ARTICLE IN PRESS

Statistics and Probability Letters xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Statistical methods and challenges in connectome genetics

Dustin Pluta [a], Zhaoxia Yu [a], Tong Shen [a], Chuansheng Chen [b], Gui Xue [c], Hernando Ombao [a,d,*]

[a] Department of Statistics, University of California, Irvine, USA
[b] Department of Psychology and Social Behavior, University of California, Irvine, USA
[c] State Key Laboratory of Cognitive Neuroscience and Learning and IDG/McGovern Institute for Brain Research, Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing, PR China
[d] Biostatistics Group, King Abdullah University of Science and Technology, Saudi Arabia

ARTICLE INFO

Article history:
Available online xxxx

ABSTRACT

The study of genetic influences on brain connectivity, known as *connectome genetics*, is an exciting new direction of research in imaging genetics. We here review recent results and current statistical methods in this area, and discuss some of the persistent challenges and possible directions for future work.

1. Introduction

The human brain is one of the most intensely studied objects in modern science, yet much about its structure and function is still poorly understood. Despite the efforts of generations of researchers, and the creation of numerous disciplines specializing in particular aspects of the brain, the complexity of neurophysiology and human behavior have stymied attempts to develop a comprehensive model to describe brain structure, brain function, and cognition. Imaging genetics is a relatively recent direction of study, which seeks to determine the genetic factors that significantly influence the brain and behavior. Initially motivated by the study of neurological diseases such as Alzheimer's disease (AD) and schizophrenia, the scope of imaging genetics has been expanded to consider nearly all aspects of the brain and cognition.

We here review recent developments in statistical methods for imaging genetics, with a specific focus on the analysis of genetic influences on brain connectivity, sometimes referred to as "connectome genetics". To date, relatively few studies have considered the genetic influences of brain connectivity, although there is growing interest in this area due to continued findings supporting the importance of brain connectivity for many aspects of cognition. An important survey of this topic was given a few years ago in Thompson et al. (2013), which introduces the essential concepts and modeling approaches for connectome genetics, and summarizes the key results in the field up to that point. Following the comprehensive exposition given there, this review presents a selection of recent methods and results in connectome genetics. Interested readers may also refer to the contemporaneous review by Nathoo et al. (2017) which considers other general statistical methods for imaging genetics.

2. Brain connectivity

Commonly used technologies in current neuroimaging studies include structural and functional magnetic resonance imaging (MRI and fMRI), electroencephalogram (EEG), magnetoencephalogram (MEG), and diffusion tensor imaging (DTI).

* Corresponding author at: Biostatistics Group, King Abdullah University of Science and Technology, Saudi Arabia.
E-mail address: hernando.ombao@kaust.edu.sa (H. Ombao).

https://doi.org/10.1016/j.spl.2018.02.048
0167-7152/

Since each of these imaging modalities measures a different set of neural features or phenomena, they may differ in their effectiveness in answering particular scientific questions. For instance, fMRI measures changes in the blood oxygen level over time in over 5×10^5 voxels throughout the brain (Lindquist, 2008), with very high spatial resolution (typically less than 3 mm), but poor temporal resolution (around 2s). In contrast, EEG measures the collective cortical activity of populations of neurons via 64–256 electrodes placed on the scalp (Ombao et al., 2016), allowing for a high temporal resolution (1 kHz) at the cost of spatial resolution. Instead of measuring dynamic brain activity, structural imaging methods provide information about the physical relationship of brain regions, such as DTI, which measures the structure and orientation of the brain's white matter fiber tracks via the diffusion of water molecules through the entire brain volume (Basser et al., 1994).

It is now generally understood that higher-level cognitive processing (e.g., in memory retrieval, decision making) critically depends on the interaction and transfer of information between many localized regions. Numerous different methods for calculating brain connectivity have been proposed, but the reliability, interpretation, and relationship of these different measures is not well established (Fiecas et al., 2013). There are three concepts of brain "connectivity" that have been of interest, namely *structural*, *functional*, and *effective* connectivities. Structural connectivity refers to the anatomical connections between brain regions, measured using DTI or structural MRI. Functional connectivity is a symmetric and undirected measure of concordant activity between brain regions, commonly calculated as Pearson's correlation, partial correlation, coherence, or partial coherence between the activation signals of two regions (Fiecas and Ombao, 2011). Modalities used for functional connectivity studies include fMRI, EEG, and MEG. By contrast, effective connectivity is a directed measure of how past activity in one region may influence the future activity of another region. Effective connectivity is closely related to Granger causality and is often estimated with a vector autoregressive (VAR) model (Gorrostieta et al., 2013; Chiang et al., 2017).

