Portfolio 2

Methods 3, Bachelor of Cognitive Science
By Laura Rahbek, Marie Frederiksen, Ida Dencker & Sofie Mosegaard

Q1: simulating data

Firstly we simulated 100 studies ($\mu = 20$, $\sigma = 10$) with a mean effect size of 0.4, average standard deviation by study of 0.4 and measurement error of 0.8. By doing so the 100 studies had their own individual effect size mean and standard error based on the population distribution. Thereafter, we specified the only intercept formula:

Sam_EffectMu| se(Sam_EffectSigma) ~ 1 + (1|Study),

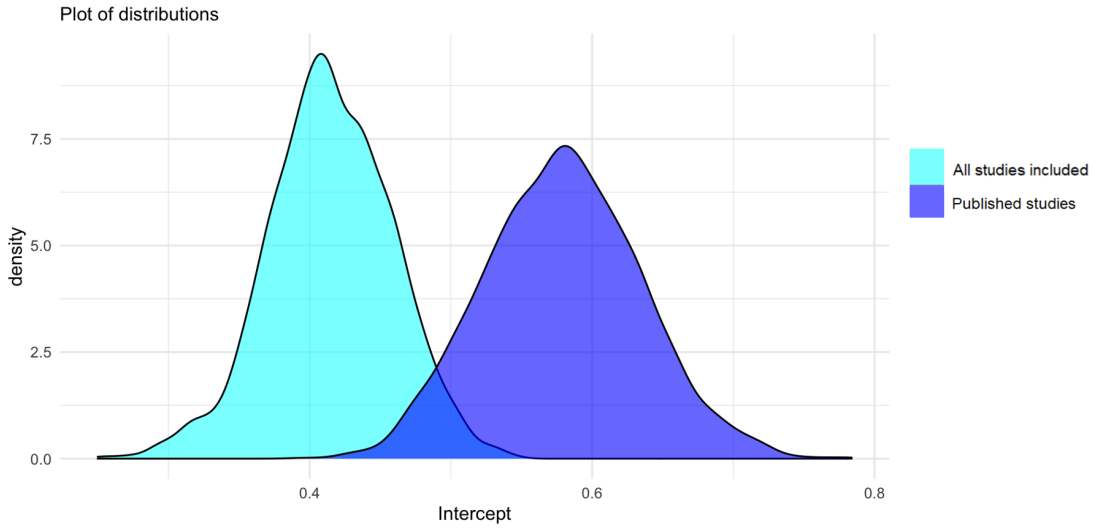
By making sure that the distribution of Sam_EffectMu is weighted by the standard error, meaning that more error will lead to less weight in each study.

When setting our priors we choose the default given by R, i.e. wide priors. In doing so, we want the data to persuade us of any possible difference in the group. This is why we set the mean to be 0 (no difference). Following our knowledge of effect sizes in the psychological field, an effect size of 0.6 is considered very large. When setting a prior of mean 0 and SD of 0.3, we get a distribution where 96% of our data falls between -0.6 and 0.6, allowing for extreme effect sizes in both ends, i.e. wide priors.

The parameters recovered from the model of all simulated data had an estimated mean effect size (intercept) of 0.41 and a SD of 0.4, corresponding to the distribution the data is simulated from, which is expected.

We then made a new dataset, with a publication bias. We choose a publication bias that depends on whether the study or not yields a significant result, i.e. two times SD of the mean subtracted from the absolute value of the mean. If the study did meet this, it is given a 90% possibility of publication, if not only 10%. This excludes roughly 40% of the simulated data. We did not specify whether or not the mean difference in participants with schizophrenia and the control group should be negatively or positively significant. If we had a hypothesis, regarding the direction of the difference, we could have included such. We then fitted a new model using only the published studies.

