Characterizing and Utilizing fMRI Fluctuations, Patterns, and Dynamics

Peter A. Bandettini*[a,b], Prantik Kundu [a],
Javier Gonzalez-Castillo [a], Masaya Misaki [a], Paul Guillod [a]

[a]Section on Functional Imaging Methods, Laboratory of Brain and Cognition, and [b]Functional MRI Core Facility, National Institutes of Health, 10 Center Dr. Bethesda, MD 20892

ABSTRACT

Functional MRI is fundamentally grounded in the hemodynamic response. With an increase in neuronal activity, blood flow increases, causing an increase in blood oxygenation, leading to an increase in transverse relaxation rate T2*. This increase in blood flow is slow and highly variable and shows a considerable spatial heterogeneity. In spite of these limitations, the hemodynamic response has been proven to be exquisitely sensitive to subtle differences in neuronal activity in time, over space, and between subjects. This paper is a brief review of my Keynote address describing some of the effort coming from my group that further demonstrates methods to robustly extract ever more information from both resting state fMRI and activation-induced fMRI. Specifically, I discuss 1) our new method to use multi-echo fMRI time series data collection to separate blood oxygen level dependent (BOLD) signal from non-BOLD signal, 2) activation of the whole brain obtained using a simple task and, importantly massive averaging and a model-free analysis approach, and 3) fMRI decoding of left vs right eye ocular dominance column activation with a timing offset as low as 100 ms.

Keywords: fMRI, resting state fMRI, functional connectivity, clustering, multi-echo EPI, physiologic fluctuations, blood oxygenation, fMRI decoding.

1. INTRODUCTION

Functional MRI has been in existence for over 20 years and yet still continues to advance and impress both those who are directly involved with developing it to those who look forward to using it clinically in a routine manner. The primary challenges with regard to fMRI are: 1. The fMRI signal is based on a sluggish, highly variable, and relatively complex hemodynamic response. 2. The time series fluctuations are relatively large and very difficult to remove sufficiently. 3. The neuronal activity as it evolves over time is likely much more complex than typically expected, even when using relatively simple paradigms. These challenges are illustrated schematically in Figure 1. In this paper, I summarize how some of our recent work addresses these issues and further pushes the sensitivity and utility of fMRI.

Medical Imaging 2013: Biomedical Applications in Molecular, Structural, and Functional Imaging,
edited by John B. Weaver, Robert C. Molthen, Proc. of SPIE Vol. 8672, 86720T
 · doi: 10.1117/12.2012737

