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

tracer uptake by axons traversing a tracer-injected site without making any synapses at that site, and due to the spread of the tracer substance around the injection site (Dyrby et al., 2007; Knosche et al., 2015). These effects can potentially explain some of the substantial variation in connection density that is evident between macro-scale connectomes mapped using tract tracing, ranging as high as 66% for the Markov-Kennedey macaque network (Markov et al., 2014) to densities well below 50% for mouse connectome(Oh et al., 2014). Also see Table 1.

The aim of this paper is to quantify the impact of FPs and FNs on the estimation of complex network properties of binary connectomes mapped at various scales. We systematically increase the number of FPs and FNs in micro-scale (worm), meso-scale (mouse) and macro-scale (human) connectomes to quantify the extent to which the estimation of three commonly used descriptors of connectome topology—network clustering, network efficiency and network modularity—are altered with the addition of each new FP and FN. We only consider binary networks here, in which case the goal of connectome reconstruction is to infer the absence or presence of connections. We first empirically quantify the relative impact of FPs to FNs on the ability to accurately estimate complex network properties in connectomes mapped at various scales. We then undertake an asymptotic analysis on small-world networks with idealized modular structure to verify our empirical findings analytically and to understand why FPs and FNs differ in the extent to which they affect graph theoretical analyses of connectomes.

Methods

We analyzed the impact of false positive and negative connections on macro-scale (human), meso-scale (mouse) and micro-sale (worm) anatomical brain networks. Erdös-Rényi (random) networks comprising 100 nodes with a connection density of 20% were also evaluated to provide a point of comparison. All networks were modeled as binary graphs. The mouse and worm networks were directed, whereas the human and random networks were undirected. This distinction is due to the inability to infer fiber directionality with diffusion tractography. Fig. 1 shows connectivity matrices for the three species (van den Heuvel et al., 2016).

Human connectome

The human network comprised 116 nodes defined by the automated anatomical labeling atlas (Tzourio-Mazoyer et al., 2002). The network was constructed as follows: Connectivity matrices were individually mapped for 100 healthy individuals using automated whole-brain fiber tracking and diffusion-weighted imaging (3 Tesla, 64 gradient-weighted volumes, 2.4 mm isotropic voxels). White matter fiber orientations were estimated with constrained spherical deconvolution (Tournier et al., 2008). One million streamlines were initiated throughout white matter (Li et al., 2012) and propagated probabilistically using a step length of 0.25 mm (Tournier et al., 2012). Each streamline was assigned to a unique pair of regions based on the location of its two endpoints. Therefore, each element of an individual's connectivity matrix enumerated the total number of streamlines assigned to the corresponding pair of regions, known as the streamline count (Zalesky et al., 2010a). Streamlines intersecting fewer than two regions were ignored. A single binary network representing a consensus across all 100 individuals was constructed such that connections were only included between pairs of regions that were interconnected by a streamline in at least 50% of the individuals (de Reus and van den Heuvel, 2013), resulting in a connection density of 22%. Alternative consensus-based thresholding methods may be able to better preserve distance effects between individuals and the consensus matrix (Mišić et al., 2015). Physical connection lengths were approximated as the Euclidean (straight line) distance between regional center of masses (Betzel et al., 2016).

Mouse connectome

The mouse network was mapped by Oh and colleagues using viral tract tracing techniques and a computational model (Oh et al., 2014). Axonal projections were mapped by injecting an anterograde viral tracer into spatially distinct brain regions for each of 469 wild-type mice. The approximate size and extent of tracer infection was determined using two-photon tomography and localized to 295 non-overlapping target regions. This yielded axonal projections from 469 injection sites to 295 target regions. A computational model was used to transform this injection-to-target matrix to a target-to-target connectivity matrix that listed the mutual connections and corresponding *p*-values among 213 regions. This connectivity matrix was rendered binary by setting connections with a *p*-value below 0.05 to one, with all other connections set to zero, thereby yielding a directed network of 213 nodes with a connection density of 6.8% representing intra- and inter-hemispheric connections.

Worm connectome

The cellular-scale nervous system of the nematode *C. elegans* was mapped by White and colleagues using serial sectioning and electron

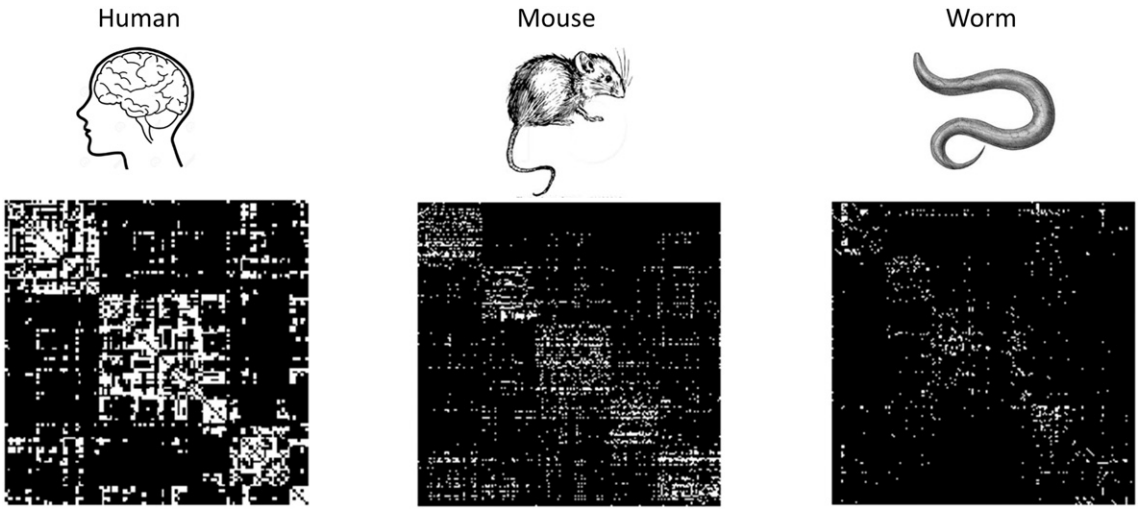

Fig. 1. Binary connectivity matrices for human, mouse and worm connectomes. Rows/columns are ordered such that network modules appear as blocks along the matrix diagonal. White cells denote the presence of a connection. The mouse and worm networks are directed, whereas the human network is undirected.

microscopy (White et al., 1986). The network comprises 279 nodes that correspond with individual neurons and each connection denotes the presence of a synaptic contact (gap junctions and chemical synapses). Inclusion of recently identified connections associated with ventral cord and other neurons yielded a connection density of 3.8% (Varshney et al., 2011).

Evaluation methodology

These empirical networks served as reference graphs to which FP and FN connections were systematically introduced to generate *perturbed networks* that modeled inaccuracies in connectome reconstruction. To this end, new connections were incrementally added to generate FPs and existing connections were incrementally removed to generate FNs, thereby affording precise and systematic control of the FP and FN rates. The extent to which each new FP and FN altered network properties—clustering, efficiency and modularity—was then quantified by comparing these properties between the empirical networks and their perturbed versions.

It is important to note that the empirical networks inevitably contained spurious connections (FPs) as well as connections that have been overlooked (FNs), due to limitations inherent to the connectome reconstruction methods employed. Therefore, the empirical networks should be understood as reference graphs with respect to our evaluation methodology, but not as biological ground truths. In other words, our approach was to systematically increase the existing (unknown) level of FPs and FNs in these empirical networks to determine the impact these increases have on the ability to accurately measure properties of complex network organization in the brain.

Which existing connections should be removed from the empirical networks to model FNs and where should new connections be added to model FPs? Ideally, this choice should be determined by the connections and pairs of nodes that are in practice most vulnerable to FPs and FNs. However, little is known about the spatial and topological distribution of FPs and FNs in brain networks, except for simple distance effects. In particular, evidence suggests that FNs at the macro-scale are more likely to be associated with long-distance ipsilateral and contralateral fibers (Jbabdi et al., 2015), due to the difficulty most tractography algorithms have in mapping long-distance fibers (Zalesky and Fornito, 2009), particularly those fibers that enter the cortex at sulcal fundi (Reveley et al., 2015). At the meso- and micro-scale, it is clear that tract tracing and microscopic connectome reconstruction techniques are also prone to FPs and FNs (Zaborszky et al., 2006), but the pairs of nodes that are most vulnerable to these effects are less well known, although long-distance axons are also prone to FNs with tract tracing (Ercsey-Ravasz et al., 2013; Markov et al., 2013).

In this study, we evaluated two distinct methods to determine which existing connections should be removed from the empirical networks to model FNs and where new connections should be added to model FPs. The first method involved placing FPs and FNs probabilistically by sampling from the set of all available connections either uniformly or from a distribution that was biased towards either short or long connections (Section 2.5). The second method involved deterministically placing FPs and FNs according to the order of streamline counts. Each new FP was added between the pair of unconnected nodes with the greatest streamline count in the unthresholded matrices and each new FN was generated by eliminating the current connection associated with the lowest streamline count (Section 2.6). The second method was only evaluated on the human connectome, given that streamline counts were not available for the mouse and worm.

Uniform and distance-based placement of FPs and FNs

FPs and FNs were positioned by sampling from either a uniform distribution or a distribution that was biased by connection distance. Under the *uniform model*, to generate each FP, we sampled a pair of nodes with uniform probability from the set of all node pairs that were disconnected in the empirical network and added a connection to that node pair. Similarly, to generate each FN, we sampled a pair of nodes with equal probability from the set of all node pairs that were connected in the empirical network and removed that connection. A connection that was removed at one iteration to model a FN was never added back at a later iteration to model a FP. In other words, once a connection was added or removed from a node pair, that node pair remained fixed and had a zero probability of being sampled again. The maximum number of FNs that could be added was therefore bounded by the number of connections comprising the empirical network.

