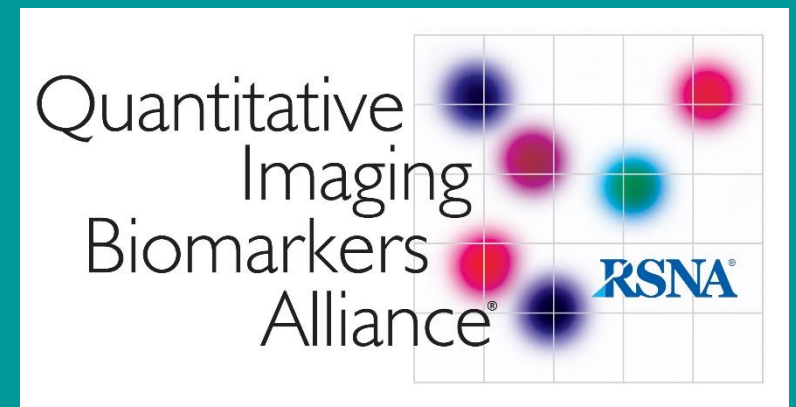


fMRI Biomarker Development: Progress Report 2016

Edgar DeYoe¹ PhD (co-chair), Cathy Elsinger² PhD, Feroze B Mohamed³ PhD, Nancy Obuchowski⁴ PhD, Jay Pillai⁵ MD, Jeffrey Petrella⁶ MD (co-chair), James Reuss⁷ PhD (co-chair), David Soltysik⁸ PhD, James Voyvodic⁶ PhD, Zhiyue Jerry Wang⁹ PhD, Yuxiang Zhou¹⁰ PhD
¹Medical College of Wisconsin, ²NordicNeuroLab, ³Thomas Jefferson University, ⁴Cleveland Clinic, ⁵Johns Hopkins University, ⁶Duke University, ⁷Prism Clinical Imaging, Inc., ⁸U.S. Food and Drug Administration, ⁹UT Southwestern Medical Center, ¹⁰Mayo Clinic Arizona



BOLD fMRI as a Quantitative Biomarker

The aim of the QIBA fMRI technical committee is to establish detailed profiles for using functional MRI as a quantitative biomarker for imaging brain function. The profile is a document describing claims for the quantitative precision that can be achieved with fMRI, along with equipment and procedure specifications for how to achieve those claims. The profile also specifies assessment procedures whereby users of the profile can assess their ability to conform with the profile's data quality conditions and specifications. The primary context of use for the fMRI biomarker profile is diagnostic fMRI to map critical brain areas for neurosurgical planning. An important aspect of the biomarker profile is identifying standardized imaging procedures for reliably obtaining reproducible quantitative fMRI results in a clinical context. The focus of Profile V1.0 is mapping of hand motor regions. Profile V2.0 will address mapping of brain regions involved in language processing.

The QIBA fMRI biomarker committee has established working groups with weekly teleconferences to focus on specific issues (i.e., reproducibility, bias, DICOM standards, profile writing). The committee has also undertaken several RSNA-funded groundwork projects to clarify specific issues critical to creating fMRI biomarker profiles for mapping motor and language function. Completed projects include:

- 1) Establishing metrics for assessing scan-rescan reproducibility and procedures for improving reproducibility of results.
- 2) Characterizing neuro-vascular uncoupling (NVU) affecting BOLD signals and QA methods for detecting NVU
- 3) Developing realistic, standardized, synthetic digital reference objects (DROs) for comparison of fMRI analysis methods in common use and for testing specific sources of variance in fMRI (e.g, head motion, behavioral performance, etc.)

Profile Status

fMRI Profile V1.0 has been released to the QIBA Community for public comment. Profile V1.0 establishes the claim that the center of mass of activation (CMA) for a motor task can be localized reproducibly in brain activation maps. Specifically:

Biomarker measurand: Local T2* MRI contrast change – commonly referred to as the BOLD fMRI signal

Context of use: Preoperative mapping of eloquent cortex for treatment planning/guidance

Cross-sectional measurement: Location of BOLD signal as a biomarker of motor cortex

Index: The center of mass of activation evoked by a hand movement task

Precision profile: If XYZ is the measured location of the weighted center-of-mass (wCMA) of a single focus of fMRI hand motor activation, then the 95% confidence interval for the true wCMA is XYZ +/-5mm (assuming no systematic bias). (The +/-5 mm precision value represents 1.96 x within-subject standard deviation).