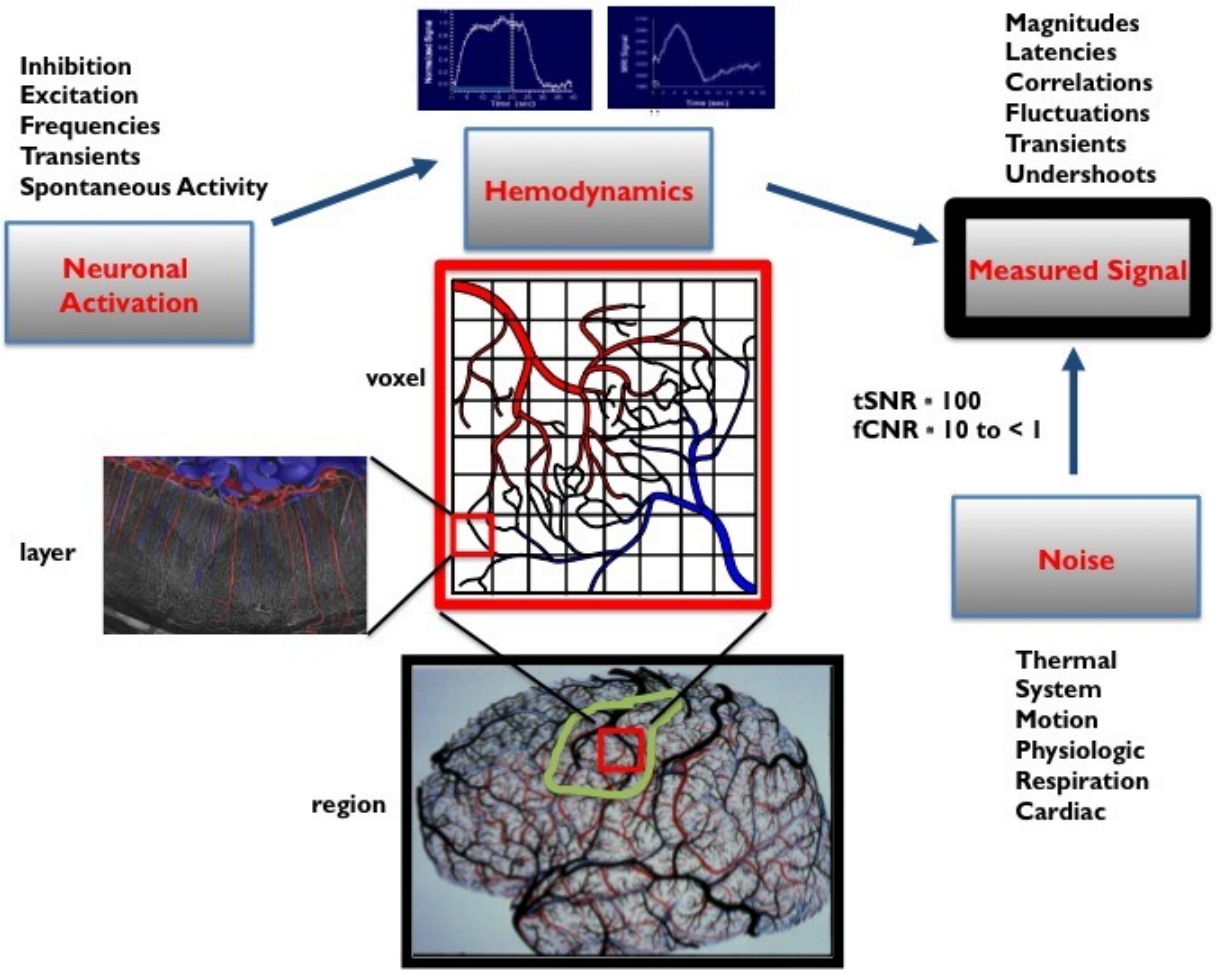

Figure 1: Illustration of the challenges presented to interpreting fMRI data. First neuronal activation can be very complex and variable, second the measured signal has many parameters that may or may not be sensitive to the type of neuronal activity being investigated. Third, the noise in the time series is prevalent, causing a temporal SNR of only 100 and a functional contrast to noise of about 10 down to less than 1.

This paper is organized into three sections, with each focusing on an attempt to increase fMRI interpretability and sensitivity. The first section is focused on resting state fluctuations and our method using multi-echo EPI, to clean up the noise (1) by separating BOLD from non-BOLD signal. The second section is focused on our recent experiment to determine the limits of fMRI sensitivity. Here we averaged over 10 hours of a subject performing a relatively simple task and were able to determine that all of gray matter and most subcortical regions are, in fact, activated by this relatively simple task. We were able to extract this information by relaxing our expected response models and performing clustering on this extremely high SNR data. A large percentage of the activated regions would have been missed if only a simple model were used (2). Third, we recently published a finding in which we were able to use decoding methods to extract, at 3T and using temporal and spatial resolution below what should be necessary, onset differences down to 100 ms between left and right ocular dominance columns. It should be noted that we did not map these, but rather used fMRI to decode this timing difference(3).

2. Multi-echo acquisition to sort ICA components and to minimize motion effects

Many methods are being actively developed to process and display resting state fMRI signal. Three of the most prominent of these include 1. The seed voxel approach in which the correlation coefficient of the time series from one voxel or region with every voxel in the brain is calculated, 2. Independent component analysis (ICA) for extracting spatially independent regions, and 3. Hub maps, the average correlation of every voxel with every other voxel in the brain is mapped. Seed voxel maps are useful in that they are hypothesis driven (i.e.: what regions specifically correlates with what other regions or voxels?). An example of a seed voxel approach that effectively maps the cross- colossal connectivity of the motor cortex is shown in Figure 2. Here, using a moving seed we demonstrate in a single subject, reduced connectivity in the hand motor region.

The method involving ICA is extremely efficient in extracting multiple independent sets of correlated regions, yet it's difficult to interpret the ICA components if time series artifacts and non-BOLD signal are present. I will discuss this topic more below.