The *distance model* was identical except that sampling was no longer uniform but was rather proportional to an exponential function of the Euclidean distance, d_{ij}, between nodes i and j (Ercsey-Ravasz et al., 2013). Specifically, the sampling probability $p_{ij} \propto \exp(-\beta d_{ij})$ ensured FPs and FNs were biased towards node pairs that were separated by short distances, where $\beta > 0$ was a constant. Note that $p_{ij} \propto 1$ under the uniform model. To model FPs and FNs biased towards node pairs that were separated by long distances, the sign of β was reversed (i.e. $\beta < 0$) so that the longest distances became the shortest. Under the distance model, the distribution of FPs and FNs was therefore determined by geometry, whereas the uniform model had no prior bias.

Denoting the binary connectivity matrices of the empirical and perturbed networks with $\mathbf{A}_{ij}$ and $\mathbf{A}'_{ij}$, respectively, the number of FPs and FNs were therefore given by,

$$\#\mathrm{FP} = \sum_{ij} 1_{\{\mathbf{A}'_{ij} - \mathbf{A}_{ij}\}}, \quad \#\mathrm{FN} = \sum_{ij} 1_{\{\mathbf{A}_{ij} - \mathbf{A}'_{ij}\}}, \tag{1}$$

with the FP and FN rates,

$$\mathrm{FPR} = \frac{\#\mathrm{FP}}{\sum_{ij}(1 - \mathbf{A}_{ij})}, \tag{2}$$

and

$$\mathrm{FNR} = \frac{\#\mathrm{FN}}{\sum_{ij} \mathbf{A}_{ij}}, \tag{3}$$

where $1_{\{x\}} = 1$, if $x = 1$, otherwise $1_{\{x\}} = 0$; and $\mathbf{A}_{ij} = 1$, if a connection is present from node i to node j, otherwise $\mathbf{A}_{ij} = 0$.

Using either the uniform or distance model, perturbed networks were generated for all possible combinations of the number of FPs and FNs, ranging between one and 10% of the total number of connections in the empirical networks. This yielded a family of perturbed networks positioned within a two-dimensional FP versus FN coordinate space. We sampled 100 perturbed networks for each coordinate in this space. Each of the 100 networks for a given coordinate had the same number of FPs and FNs, but the pairs of nodes between which the FPs and FNs were positioned differed due to the probabilistic nature of the two sampling models. For each of the 100 perturbed networks, measures of network clustering, network efficiency and network modularity were computed and compared with the empirical networks using various test statistics (see Section 2.7).

Streamline-based placement of FPs and FNs

For binary networks derived from thresholding dense connectomes mapped with tractography, the most likely location of FPs and FNs was also inferred from the order of streamline counts. In particular, each

new FP was added between the pair of unconnected nodes with the greatest streamline count and each new FN was generated by eliminating the current connection associated with the lowest streamline count. Ties in streamline counts were broken randomly. This deterministic method was motivated by the assumption of a monotonic relationship between connection probability and streamline count, in which case the most likely position for a new FP was the unconnected pair of nodes with the greatest streamline count.

We thresholded the human connectome to generate binary networks with connection densities ranging between 8% and 100%. Densities below 8% resulted in fragmented networks. Each of these binary networks served as an empirical network to which FPs were progressively added in a deterministic sequence until a fully-connected binary network was attained. Similarly, to evaluate the effect of decreases in connectome sensitivity, FNs were progressively added to each of the binary networks by eliminating connections with the lowest streamline counts until the connection density was reduced to 8% and the network fragmented. The addition of each new FP or FN resulted in a new perturbed network with worsening specificity or sensitivity, respectively. Network measures were computed in each perturbed network and compared with the empirical network from which they were derived.

Test statistics

To assess the impact of FPs and FNs on nodal network measures, the Pearson correlation coefficient was computed between the nodal values in the perturbed and empirical networks. This quantified the extent to which variations across nodal values were preserved between the two networks. To assess the impact of FPs and FNs on global network measures, the relative error between the empirical and perturbed networks was determined as $(x' - x)/x$, where x and x' denote the measures computed in the empirical and perturbed networks, respectively. Global measures were also plotted as a function of the number of FPs and FNs. The slope of the best fitting lines for each plot was determined, giving the amount of error that was introduced per each new FP or FN, assuming a linear fit was justified. The absolute value of the ratio of the FP to FN slopes was then used to quantify how many times each FP was more detrimental than each FN. For example, a ratio of two indicates that each FP introduces twice as much error as each new FN, a ratio of 0.5 indicates the converse and a ratio of one means FPs and FNs are equally detrimental.

Results

Network clustering, network efficiency and network modularity were measured in empirical brain networks that were perturbed by incrementally adding FP or FN connections. We report on the extent to which these network properties were misestimated when the number of FPs or FNs was increased to model errors in connectome reconstruction. We begin with uniform placement of FPs and FNs, as dictated by the uniform model. The results of our evaluation are summarized in Table 2 and described in detail below. We then position FPs and FNs according to the distance model and based on the order of streamline counts in the human connectome.

Table 2
How many times FPs were more detrimental than FNs when estimating key measures of complex network organization in human, mouse and worm connectomes. Each measure was plotted as a function of the number of FPs or FNs and the slope of the best fitting line determined. The absolute value of the ratio of the FP to FN slopes is tabulated here for the three species investigated and random networks. Confidence intervals for all estimates are less than ±0.1, except for modularity, where ±3 (human), ±9 (mouse) and ±8 (worm). All results for the uniform model.

	Human	Mouse	Worm	Random
Clustering	1.4	1.5	1.7	1.0
Efficiency	2.3	3.6	3.5	0.9
Modularity	9	76	33	1

Network clustering

Network clustering was measured with the clustering coefficient, a classic measure of functional specialization in brain networks (Rubinov and Sporns, 2010). A node's clustering coefficient is the ratio of its closed triangles to all possible triangles. In other words, it is the proportion of a node's immediate neighbors that are connected with each other (Bullmore and Sporns, 2009). The clustering coefficient was computed in the directed networks based on a widely accepted definition (Watts and Strogatz, 1998). Global clustering was measured by averaging the nodal clustering coefficients over all nodes.

Fig. 2 (panel A) shows the impact of FPs and FNs on the clustering coefficient when FPs and FNs are uniformly distributed. The three rows correspond to different species and the columns correspond to different test statistics for quantifying misestimation of the global and nodal clustering coefficient as a function of the number of FPs or FNs.

Species are illustrated in the first column. The second column shows the global clustering coefficient as a function of the number of FPs (blue lines) or FNs (red lines). The leftmost data point in these plots corresponds to the empirical network; that is, the case of zero FPs and zero FNs (e.g. no FPs or FNs added to the empirical connectomes). Introducing FPs or FNs resulted in an approximately linear decrease in clustering, thus leading to a gradual underestimation of this measure. FPs were more detrimental than FNs across all species. As quantified by the ratio of the FP to FN slopes of the best fitting lines, each new FP resulted in a decrease in clustering that was on average 1.4 (human), 1.5 (mouse) and 1.7 (worm) times greater than each new FN (Table 2).

The third column of Fig. 2 shows the relative error in measurements of the global clustering coefficient between the empirical and perturbed networks. The fourth column shows the correlation between the empirical and perturbed networks in the nodal clustering coefficients. The consensus across all test statistics and species is that the presence of FPs or FNs leads to underestimation of the clustering coefficient, with each FP approximately 1.5 times more detrimental than each FN.

Perturbed networks comprising various combinations of FPs and FNs were generated to identify potential interactions between FPs and FNs. To this end, the final column of Fig. 2 shows a two-dimensional FP versus FN coordinate space where each coordinate is colored with the correlation in nodal clustering between the empirical and perturbed networks (average of 100 perturbed network samples). The correlation at the bottom, leftmost coordinate is one (white color), since the nodal clustering at this coordinate is necessarily identical between the perturbed and empirical networks. The correlation at the top, rightmost coordinate is approximately 0.6 (black color) and corresponds to perturbed networks with the maximum number of FPs and FNs. The change in color is more pronounced along the vertical axis (FPs), indicating that FPs are more detrimental than FNs. For example, with 10 FPs and 100 FNs, the correlation between the perturbed and empirical networks in the human and mouse networks remains relatively high (~0.9, yellow zone), but for the converse case (10 FNs and 100 FPs) the correlation is significantly reduced (~0.7, red zone). This effect in correlation is diminished in the worm network, where FPs were only slightly more detrimental to estimating nodal clustering.

Network efficiency

Network efficiency is a popular measure of integration that assumes neural information traverses the shortest polysynaptic paths between pairs of nodes (Sporns et al., 2004; Sporns, 2011). Brain network efficiency has been widely studied (Bullmore and Sporns, 2012) and is closely related to the concept of small-world networks (Latora and

