NeuroImage 66 (2013) 543–552

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Bayesian inference of structural brain networks

Max Hinne [a,b,*], Tom Heskes [a], Christian F. Beckmann [b], Marcel A.J. van Gerven [b]

[a] Radboud University Nijmegen, Institute for Computing and Information Sciences, Nijmegen, The Netherlands
[b] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Accepted 28 September 2012
Available online 4 October 2012

Keywords:
Structural connectivity
Probabilistic tractography
Hierarchical Bayesian model

ABSTRACT

Structural brain networks are used to model white-matter connectivity between spatially segregated brain regions. The presence, location and orientation of these white matter tracts can be derived using diffusion-weighted magnetic resonance imaging in combination with probabilistic tractography. Unfortunately, as of yet, none of the existing approaches provide an undisputed way of inferring brain networks from the streamline distributions which tractography produces. State-of-the-art methods rely on an arbitrary threshold or, alternatively, yield weighted results that are difficult to interpret. In this paper, we provide a generative model that explicitly describes how structural brain networks lead to observed streamline distributions. This allows us to draw principled conclusions about brain networks, which we validate using simultaneously acquired resting-state functional MRI data. Inference may be further informed by means of a prior which combines connectivity estimates from multiple subjects. Based on this prior, we obtain networks that significantly improve on the conventional approach.

Introduction

Human behavior ultimately arises through the interactions between multiple brain regions that together form networks that can be characterized in terms of structural, functional and effective connectivity (Penny et al., 2006). Structural connectivity presupposes the existence of white-matter tracts that connect spatially segregated brain regions which constrain the functional and effective connectivity between these regions. Hence, structural connectivity provides the scaffolding that is required to shape neuronal dynamics. Changes in structural brain networks have been related to various neurological disorders. For this reason, optimal inference of structural brain networks is of major importance in clinical neuroscience (Catani, 2007). Inference of these networks entails two steps. First is the estimation of the white matter tracts. The second step consists of obtaining the network that captures which regions are connected, based on the earlier identified fiber tracts. In this paper, we focus on the latter step.

For the first step, we use diffusion-weighted imaging (DWI), which is a prominent way to estimate structural connectivity of whole-brain networks in vivo. It is a variant of magnetic resonance imaging (MRI) which measures the restricted diffusion of water molecules, thereby providing an indirect measure of the presence and orientation of white-matter tracts. By following the principal diffusion direction in individual voxels, streamlines can be drawn that represent the structure of fiber bundles, connecting separate regions of gray matter. This process is known as deterministic tractography (Chung et al., 2010; Conturo et al., 1999; Shu et al., 2011). Alternatively, fibers may be estimated using probabilistic tractography (Behrens et al., 2003, 2007; Friman et al., 2006; Jbabdi et al., 2007). This comprises a model for the principal diffusion direction that is then used to sample distributions of streamlines. Ultimately, the procedure results in a measure of uncertainty about where a hypothesized connection will terminate. A benefit of the probabilistic approach is that it explicitly takes uncertainty in the streamlining process into account.

Apart from studies focusing on particular tracts, much research has been devoted to the derivation of macroscopic connectivity properties, that is, whole-brain structural connectivity. Several approaches have been suggested to extract whole-brain networks from probabilistic tractography results (Gong et al., 2009; Hagmann et al., 2007; Robinson et al., 2008). Unfortunately, inference of whole-brain networks from probabilistic tractography estimates remains somewhat ad hoc. Typically the underlying brain network is derived by thresholding the streamline distribution such that counts above or below threshold are taken to reflect the presence or absence of tracts, respectively. This approach is easy to implement but it has a number of issues. First, the threshold is arbitrarily chosen to have a particular value. In a substantial part of the literature, the threshold that is used to transform the streamline distribution into a network is actually set to zero (Chung et al., 2011; Hagmann et al., 2007, 2008; Vaessen et al., 2010; Zalesky et al., 2010). However, probabilistic streamlining

* Corresponding author at: Radboud University Nijmegen, Faculty of Science, Institute for Computing and Information Sciences, Postbus 9010, 6500 GL Nijmegen, The Netherlands. Fax: +31 24 365 27 28.
E-mail address: mhinne@cs.ru.nl (M. Hinne).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2012.09.068

depends on the arbitrary number of samples that are drawn per voxel. This implies that, as more samples are drawn, more brain regions are likely to eventually become connected given a threshold at zero. Alternatively, the number of streamlines can be interpreted as connection weight (Bassett et al., 2011; Robinson et al., 2010; Zalesky et al., 2010), or a relative threshold can be applied (Kaden et al., 2007). This way, the relative differences between connections remain respected. Unfortunately, the connection weights do not have a straightforward (probabilistic) interpretation. Simply normalizing these weights does not yield a true notion of connection probability. At most, it can be regarded as the conditional probability that a streamline ends in a particular voxel given the starting point of the streamline. In the case of a streamline distribution with, say, half of the streamlines starting at node A ending in node B, and the other half ending in node C, normalized streamline counts cannot distinguish between one edge with an uncertain end point, or two edges with definite end points. Finally, several graph-theoretical measures such as characteristic path length and clustering coefficient are ill-defined for non-binary networks.