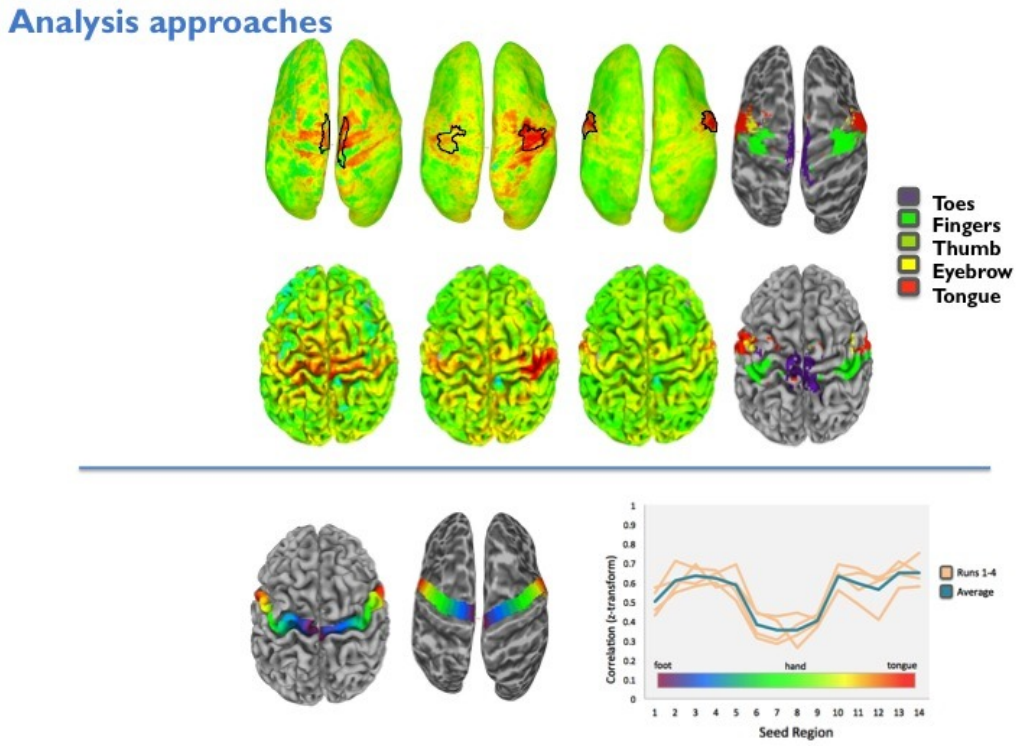

Figure 2: Cross-hemisphere resting state correlations and areas of activation with specific movements in flattened and non-flattened display. Color-coded motor strip locations and corresponding cross-hemisphere resting correlation magnitude showing reduced connectivity in the motor hand region. This shows a utility of the seed voxel approach.

An ongoing challenge has been to remove the large fraction of non-neuronal fluctuations rom rs-fMRI time series. Work on characterizing and removing non-neuronal fluctuations has focused on time series modeling based on external measures of physiologic processes. In this study, we fully embrace and carry forward the proposed work in our last BSC report from 2007 on using TE-dependence to separate BOLD from non-BOLD time signal so that non-BOLD - thus non-neuronally relevant - fluctuations can be removed. BOLD signal changes are manifest as changes in T2*, which can be characterized as showing a linear increase in fractional signal change with echo time (TE). Motion, system instabilities, and inflow effects can be manifest as changes in longitudinal relaxation, T_1 or proton density, S_0 but not typically T2*.

Here we show the utility of collecting multi-echo EPI (me-EPI) to clearly separate BOLD signal fluctuations from non-BOLD signal fluctuations. We collect three-echo me-EPI data and then apply independent component analysis (ICA) on the echo-concatenated time series data resulting in each ICA map consisting of three images, each collected at a different TE. On a voxel-wise basis, the change of the component's percent signal change with TE is fit separately to ΔR_2* and

ΔS_0 signal change models, and the fit quality is computed as an F-statistic. Each component is characterized by two metrics, κ for its BOLD likeness and ρ for its non-BOLD likeness. Figure 3 shows that that resting state BOLD components and non-BOLD noise components are differentiable according to these metrics.

When the κ values of all of the ICA components are sorted and plotted as a spectrum, high κ and low κ value regimes are clearly evident, separated by a boundary (an elbow in the curve). When the ρ values corresponding to κ values are given, it is clear that the majority of components have high κ and low ρ, or vice-versa, thus distinguishing BOLD and non-BOLD components. The differentiation of BOLD and non-BOLD components not only enables sorting of ICA components for subsequent clustering analysis but also allows for the use of low κ component time series for de-noising in task-based fMRI or seed based rs-MRI.

We demonstrate in our manuscript introducing this approach, that de-noising performed by removing the regressors identified as those with low κ components is more effective than standard methods involving modeling and regressing out the noise - particularly in problematic regions such as hippocampus, thalamus, and brainstem.

