Learnings and recommendations for performing multi-center clinical drug trials with resting state functional MRI in Alzheimer patient population: data processing pipelines and functional connectivity metrics

Alexandre Coimbra[1], Farshid Faraji[1], Alex De Crespigny[1], and David Clayton[1]
[1]*Genentech, Inc., South San Francisco, California, United States*

Introduction: In single-site studies, brain functional connectivity metrics (FCMs) measured by resting-state functional MRI (rs-fMRI) have been reported to be sensitive to Alzheimer Disease (AD, [1]), beta-amyloid burden [2], and therapeutic effect [3]. Exploratory RS-fMRI was included in two global Phase II studies of a novel AD amyloid drug. We studied different types of data processing pipelines and FCMs make recommendations based on ensuing FCMs test-retest performance and ability to differentiate cohorts scanned at 1.5T vs. 3T magnets. A discussion on FCMs and statistical power to detect changes due to disease progression and/or therapeutic effect is also provided.

Materials & Methods: Data Acquisition: A total of 180 scans were acquired from 99 patients across 18 sites. Seven sites used 1.5T and the remaining 11 sites used 3T magnets from three major manufacturers. For 48 patients, two or three rs-fMRI datasets were acquired in 7-week time intervals, and the assumption was made that consecutive intra-subject scans were equivalent to test-retest data sets. There were 68 such consecutive pairs of rs-fMRI datasets. GRE-EPI and high-resolution anatomical T1-weighted scans were acquired using parameters similar to those used for ADNI2 (http://adni-info.org/Scientists/ProceduresManuals.aspx#); local optimizations were made for different magnet manufacturer and field strength. Data Pre-processing: All data were processed using FSL (FSL, FMRIB, Oxford, UK) and in-house developed Matlab (Mathworks, Natik MA) and Perl (www.perl.org) functions. EPI data was brain-extracted, motion-corrected, spatially (5 mm FWHM kernel) and temporally (100 s high pass cutoff) filtered. EPI datasets were registered to the brain extracted T1 scans, and subsequently registered to MNI152 standard space using both linear and non-linear registration algorithms. The following nuisance variables were regressed out of the EPI data by general linear model (GLM) fitting: mean of predefined CSF and WM seed, global brain signal, 6 motion correction parameters. FCMs: FCMs were computed for both linear and non-linear MNI152-registered data. For each dataset (1) goodness-of-fit was computed for the Default Mode Network (GOF-DMN) [4], following independent component analysis (ICA) performed using FSL/MELODIC with automatic dimensionality estimation. (2) Degree of functional connectivity between the Posterior Cingulate Cortex (PCC) and gray matter was assessed by averaging the z-score of GLM fitting of the mean PCC signal (Seed-Z-score). Finally, (3) the temporal correlation between mean PCC and precuneus signals was computed with Pearson correlation coefficients (R-PCCxPrecu). Assessments: processing pipelines were assessed with (1) test-retest analysis of ensuing FCMs using intra-class correlation (ICC) and coefficient of variation (CoV), and with (2) two-sample t-test comparison of FCMs measured from data acquired at 1.5T and 3T magnets. Sample size curves were derived to infer minimal detectable FCM changes.

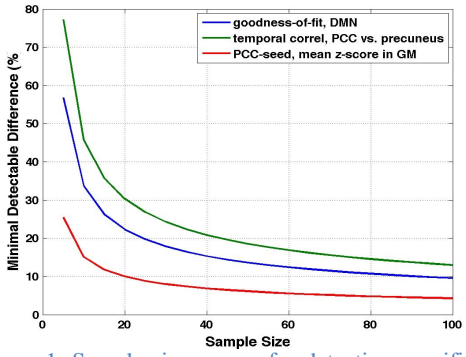

Figure 1: Sample size curves for detecting specific % changes with 80% power at 0.05 significance level across the three FCMs investigated in this work

Results: Table 1 shows test-retest and field strength comparison between FCMs computed with linear and non-linear registration pipelines at 1.5T and 3T. Test-retest performance is comparable for linear and non-linear registration pipelines. However, the non-linear registration pipeline rendered larger values for FCMs. Also, significant difference in connectivity between PCC and precuneus as measured by Pearson R, was observed with non-linear registration pipeline, but not linear pipeline. Regressing nuisance variables out of the EPI data also significantly enhanced both FCM strength and reproducibility (not shown). Finally, Figure 1 shows sample size curves for detecting specific % changes with 80% power at 0.05 significance level across the three FCMs investigated in this work. Note that sample sizes 25-15 patients are needed to detect 20-30% change in GOF-DMN (corresponding to ~25% of the mean GOF difference observed between AD patients and age matched controls [1]).

	Linear registration					Non-linear registration				
	1.5T vs 3.0T			Test-retest		1.5T vs 3.0T			Test-retest	
	Mean (SD) 1.5T	Mean (SD) 3T	p	ICC [c.i.]	CoV (%)	Mean (SD) 1.5T	Mean (SD) 3T	p	ICC [c.i.]	CoV (%)
GOF-DMN	0.71 (0.34)	1.16 (0.38)	<0.0001	0.60 [0.42 0.73]	35.7	0.77 (0.39)	1.23 (0.39)	<0.0001	0.61 [0.43 0.74]	33.7
R-PCCxPrecu	0.47 (0.28)	0.52 (0.27)	**=0.2205**	0.42 [0.20 0.60]	48.6	0.45 (0.29)	0.56 (0.23)	<0.005	0.46 [0.25 0.63]	45.8
Seed-Z-score	4.71 (0.57)	5.10 (0.58)	<0.00001	0.42 [0.20 0.60]	15.1	4.74 (0.53)	5.14 (0.57)	<0.00001	0.40 [0.18 0.58]	15.5

Table 1: Test-retest metrics and comparison between 1.5T and 3.0T FCMs. For data processed with both linear and nonlinear registration, FCMs mean and standard deviation at 1.5T and 3.0T are shown together with two sample t-test p-levels for the field strength comparison; reproducibility is characterized by ICC [confidence interval] and coefficient-of-variation. Larger FCMs are observed with non-linear registration pipeline. Nonlinear registration also better differentiates between FCMs computed from 1.5T and 3.0T data.

Conclusion: In the context of a multicenter clinical trial, we successfully tested and quantitatively characterized a number of data processing pipelines and rs-fMRI FCMs. Pipelines that include non-linear registration and regress nuisance variables out of the data render stronger and more reliable FCMs that better distinguish between 1.5T and 3.0T data. Sample sizes for detection of 25% FMC changes are within the range of typical sample size for a Phase II study. The potential clinical implications of an observed change of 20-30% in FCMs will need to be interpreted in the context of associated clinical data at the end of the trials.

References:

[1] Greicius et al 2004, *PNAS*, 101(13):4637-42; [2] Sheline et al 2010, *Biol Psychiatry*, 67:584–7; [3] Li et al 2012, *NeuroImage*, 60(2):1083-91.

Acknowledgments: We would like to thank M. Greicius and W. Shirer from Stanford University, Stanford CA, USA, for their valuable comments and suggestions to this work.