In general, it is problematic to use thresholding since it ignores the relative differences between streamline counts. Intuitively, one would expect that if, say, 90% of the streamlines connect from voxel A to voxel B, and 10% connect voxel A to voxel C, then at the least the former has a higher probability of having a corresponding edge in the network than the latter, but both edges are possible as well. This is related to the *burstiness* phenomenon of words in document retrieval, where the occurrence of a rare word in a document makes its repeated occurrence more likely (Xu and Akella, 2010). Summarizing, the issue with thresholding approaches is that they consider each tract in isolation. This ignores the information that can be gained from the possible symmetry in streamline counts, as well as from the relative differences within a streamline distribution.

Another important observation is that the mentioned approaches do not easily support the integration of probabilistic streamlining data with other sources of information. Data is often not collected in isolation but rather acquired for multiple subjects, potentially using a multitude of imaging techniques. Multi-modal data fusion is needed in order to provide a coherent picture of brain function (Groves et al., 2011; Horwitz and Poeppel, 2002). The integration of multi-subject data is required for group-level inference, where the interest is in estimating a network that characterizes a particular population, for example, when comparing patients with controls in a clinical setting (Simpson et al., 2011).

In the following, we provide a Bayesian framework for the inference of whole-brain networks from streamline distributions. In our approach, we consider the distribution of (binary) networks that are supported by our data, instead of generating a single network based on an arbitrary threshold. Our approach relies on defining a generative model for whole-brain networks which extends recent work on network inference in systems biology (Mukherjee and Speed, 2008) and consists of two ingredients. First, a network prior is defined in terms of the classical Erdős–Rényi model (Erdős and Rényi, 1960). This prior is later extended to handle multi-subject data, capturing the notion that different subjects' brains tend to be similar. Second, we propose a forward model based on a Dirichlet compound multinomial distribution which views the streamline distributions produced by probabilistic tractography as noisy data, thus completing the generative model.

In order to validate our Bayesian framework we make use of the often reported observation that resting-state functional connectivity reflects structural connectivity (Damoiseaux and Greicius, 2009; Greicius et al., 2009; Honey et al., 2009; Koch et al., 2002; Lv et al., 2010; Park et al., 2008; Skudlarski et al., 2008). We show that structural networks that derive from our generative model informed by the connectivity for other subjects provide a better fit to the (in)dependencies in resting-state functional MRI (rs-fMRI) data than the standard thresholding approach.

Material and methods

Data acquisition

Twenty healthy volunteers were scanned after giving informed written consent in accordance with the guidelines of the local ethics committee. A T1 structural scan, resting-state functional data and diffusion-weighted images were obtained using a Siemens Magnetom Trio 3 T system at the Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen, The Netherlands. The rs-fMRI data were acquired at 3 Tesla using a multi echo-echo planar imaging (ME-EPI) sequence (voxel size 3.5 mm isotropic, matrix size 64×64, TR = 2000 ms, TEs = 6.9, 16.2, 25, 35 and 45 ms, 39 slices, GRAPPA factor 3, 6/8 partial Fourier). A total of 1030 volumes were obtained. An optimized acquisition order described by Cook et al. (2006) was used in the DWI protocol (voxel size 2.0 mm isotropic, matrix size 110×110, TR = 13,000 ms, TE = 101 ms, 70 slices, 256 directions at $b = 1500$ s/mm^2 and 24 directions at $b = 0$).

Preprocessing of resting-state data

The multi-echo images obtained using the rs-fMRI acquisition protocol were combined using a custom Matlab script (MATLAB 7.7, The MathWorks Inc., Natick, MA, USA) which implements the procedure described by Poser et al. (2006) and also incorporates motion correction using functions from the SPM5 software package (Wellcome Department of Imaging Neuroscience, University College London, UK). Of the 1030 combined volumes, the first six were discarded to allow the system to reach a steady state. Tools from the Oxford FMRIB Software Library (FSL, FMRIB, Oxford, UK) were used for further processing. Brain extraction was performed using FSL BET (Smith, 2002). For each subject, probabilistic brain tissue maps were obtained using FSL FAST (Zhang et al., 2001). A zero-lag 6th order Butterworth bandpass filter was applied to the functional data to retain only frequencies between 0.01 and 0.08 Hz. After preprocessing, the fMRI data were parcellated according to the Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002). Regions without voxels with gray-matter probability ≥ 0.5 were discarded. This resulted in an average region count of 115 ± 0.1. For these regions the functional data was summed and then standardized to have zero mean and unit standard deviation. The resulting data were used to compute the empirical covariance matrix $\hat{\Sigma}$. Example covariance matrices are shown in Fig. 1a.

Preprocessing of diffusion imaging data

The preprocessing steps for the diffusion data were conducted using FSL FDT (Behrens et al., 2003) and consisted of correction for eddy currents and estimation of the diffusion parameters. Raw color-coded fractional anisotropy maps are shown in Fig. 1b. To obtain a measure of white-matter connectivity, we used FDT Probtrackx 2.0 (Behrens et al., 2003, 2007). As seed voxels for tractography we used those voxels that live on the boundary between white matter and gray matter. For each of these voxels 5000 streamlines were drawn, with a maximum length of 2000 steps. The streamlines were restricted by the fractional anisotropy to prevent them from wandering around in gray matter. Streamlines in which a sharp angle ($>80°$) occurred or that had a length less than 2 mm were discarded. The output thus obtained is a matrix $\mathbf{N}$ with n_{ij} the number of streamlines drawn from voxel i to voxel j. To transform this into the parcellated scheme as dictated by the AAL atlas, the streamlines were summed over all voxels per region, resulting in an aggregated connectivity matrix which ranges over regions instead of voxels. Regions that had been removed after preprocessing the fMRI data were removed from the aggregated connectivity matrix as well.