

# FMRI data analysis: How does analytical variability vary with sample size?

## Introduction

Neuroimaging workflows are highly flexible leaving researchers with many possible choices to make at each step of their analysis (Carp, 2012).

However, different analytical choices can cause variation in the results (Botvinik-Nezer et al., 2020), leading to a "vibration of effects" (Ioannidis, 2008).

In psychology, (Klau et al., 2020) showed that the vibration of effects decreases and stabilizes as sample size increases.

### ➤ How varying sample sizes impact the vibration of effects in fMRI data analyses?

## Methods

We used the dataset from (Botvinik-Nezer et al., 2020), an fMRI "many analyst" study in which 70 teams used their favorite pipeline to analyze the same dataset and answer 9 hypotheses.

### 1 – Reproduction of teams results & validation

Using descriptions provided in the original study, we reproduced the pipelines of 8 teams using FSL and SPM.

To validate the reproductions, we compared our results with the original statistical maps published on NeuroVault (Gorgolewski et al., 2015).

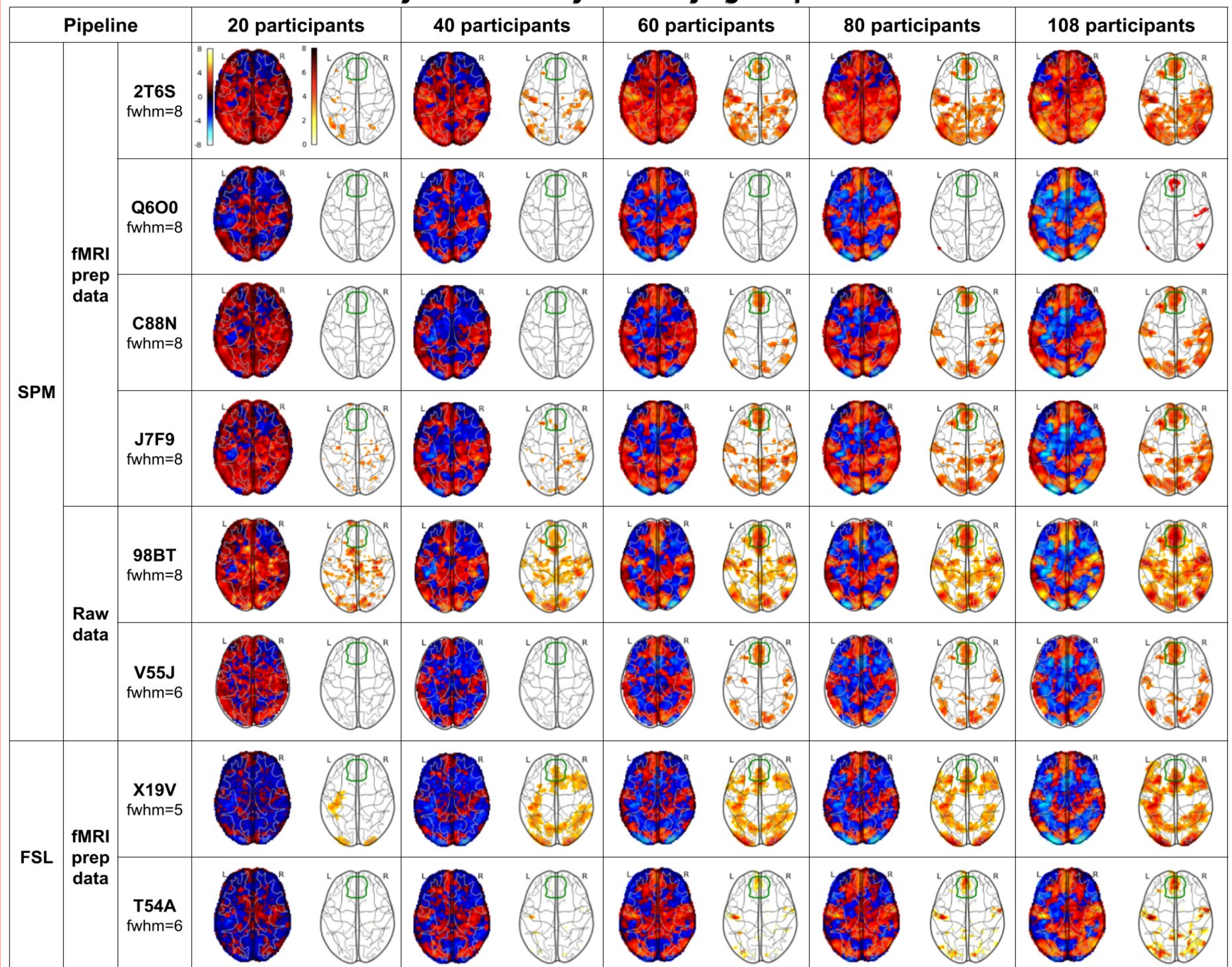
### 2 – Study of the vibration of effects under different sample sizes

