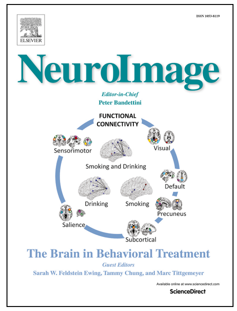
Accepted Manuscript



Discovering dynamic brain networks from big data in rest and task

Diego Vidaurre, Romesh Abeysuriya, Robert Becker, Andrew J. Quinn, Fidel Alfaro-Almagro, Stephen M. Smith, Mark W. Woolrich

PII: S1053-8119(17)30548-7

DOI: 10.1016/j.neuroimage.2017.06.077

Reference: YNIMG 14159

To appear in: *NeuroImage*

Received Date: 21 March 2017

Revised Date: 12 June 2017

Accepted Date: 28 June 2017

Please cite this article as: Vidaurre, D., Abeysuriya, R., Becker, R., Quinn, A.J., Alfaro-Almagro, F., Smith, S.M., Woolrich, M.W., Discovering dynamic brain networks from big data in rest and task, *NeuroImage* (2017), doi: 10.1016/j.neuroimage.2017.06.077.

Title: Discovering dynamic brain networks from big data in rest and task

Authors: Diego Vidaurre[1], Romesh Abeysuriya[1], Robert Becker[1], Andrew J. Quinn[1], Fidel Alfaro-Almagro[2], Stephen M. Smith[1] and Mark W. Woolrich[1,2]

[1] Oxford Centre for Human Brain Activity (OHBA), University of Oxford, UK
[2] Oxford University Centre for Functional MRI of the Brain (FMRIB), University of Oxford, UK

Abstract: Brain activity is a dynamic combination of the responses to sensory inputs and its own spontaneous processing. Consequently, such brain activity is continuously changing whether or not one is focusing on an externally imposed task. Previously, we have introduced an analysis method that allows us, using Hidden Markov Models (HMM), to model task or rest brain activity as a dynamic sequence of distinct brain networks, overcoming many of the limitations posed by sliding window approaches. Here, we present an advance that enables the HMM to handle very large amounts of data, making possible the inference of very reproducible and interpretable dynamic brain networks in a range of different datasets, including task, rest, MEG and fMRI, with potentially thousands of subjects. We anticipate that the generation of large and publicly available datasets from initiatives such as the Human Connectome Project and UK Biobank, in combination with computational methods that can work at this scale, will bring a breakthrough in our understanding of brain function in both health and disease.

Introduction

Understanding the nature of temporal dynamics of brain activity at a range of temporal and spatial scales is an important challenge in neuroscience. When studying task data, the aim is to discover the neural underpinnings and brain mechanisms elicited by the task, for which one relates the time course of the measured data to behaviour as comprehensively as possible. That is to say, we are interested in the dynamics evoked by the task. In this case, many repetitions of the same task are typically considered in the hope of characterising and interpreting the differences with respect to some baseline condition. Presumably, the brain adapts to the task at different time scales and in an online fashion, and we would like to capture these changes at as high a temporal resolution as the imaging modality will allow. When studying rest data, where the brain is not engaged in a predefined task, the brain will still process information dynamically, adapting its activity to the current perception of the environment combined with the products of its own spontaneous activity. In this case, then, we are interested in characterising the spontaneous dynamics. Either case, being able to characterise the temporal trajectories of whole-brain network activity at different time scales is, considering the complex and deeply integrated nature of the brain, crucial to understand the ultimate underpinnings of cognition [1,2].

The most common analysis technique used for describing brain network dynamics in both task and rest is the use of sliding windows [3,4]. The sliding windows approach (and methods built upon it) suffers from a number of limitations [5] that can undermine any conclusions. In particular, they need a pre-specification of the time scale at which the neural processes of interest occur, i.e. the temporal width of the window. This choice is crucial and is a trade-off between two conflicting criteria: too long a window will miss fast dynamics, whereas too short a window will have insufficient data to provide a reliable network