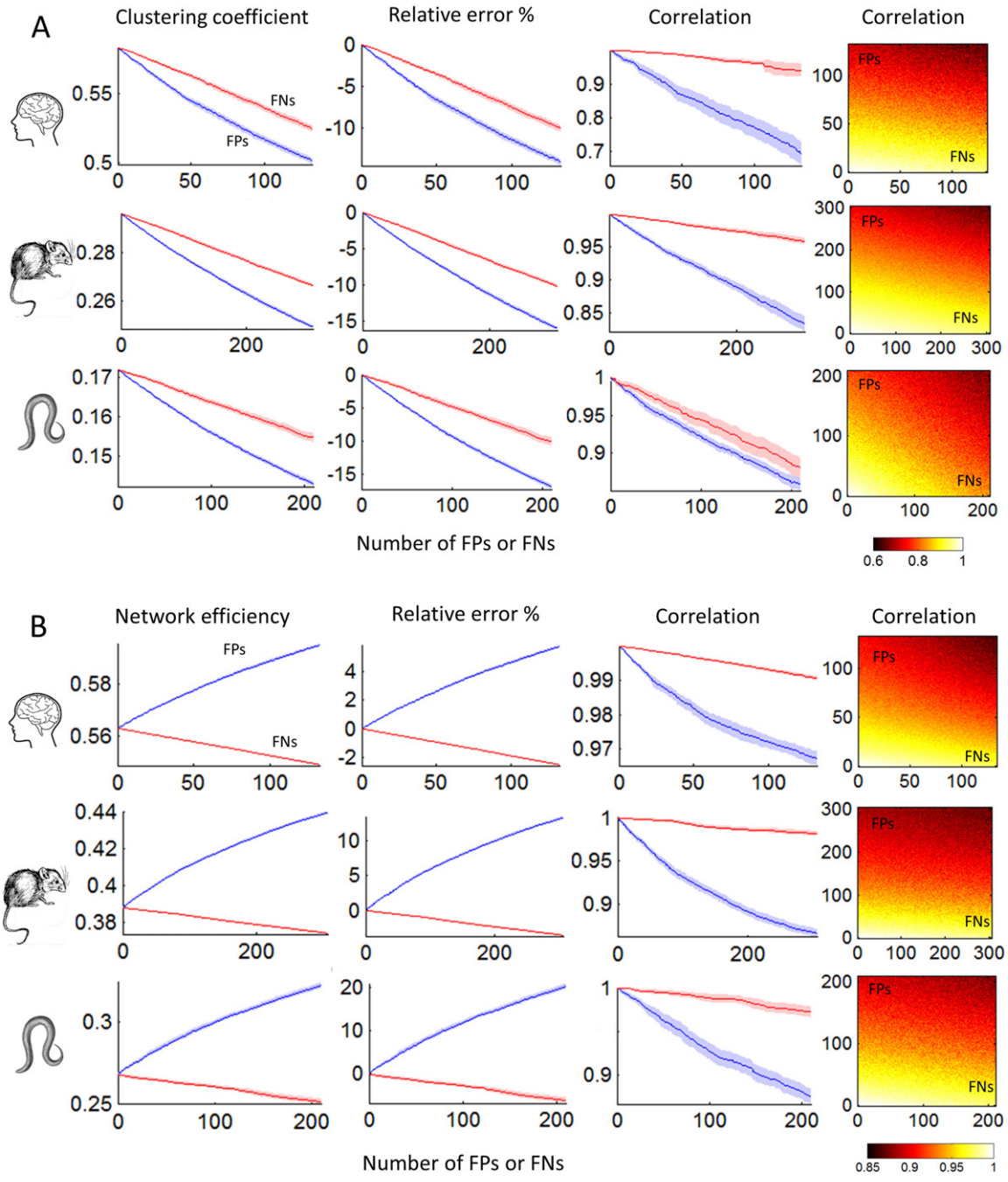

Fig. 2. Impact of false positive (FP, blue lines) and false negative (FN, red lines) connections on the measurement of the clustering coefficient (A) and network efficiency (B). FPs or FNs were progressively added to human, mouse and worm connectomes in a uniform manner to determine the extent to which they impaired the estimation of clustering and efficiency. The shaded areas encapsulate bootstrap-estimated 95% confidence intervals. Each data point represents an average across 100 perturbed networks with uniformly distributed FPs or FNs. Correlation is the Pearson correlation in nodal clustering and nodal efficiency values between the empirical and perturbed networks. The heat maps show this correlation for combinations of FPs and FNs, where the horizontal axis is the number of FNs and the vertical axis is the number of FPs.

Marchiori, 2001) and network communication (Goñi et al., 2014; Mišić et al., 2015). To compute efficiency, the shortest path lengths were first determined between all possible pairs of nodes. Nodal efficiencies were computed as the reciprocal of the harmonic mean of the shortest path lengths from a given node to all other nodes (Rubinov and Sporns, 2010). Global efficiency was computed by averaging nodal efficiencies.

Fig. 2 (panel B) shows the impact of FPs and FNs on nodal and global efficiency. Introducing FPs to the empirical networks resulted in an approximately linear increase in network efficiency, whereas FNs resulted in a gradual linear decrease. Each new FP resulted in a change in efficiency that was on average 2.3 (human), 3.6 (mouse) and 3.5 (worm) times greater than each new FN. For example, the relative error was <2% when 100 FNs were added to the human connectome, but introducing the same number of FPs resulted in the relative error exceeding 4%. As evident from the heat maps, for all species, the correlation in nodal efficiencies between the empirical and perturbed networks was better preserved with the addition of FNs than FPs.

Network modularity

Brain networks are modular (Bassett et al., 2010; Meunier et al., 2010; Sporns and Betzel, 2015), meaning that they can be decomposed into subsets of nodes that are more densely connected than expected by

chance (Betzel et al., 2013). The extent to which a network is modular is typically determined by comparing the number of connections within each module to the expected number of these connections in appropriately matched random networks (Newman, 2006).

We used the Louvain algorithm (Blondel et al., 2008) to decompose the empirical and perturbed networks into sets of nodes that defined mutually exclusive modules. For each network, we generated a co-classification matrix that summarized 20 independent runs of the Louvain algorithm applied to the same network. Different runs of the Louvain algorithm did not necessarily yield identical solutions, due to the stochastic nature of the algorithm (Good et al., 2010). The co-classification matrix had the same number of rows and columns as nodes. Element (i,j) stored the proportion of runs for which nodes i and j resided within a common module. To determine a single set of modules that represented a consensus over all 20 runs, elements in the co-classification matrix below 0.4 were set to zero, after which the Louvain algorithm was applied to this thresholded co-classification matrix to yield a final consensus decomposition. It is typical for this consensus procedure to be iterated many times based on multiple runs of the Louvain algorithm applied to the co-classification matrix (Lancichinetti and Fortunato, 2012). Given that we needed to decompose each of hundreds of perturbed networks (i.e. a new modular decomposition was computed each time a FP or FN was added), for reasons of computational tractability, only one iteration was performed and we did not generate a two-dimensional FP-FN space. Note that we computed a new modular decomposition for each perturbed network. We did not impose the decomposition for the empirical networks onto their perturbed versions because this is not how a modularity analysis proceeds in practice.

The Q-index of modularity was computed for the final consensus decomposition for each of the empirical and perturbed networks. The Q-index is a value in the range $[-1,1]$ that quantifies the extent to which a network is modular, with positive values indicating greater modular organization than expected by chance (Newman, 2006). Fig. 3 shows the Q-index plotted as a function of the number of FPs (blue lines) or FNs (red lines). Introducing FNs to the human connectome resulted in a slight increase in the Q-index, but FNs did not significantly affect the mouse and worm connectomes. In other words, the presence of FNs did not impair the ability to accurately estimate the extent to which the empirical networks were modular, as quantified with the Q-index. In contrast, the Q-index decreased substantially with the addition of FPs. As quantified by the ratio of the FP to FN slopes of the best fitting lines, each new FP resulted in a change in the Q-index that was on average 9 (human), 76 (mouse) and 33 (worm) times greater than each new FN. The near-zero FN slopes resulted in wide confidence intervals for these estimates: ± 3 (human), ± 9 (mouse) and ± 8 (worm).

The relative error in the Q-index is also shown in Fig. 3. This confirms the substantially greater impairment FPs have on estimation of the Q-index. For example, when 100 FPs were introduced to the human connectome, the relative error in the Q-index was approximately 8%, but when the same number of FNs were introduced, the relative error was $<1\%$.

The Q-index is not informative with respect to the composition of modules. For example, it is possible that the composition of modules in the empirical and perturbed networks was very different, despite the two partitions being associated with a similar Q-index. The Rand index was therefore used to quantify the overlap in the composition of modules between the empirical and perturbed networks. To this end, binary co-classification matrices were determined for the empirical and perturbed networks. For a given modular decomposition, if nodes i and j resided within a common module, then element (i,j) was set to one, otherwise (i,j) was zero. For a pair of co-classification matrices, the Rand index was then computed as the ratio of the number of elements that were equal between the pair of matrices, either both zero or both one, to the total number of elements in the matrix, excluding the matrix diagonal (Rand, 1971). Identical modular decompositions yield a Rand index of one, while decompositions with no similarity yield zero.

Fig. 3 shows the Rand index as a function of the number of FPs or FNs. The Rand index for the leftmost data point corresponds to the case of zero FPs and zero FNs, and thus quantifies the agreement in module composition between two independent runs of the Louvain

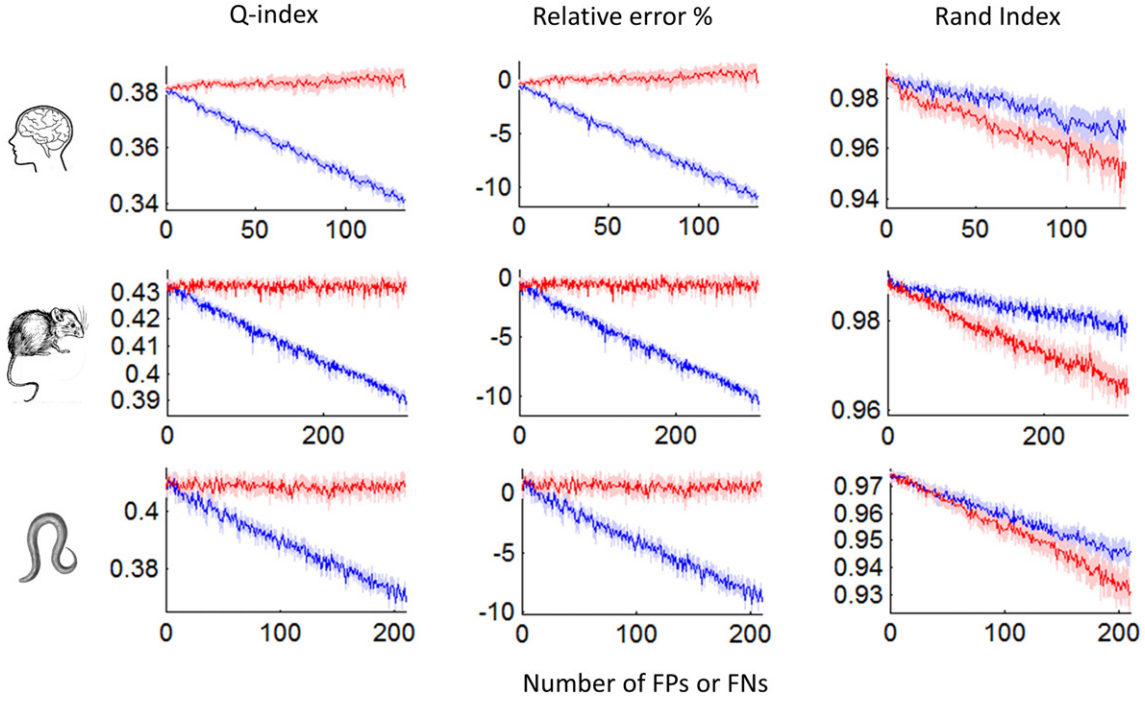

Fig. 3. Impact of false positive (FP, blue lines) and false negative (FN, red lines) connections on the measurement of network modularity. FPs or FNs were progressively added to human, mouse and worm connectomes in a uniform manner to determine the extent to which they impaired the estimation of network modularity (Q-index) and the composition of modules (Rand index) between the empirical networks and their perturbed versions. The relative error is with respect to the Q-index.

algorithm applied to the same empirical network. This agreement exceeded 98% for all species, with the residual dissimilarity due to degeneracy of the solution space and the fact that only one iteration of the consensus procedure was performed. The Rand index between the empirical and perturbed networks decreased linearly with an increasing number of FPs or FNs. This decrease was however slight, with less than a 0.02% decrease in the Rand index per each FP or FN in the human connectome. For all species, FNs were slightly more detrimental with respect to the Rand index.