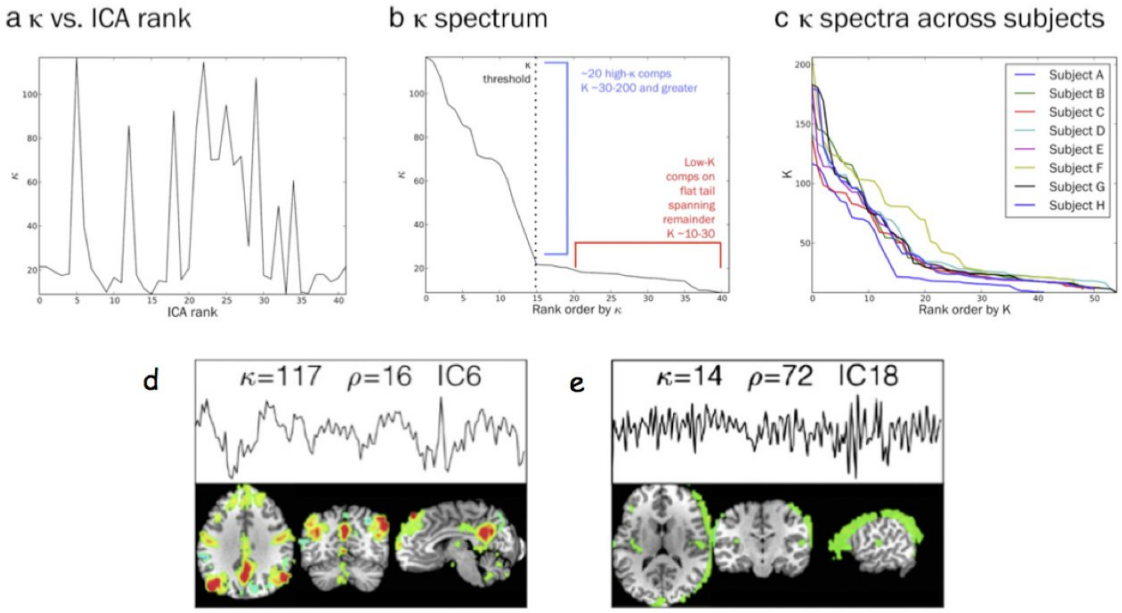

Figure 3: For a representative subject, κ score vs (a) ICA rank (variance explained), and (b) rank by κ (κ spectrum). The κ spectrum, is an L-curve with two distinct regimes: high κ (κ>20) and low κ (κ<20), with low κ components on a linear tail. (c) κ spectra for 8 subjects. (d) Example of high κ and low κ component maps. Each panel shows the time course and thresholded ΔR2* map. Components are annotated with κ-score, ρ-score, and ICA component number. Modified from Ref 1.

A major goal in rs-fMRI is the whole-brain mapping of functional cortical networks down to the scale of voxels from a single dataset. Promising results have been obtained on highly subject-averaged data, however, our ultimate goal is to use the characteristics of these networks as individual biomarkers. Towards this goal, we have developed an approach, based on our me-ICA method that converges on stable and consistent areas with very few averages. The maps that we produce can also predict subject-specific activation patterns associated with activation-based paradigms.

Seven subjects were scanned with 10 min rest runs and with several task-activation runs consisting of the following paradigms: face vs. objects differentiation, sense vs. nonsense sentence differentiation, tongue vs. toe movement. Imaging was performed on a General Electric (GE) 3 Tesla GE Signa HDx MRI scanner with a GE 8-channel receive-only head coil. Anatomical images were acquired using a T1-weighted MPRAGE sequence. Functional images were acquired with a *multi-echo EPI* sequence (TR 2.5 s, flip angle 90, matrix size 64x64, in-plane resolution 3.75 mm, FOV 240 mm, 31 axial slices, slice thickness 4.2 mm with 0.3mm gap, acceleration factor 2). Three echoes were acquired with the shortest possible echo times, TE = 15ms, 39ms, and 63ms. Multi-echo data were analyzed with me-ICA to

reveal BOLD ICA components and remove non-BOLD ICA components. This non-BOLD component removal is fundamentally important in the success of these results as the clustering method was based on the voxel-wise grouping and differentiation based on its unique belonging "signature" to each ICA component.

This exquisitely detailed hierarchical structure is shown in Figure 4. Large modules (from low level clustering, i.e. k=10 clusters) reflect classical neuroanatomy. These modules correspond to regions such as superior, middle, and inferior frontal and temporal gyrii; premotor, motor, sensory, and insular cortices; cingulate and visual areas, etc. Smaller modules (from high level clustering, i.e. k=350 clusters) reflect functionally specific areas. These modules correspond, to regions in the motor cortex such as lateralized hand, foot areas, as well as language areas (Broca's and Wernicke's), visual areas (V1/V2), and many others. The novel implementation of hierarchical clustering approach carried the organization of low level clustering to high level clustering in terms of cluster index ranges. This means that color proximity (i.e. on a color bar) of cortical areas indicates functional relatedness, and color separation indicates functional distinctness. This coloring scheme greatly aids in interpretation of high level clustering, especially in mapping major functional boundaries.

It was found that cluster maps converged as the number of subjects used in group analysis was increased. A stable organization was achieved using only 7 subjects. This indicates a large increase in statistical power of this method over prior methods that do not use me-ICA, ICA-based clustering, and cortical surface-based averaging. Since analysis was conducted on the cortical surface, the rapid convergence also suggests strong conservation of functional organization on the cortical sheet.

This novel implementation of hierarchical clustering was able to create a community structure of brain organization, showing how large modules at low levels of clustering are related to smaller modules at high levels of clustering. We found that the 30-50 BOLD components from individual subject high dimensional me-ICA have localization to specific, finely delineated functional areas. The ability to consistently identify functional areas without functional localization tasks could also be of value as a biomarker for individual differences or disorder characterization.

