

**Fig S2**. Category percentages for the phrases describing non-significant results with the number of manually extracted P values with occurrences between 30 and 100 times in our manual analysis. Error bars represent the proportional 95% confidence interval. The associated median P value (with 25% and 75% quantiles) is presented in the upper left corner of each phrase.