Finally, the above analyses were repeated for the null case in which the empirical brain networks were replaced with random networks. Supplementary Fig. 1 shows the impact of FPs and FNs on clustering, efficiency and modularity for this null case. Unlike the empirical brain networks, it can be seen that the relative error increased at the same rate for FPs and FNs, and thus these two kinds of errors were equally detrimental in random networks. This suggests that one or more of the non-random topological properties of a nervous system might explain the disproportional impact of FPs. We show in Section 4 that modular organization is the main property that results in the disproportional impact of FPs in brain networks.

Distance model

The results presented thus far pertain to the uniform model where FPs were sampled uniformly from the set of all disconnected pairs of nodes and FNs were sampled uniformly from the set of all connections in the empirical networks. Fig. 4 shows relative errors in the estimation of the clustering coefficient, network efficiency and modularity in the human connectome under the distance model. Relative errors are shown as a function of the number of FPs (blue lines) and FNs (red lines). Long FPs are cases where the sampling of FPs was biased towards pairs of disconnected nodes that were physically separated by long distances ($\beta = -0.5$, see Methods), whereas short FPs were biased conversely using the same value of β but with opposite sign. Similarly, long FNs were cases where the sampling of FNs was biased towards connections that were short in length, and conversely for short FNs. The lines labeled only with FN or FP correspond to the uniform model.

Fig. 4 shows that long FNs and short FPs resulted in the smallest relative error in the estimation of the clustering coefficient, although with this specific combination, FPs still remained significantly more detrimental than FNs. In cases where both FNs and FPs were short, FNs became slightly more detrimental than FPs at high error rates. Short FPs reduced the relative error in estimating the *Q*-index. Modeling distance biases did not have a substantial impact on the estimation of network efficiency.

Streamline-based placement of FPs and FNs

Fig. 5 (panel A) shows the relative error in estimating the clustering coefficient and network efficiency in the human connectome when FPs or FNs were added based on the order of streamline counts in the unthresholded matrices. Each of the blue curves corresponds to an empirical network with a distinct connection density that was generated by thresholding the human connectome to yield a binary network. The connection densities of the empirical networks are specified by the point at which each blue curve intersects the horizontal line corresponding to a relative error of zero. Each time a FP was added to the unconnected node pair with the greatest streamline count, the connection density of the perturbed network increased, thereby progressively extending the corresponding blue curve to the right. Each time a FN was added by removing the connection with the lowest streamline count, the connection density of the perturbed network decreased, thereby progressively extending the blue curve to the left.

Let us assume for argument's sake that the true connection density of the human connectome is 10%, in which case we are interested in the blue curve that attains a relative error of zero for a connection density of 10%. Tracing this specific blue curve to the right quantifies the increase in relative error that can be expected as more FPs are added and the connections density is overestimated. Similarly, tracing this curve to the left reveals the increase in relative error that can be expected as more FNs are added and the connection density is underestimated. Fig. 5 shows that the relative error increases faster when adding FPs (right of zero-crossing) compared to FNs (left of zero-crossing).

This effect is further quantified by the heat maps shown in Fig. 5 (panel B). If the true connection density of the human connectome is 10% (horizontal axis of the heat maps), but we threshold to achieve an assumed connection density of 30% (vertical axis), the relative error is high and resides within the red zone of the heat maps. In contrast, if the true connection density is 30%, but we threshold to achieve an assumed connection density of 10%, the relative error is smaller and remains within the yellow zone. In the first case, our threshold was

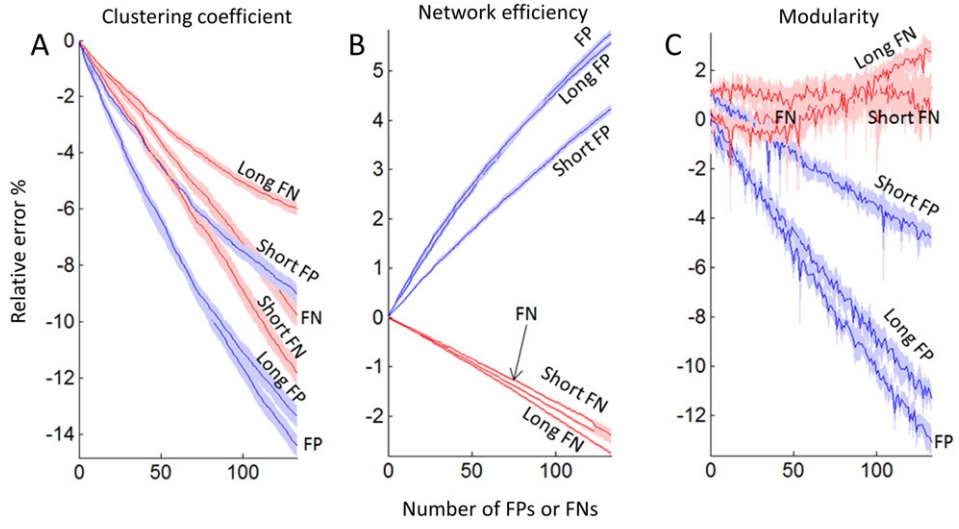

Fig. 4. Impact of distance-biased false positive (FP, blue lines) and false negative (FN, red lines) connections when estimating the clustering coefficient (A), network efficiency (B) and modularity (C) in the human connectome. FPs and FNs were progressively added by sampling connections with probabilities that either increased or decreased with connection distance. Long FPs/FNs denote a preference towards the placement of FPs or FNs at long connections. Short FPs/FNs denote a preference towards the placement of FPs or FNs at short connections.

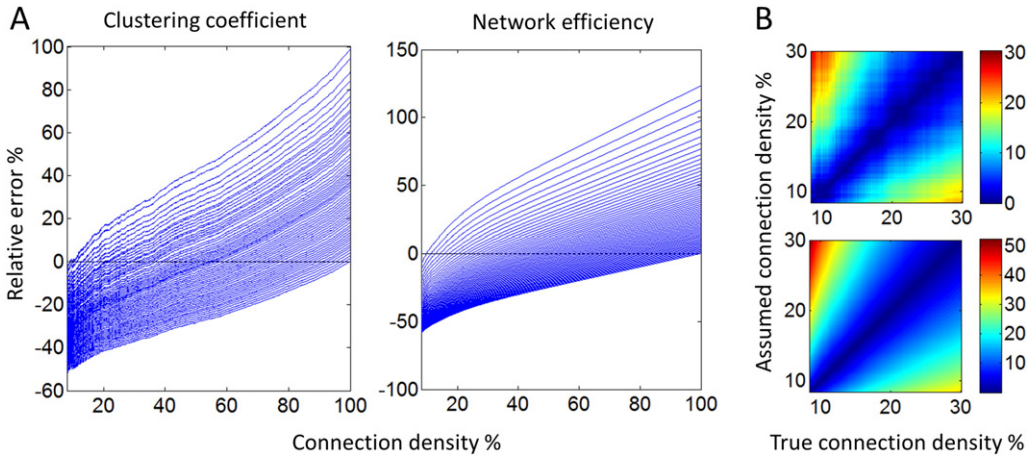

Fig. 5. Relative error in the estimation of the clustering coefficient and network efficiency in the human connectome when false positive (FP) and false negative (FN) connections were added based on the order of streamline counts. **A)** Each of the blue curves corresponds to an empirical network with a distinct connection density that was produced by thresholding the human connectome to yield a binary network. The portion of the curve to the right of the zero-crossing shows the increase in relative error as FPs were progressively added to the unconnected node pairs with the greatest streamline counts, whereas the portion to the left of the zero-crossing is where FNs were added by eliminating the connections with the lowest streamline counts. **B)** The heat maps quantify the relative error associated with thresholding the human connectome either too leniently (upper triangle) or too stringently (lower triangle). Note that the upper triangle has a more expansive red zone, indicating greater relative error with lenient compared to stringent thresholding. The upper heat map is for the clustering coefficient and the lower is for network efficiency.

too lenient, resulting in FPs, whereas our threshold was too stringent in the second case, resulting in FNs. The heat maps show that excessively stringent thresholding (i.e. many FNs) is more desirable than excessively lenient thresholding (i.e. many FPs), even when connections are added or removed by selecting the ones that are most likely to be absent or present based on the number of streamlines. In other words, FPs are more detrimental than FNs to estimating the clustering coefficient and network efficiency.

Why are FPs more detrimental than FNs?

It remains to be determined *why* FPs are more detrimental than FNs to estimating topological properties of connectomes. To address this question, we consider a network model that is typical of brain networks. The model is a small-world network with idealized modular structure and a rich club (Fig. 6). The network comprises N modules. Each module comprises n nodes that are fully connected. The N modules are interconnected by a fully-connected rich club, with one rich-club node residing in each module. The symmetry of this network makes it amenable to straightforward analysis under the assumption that FPs and FNs occur uniformly (uniform model; Section 2.5).

