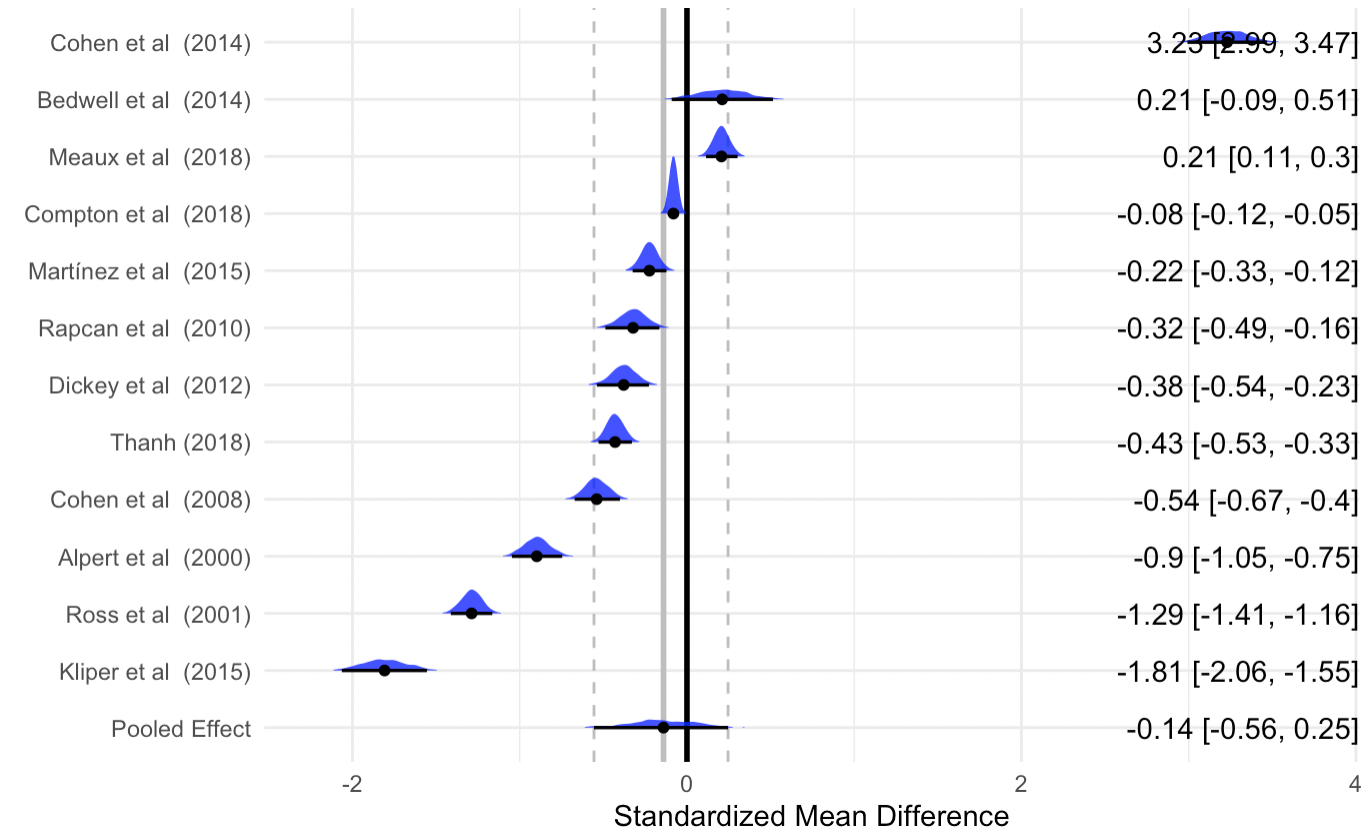
The posteriors are okay within the range of our priors.

When looking at the numbers from the model, we get a population level effect size of -0.14. However, it’s evident that our number of observations is a bit problematic, since it’s down to 15, which is a big reduction from the 48 studies, we started with. But the dataset contained many NAs that were removed, when during the calculation of cohens d (D) and sampling variances (SE).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | Family: gaussian | | Links: mu = identity; sigma = identity | | Formula: StudyEffect | se(ObservedSigma) ~ 1 + (1 | Article) | | Data: PitchMeta (Number of observations: 15) | | Draws: 2 chains, each with iter = 1000; warmup = 100; thin = 1; | | total post-warmup draws = 1800 | |  | | Group-Level Effects: | | ~Article (Number of levels: 12) | | Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS | | sd(Intercept) 0.80 0.09 0.66 0.99 1.01 387 691 | |  | | Population-Level Effects: | | Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS | | Intercept -0.14 0.18 -0.47 0.23 1.00 303 497 | |  | | Family Specific Parameters: | | Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS | | sigma 0.00 0.00 0.00 0.00 NA NA NA | |  | | Draws were sampled using sample(hmc). 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). | |

The population-level effect size of -0.14 is in general not reflected very well across the studies. From the Forest plot the estimate of the “true” effect size (-0.14) of the studies based on our Bayesian model is not captured by many studies.

it’s evident that many studies get results that are very different. We have some very influential studies fx especially Cohen et al (2014) that has an effect of 3.23, as well as Kliper et al (2015) with an effect of -1.81.



We also know from assignment 1 that we should be skeptical about the finding of the effect sizes, since a potential publication bias is very probable and therefore could influence the results, so it could be possible that there actually in reality isn’t a proper effect of pitch variability.