We used subsets of size 20, 40, 60, 80 (each being a superset of the previous one) and the full dataset of 108 participants. All maps were converted to z-statistic maps for comparison.

Gitlab repository of all analyses, including reproduced pipelines: [https://gitlab.inria.fr/egermani/analytic\\_variability\\_fmri](https://gitlab.inria.fr/egermani/analytic_variability_fmri)

## Results

### Analytical variability with varying sample sizes



### Validation of reproductions

→ All reproductions were validated using qualitative and quantitative comparisons. We focused on the results of hypothesis 5 (H5).

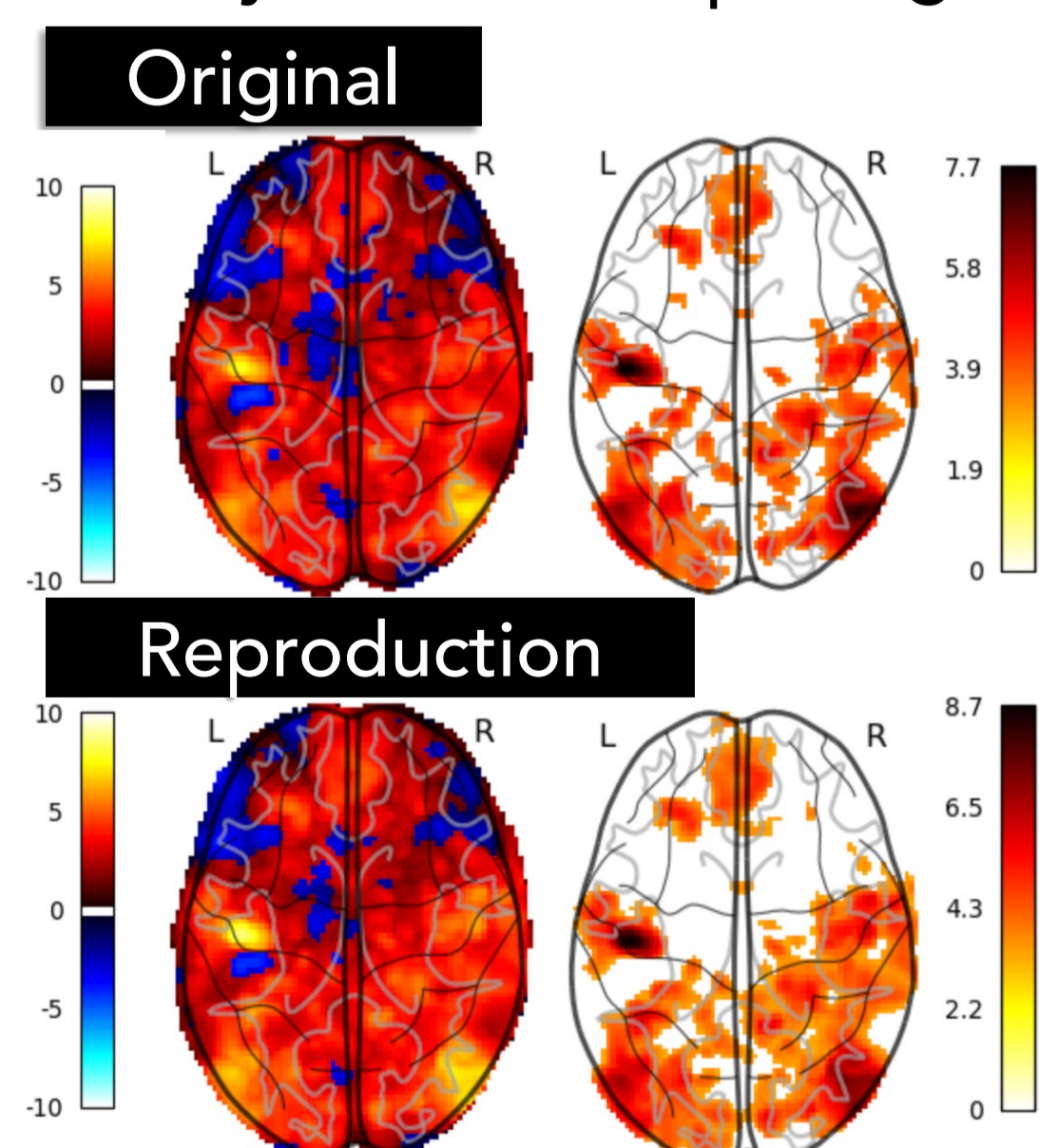
✓ Similar activation patterns were found in original and reproduced maps.