In modular small-word networks such as the brain, FPs are most likely to occur *between* modules because inter-modular node pairs are more likely to be unconnected than their intra-modular counterparts, due to the very definition of modularity. These unconnected node pairs provide the necessary openings for FPs to emerge. Indeed, the idealized modular structure of our network model ensures FPs are necessarily inter-modular. Conversely, FNs are most likely to occur *within* modules because the majority of connections in a modular network are intra-modular. FNs can also occur at rich-club connections, but the number of intra-modular connections usually dwarfs the number of rich-club connections in most modular networks, making this occurrence rare.

For our network model, FPs and FNs can be classified as either type 1 or type 2 (Fig. 6). FPs and FNs that occur between peripheral nodes are labeled *type 1*, whereas those that occur between peripheral and rich-club nodes are labeled *type 2*. A third type of FN that occurs exclusively between rich-club nodes can also be defined, but we ignore this type of FN because it becomes an increasingly unlikely occurrence in typical modular networks where $n \gg N$. Note that this classification of FPs and FNs is not in any way related the classical distinction between type I (FP) and type II (FN) errors.

In this section, we analytically determine the change in the clustering coefficient, C, and network efficiency, E, after the addition of a single FP or FN to the network model shown in Fig. 6. Our analysis is made amenable by the symmetry of the model, which ensures that it is sufficient to analyze the effect of an arbitrarily located FP and FN of type 1 and 2. The notation Δ_x^y, $x \in \{FP,FN\}$, $y \in \{1,2\}$, is used to denote the amount by which a network measure changes after adding a FP or FN that is of type 1 or 2. For example,

$$\Delta_{FP}^1 C = C - \hat{C}, \tag{4}$$

denotes the amount by which the clustering coefficient changes with the addition of a type 1 FP, where C is the clustering coefficient of the original network and $\hat{C}$ is the clustering coefficient of the perturbed network with a single FP of type 1 added. Analogously,

$$\Delta_{FP}^1 E = E - \hat{E}, \tag{5}$$

denotes the amount by which the network efficiency changes with the addition of a type 1 FP. We also use

$$\Delta_x = p_x^1 \Delta_x^1 + p_x^2 \Delta_x^2 \tag{6}$$

to denote the *average* change, taking into account the probability of type 1, p_x^1, and type 2, p_x^2, connections.

Under the uniform model, the probabilities that a FP is type 1 or 2 are

$$p_{FP}^1 \propto \frac{N(N-1)(n-1)}{2}, \qquad p_{FP}^2 \propto N(N-1). \tag{7}$$

Similarly, the probabilities that a FN is type 1 or 2 are

$$p_{FN}^1 \propto \frac{(n-1)(n-2)}{2}, \qquad p_{FN}^2 \propto n-1. \tag{8}$$

In the following, we establish two theorems quantifying how many times FPs are more detrimental than FNs when estimating the clustering

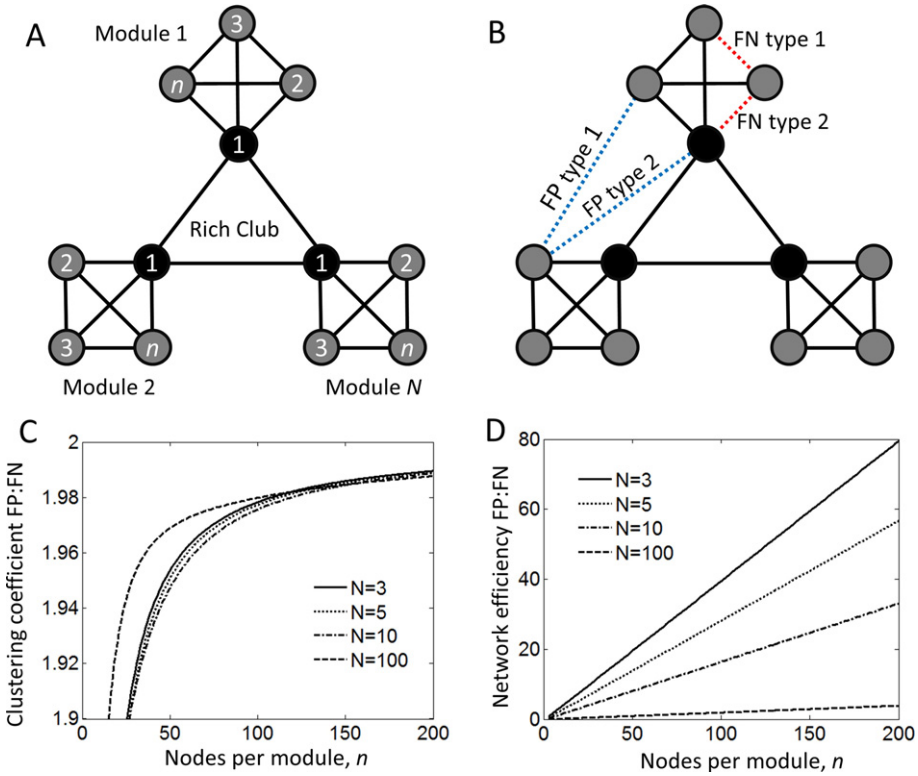

Fig. 6. Asymptotic analysis of the tradeoff between sensitivity and specificity in small-world networks with idealized modular structure. **A)** The network model comprises N modules. Each module contains n nodes that form a fully-connected subgraph. The N modules are interconnected by a rich club, with each module comprising a single rich-club node (black) and $n-1$ peripheral nodes (grey). **B)** False positive (FP) and false negative (FN) connections are either of type 1 (peripheral-to-peripheral) or type 2 (peripheral-to-rich-club). **C)** $C_{FP:FN}$ converges to a limiting value of 2 as the number of nodes per module is increased, indicating that FPs are twice as detrimental than FNs to estimating the clustering coefficient in large networks with idealized modular structure. **D)** $E_{FP:FN}$ increases approximately linearly with the number of nodes per module.

coefficient (Theorem 1) and network efficiency (Theorem 2) in small-world networks with idealized modular structure. We specifically focus on the limiting behavior of the ratio $FP:FN = |\Delta_{FP}/\Delta_{FN}|$ when the number of nodes, n, tends to infinity for fixed N. Detailed proofs for Theorems 1 and 2 are relegated to the Supplementary Material.

Theorem 1. *As the number of nodes in small-world networks with idealized modular structure grows* (i.e. $n \to \infty$), *FPs become exactly twice as detrimental than FNs to estimating the clustering coefficient. Specifically, for* $N > 1$,

$$\lim_{n\to\infty} C_{FP:FN} = 2.$$

Proof. The clustering coefficient of the network shown in Fig. 6 is

$$C = \frac{(N-1)(N-2)+(n-1)(n-2)}{n(n+N-2)(n+N-3)} + \frac{n-1}{n}, \tag{9}$$

with $C \to 1$ as $N \to \infty$ and $n \to \infty$. How does C change with the addition of a FP or FN? Trivial yet tedious calculations reveal that the change in clustering due to a type 1 FP is

$$\Delta^1_{FP}C = \frac{4}{Nn^2}, \tag{10}$$

and for a type 2 FP,

$$\Delta^2_{FP}C = \frac{1}{Nn}\left(\frac{2n^2-2n+4}{N+n-1} - \frac{4(n-1)^2}{N+n-2} + \frac{2n^2-6n+2}{N+n-3} + \frac{2(n-2)}{n(n-1)}\right). \tag{11}$$

Similarly, the change in clustering due to a type 1 FN is

$$\Delta^1_{FN}C = \frac{1}{Nn}\left(\frac{2(n-3)}{(n-1)(n-2)} + \frac{2}{(n+N-2)(n+N-3)}\right), \tag{12}$$

and for a type 2 FN,

$$\Delta^2_{FN}C = \frac{1}{Nn}\left(\frac{2(n-2)}{(n-1)(n-2)} - \frac{2(N-n)(N-1)}{(n+N-2)(n+N-3)(N+n-4}\right) \tag{13}$$

To quantify how many times FPs are more detrimental than FNs to the estimation of the clustering coefficient, we combine Eqs. (10)–(13) to compute the ratio,

$$C_{FP:FN} = \left|\frac{\Delta_{FP}C}{\Delta_{FN}C}\right| = \left|\frac{p^1_{FP}\Delta^1_{FP}C + p^2_{FP}\Delta^2_{FP}C}{p^1_{FN}\Delta^1_{FN}C + p^2_{FN}\Delta^2_{FN}C}\right|. \tag{14}$$

The ratio $C_{FP:FN}$ is equivalent to the empirically derived ratios shown in Table 2. For example, $C_{FP:FN} = 2$ indicates that each FP introduces twice as much error as each FN, a ratio of 0.5 indicates the converse and a ratio of one means FPs and FNs are equally detrimental. Repeated application of L'Hôpital's rule to $C_{FP:FN}$ gives the required result. ■

Theorem 2. *As the number of nodes in small-world networks with idealized modular structure grows* (i.e. $n \to \infty$), *the detriment of FPs relative to FNs to estimating network efficiency becomes arbitrarily large. Specifically, for* $N > 1$,

$$\lim_{n\to\infty} E_{FP:FN} = \infty.$$

Proof. The efficiency of the network shown in Fig. 6 is

$$E = \frac{N-4n+Nn+Nn^2+2n^2-1}{3n(Nn-1)}, \tag{15}$$

with $E \to 1/3$ as $N \to \infty$ and $n \to \infty$. It can be shown that the change in efficiency due to a FP is

$$\Delta^1_{FP}E = \frac{-2}{3Na}, \quad \Delta^2_{FP}E = \frac{-(n+2)}{3Nna}, \tag{16}$$

where $a = (Nn-1)$. The change in efficiency due to a type 1 FN is

$$\Delta^1_{FN}E = \frac{1}{Nna}, \tag{17}$$

and for a type 2 FN,

$$\Delta^2_{FN}E = \frac{(N-1)n+N+5}{6N^2n^2-6Nn}. \tag{18}$$

Combining Eqs. (16)–(18) gives the ratio,

$$E_{FP:FN} = \left|\frac{p^1_{FP}\Delta^1_{FP}E + p^2_{FP}\Delta^2_{FP}E}{p^1_{FN}\Delta^1_{FN}E + p^2_{FN}\Delta^2_{FN}E}\right| = \frac{2n(n^2+2)}{(n+1)(N+2n+Nn-1)}. \tag{19}$$

It can be seen that for fixed $N > 0$ and $n \to \infty$, $E_{FP:FN} \to \infty$ because the highest order term in the numerator is n^3, whereas the highest order term in the numerator is only n^2. This establishes the theorem. ■

It can also be seen from Eq. (19) that $\lim_{N\to\infty} E_{FP:FN} = 0$ for fixed n. Solving the equation $E_{FP:FN} = k$ for the number of modules, N, reveals that $N \sim 2n/k$. This means that FPs are twice (i.e. $k = 2$) as detrimental than FNs to estimating network efficiency when $N = n$. When $N = 2n$, FPs and FNs are equally detrimental, and when $N > 2n$, FNs introduce more error than FNs. The last case is unlikely in practice, since the number of modules in most brain networks does not exceed twice the number of nodes per module.