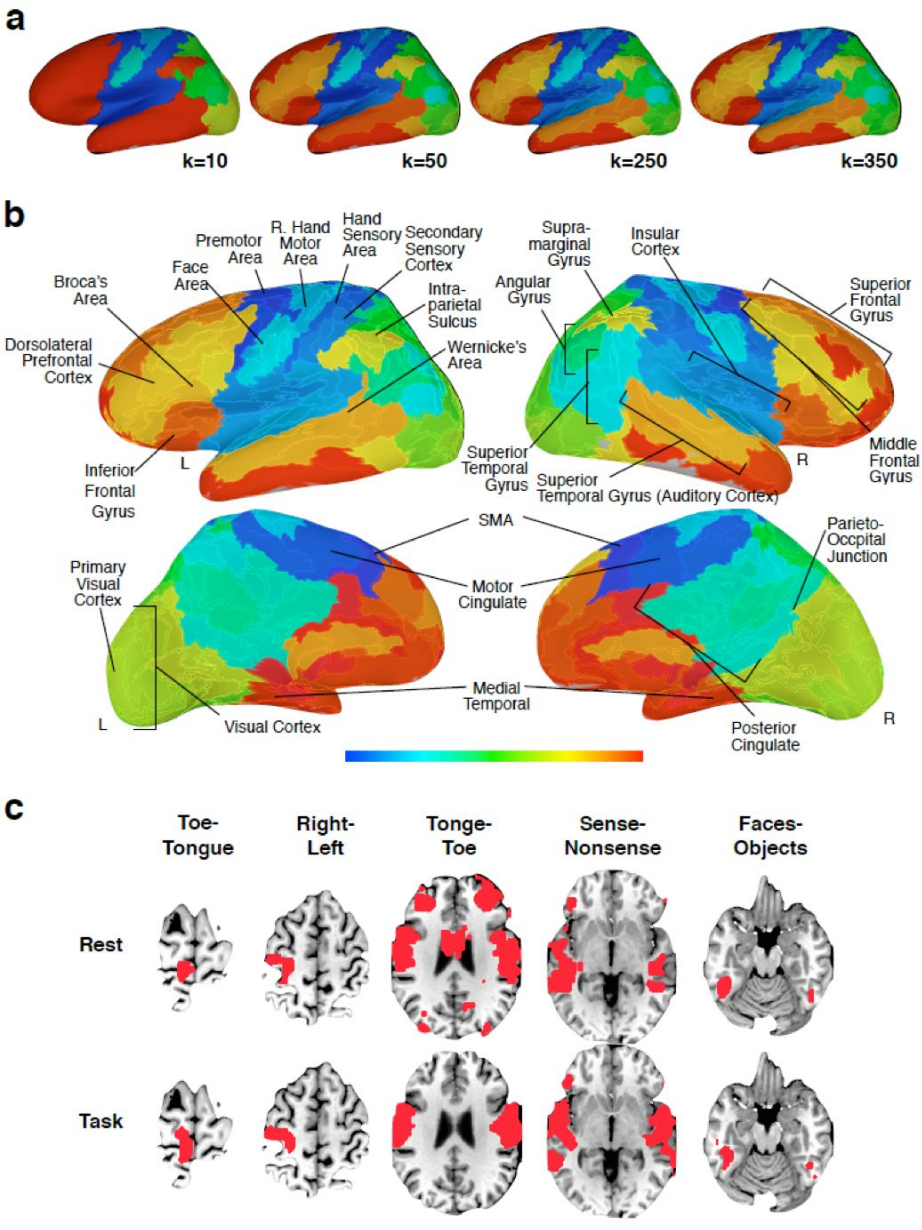

Figure 4: (A) Functional organization (n=7) at increasingly high levels of clustering. Lower cluster levels show large modules. Higher cluster levels show increasing differentiation of functional areas within color ranges determined by lower level clustering. (B) Labeled map of the k=350 functional organization. (C) Comparison of task activated regions and resting state clusters in a single subject, showing a close correspondence.

3. Finding activation in almost the entire brain with relatively simple tasks

In the past, task-based functional MRI (fMRI) studies have supported a localizationist view of brain function, as typically, only a handful of regions have shown significant activation with a task or stimulus. Here, this view is

challenged with evidence that under high contrast to noise conditions, fMRI activations extend well beyond areas of primary relationship to the task - *appearing in over 95% of the brain* for a simple visual stimulation/attention control task[1]. Moreover, we show that response shape varies across regions and can deviate substantially from a typical canonical response. Whole-brain parcellations based on those response shape differences produce distributed clusters that are anatomically and functionally meaningful, symmetrical across hemispheres, and reproducible across subjects. This result has the potential to fundamentally alter how we understand and model brain activation.