✓ Quantitative metrics showed the high quality of the reproduction (e.g. correlations original vs reproduced unthresholded maps close to 1).

✓ Larger differences observed between thresholded maps.

Pipeline	Pearson's correlation coefficient	Dice score
2T6S	0,98	0,77
V55J	0,98	0,75
Q6O0	0,99	0,63
C88N	0,99	0,81
X19V	0,96	0,80
98BT	0,97	0,60
J7F9	1	1
T54A	1	0.68

→ Here, we present the original and reproduced maps of the pipeline of team 2T6S for H5. This team used the dataset preprocessed with fMRIPrep and SPM12 as analysis software package.



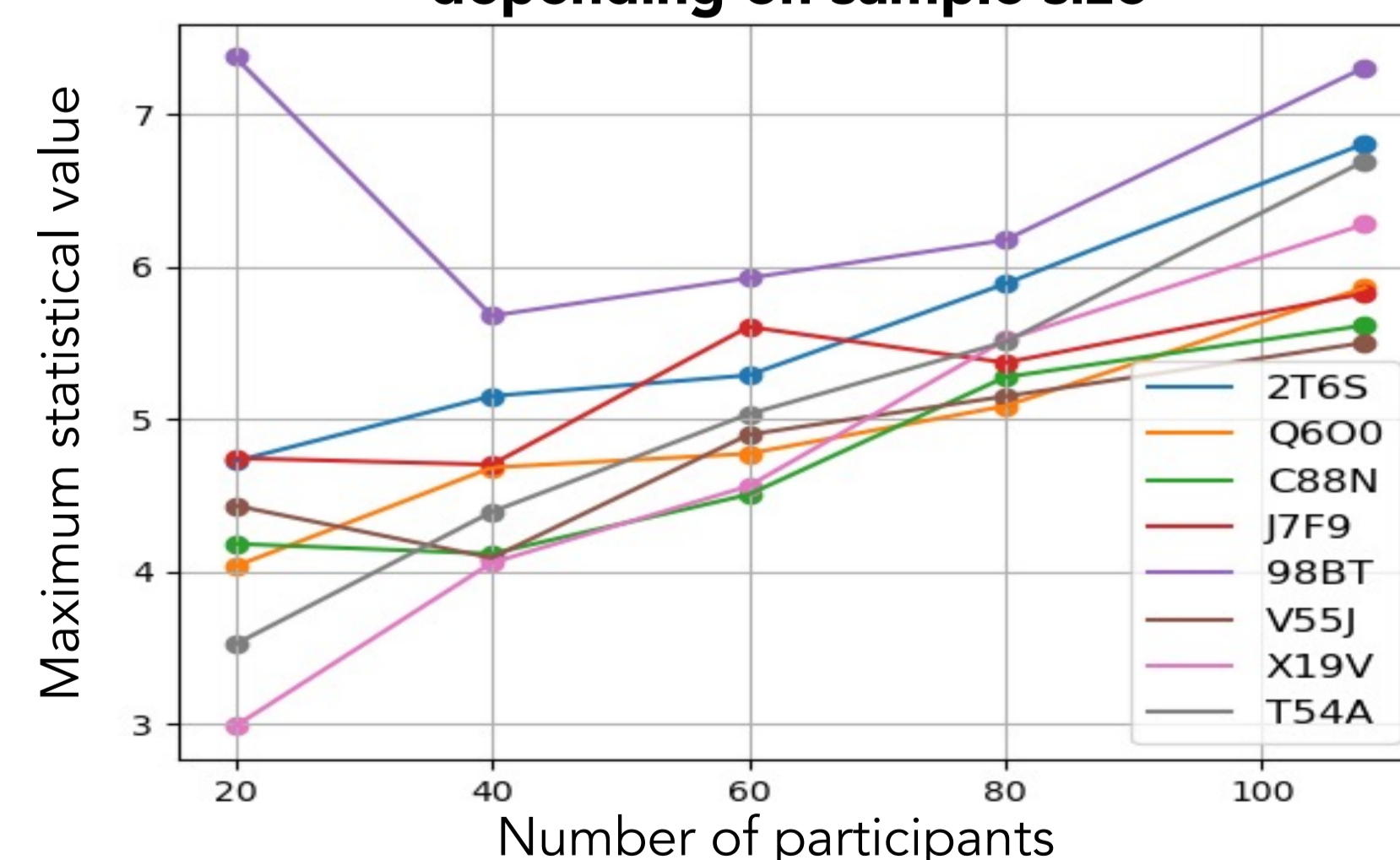
✓ Lots of variability between the maps of different pipelines.

✓ With growing sample sizes, maximum statistical values seemed to converge, consistent with the previous observation that the vibration of effects reduces and stabilizes with larger sample sizes.