To understand the asymptotic behavior of the limiting regimes underlying Theorems 1 and 2, $C_{FP:FN}$ and $E_{FP:FN}$ are plotted as a function of n (nodes per module) for various values of N (number of modules) in Fig. 6. It can be seen that convergence of $C_{FP:FN}$ to its limiting value of 2 is fastest when the number of modules is small. The increase of $E_{FP:FN}$ as a function of n is approximately linear.

It useful to draw a parallel with social networks to make clear the increased detriment of FPs relative to FNs when estimating network efficiency. Connections in social networks represent friendships between pairs of individuals. Modules in these networks might represent groups of individuals with common interests, for example. If a friendship is broken between a pair of individuals in the same module, network integration is unlikely to suffer a substantial reduction, since the cohesiveness of the group is ensured by its many other social ties. Friends within a social group are shared, ensuring any individual friendship within a group is not crucial to the cohesiveness of the social group as a whole. In contrast, if a new friendship is created between a pair of individuals in distinct modules, network integration is likely to increase substantially, since this new friendship establishes opportunities for tighter integration between two large groups of individuals that were previously only weakly tied to each other. In other words, adding inter-modular connections (i.e. FPs) is likely to have a greater impact on network integration than removing intra-modular connections (i.e. FNs).

In summary, connectome specificity is more important than connectome sensitivity because connectomes are modular networks. In modular networks, FPs are more likely to be inter-modular connections and FNs are more likely to be intra-modular. This is because inter-modular node pairs are more likely to be unoccupied (FPs can only emerge at unoccupied node pairs), whereas intra-modular node pairs are densely connected (FNs can only emerge at genuine connections). Overlooking genuine intra-modular connections is less detrimental than including spurious inter-modular connections. Intra-modular connections are abundant in modular networks. Overlooking a few is not very detrimental because there are so many of them. Inter-modular connections are by definition rarer, and thus including spurious connections between modules is more likely to impact on network structure.

Discussion

Developing axonal fiber tracking methods that provide high sensitivity to resolve intricate fiber architectures and navigate crossing-fiber geometries with ease is crucially important across all axonal scales, from individual axons to large axonal fiber bundles. However, when these methods are applied to reconstruct connectomes for the purpose of studying the topological properties of the brain (Bullmore and Sporns, 2009), upholding specificity becomes paramount and significantly *more* important than the need to maximize sensitivity. In other words, guarding against the detection of spurious fibers (FPs) is more important than failure to detect true fibers (FNs). Connectomes with both high sensitivity and high specificity are obviously most advantageous, but they are unattainable with current state of the art, particularly at the macro-scale (Azadbakht et al., 2015; Knosche et al., 2015; Thomas et al., 2014, also see Fig. 7). Globally optimal tractography algorithms (Jbabdi et al., 2007; Sotiropoulos et al., 2010; Zalesky and Fornito, 2009), precision measures of connectivity (Iturria-Medina et al., 2008), multi-shell models and improved imaging acquisitions (Glasser et al., 2013) and streamline filters (Smith et al., 2013; Pestilli et al., 2014) hold much promise towards improvements in this regard. In the interim though, we argue that connectome specificity should be emphasized over connectome sensitivity when analyzing graph theoretic properties, particularly the clustering coefficient, network efficiency and modularity. This can be achieved with stringent thresholding or the use of conservative fiber tracking methods.

At the macro-scale, fitting advanced crossing-fiber models to low b-value and low angular resolution data can result in poorly estimated fiber orientation distribution functions (fODFs). The lobes of the fODFs become 'plump' and imprecise. This offers probabilistic streamlines greater autonomy to follow a broader range of directions, some of which might represent spurious fiber orientations. What are the consequences of following these spurious orientations? Tractography has classically been performed and evaluated by seeding streamlines from circumscribed regions of interest (Fillard et al., 2011), in which case a few spurious fiber directions might only result in a few inconsequential streamlines taking 'wrong turns'. In contrast, connectome reconstruction demands for streamlines to be seeded from all white matter

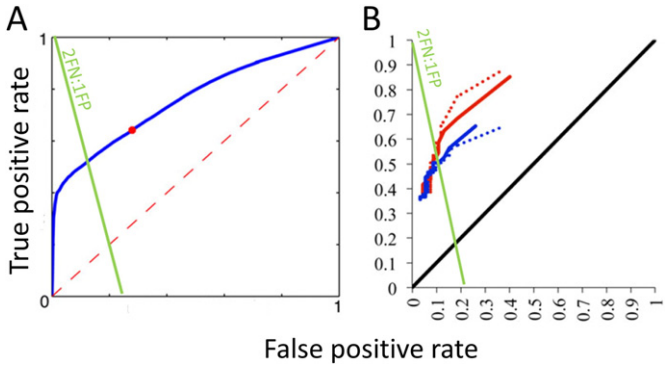

Fig. 7. Receiver operating characteristic (ROC) curves for mouse (A) and macaque (B) connectomes reconstructed with tractography and compared to axonal tract tracing to determine false positive and false negative connections. ROC curves are parameterized by a threshold that determines connection density. Lenient thresholds yield dense connectomes with high sensitivity (rightmost points), whereas sparse connectomes with high specificity are attained with stringent thresholds (leftmost points). The 2:1 rule is satisfied at the intersection of the green line with the ROC curve. In panel B, red/blue denotes left/right hemispheres and solid/dashed lines denote distinct diffusion imaging data sets. Panels A and B reproduced from Calabrese et al. (2015) and Azadbakht et al. (2015), respectively, with green lines superimposed.

voxels (Li et al., 2012). It can be argued that this inherently creates a need to correct for multiplicity in spuriously estimated orientations across the set of all white matter voxels. In particular, the chance that streamlines take multiple 'wrong turns' is clearly much greater than in the case when seeding is performed from a region of interest. This is why poorly estimated fODFs are more detrimental to connectome reconstruction than classic tractography applications that only seed streamlines from a circumscribed region of interest.

A comprehensive evaluation of popular deterministic, probabilistic and global tractography algorithms was undertaken by Bastiani et al. (2012). They empirically evaluated the sensitivity of these algorithms to reconstruct three well-known white matter fiber tracts and suggested that global approaches may provide the best tradeoff. Specificity was quantified with respect to a pair of distance-matched regions that were assumed to be anatomically disconnected. In the current study, we have focussed on evaluating the relative detriment of FPs and FNs in graph theoretical analyses of brain networks, whereas Bastiani et al. (2012) provide an excellent evaluation of different tractography algorithms with respect to connectome reconstruction.

Recommendations

Graph theoretical analyses of brain networks should ideally be performed on connectomes with at least twice the number of missing connections (FNs) as spurious connections (FPs). We call this the 2:1 rule of thumb. In other words, if a connectome comprises 100 erroneous connections (FPs + FNs), then about 33 of these should be FPs, with the others being FNs. The 2:1 rule is exact when estimating the clustering coefficient in small-world networks with many nodes and an idealized modular structure (Theorem 1). For estimation of network efficiency, the 2:1 rule is exact if the number of nodes per module is equal to the total number of modules. As the number of nodes per module exceeds the total number of modules, the detriment of FPs relative to FNs grows linearly as a function of the number of nodes (Theorem 2). Empirically, the 2:1 rule comes about by rounding down to the nearest integer the median of the three values in the first column of Table 2. For estimation of network modularity, particularly the Q-index, connectomes with more than twice as many FNs as FPs are likely to be appropriate, making the 2:1 rule conservative in this case. Recommending two FNs per each FP achieves a simple and conservative compromise across network measures and species.

A few caveats should be noted when considering these guidelines. First, the 2:1 rule assumes that minimizing the absolute error associated with estimates of network measures is paramount. In circumstances demanding upper or lower bounds rather than minimum absolute errors, sensitivity can be more important than specificity. For example, applications that mandate an optimistic estimate (upper bound) of network efficiency require connectomes with greater sensitivity, since FPs increase efficiency, leading to overestimation, whereas FNs decrease efficiency, causing underestimation.

Second, the empirical evaluations undertaken in this study were not exhaustive. While consensus about the importance of connectome specificity was apparent across the species and network measures evaluated, it is possible that accurate estimation of some niche properties does not follow this rule of thumb. Further work is required to evaluate the relative impact of FPs and FNs on other topological measures, such as rich-club measures (Van den Heuvel et al., 2012), network motifs and participation coefficients as well as the properties of *weighted* brain networks (Rubinov and Sporns, 2010; van den Heuvel et al., 2015). Evaluating the impact of FPs and FNs on weighted networks requires ground truths for both the absence and presence of connections and the weight of each connection that is present. Nevertheless, our analysis of small-world networks with idealized modular structure indicated that the greater importance of specificity relative to sensitivity is a general feature of all modular small-world networks, and perhaps networks with complex topology in general. Given that most nervous systems display small-world organization and modularity, it is likely that the current findings generalize to species other than the three considered here. For cost-integrated network measures (Ginestet et al., 2011), the greater detriment of FPs motivates constraining the integration domain to sparse and specific networks, rather than integrating over all network costs. Our findings suggest that integrating over dense, high-cost networks is likely to introduce greater error to cost-integrated measures (i.e. area under curve) than sparse networks.

