Updates on the Resampling Analysis

April 9, 2021

**The resampling analysis**

We have identified 20 published meta-analysis data set for the resampling analysis (see the attached Excel spreadsheet for a brief summary of the papers). The resampling analysis consists of the following steps:

1. Randomly choose 20 meta-analysis data sets with replacement from the 20 data sets we compiled (some data sets may be chosen multiple times);
2. For each data set, use the same meta-analysis model but add a grouping variable that randomly allocates each paper within the dataset into one of the two groups, obtain the p-value testing the difference between the two groups;
3. Perform step (2) for each of the 20 datasets and obtain 20 p-values. If all the meta-analysis methods are appropriate, the 20 p-values should form a uniform distribution between 0 and 1. Use Kolmogrov-Smirnov test to see if the 20 p-values differs from a uniform distribution.
4. Repeat step 1-3 many times. If the meta-analysis method is appropriate, we expect to observe only 5% significant deviation from uniform distribution.

First, we look at the p-values testing the random groups for the 20 meta-analysis data set (9 panels are 9 iterations of testing the random groups). Note this is different from step 1 (Each panel in the figure is based on the same 20 meta-analysis datasets we compiled while step 1 is a resample with replacement of the 20 datasets). This is not formally part of the analysis, but can help us get a rough idea of how many significant comparison occurs. In the figure, meta-analysis papers were color coded to indicate whether the analysis methods originally accounts for non-independence in some way.

A picture containing diagram

Description automatically generated

The 9 panels above shows 9 iterations of testing the difference of a random group. We can also plot the “average” histogram. Here, an average histogram is a histogram based on all 3000 iterations of the resampling for all 20 papers. The frequency is divided by the number of iterations so that it is directly comparable to each individual iteration.

Chart, histogram

Description automatically generated

After performing the resampling analysis described in steps 1-4 3000 times, we found that significant deviation from uniform distribution occurs 73.03% times (95% confidence interval 71.4%–74.6%). If the original meta-analysis methods are appropriate, we expect that deviation from uniform distribution occurs 5% times. Thus, this result suggests that some of the methods used in the 20 meta-analyses are not appropriate, resulting in non-uniform distribution of the p-values testing the random groups. Figure below shows the frequency distribution (3000 iterations) of the p-value resulting from the Kolmogrov-Smirnov test for uniform distribution of the 20 p-values. If the meta-analyses methods are appropriate, this distribution should be uniform as well and there should be about 5% times when p<0.05.

Chart, histogram

Description automatically generated

Deviation from uniform distribution for the 20 p-values tells us that there is problem with the meta-analysis methods. It, however, does not tell us the magnitude of the problem. To characterize the magnitude of the issue, we can obtain the proportion of significant group comparison from each round of resampling (step 1–3) and then obtain a mean and confidence interval through bootstrapping (step 4). We found that using the same method as in the published meta-analysis result in on average 31.43% significant comparison (95% CI 17.72%-58.29%). Figure below shows the frequency distribution of the proportion of significant comparison in 3000 iterations of bootstrapping.

Chart, histogram

Description automatically generated

**Does adding a random paper effect address the issue?**

The analyses above suggest that the method used in the 20 meta-analyses do not properly address the “paper effect”. Here we explore whether adding a simple random paper effect in the model could address the issue. We performed the same analysis as outlined above, but add a random paper effect in the meta-analysis model.

Adding a random paper effect alleviate the issue. Here, we present the same set of figures as in the original resampling analysis. First, 9 examples of resampling analysis with an added random paper effect using the 20 meta-analysis data set. Comparing to the original resampling analysis, frequency of significant results reduced and the distribution appears to be more uniform.

A picture containing chart

Description automatically generated

We can also combine all 3000 iterations of resampling to visualize an average histogram of p-values testing a random group after adding a random paper effect in the model.

Chart, histogram

Description automatically generated

Adding a random paper effect also reduces the occurrence of deviation of the 20 p-values from a uniform distribution (21.1%, 95% CI 19.6%–22.6%) although it still exceeds the 5% expected when the meta-analysis method correctly models the data. This suggest that adding a random paper effect help alleviate, but does not fully correct the issue. Figure below shows the frequency distribution for the Kolmogrov-Smirnov test for uniform distribution. Note the distribution is not uniform, but appears more “uniform” compared to the corresponding figure using the original method in the meta-analysis.

Chart, histogram

Description automatically generated

After adding a random paper effect, on average, 13.63% (95% CI 14.07–38.44%) of the 20 meta-analysis shows significant differences between the random groups. This is a marked decrease from when using the original method in the published meta-analysis. Figure below shows the proportion of significant results from bootstrapping.

Chart, histogram

Description automatically generated

The meta-analysis datasets included in our analysis all contained some sort of gropu comparison. To illustrate how adding a random paper effect influences the p-value for the original group comparison, we compared the p-value for the original group comparison in the meta-analysis with or without adding a random paper effect. The right panel is zoomed in between 0-0.01 of the left panel for clarity of visualization.

Chart, scatter chart

Description automatically generated

**Relationship between Number of studies per paper and proportion of significant comparison from the resampling analysis**

For each of the 20 meta-analysis datasets, we can repeatedly perform the resampling analysis and examine how often significant results occur. While a higher than 5% occurrence of significant result does not definitely proves that the method is inappropriate, examining whether the proportion of significant results relates to variables indicative of the existence of paper effect could be informative. Here, we looked at how proportion of significant results for each meta-analysis dataset relates to the number of studies per paper. Points are color coded based on whether the meta-analysis accounts for non-independence in some way.

Chart, scatter chart

Description automatically generated

**Contrasting meta-analysis accounting for or not accounting for nonindependence**

We divided the 20 meta-analysis into two subsets, one that accounts for nonindependence in some way and one that does not. We performed the same resampling analysis for each subset. Difference between the two subsets may imply the potential influence of nonindependence. It is clear that meta-analysis that does not account for non-independence results in far more deviation from uniform distribution of p-values in the resampling analysis.

Chart, histogram

Description automatically generated

For each round of resampling analysis, meta-analyses that does not account for non-independence resulted in much more significant result than those that account for non-independence in some way. But both are more than 5% (15.8% for those accounting for nonindependence and 42.6 for those not accounting for non-independence)

Chart, histogram

Description automatically generated