Three subjects were scanned on a General Electric 3T MRI scanner. All subjects underwent 100 functional runs, which consisted of five blocks of stimulation. Each block consisted of 20 s flickering checkerboard at 8 Hz + letter/number discrimination task and 40 s of rest. For the letter/number discrimination task, subjects responded using a response box with their right hand. Data were analyzed with AFNI (pre-preprocessing, and statistical analysis), and MATLAB (clustering).

We used three response models for activation: 1. A sustained response model (SUS) consisting of the convolution of a gamma-variate function with a boxcar function that follows the experimental paradigm; 2. An onset+sustained+offset model (OSO) that includes transitory responses at blocks onset/offsets in addition to the sustained response; and 3. An unconstrained model (UNC) consisting of 30 impulse functions spanning the duration of a single on/off cycle (60 s) and therefore setting no a priori constraints on response shape other than agreement with task periodicity.

We found that the extent of activations increased significantly with number of runs (Nruns) inputted to the analysis as well as with a relaxation of the predictive BOLD response shape constraints. Figure 5 shows how activation extent increased with increases in number of runs. For all subjects, significantly active voxels at Nruns = 100 represent on average over 71% of the imaged brain for the SUS model; and over 89% for the other two models at $P_{FDR} < 0.05$. Conversely, for Nruns = 5, which represents a typical number of runs per condition in fMRI experimentation, activated voxels represented ~20% of the imaged volume at PFDR < 0.05 for the SUS analysis and between 35 and 44% for the other two analyses. Active voxels are defined as those where the model accounts for a significant amount of variability in the data (F-stat) at $P_{FDR} < 0.05$ or $P_{Bonf} < 0.05$. Within-subject averaging reduced random noise while keeping non-random signal levels unaffected. Statistically significant signal changes (sometimes less than 0.2%) time-locked with the task could be observed in almost every location of the brain when 100 runs were averaged. Response shape and magnitude varied significantly across regions. Some regions responded in a sustained positive manner for the whole duration of the task epochs (e.g., occipital, insular and left motor cortex), but others responded more prominently during task-switching periods (e.g., occipito-parietal junction). Several regions responded with negative deflections during active epochs (e.g., some parietal locations, right motor cortex). In regions that responded similarly (e.g., sustained), there were differences in onset, offset, and steady-state.

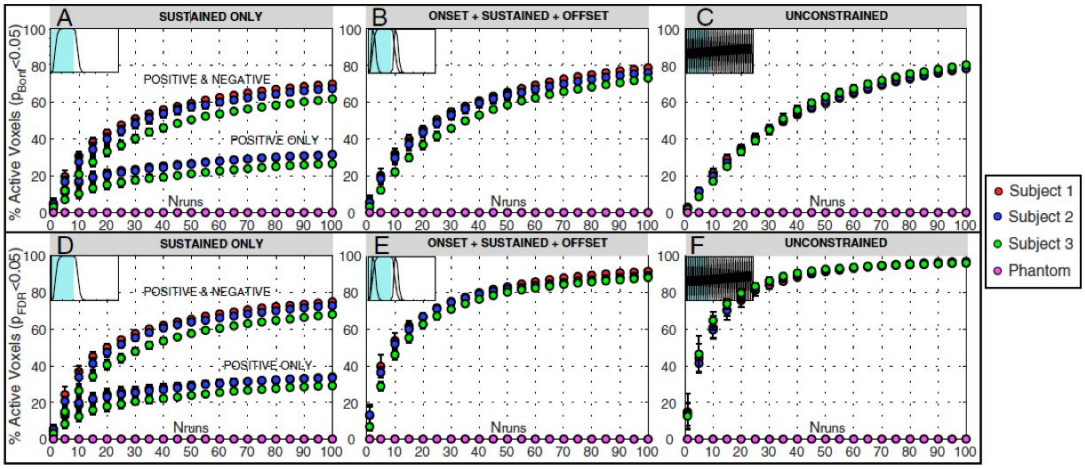

Figure 5: Activation extent results for the three response models. A-C show results for $P_{bonf} < 0.05$ and D-F for $P_{FDR} < 0.05$. The significantly active volume increased with number of runs in all subjects and models. Modified from Ref 2.

Parcellation of the whole brain activation data was accomplished as follows: Voxels with similar response profiles were spatially clustered using both k-means and hierarchical clustering on the voxel-wise BOLD responses calculated using all 100 runs per subject. Parcellations were computed only for cortical and subcortical gray-matter voxels, excluding the cerebellum. Figure 6 shows the k-means decomposition for a representative subject and k=20. The resulting topography is symmetrical across hemispheres, anatomically meaningful, and reproducible across subjects. Hemispheric symmetry is evident in the occipital cortex, superior temporal cortex, anterior insula, hippocampus, and in subcortical structures, such as the thalamus and the putamen.