Practical implications

How might the 2:1 rule be used to practically guide connectome reconstruction? Connectome reconstruction methods cannot be straightforwardly controlled to ensure FNs are approximately twice more abundant than FPs. While thresholding methods afford precise control of the tradeoff between sensitivity and specificity in dense connectomes, the exact ratio of FPs to FNs is usually unknown for a given threshold or connection density, except in circumstances where a ground truth is available.

Axonal tract tracing is often used as a gold standard to benchmark the accuracy of macro-scale connectomes reconstructed with diffusion-weighted imaging and tractography. In this way, a receiver operating characteristic (ROC) curve can be plotted to quantify the tradeoff between sensitivity and specificity as a function of the threshold applied. Example ROC curves are shown in Fig. 7 for mouse (Calabrese et al., 2015) and macaque (Azadbakht et al., 2015) connectomes.

Based on these ROC curves, we can determine the threshold that must be applied to the mouse and macaque connectomes to satisfy the 2:1 rule. While this threshold is unlikely to generalize to other connectomes and species, it provides a ballpark estimate of the extent to which connectomes should be thresholded to ensure FNs are twice more abundant than FPs. To achieve two FNs for each FP, the false positive rate, FPR, and true positive rate, TPR, must satisfy,

$$\frac{\#\mathrm{FN}}{\#\mathrm{FP}} = 2 = \frac{1-\mathrm{TPR}}{(1/d-1)\mathrm{FPR}}, \tag{20}$$

where d denotes the connection density of the ground truth connectome. Solving this equation yields,

$$\mathrm{TPR} = 1-2(1/d-1)\mathrm{FPR}. \tag{21}$$

The line defined by this relation is superimposed on the ROC curves shown in Fig. 7 (green lines), assuming a nominal connection density of $d = 0.3$. The coordinate at which this line intersects each ROC curve corresponds to the threshold where FNs are twice more abundant as FPs. In other words, the 2:1 rule of thumb is satisfied when the threshold is chosen based on the intersection of the green line with the ROC curve. This intersection corresponds to the point along the ROC curve that is closest in Euclidean distance to the $(FPR = 0, TPR = 1)$ coordinate, subject to the constraint that the ratio of FNs to FPs is 2:1.

The intersection points for both of the ROC curves shown in Fig. 7 occur at stringent thresholds that demand elimination of more than 50–60% of all connections. While this might seem excessive, it is in line with the recent findings of Azadbakht et al. (2015), where connectome accuracy was found to be maximized for extremely stringent thresholds that eliminated more than 90% of all connections reconstructed with a well-known probabilistic tractography algorithm. In their study, the optimum threshold was determined by maximizing accuracy, which was classically defined as, $(TP + TN)/(TP + TN + FP + FN)$. Optimizing the choice of threshold to maximize accuracy does not necessarily ensure the 2:1 rule is satisfied, but rather assumes that FPs and FNs are equally detrimental.

Limitations

In addition to the important caveats described above, some general limitations of this study should be considered. First, for reasons of computational tractability, the maximum number of FPs or FNs evaluated did not exceed 10% of the total number of connections in the empirical networks. This limited the maximum rate of simulated FNs to 10%. The validity of our recommendations might therefore deteriorate for highly inaccurate connectomes where error rates significantly exceed 10%. To address this uncertainty, we extended our evaluation to include error rates up to 50% for the human connectome. This was performed using coarser increments with only 20 perturbed networks sampled at each increment. When including these very high error rates, the linearity of the clustering coefficient and efficiency as a function of the number of FPs or FNs decreased, but FPs still remained very significantly more detrimental than FNs.

Second, we intentionally did not normalize any of the network measures. Normalization is typically performed to determine the extent to which empirical networks are more segregated or integrated than appropriately matched random networks. In contrast, the objective here was to determine the absolute change in clustering or efficiency relative to the empirical connectomes. In other words, the empirical networks were the benchmark. Given that adding FPs or FNs increased/decreased the connection density, it might be argued that changes in the clustering coefficient and efficiency simply reflect the dependency of these measures on connection density. However, this cannot be the case. For example, the clustering coefficient generally increases with connection density, whereas Fig. 2 shows that the clustering coefficient decreased with an increasing number of FPs.

Conclusion

Connectome specificity is important—at least twice as important as connectome sensitivity, particularly with respect to analysis of the topological properties of brain networks. In this study, the greater detriment of FPs relative to FNs was demonstrated with detailed empirical evaluations performed in human, mouse and worm connectomes, and verified with asymptotic analysis of small-world networks with idealized modular structure. In modular small-world networks such as the brain, FPs are most likely to emerge between modules, whereas FNs are most likely to emerge within modules. The greater detriment of FPs is due to the relative rarity of inter-modular connections and the greater impact they have on shaping network topology. Connectome reconstruction and thresholding methods should aim to satisfy a 2:1 rule of thumb in which false negatives are twice as abundant as false positives, although this rule may be difficult to achieve in practice given that the ratio of FPs to FNs is typically unknown. Future work should focus on developing methods to estimate this ratio as well as statistical thresholding techniques that are informed by prior information (Meskaldji et al., 2015). In the meantime, we contend that densely reconstructed probabilistic connectomes should be thresholded very stringently to ensure the impact of FPs on graph theoretical analyses of brain networks is minimized. Efforts to maximize the sensitivity of connectome reconstruction should be realigned with the need to map brain networks with high specificity. Cautious application of advanced probabilistic fiber tracking methods and crossing-fiber models to high quality data is recommended, bearing in mind the 2:1 rule of thumb, but blind application of sophisticated models to data of low quality is ill-advised when performing graph theoretical analyses of brain networks.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2016.06.035.

Acknowledgements

A.Z., A.F., L.C., L.L.H and M.B. are supported by the Australian National Health and Medical Research Council (grant identifiers: 1050504, 1066779, 1047648, 1037196, 1103252 and 1099082) and the Australian Research Council (ID: FT130100589, CE140100007). M.P.vdH. is supported by a VENI (451-12-001) grant from the Netherlands Organization for Scientific Research and a fellowship from MQ.

References

Alexander, A.L., Lee, J.E., Lazar, M., Field, A.S., 2007. Diffusion tensor imaging of the brain. Neurotherapeutics 4 (3), 316–329.
Azadbakht, H., Parkes, L.M., Haroon, H.A., Augath, M., Logothetis, N.K., de Crespigny, A., D'Arceuil, H.E., Parker, G.J., 2015. Validation of high-resolution tractography against in vivo tracing in the macaque visual cortex. Cereb. Cortex 25 (11), 4299–4309.
Bassett, D.S., Greenfield, D.L., Meyer-Lindenberg, A., Weinberger, D.R., Moore, S.W., Bullmore, E.T., 2010. Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. PLoS Comput. Biol. 6 (4), e1000748.
Bastiani, M., Shah, N.J., Goebel, R., Roebroeck, A., 2012. Human cortical connectome reconstruction from diffusion weighted MRI: the effect of tractography algorithm. NeuroImage 62 (3), 1732–1749.
Behrens, T.E., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., et al., 2003. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magn. Reson. Med. 50, 1077–1088.
Behrens, T.E., Berg, H.J., Jbabdi, S., Rushworth, M.F., Woolrich, M.W., 2007. Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? NeuroImage 34 (1), 144–155.
Betzel, R.F., Griffa, A., Avena-Koenigsberger, A., Goñi, J., Thiran, J.-P., Hagmann, P., Sporns, O., 2013. Multi-scale community organization of the human structural connectome and its relationship with restingstate functional connectivity. Netw. Sci. 1 (3), 353–373.
Betzel, R.F., Avena-Koenigsberger, A., Goñi, J., He, Y., de Reus, M.A., Griffa, A., Vertes, P.E., Mišić, B., Thiran, J.P., Hagmann, P., van den Heuvel M, Zuo, X.N., Bullmore, E.T., Sporns, O., 2016. Generative models of the human connectome. NeuroImage 124 (Pt A), 1054–1064.
Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. 10, P10008.
Bota, M., Dong, H.-W., Swanson, L.W., 2012. Combining collation and annotation efforts toward completion of the rat and mouse connectomes in BAMS. Front. Neuroinform. 6, 2.
Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10 (3), 186–198.
Bullmore, E., Sporns, O., 2012. The economy of brain network organization. Nat. Rev. Neurosci. 13 (5), 336–349.
Calabrese, E., Badea, A., Cofer, G., Qi, Y., Johnson, G.A., 2015. A diffusion MRI tractography connectome of the mouse brain and comparison with neuronal tracer data. Cereb. Cortex 25 (11), 4628–4637.
Conturo, T.E., Lori, N.F., Cull, T.S., Akbudak, E., Snyder, A.Z., et al., 1999. Tracking neuronal fiber pathways in the living human brain. Proc. Natl. Acad. Sci. U. S. A. 96, 10422–10427.
de Reus, M.A., van den Heuvel, M.P., 2013. Estimating false positives and negatives in brain networks. NeuroImage 70, 402–409.
Dyrby, T.B., Sogaard, L.V., Parker, G.J., Alexander, D.C., Lind, N.M., Baare, W.F., Hay-Schmidt, A., Eriksen, N., Pakkenberg, B., Paulson, O.B., Jelsing, J., 2007. Validation of in vitro probabilistic tractography. NeuroImage 37, 1267–1277.
Ercsey-Ravasz, M., Markov, N.T., Lamy, C., Van Essen, D.C., Knoblauch, K., Toroczkai, Z., Kennedy, H., 2013. A predictive network model of cerebral cortical connectivity based on a distance rule. Neuron 80, 184–197.

