NeuroImage 142 (2016) 407–420

Contents lists available at ScienceDirect

NeuroImage

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

Connectome sensitivity or specificity: which is more important?

Andrew Zalesky [a,*], Alex Fornito [b], Luca Cocchi [c], Leonardo L. Gollo [c], Martijn P. van den Heuvel [d], Michael Breakspear [c,e]

[a] Melbourne Neuropsychiatry Centre and Melbourne School of Engineering, The University of Melbourne, Australia
[b] Brain and Mental Health Laboratory, Monash Institute of Cognitive and Clinical Neuroscience, School of Psychological Sciences and Monash Biomedical Imaging, Monash University, Melbourne, Victoria, Australia
[c] QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia
[d] Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands
[e] Metro North Mental Health Service, The Royal Brisbane and Women's Hospital, Herston, Queensland, Australia

ARTICLE INFO

Article history:
Received 2 April 2016
Accepted 18 June 2016
Available online 28 June 2016

Keywords:
Connectome
Complex networks
False positives
False negatives
Sensitivity
Specificity
Clustering coefficient
Modularity
Network efficiency
Tractography

ABSTRACT

Connectomes with high sensitivity and high specificity are unattainable with current axonal fiber reconstruction methods, particularly at the macro-scale afforded by magnetic resonance imaging. Tensor-guided deterministic tractography yields sparse connectomes that are incomplete and contain false negatives (FNs), whereas probabilistic methods steered by crossing-fiber models yield dense connectomes, often with low specificity due to false positives (FPs). Densely reconstructed probabilistic connectomes are typically thresholded to improve specificity at the cost of a reduction in sensitivity. What is the optimal tradeoff between connectome sensitivity and specificity? We show empirically and theoretically that specificity is paramount. Our evaluations of the impact of FPs and FNs on empirical connectomes indicate that specificity is at least twice as important as sensitivity when estimating key properties of brain networks, including topological measures of network clustering, network efficiency and network modularity. Our asymptotic analysis of small-world networks with idealized modular structure reveals that as the number of nodes grows, specificity becomes exactly twice as important as sensitivity to the estimation of the clustering coefficient. For the estimation of network efficiency, the relative importance of specificity grows linearly with the number of nodes. The greater importance of specificity is due to FPs occurring more prevalently between network modules rather than within them. These spurious inter-modular connections have a dramatic impact on network topology. We argue that efforts to maximize the sensitivity of connectome reconstruction should be realigned with the need to map brain networks with high specificity.

Introduction

Methods for mapping connectomes are imperfect. Structural connections can be erroneously inferred between pairs of nodes that are truly disconnected, giving rise to spurious connections known as *false positives* (FPs) and reducing the *specificity* of connectome reconstructions. Conversely, genuine connections can be overlooked, resulting in *false negatives* (FNs) and reducing connectome *sensitivity*. Despite current state of the art, it remains challenging to reconstruct micro-, meso and macro-scale connectomes that display both high sensitivity and high specificity (Azadbakht et al., 2015; Bastiani et al., 2012; Calabrese et al., 2015; Knosche et al., 2015; Reveley et al., 2015; Thomas et al., 2014).

This study primarily focuses on the sensitivity and specificity of *macro-scale* connectomes, which are most often mapped with automated fiber tracking methods (tractography) performed on diffusion-weighted magnetic resonance imaging data (Hagmann et al., 2008; Sporns et al., 2005). A considerable variety of tractography algorithms has been developed to reconstruct axonal fiber bundles and thereby infer where connections should be placed in network models of the brain. Typically, millions of streamlines that follow the trajectories of all major neural white matter pathways are initiated throughout the brain and the number of streamlines interconnecting pairs of brain regions comprising a predefined parcellation atlas are enumerated to yield a connectivity matrix of streamline counts (Li et al., 2012). Deterministic tractography algorithms (Conturo et al., 1999; Mori et al., 1999) guided by the diffusion tensor are criticized for their failure to resolve crossing-fiber geometries (Alexander et al., 2007). This failure predominantly results in FNs, but can also yield FPs as well, depending on the specific method, data quality and parcellation resolution (Zalesky et al., 2010a). Connectome sensitivity can be substantially improved with probabilistic tractography algorithms (Behrens et al., 2003; Koch

* Corresponding author.
E-mail address: azalesky@unimelb.edu.au (A. Zalesky).

http://dx.doi.org/10.1016/j.neuroimage.2016.06.035
1053-8119/

Table 1
Connectome connection density variation across species and reconstruction methodologies.

Species	Investigators	Method	Density	Nodes
Human	E.g. Roberts et al. (2016)	Probabilistic/X-fiber	>50%	~50–5000
Human	E.g. Zalesky et al. (2010a)	Deterministic/tensor	1–40%[a]	~50–5000
Rat	Bota et al. (2012)	Axonal tracer	45%	73
Caenorhabditis elegans	Varshney et al. (2011), White et al. (1986)	Electron microscopy	3.8%	279
Macaque	Markov et al. (2014)	Reterograde axonal tracer	66%	29
Mouse	Oh et al. (2014)	Anterograde axonal tracer & model	5–20%[b]	213
Fruit fly	Shih et al. (2015)	Genetic labeling and light microscopy	82%	49
Mouse	Zingg et al. (2014)	Reterograde & anterograde tracers	~50%[c]	49

[a] Density depends on data quality, number of nodes and tractography algorithm.
[b] Density depends on *p*-value threshold.
[c] Density not reported. Estimate based on connectivity matrix.

et al., 2002) that are combined with sophisticated crossing-fiber models (Behrens et al., 2007; Tournier et al., 2008), but probabilistic methods can produce FPs.