In studying the etiology of neurological diseases such as Alzheimer's, schizophrenia, and autism spectrum disorders, significant links with these forms of brain connectivity have been consistently shown (Woodward and Cascio, 2015). There is also evidence that dynamic characteristics of brain connectivity, and not just global measures of connectivity over an entire experiment, are also important for understanding brain function. Formally defining and modeling dynamic connectivity has motivated the development of new statistical methods. Current approaches for modeling dynamic connectivity include sliding window methods (e.g., Chang and Glover, 2010), graphical Bayesian modeling (Warnick et al., 2017) change-point detection via VAR models (Kirch et al., 2015), and hidden Markov switching-VAR models (Samdin et al., 2017). A full treatment of current dynamic connectivity methods is outside the scope of this review, but we briefly describe recent work in this area by our collaborators. Ting et al. (2017) proposed modeling dynamic connectivity with a so-called "regime-switching" factor model (SVAR). A three-step procedure for fitting this model (for a single subject) is described. (1) Compute initial estimates of connectivity subspaces shared across regimes using a stationary factor model; (2) apply a factor SVAR model, and identify regime boundaries with a switching Kalman filter to partition the neural signal into a small number of distinct states; (3) finally, use the low-dimensional factor representations for each regime to estimate within-regime effective connectivity. This model is able to detect abrupt changes in mental state, such as might occur in an experiment with varied cognitive demands, and is also able to estimate recurring states over the course of the experiment.

3. Imaging genetics

Human traits, including neural, cognitive, and behavioral traits, are shaped by genetics and environmental factors through a number of complex biological processes. In the past few decades, single nucleotide polymorphisms (SNPs) have become one of the most commonly measured forms of genetic data, as they are abundant (about 10 million in the human genome), stable, and easy to measure. Each SNP is a single DNA location (nucleotide) that exhibits relatively high variation between individuals, and are measured as *alleles*, which are typically encoded as a 0, 1, and 2 respectively, and treated numerically. Many complex traits are believed to depend on higher-order interactions between collections of SNPs, but determining these relationships is often difficult, and in practice it is often necessary to restrict attention to only the additive effects of SNPs. The proportion of variance of an observed trait that is accounted for by the additive effect of genome-wide SNPs is referred to as *narrow-sense heritability*, denoted h^2.

Some hallmark successes in genetics have come from the study of prominent neurological and psychological diseases, such as Parkinson's disease, Alzheimer's disease, and schizophrenia. For each of these diseases, many significant genetic risk factors of relatively large effect have been found, which have in turn led to improved understanding of the mechanisms of action for these diseases, and have even resulted in promising gene therapies for Parkinson's disease (Palfi et al., 2014). Initially motivated by these promising results, the field of imaging genetics has expanded to now consider a number of different neural disorders and many facets of cognition in healthy-individuals. We mention a few recent results to illustrate the breadth of current research. The study by Bohlken et al. (2016) found that global reductions in white matter, which underpin structural connectivity, are largely explained by genetic risk factors for schizophrenia. In very recent work by Sudre et al. (2017), a study of families with histories of attention-deficit/hyperactivity disorder (ADHD) found significant genetic heritability for the default mode, cognitive control, and ventral attention networks. An analysis of 161 twin pairs found significant heritability of anatomical connectivity measured with DTI, and identified significant differences in heritability across subnetworks (with average heritability over all white matter tracts estimated at around 30% (Shen et al., 2014). With data from 1320 unrelated, young, healthy adults, Ge et al. (2016) demonstrated that several characteristics of brain structure are genetically heritable, such as brain volume and neuroanatomical shape. Using resting-state fMRI data from twins and non-twin siblings, Vidaurre et al. (2017) found evidence of significant heritability for certain characteristics of dynamic connectivity patterns; this is one of the few studies to date to have considered the heritability of dynamic connectivity.

Please cite this article in press as: Pluta D., et al., Statistical methods and challenges in connectome genetics. Statistics and Probability Letters (2018), https://doi.org/10.1016/j.spl.2018.02.048.