Fillard, P., Descoteaux, M., Goh, A., Gouttard, S., Jeurissen, B., Malcolm, J., Ramirez-Manzanares, A., Reisert, M., Sakaie, K., Tensaouti, F., Yo, T., Mangin, J.F., Poupon, C., 2011. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom. NeuroImage 56 (1), 220–234.
Fornito, A., Zalesky, A., Breakspear, M., 2013. Graph analysis of the human connectome: promise, progress, and pitfalls. NeuroImage 80, 426–444.
Fornito, A., Zalesky, A., Bullmore, E., 2016. Fundamentals of Brain Network Analysis. Elsevier.
Ginestet, C.E., Nichols, T.E., Bullmore, E.T., Simmons, A., 2011. Brain network analysis: separating cost from topology using cost-integration. PLoS One 6 (7), e21570.
Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., WU-Minn HCP Consortium, 2013. The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage 80, 105–124.
Goñi, J., van den Heuvel, M.P., Avena-Koenigsberger, A., Velez de Mendizabal, N., Betzel, R.F., Griffa, A., Hagmann, P., Corominas-Murtra, B., Thiran, J.P., Sporns, O., 2014. Resting-brain functional connectivity predicted by analytic measures of network communication. Proc. Natl. Acad. Sci. U. S. A. 111 (2), 833–838.
Good, B.H., de Montjoye, Y.A., Clauset, A., 2010. Performance of modularity maximization in practical contexts. Phys. Rev. E 81, 046106.
Griffa, A., Baumann, P.S., Thiran, J.P., Hagmann, P., 2013. Structural connectomics in brain diseases. NeuroImage 80, 515–526.
Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., Sporns, O., 2008. Mapping the structural core of human cerebral cortex. PLoS Biol. 6, e159.
Iturria-Medina, Y., Sotero, R.C., Canales-Rodriguez, E.J., Aleman-Gomez, Y., Melie-Garcia, L., 2008. Studying the human brain anatomical network via diffusion-weighted MRI and Graph Theory. NeuroImage 40 (3), 1064–1076.
Jbabdi, S., Woolrich, M.W., Andersson, J.L., Behrens, T.E., 2007. A Bayesian framework for global tractography. NeuroImage 37 (1), 116–129.
Jbabdi, S., Sotiropoulos, S.N., Haber, S.N., Van Essen, D.C., Behrens, T.E., 2015. Measuring macroscopic brain connections in vivo. Nat. Neurosci. 18 (11), 1546–1555.
Jones, D.K., 2010. Challenges and limitations of quantifying connectivity in the human brain in vivo with diffusion MRI. Imaging Med. 2, 341–355.
Jones, D.K., Knosche, T.R., Turner, R., 2013. White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. NeuroImage 73, 239–254.
Kaden, E., Knosche, T.R., Anwander, A., 2007. Parametric spherical deconvolution: inferring anatomical connectivity using diffusion MR imaging. NeuroImage 37, 474–488.
Knosche, T.R., Anwander, A., Liptrot, M., Dyrby, T.B., 2015. Validation of tractography: comparison with manganese tracing. Hum. Brain Mapp. 36 (10), 4116–4134.
Koch, M.A., Norris, D.G., Hund-Georgiadis, M., 2002. An investigation of functional and anatomical connectivity using magnetic resonance imaging. NeuroImage 16, 241–250.
Lancichinetti, A., Fortunato, S., 2012. Consensus clustering in complex networks. Sci. Report. 2, 336.
Latora, V., Marchiori, M., 2001. Efficient behavior of small-world networks. Phys. Rev. Lett. 87 (19), 198701.
Li, L., Rilling, J.K., Preuss, T.M., Glasser, M.F., Hu, X., 2012. The effects of connection reconstruction method on the interregional connectivity of brain networks via diffusion tractography. Hum. Brain Mapp. 33 (8), 1894–1913.
Markov, N.T., Ercsey-Ravasz, M., Lamy, C., Gomes, A.R.R., Magrou, L., Misery, P., Giroud, P., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D.C., Kennedy, H., 2013. The role of long-range connections on the specificity of the macaque interareal cortical network. Proc. Natl. Acad. Sci. U. S. A. 110, 5187–5192.
Markov, N.T., Ercsey-Ravasz, M.M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D.C., Kennedy, H., 2014. A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cereb. Cortex 24, 17–36.
Meskaldji, D.E., Vasung, L., Romascano, D., Thiran, J.P., Hagmann, P., Morgenthaler, S., Van De Ville, D., 2015. Improved statistical evaluation of group differences in connectomes by screening-filtering strategy with application to study maturation of brain connections between childhood and adolescence. NeuroImage 108, 251–264.
Meunier, D., Lambiotte, R., Bullmore, E.T., 2010. Modular and hierarchically modular organization of brain networks. Front. Neurosci. 4, 200.
Mišić, B., Betzel, R.F., Nematzadeh, A., Goñi, J., Griffa, A., Hagmann, P., Flammini, A., Ahn, Y.-Y., Sporns, O., 2015. Cooperative and competitive spreading dynamics on the human connectome. Neuron 86 (6), 1518–1529.
Mori, S., Crain, B.J., Chacko, V.P., van Zijl, P.C.M., 1999. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Ann. Neurol. 45, 265–269.
Newman, M.E.J., 2006. Modularity and community structure in networks. Proc. Natl. Acad. Sci. U. S. A. 103 (23), 8577–8582.
Oh, S.W., Harris, J.A., Ng, L., et al., 2014. A mesoscale connectome of the mouse brain. Nature 508, 207–214.
Pestilli, F., Yeatman, J.D., Rokem, A., Kay, K.N., Wandell, B.A., 2014. Evaluation and statistical inference for human connectomes. Nat. Methods 11 (10), 1058–1063.
Rand, W.M., 1971. Objective criteria for the evaluation of clustering methods. J. Am. Stat. Assoc. 66 (336), 846–850.
Reveley, C., Seth, A.K., Pierpaoli, C., Silva, A.C., Yu, D., Saunders, R.C., Leopold, D.A., Ye, F.Q., 2015. Superficial white matter fiber systems impede detection of long-range cortical connections in diffusion MR tractography. Proc. Natl. Acad. Sci. U. S. A. 112 (21), E2820–E2828.
Roberts, J.A., Perry, A., Lord, A.R., Roberts, G., Mitchell, P.B., Smith, R.E., Calamante, F., Breakspear, M., 2016. The contribution of geometry to the human connectome. NeuroImage 124 (Pt A), 379–393.
Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. NeuroImage 52 (3), 1059–1069.
Shih, C.T., Sporns, O., Yuan, S.L., Su, T.S., Lin, Y.J., Chuang, C.C., Wang, T.Y., Lo, C.C., Greenspan, R.J., Chiang, A.S., 2015. Connectomics-based analysis of information flow in the Drosophila brain. Curr. Biol. 25 (10), 1249–1258.
Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2013. SIFT: spherical-deconvolution informed filtering of tractograms. NeuroImage 67, 298–312.
Sotiropoulos, S.N., Bai, L., Morgan, P.S., Constantinescu, C.S., Tench, C.R., 2010. Brain tractography using Q-ball imaging and graph theory: improved connectivities through fibre crossings via a model-based approach. NeuroImage 49 (3), 2444–2456.
Sporns, O., 2011. Networks of the Brain. MIT Press.
Sporns, O., Betzel, R.F., 2015. Modular brain networks. Annu. Rev. Psychol. 67 (1).
Sporns, O., Chialvo, D.R., Kaiser, M., Hilgetag, C.C., 2004. Organization, development and function of complex brain networks. Trends Cogn. Sci. 8 (9), 418–425.
Sporns, O., Tononi, G., Kotter, R., 2005. The human connectome: a structural description of the human brain. PLoS Comput. Biol. 1, e42.
Thomas, C., Ye, F.Q., Irfanoglu, M.O., Modi, P., Saleem, K.S., Leopold, D.A., Pierpaoli, C., 2014. Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. Proc. Natl. Acad. Sci. U. S. A. 111 (46), 16574–16579.
Tournier, J.D., Yeh, C.H., Calamante, F., Cho, K.H., Connelly, A., Lin, C.P., 2008. Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data. NeuroImage 42 (2), 617–625.
Tournier, J.-D., Calamante, F., Connelly, A., 2012. MRtrix: diffusion tractography in crossing fiber regions. Int. J. Imaging Syst. Technol. 22, 53–66.
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15 (1), 273–289.
Van den Heuvel, M.P., Kahn, R.S., Goni, J., Sporns, O., 2012. High-cost, high-capacity backbone for global brain communication. Proc. Natl. Acad. Sci. U. S. A. 109 (28), 11372–11377.
van den Heuvel, M.P., de Reus, M.A., Feldman Barrett, L., Scholtens, L.H., Coopmans, F.M., Schmidt, R., Preuss, T.M., Rilling, J.K., Li, L., 2015. Comparison of diffusion tractography and tract-tracing measures of connectivity strength in rhesus macaque connectome. Hum. Brain Mapp. 36 (8), 3064–3075.
van den Heuvel, M.P., Bullmore, E.T., Sporns, O., 2016. Comparative connectomics. Trends Cogn. Sci. 20 (5), 345–361.
van Wijk, B.C., Stam, C.J., Daffertshofer, A., 2010. Comparing brain networks of different size and connectivity density using graph theory. PLoS One 5 (10), e13701.
Varshney, L.R., Chen, B.L., Paniagua, E., Hall, D.H., Chklovskii, D.B., 2011. Structural properties of the *C. elegans* neuronal network. PLoS Comput. Biol. 7, e1001066.
Watts, D.J., Strogatz, S., 1998. Collective dynamics of small-world networks. Nature 393 (6684), 440–442.
White, J.G., Southgate, E., Thomson, J.N., Brenner, S., 1986. The structure of the nervous system of the nematode *Caenorhabditis elegans*. Philos. Trans. R. Soc. Lond. B 314, 1–340.
Zaborszky, L., Wouterlood, F.G., Lanciego, J.L. (Eds.), 2006. Neuroanatomical Tract-Tracing: Molecules, Neurons, and Systems. Springer.
Zalesky, A., Fornito, A., 2009. A DTI-derived measure of cortico-cortical connectivity. IEEE Trans. Med. Imaging 28 (7), 1023–1036.
Zalesky, A., Fornito, A., Harding, I.H., Cocchi, L., Yucel, M., Pantelis, C., Bullmore, E.T., 2010a. Whole-brain anatomical networks: does the choice of nodes matter? NeuroImage 50 (3), 970–983.
Zingg, B., Hintiryan, H., Gou, L., Song, M.Y., Bay, M., Bienkowski, M.S., Foster, N.N., Yamashita, S., Bowman, I., Toga, A.W., Dong, H.W., 2014. Neural networks of the mouse neocortex. Cell 156 (5), 1096–1111.