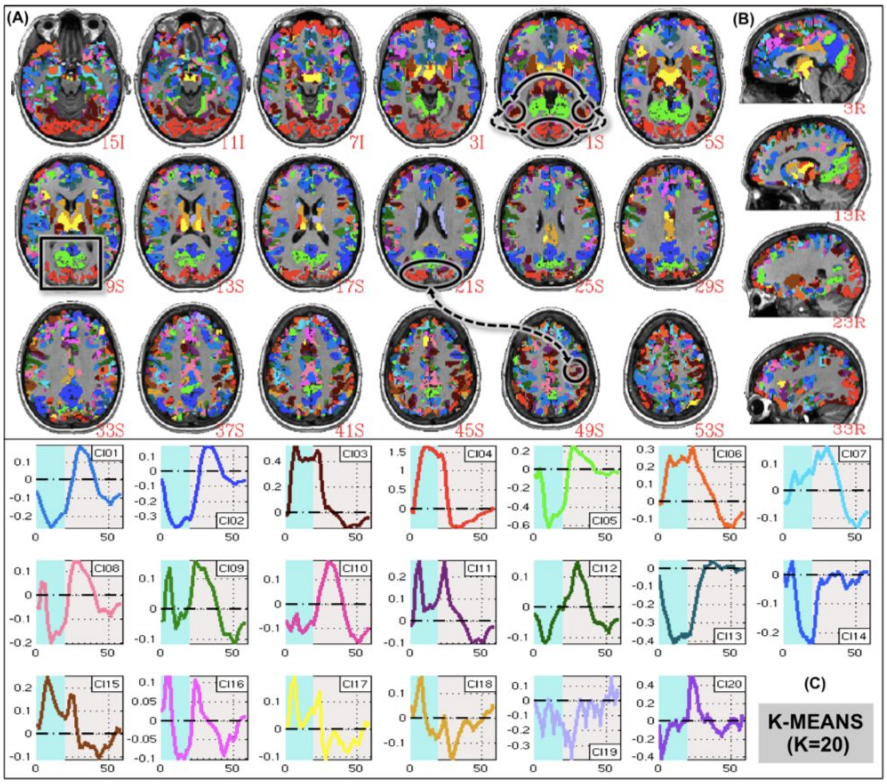

Figure 6: Color-coded (A) axial and (B) sagittal views of k-means clusters (k=20) for subject 3. (C) Color-coded hemodynamic responses for each cluster. Modified from Ref 2.

Primary visual and primary hand motor cortices correspond to different clusters. The visual cortex is segmented into several regions both in the anterior-posterior (A-P) and medial-lateral (M-L) directions. For example, in the M-L direction V1 and V5 are segregated. In the A-P direction, V1 and higher visual processing areas closer to the parieto-occipital junction are also part of different clusters. In most cases, clusters did not appear in the form of a single contiguous agglomeration of voxels but as distributed sets of nodes. Grouping patterns go beyond hemispheric symmetry, and in some cases resemble connectivity patterns similar to those present in resting-state data. For example, the area CL03 resembles a motor control network with nodes in the left primary motor hand cortex, medial supplementary motor cortex, and postero-lateral thalamus. Cluster CL02, with nodes in the bilateral infero-lateral parietal cortex, posterior cingulate, and ventro-medial frontal cortex, resembles the default-mode network. Figure 6 C shows cluster-averaged responses. All clusters display responses time-locked with the experimental paradigm.

These findings highlight the exquisite detail in fMRI signals beyond what is normally examined, and emphasize both the pervasiveness of false negatives, and how the sparseness of fMRI maps is not a result of localized brain function, but a consequence of high noise and overly strict predictive response models. It also shows how inter-regional variability in the hemodynamic response to tasks can be used to functionally parcellate the whole brain in a manner similar to resting-state scans. Differences in functional parcellations between task-state and resting-state may provide good markers for identification of clinical populations. We plan to continue this work at high field, using me-EPI, and with higher

sensitivity RF coils in hope that we can achieve a temporal signal to noise such that the necessary scan duration to obtain these results is more realistic.

4. Decoding 100 ms timing differences between ocular dominance columns

We investigated the decoding of millisecond-order timing information in columnar-level neural activation from the blood oxygen level dependent (BOLD) signal in human functional magnetic resonance imaging (fMRI) . In the experiment, ocular dominance columns were activated by monocular visual stimulation with 500- or 100- ms onset differences. We observed that the event-related hemodynamic response (HDR) in the human visual cortex was sensitive to the subtle onset differences, but these activation onset differences *were not reflected in hemodynamic latencies*.

We examined decoding success based on various characteristics of HDR including response amplitude, time to peak, full width at half-maximum response. This approach is shown schematically in Figure 7.

