Project Proposal

In recent years there has been a growing focus on evaluating the reproducibility of scientific findings. With a large, confused, and growing lexicon of definitions in this space, reproducibility can briefly be described as the ability for a researcher to successfully reproduce their own findings, while replicability refers to the ability of other researchers using similar means to arrive at the same conclusions [1]. Meta-analyses have been performed across several domains of science [2]–[4] which have shown that a significant proportion of claims fail to replicate, posing a significant problem for science.

In neuroimaging this issue has been explored with respect to the impact that analysis tools themselves have on claims and their replicability. In functional Magnetic Resonance Imaging (fMRI) selection of software tool has been shown to produce different results [5], and in structural MRI the stability of two software packages with respect to minor data perturbations has been evaluated [6]. Currently, there lacks a joint evaluation of tool stability and it relative impact to selection.

In my Ph.D. thesis I plan to characterize the stability of neuroimaging tools in the context of diffusion and functional MRI. Leveraging the publicly-available Consortium of Reproducibility and Reliability (CoRR) dataset [7] consisting of similar data collected across multiple sites, I will evaluate the stability of commonly-used analysis pipelines (FSL [8], MRTrix [9], Dipy [10]) with respect to perturbations in input data.

In numerical analysis, a condition number describes the stability of a function or matrix, where a larger value indicates a more significant change in output with respect to input, or, lower stability. This is represented functionally as $C = \frac{x f'(x)}{f(x)}$, where x is the input data, $f(x)$ is the function being applied to the data, and $f'(x)$ is the derivative of the function applied at the same point. While this conditioning can be computed for differentiable functions, in the case of complex processing pipelines performing multiple independent steps on high dimensional data, obtaining closed-form solutions is intractable. However, by performing known perturbations to unprocessed data, we can obtain an empirical estimate of tool conditioning based on the relative variance of input data and produced derivatives.

While this conditioning can provide insights into the stability of tools, it can also be applied across independent datasets and tools to serve as a proxy for the generalizability of derivatives between selections, and importantly identify the impact that a given dataset or tool may have on the obtained results. During this project abroad I will:

1. Identify target algorithms and tools in f-MRI for evaluation of numerical stability. These algorithms should be both commonly-used and differentiable, such that theoretical conditioning of the functions can be performed and compared to empirical estimates.

2. Characterize the significance of result instability in the within- and across-tool settings. This will be evaluated using test-retest reliability as well as techniques demonstrated and developed by Dr. Maumet.

The successful completion of my proposed project has the potential to shed light on the effect that numerical instabilities have on neuroimaging analyses, and identify a significant dependence between scientific claims and the tools used to generate them.

Benefits From Study Abroad

If funding, working with Dr. Camille Maumet's team at the Institut national de recherche en informatique et en automatique (INRIA, Rennes, France) will have a hugely positive impact on the trajectory and outputs of my Ph.D. Dr. Maumet has significant experience with both provenance-tracing and the analysis and comparison of popular fMRI pipelines. I will be able to learn state-of-the-art approaches to managing computational experiments and their results, and methods for evaluating the significance of variability among outputs. Without a collaboration with Dr. Maumet, I would take significantly longer to learn these methods and would not be able to easily translate my work from diffusion MRI to functional MRI applications, which is the dominant modality in the field.

Timeline for project including milestones

While at INRIA, the timeline for research will be as follows:

2018/04/01 - 05/01: My project will begin by learning the fMRI tool landscape. Dr. Maumet's team has worked extensively with the execution and evaluation of fMRI-based processing tools. The first month of this exchange will involve a large degree of knowledge exchange in which I am introduced thoroughly to processing tools, the steps involved, common failure-modes, implicit assumptions, and current exploration of the stability and reproducibility within this space.

05/01 - 06/01: Alongside learning about the fMRI processing pipelines and their steps, we will identify potential algorithms to compute closed-form solutions for condition and investigation for sources of instability. We will then compute the closed form solutions for these tools, and compute their empirical stability estimates.

06/01 - 07/01: Finally, we will compute and compare the stability of tools with that of subsequent claims made across both a) different datasets using the same tool, and b) the same dataset using a different tool. These two questions will lend themselves towards understanding the generalizability or transitivity of findings across both data and software selections.

References

[1] P. Patil, R. D. Peng, and J. Leek, “A statistical definition for reproducibility and replicability,” *bior*Xiv, Jul. 2016.

[2] Open Science Collaboration, “PSYCHOLOGY. Estimating the reproducibility of psychological science,” *Science*, vol. 349, no. 6251, p. aac4716, Aug. 2015.

[3] K. A. Baggerly and K. R. Coombes, “DERIVING CHEMOSENSITIVITY FROM CELL LINES: FORENSIC BIOINFORMATICS AND REPRODUCIBLE RESEARCH IN HIGH-THROUGHPUT BIOLOGY,” *Ann. Appl. Stat.*, vol. 3, no. 4, pp. 1309–1334, 2009.

[4] J. P. A. Ioannidis, “Why most published research findings are false,” *PLoS Med.*, vol. 2, no. 8, p. e124, Aug. 2005.

[5] A. Bowring, C. Maumet, and T. E. Nichols, “Exploring the Impact of Analysis Software on Task fMRI Results,” *bior*Xiv, Mar. 2018.

[6] A. Salari, L. Scaria, G. Kiar, and T. Glatard, *Numerical error propagation in the HCP structural pre-processing pipelines*. 2018.

[7] X.-N. Zuo et *al.*, “An open science resource for establishing reliability and reproducibility in functional connectomics,” *Sci Data*, vol. 1, p. 140049, Dec. 2014.

[8] M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, and S. M. Smith, “FSL,” *NeuroImage*, vol. 62, no. 2, pp. 782–790, Aug. 2012.

[9] J. D. Tournier and F. Calamante, “MRtrix: diffusion tractography in crossing fiber regions,” *International Journal of*, 2012.

[10] E. Garyfallidis et *al.*, “Dipy, a library for the analysis of diffusion MRI data,” *Front. Neuroinform.*, vol. 8, p. 8, Feb. 2014.