These issues are most clearly borne out when considering the discrepancy in connection density between tractography methods. When reconstructed with tensor-guided deterministic tractography, the human connectome typically has a connection density ranging between 1% and 40% (e.g. Van den Heuvel et al., 2012; Zalesky et al., 2010a), whereas most probabilistic methods that model crossing fibers yield densities exceeding 50–60% and can even be as high as 99–100% (e.g. Roberts et al., 2016). In other words, probabilistic streamlines can be found between more than half of all pairs of brain regions. How can estimates of such a basic connectome property differ so drastically between these methods? Is it that tensor-based methods yield many FNs and are probabilistic crossing-fiber methods confounded by FPs (Thomas et al., 2014)?

One way to reconcile this discrepancy is to adopt a Bayesian view and assume probabilistic tractography provides an estimate of the likelihood of a connection. Whereas a single *deterministic* streamline might be considered adequate to indicate the presence of a connection, a single *probabilistic* streamline is unlikely to provide sufficient evidence to make such inference and might therefore be thresholded away when forming a binary network. However, the difficulty with Bayesian inference is that streamline counts and other tractography outputs do not differentiate between connection *probabilities* and connection *strengths* (Jones, 2010; Kaden et al., 2007). Does a high streamline count indicate a strong connection comprising many axonal projections, or a highly probable yet weak connection comprising few axons (Jones et al., 2013)? The difficulty in divorcing connection probability from connection strength challenges simple applications of Bayesian inference.

Despite these concerns, it is common to assume a monotonic relationship between connection probability and streamline count. This assumption enables the use of thresholding methods to eliminate likely FPs from dense connectomes reconstructed with probabilistic tractography. Thresholding involves progressively eliminating connections with the lowest streamline count until a desired connection density is attained (Fornito et al., 2013, 2016; van Wijk et al., 2010). While eliminating connections with a low streamline count can improve connectome specificity, not all eliminated connections are necessarily FPs, and thus any gain in specificity is inevitably traded for a loss in sensitivity (Azadbakht et al., 2015; Knosche et al., 2015; Thomas et al., 2014). Therefore, while thresholding methods *cannot* yield connectomes displaying both high sensitivity and high specificity, they may allow a tradeoff to be achieved between these two measures, assuming a monotonic relation between streamline counts and connection probabilities. Lenient thresholds produce dense connectomes with high sensitivity, whereas stringent thresholds yield sparse connectomes with high specificity.

Thresholding is however in many senses an unsettling approach; all the finesse of a sophisticated crossing-fiber model and probabilistic tractography is largely overridden by a simple and arbitrary threshold, which ultimately determines the most fundamental property of a connectome—its connection density. In this way, the burden of connectome reconstruction is precariously balanced on a single threshold, with less faith placed in the accuracy of the reconstruction process itself.

An important choice must therefore be made between sensitivity and specificity. Should the dense and highly sensitive reconstructions yielded by cutting-edge crossing-fiber models and probabilistic tractography be favored over the sparse and specific reconstructions that are characteristic of tensor-guided deterministic methods? Moreover, should thresholding be used to strike a balance between these two extremes of the sensitivity-specificity continuum? And if so, where along this continuum is the optimal tradeoff between sensitivity and specificity? These questions can be addressed by quantifying the relative detriment of FPs versus FNs. Are FPs worse than FNs to connectome accuracy, and if so, by how much?

The answer to these questions depends on the application at hand. For example, sensitivity is vital in neurosurgical planning, in order to minimize the risk of injury to axonal connections that would result in postoperative deficits. Statistical connectomics is another important application where this question manifests. When statistically comparing connectomes between groups (Griffa et al., 2013), FPs lead to a linear increase in the number of multiple comparisons, whereas FNs can conceal genuine group difference.

The analysis of connectome topology with the use of graph theory is the focus of this study and represents an important application (Bullmore and Sporns, 2009) for which little is known about the impact of connectome sensitivity and specificity. Is it FPs or FNs that lead to poorer estimation of the topological properties of a complex network, such as its efficiency, modularity and small-world organization? Addressing this question is crucial to determine the most appropriate connectome reconstruction methodology for maximizing the accuracy of graph theoretical analyses of brain networks.

It is trivial to see that FPs and FNs are equally detrimental to the measurement of some network properties. An example is the average nodal degree, which for a binary, undirected network is given by $\sum_i d_i/N$, where N is the total number of nodes and d_i is the number of connections incident to the ith node (Rubinov and Sporns, 2010). It can be seen that each FP *increases* the average nodal degree by $2/N$, since the degree of exactly two distinct nodes is increased by unity with the addition of a new connection, whereas each FN *decreases* the average nodal degree by the same amount. FPs and FNs are therefore equally detrimental to the average nodal degree because they introduce identical amounts of absolute error. As we will demonstrate here, this parity between sensitivity and specificity does not hold for most measures of complex network organization. The purpose of this study is to determine whether sensitivity or specificity is more important in these cases.

Connectome sensitivity and specificity is also an important concern at the micro- and meso-scale. While tract tracing techniques (Zaborszky et al., 2006) are often considered a gold standard, they are not without problems. FNs can arise due to distance dependencies of some tracers (Ercsey-Ravasz et al., 2013; Markov et al., 2013). FPs can arise due to