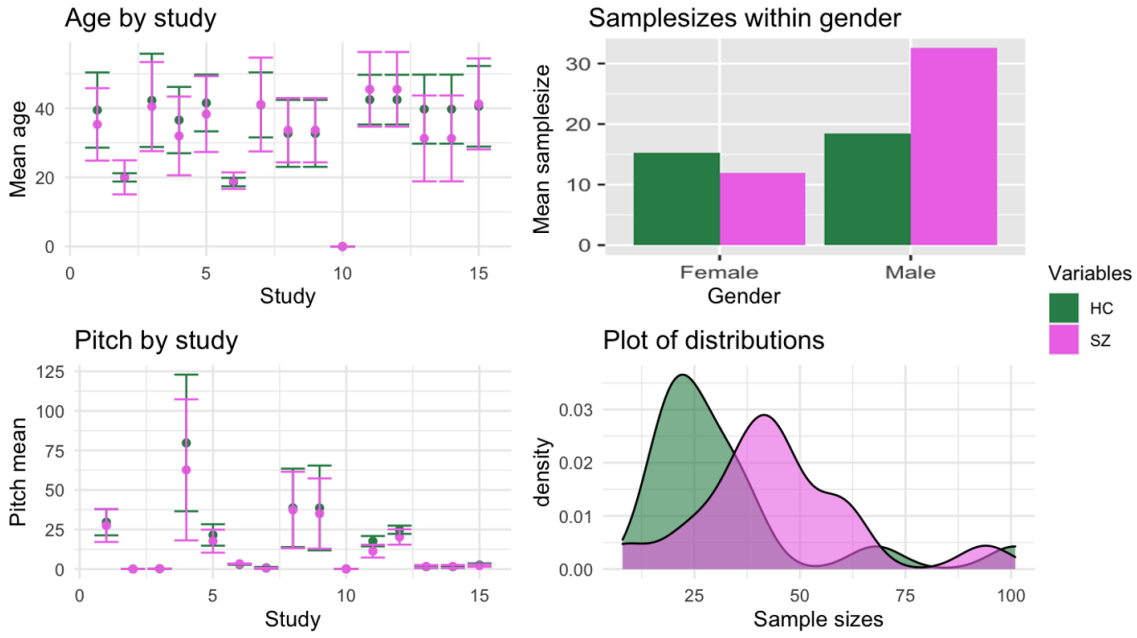
The parameters recovered from the model based on only the published studies had a much higher estimated mean effect size (Sam_EffectMu = 0.6), than the effect size found in the first model based on all 100 studies (Sam_EffectMu = 0.4). The difference is visualized in plot 1. This new effect size is .2 higher than the effect size estimated when taking all 100 studies into account. The mean effect size for the model that includes all studies is much closer to the defined true mean effect size (0.4). This points to the fact that publication bias skews results even when making a meta analysis, and to the fact that it is very important to take publication bias into account when making conclusions.



Plot 1: the two distributions based on the model 1 and model 2.

Q2: What is the current evidence for distinctive vocal patterns in schizophrenia?

We started by creating a data set extracting relevant data, including studies investigating pitch. To get an idea about the demographics of the remaining 15 studies we investigated gender, age, sample size and pitch for both groups.

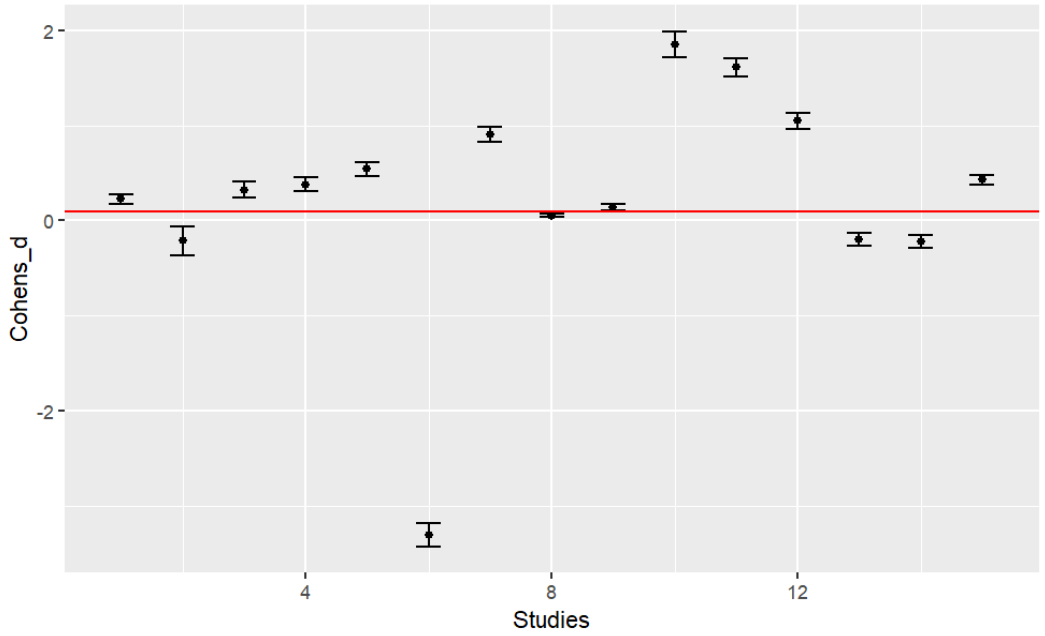


Plot 4: descriptive features of the variables

Plot 4 illustrates that the sample sizes within gender are different for the healthy control group and the schizophrenic group, where there are more males diagnosed with schizophrenia than healthy controls, which could resemble the real world. Furthermore, in general there are more males than females.