Assessing Conformance – Calculating CMA

To establish conformance with the preceding precision profile claim, a site should:

1. Obtain multiple test-retest pairs of fMRI datasets using the profile hand movement task.
2. Compute statistical parametric brain maps displaying the fMRI amplitude T-statistic.
3. Select an objective statistical threshold for identifying responsive voxels.
4. Identify an ROI containing an fMRI activation focus in primary motor cortex.
5. Compute weighted center-of-mass of active voxels (CMA) for each motor ROI.
6. Calculate test-retest difference in CMA location for each pair = reproducibility metric.
7. If the mean reproducibility metric across subjects <= precision, then claim is conformant.

DICOM WG-16 Collaboration - Update

Goal: Add storage and transmission of fMRI data to the DICOM standard.

Progress in 2016: WG-16 (fMRI) finalized CP 1584 to add color palettes to DICOM Parametric Maps with Real World Values, the basis for storing fMRI activation maps. It has drafted Supplement 189 to address the unique blending requirements found in color visualization of multi-series brain mapping including fMRI and is working to reconcile that with other DICOM blending methods.

The group met face-to-face in Arlington, VA in September to complete these efforts and presented its work in progress for review at the WG-06 meeting the following week.

Current Plans: WG-16 (fMRI) will complete Sup 189 and submit for public comment. Next it will address other fMRI data requirements including representation of task paradigms, task results, and processing results.

Round 5 Project - Quantitating clinical fMRI mapping of language

Our Round 5 QIBA groundwork project focused on assessing reproducibility of clinical fMRI mapping of language function. To do so, we performed retrospective image analysis of fMRI sessions in which patients or healthy control subjects underwent multiple language mapping fMRI tasks. Analyses were carried out independently on subject populations at 2 different sites (Johns Hopkins and Duke). All language mapping was performed in English in fluent English speakers. Subjects performed one or more silent sentence completion (SC) tasks, and/or silent word-generation (WG) tasks (task choice was based on clinical need). Results presented here are all based on multiple language task scans within a single session.

Johns Hopkins sub-project: 63 patients performed 2 or more runs of the SC task and 98 performed 2 or more runs of the WG task. We computed holohemispheric language lateralization indices encompassing all supratentorial regions and excluding the cerebellum. The lateralization indices (LIs) were computed using two different methods: 1) a threshold-independent approach taking into account the t-value-weighted distributions of all positively correlated voxels in each hemisphere [Pillai et al., Neuroimage 2010; 54:Suppl 1: S136-145], and 2) a threshold-dependent approach utilizing a 50% AMPLE threshold.

Task		Sentence completion (SC)	Word generation (WG)
# Subjects		63	98
# Tasks		153	149
Laterality Index (Threshold-independent)	Correl. Coeff	0.59	0.82
	Mean LI	0.14	0.22
	StdDev	0.11	0.14
Laterality Index (AMPLE Thresholded)	Correl. Coeff	0.67	0.84
	Mean LI	0.31	0.41
	StdDev	0.22	0.22

Note higher mean LI and correlation coefficients with the AMPLE approach when compared to the threshold-independent approach. Overall mean LI and correlation coefficients were higher for the WG task than for the SC task.



QIBA Projects and activities have been funded in whole or in part with Federal funds from the National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, Department of Health and Human Services, under Contracts: **HHSN268201000050C**, **HHSN268201300071C** and **HHSN268201500021C**. Thanks to RSNA for their valuable support and to other QIBA committees for their valuable inputs and discussions.

Duke sub-project: Test-retest repeatability was assessed in subjects who performed the same language task more than once, and reproducibility across different tasks was assessed in subjects who performed 2 different types of language tasks. Of 350 subjects who performed multiple language task scans, here we only present results from subjects scanned on Siemens 3T scanners. Of those, 60 performed multiple SC tasks, and 54 performed both a SC task and a WG task. Activation maps were generated using either a general linear model (GLM) or simple t-test analysis. Language ROIs were generated automatically from a custom template, AMPLE-normalized, and then laterality indices (LI), activation center (CMA) and peak locations were calculated. Differences in LI in frontal and temporal language areas, and in CMA and peak XYZ locations were calculated for every pair of scans. The figure shows frequency histograms of difference metrics; most pairs fall in the lower bins, indicating good reproducibility. LI and CMA results agreed well for all subjects.