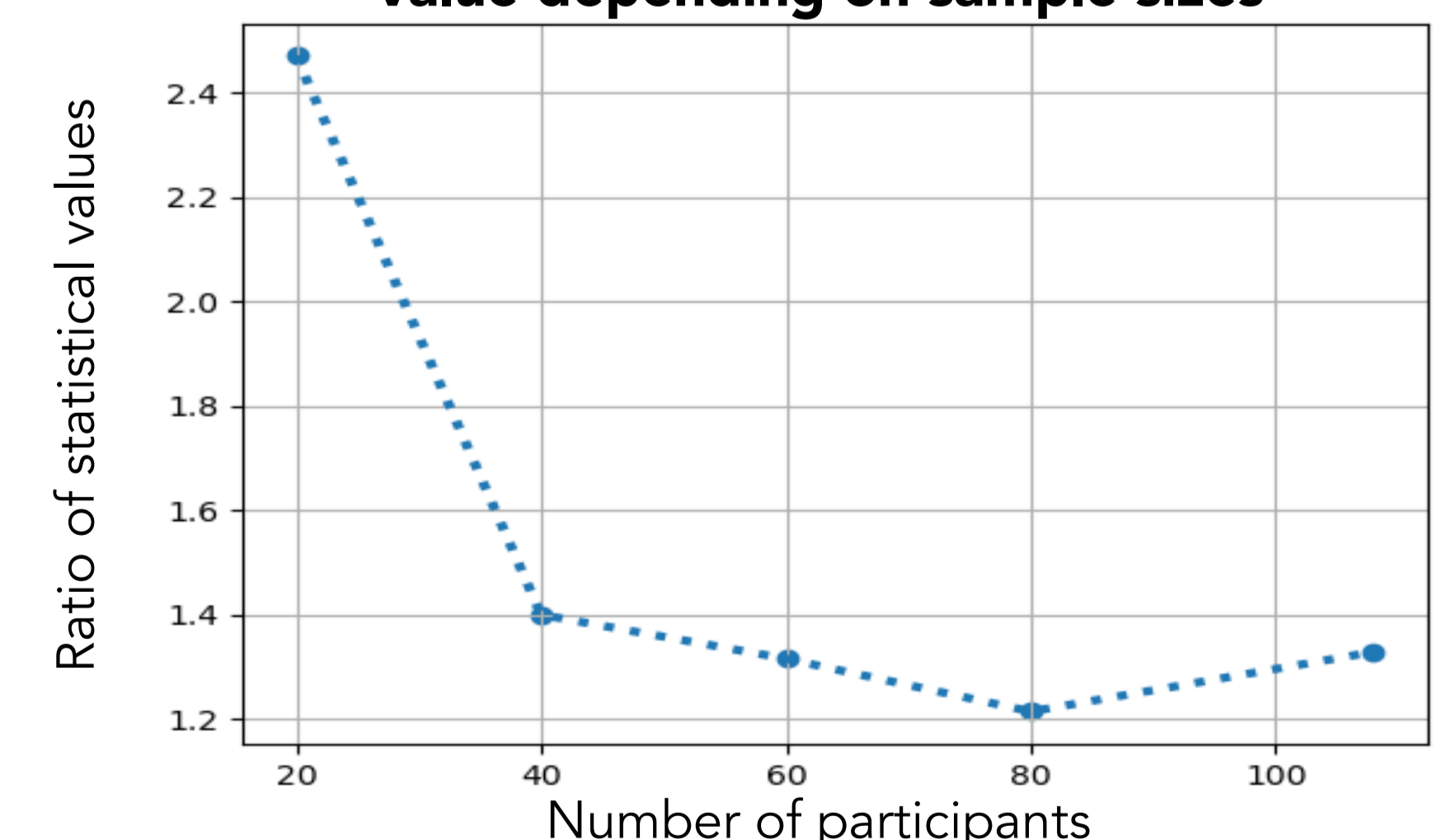
✓ Confirmed by the stabilization and convergence to 1.3 of the ratio between the largest and the smallest maximum statistical values among all teams with sample size.

✓ But still, variability remains for large sample sizes.

Maximum statistical value inside ROI depending on sample size



Ratio largest/smallest maximum statistical value depending on sample sizes



## Conclusion

Our findings show that, in fMRI data analysis, the vibration of effects decreases with sample size. Our results also suggest that some variability remain even for large sample sizes. Further work will be needed in order to include more pipelines and investigate which part of the pipelines are the most impactful.

## Acknowledgments

Elodie Germani was supported by Region Bretagne (ARED MAPPIS).

## References

- Botvinik-Nezer, R. et al. (2020), 'Variability in the analysis of a single neuroimaging dataset by many teams', Nature. ;
- Carp, J. et al. (2012), 'On the Plurality of (Methodological) Worlds: Estimating the Analytic Flexibility of fMRI Experiments', Frontiers in Neuroscience. ;
- Gorgolewski, K.J. et al. (2015), 'NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain' Frontiers in Neuroinformatics. ;
- Ioannidis, J.P.A. (2008), 'Why Most Discovered True Associations Are Inflated', Epidemiology. ;
- Klau, S. et al. (2020), 'Comparing the vibration of effects due to model, data pre-processing and sampling uncertainty on a large data set in personality psychology', Department of Statistics: Technical Reports.