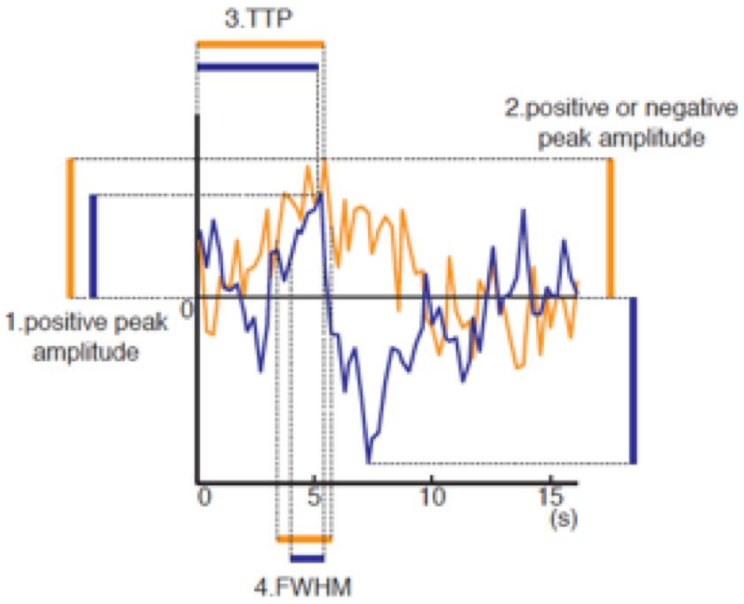

Figure 7: Responses of event-related HDR for two conditions. The explored characteristics are: 1. Positive peak amplitude 2. Pos or neg peak amplitude. 3. Time-to-peak (TTP). 4. Full-width of half maximum (FWHM). Modified from Ref 3.

We found that the timing difference was most accurately differentiated by comparison of the signal change amplitude (positive or negative). When using either positive or negative peak amplitude of the deconvolved HDR, high decoding performance could be achieved for both the 500ms and the 100ms onset differences. Figure 8 shows these results.

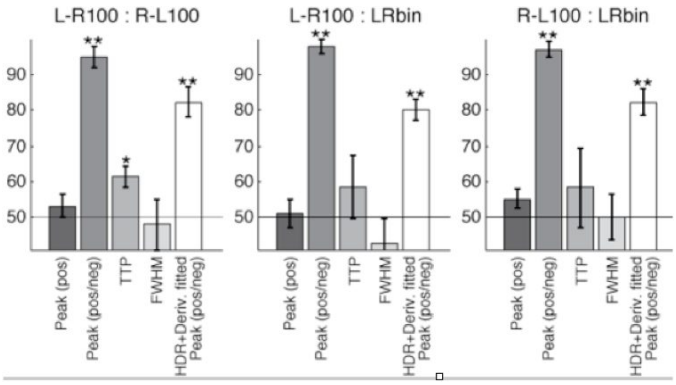

Figure 8: Decoding accuracies and their standard errors across six subjects for the extracted features of hemodynamic response shape. Asterisk (*) and double asterisk (**) indicate that the accuracy is higher than the chance level (50 %) by $p < 0.05$ and $p < 0.01$ respectively with one-sample t test. Modified from Ref 3.

The high accuracy even for the 100ms difference, given that the signal was sampled at a TR of 250 ms and 2x2x3-mm voxels, indicates that multivariate classification analysis could decode higher resolution information than actual sampling rate. Also, both down-sampling and smoothing did not affect the decoding accuracies, suggesting that high-frequency neuronal activation information is represented as a complex spatiotemporal response pattern of low frequency BOLD signal. We plan to more fully explore the temporal and spatial "signatures" of extremely brief and small activations to determine how temporal delay is transferred to amplitude with regard to the BOLD response. It's clear from these studies that, at least at this spatial scale, relatively latency distribution is extremely complex to the point of being useless, and neuronal latency is better reflected in total magnitude differences. We plan to investigate this further.

5. Conclusions

This manuscript summarizes some of the studies from our group that were described in the Keynote address. These findings were chosen to illustrate that more subtle and often surprising information can be extracted from the fMRI time series when cleaning up the time series, increasing the signal to noise ratio, or relaxing the models that we have come to assume are the only ones that reflect neuronal activity. This is ongoing work that opens up not only new and rich avenues of investigation but new hope for the application of fMRI for a wider range of clinical and other applications.

REFERENCES

[1] 1. P. Kundu, S. J. Inati, J. W. Evans, W. M. Luh and P. A. Bandettini, "Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI," NeuroImage 60(3), 1759-1770 (2012)

[2] 2. J. Gonzalez-Castillo, Z. S. Saad, D. A. Handwerker, S. J. Inati, N. Brenowitz and P. A. Bandettini, "Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis," Proceedings of the National Academy of Sciences 109(14), 5487-5492 (2012)

[3] 3. M. Misaki, W.-M. Luh and P. A. Bandettini, "Accurate decoding of sub-TR timing differences in stimulations of sub-voxel regions from multi-voxel response patterns.," NeuroImage 66(623-633 (2013)