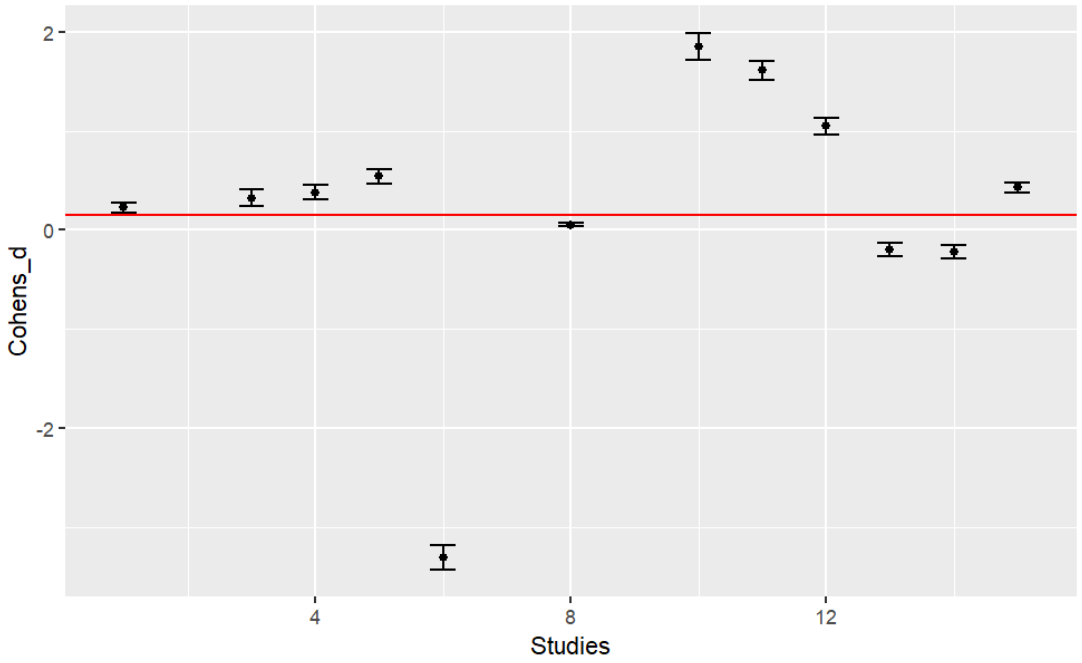
However, age, pitch and distribution between studies shows that the genders are more equally divided within the individual studies.

We calculated effect sizes of the studies to investigate vocal patterns in the two groups. In the analysis, we fitted a model using the dataset, the model's syntax is the same as the model used for the simulated data set in question 1. The estimated effect size of 0.1 reveals only a small difference in the two groups. The standard deviation of 0.94 indicates a relatively large spread, affected by outliers which are visualized in plot 5:



Plot 5: population level effect size and cohen's d per study with error bars

Applying the same publication bias as in the simulation in question 1, three studies were excluded (this varies from run to run). The number of studies excluded are low, because all the studies are published already, meaning they have been through the selection process once.



Plot 6: population level effect size and cohen's d per study with error bars after publication bias is applied

When investigating parameters of the new model including the set publication bias, we find that the parameters are not affected much. This is due to the fact that the studies that are excluded mainly are studies with large standard deviation, these are weighted less in the model. The new parameters show a mean effect size of 0.11 and a standard deviation of 0.93.

We generally find the current evidence for distinctive vocal patterns in schizophrenia to be weak. Models based on empirical data estimate a small effect size of approx. 0.1, meaning that healthy control groups in most cases have slightly more variability in pitch as they have a positive cohen's d. But due to the high deviation and outliers the effect comes with much uncertainty. Furthermore we found in question 1 that a lot of studies never reach publication due to too small effect sizes. In question 1 we found that the effect size of the studies that were 'published' was 0.2 larger than when all studies were included. From this we could argue that there actually is not a population level difference in the two groups, when the estimated effect size in the meta analysis is 0.1.