NeuroImage 145 (2017) 137–165

Contents lists available at ScienceDirect

NeuroImage

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

Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls

Mohammad R. Arbabshirani [a,b,*], Sergey Plis [a], Jing Sui [a,c], Vince D. Calhoun [a,d]

[a] The Mind Research Network, Albuquerque, NM 87106, USA
[b] Geisinger Health System, Danville, PA 17822, USA
[c] Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[d] Department of ECE, University of New Mexico, Albuquerque, NM, USA

ARTICLE INFO

Article history:
Accepted 25 February 2016
Available online 21 March 2016

Keywords:
Neuroimaging
Machine learning
Classification
Brain disorders
Prediction

ABSTRACT

Neuroimaging-based single subject prediction of brain disorders has gained increasing attention in recent years. Using a variety of neuroimaging modalities such as structural, functional and diffusion MRI, along with machine learning techniques, hundreds of studies have been carried out for accurate classification of patients with heterogeneous mental and neurodegenerative disorders such as schizophrenia and Alzheimer's disease. More than 500 studies have been published during the past quarter century on single subject prediction focused on a multiple brain disorders. In the first part of this study, we provide a survey of more than 200 reports in this field with a focus on schizophrenia, mild cognitive impairment (MCI), Alzheimer's disease (AD), depressive disorders, autism spectrum disease (ASD) and attention-deficit hyperactivity disorder (ADHD). Detailed information about those studies such as sample size, type and number of extracted features and reported accuracy are summarized and discussed. To our knowledge, this is by far the most comprehensive review of neuroimaging-based single subject prediction of brain disorders. In the second part, we present our opinion on major pitfalls of those studies from a machine learning point of view. Common biases are discussed and suggestions are provided. Moreover, emerging trends such as decentralized data sharing, multimodal brain imaging, differential diagnosis, disease subtype classification and deep learning are also discussed. Based on this survey, there is extensive evidence showing the great potential of neuroimaging data for single subject prediction of various disorders. However, the main bottleneck of this exciting field is still the limited sample size, which could be potentially addressed by modern data sharing models such as the ones discussed in this paper. Emerging big data technologies and advanced data-intensive machine learning methodologies such as deep learning have coincided with an increasing need for accurate, robust and generalizable single subject prediction of brain disorders during an exciting time. In this report, we survey the past and offer some opinions regarding the road ahead.

Introduction

Neuroimaging has opened up an exciting non-invasive window into the human brain over the past few decades. This interdisciplinary field has attracted scientists from areas such as medicine, engineering, mathematics, physics, statistics, computer science, and psychology (Epstein et al., 2001). Imaging modalities such as magnetic resonance imaging (MRI) and magnetoencephalography (MEG) along with more traditional methods such as electroencephalography (EEG) have made it possible to non-invasively study various aspects of the human brain with unprecedented accuracy. MRI-related techniques such as structural MRI (sMRI), functional MRI (fMRI) and diffusion MRI (dMRI) have the benefit of providing localized spatial information about the brain structure and function as well as detailed functional and structural connectivity maps. These techniques have provided new insight into the human brain and have brought hope to researchers trying to unravel the secrets of one of the most complex systems in the universe, the human brain.

Structural MRI has made it possible to visualize the brain at high spatial resolution (one cubic millimeter or less) (Liang and Lauterbur, 2000). sMRI high resolution images of the brain are ideal for studying various brain structures and also for detecting physical abnormalities, lesions and damages. dMRI is an imaging technique for visualization of anatomical connections between different brain regions (Le Bihan et al., 2001; Merboldt et al., 1985). Functional MRI measures brain activity by detecting changes in the blood oxygenation (DeYoe et al., 1994; Ogawa et al., 1990). fMRI makes it possible to study

* Corresponding author at: The Mind Research Network, 1101 Yale Blvd., NE, Albuquerque, NM 87106, USA.

http://dx.doi.org/10.1016/j.neuroimage.2016.02.079
1053-8119/

functional regions and networks of the brain as well as temporal associations among them.

Unfortunately, brain disorders are major health problems in the US and the rest of the world that not only impair the lives of millions of people but also impose huge financial burdens on societies (DiLuca and Olesen, 2014; Ernst and Hay, 1994; Rice, 1999). Moreover, there are no clinical tests to identify many brain disorders such as schizophrenia. One of the major hopes underlying the advanced neuroimaging tools mentioned above is to provide new understanding of brain disorders such as schizophrenia, bipolar disorder, autism spectrum disorder (ASD), Alzheimer's disease (AD), major depressive disorders, attention-deficit hyperactivity disorder (ADHD) and mild cognitive impairment (MCI). Brain disorder research aims at understanding the impact of each disease on the brain's function and structure from the cellular to system level, as well as the pathogenesis of these complex disorders. As a result, thousands of studies have been published on different aspects of brain disorders to show aberrations of some features (structural or functional) in a patient group usually in comparison with a healthy cohort (Jack et al., 1997; Jafri et al., 2008; Lorenzetti et al., 2009; McAlonan et al., 2005). While these studies are valuable in terms of finding relevant disease biomarkers, they are not sufficient for direct clinical diagnostic/prognostic adoption. The main reason is that many of these findings are statistically significant at the group level, but the individual discrimination ability of the proposed biomarkers is not typically evaluated. Since classification provides information for each individual subject, it is considered a much harder task than reporting group differences.

In recent years, there has been a growing trend in designing neuroimaging-based prognostic/diagnostic tools. As a result, there have been a lot of efforts using neuroimaging methods to automatically discriminate patients with brain disorders from healthy control or from each other (Klöppel et al., 2012). Many of these studies have reported promising prediction performances with the claim that complex diseases can be diagnosed robustly, accurately and rapidly in an automatic fashion. However, until now, these tools have not been integrated into the clinical realm. We believe that the main reason for this is that many of the studies of this nature, despite the promising results on a specific research dataset, are not designed to generalize to other datasets, specifically the clinical ones.

The purpose of this study was two-fold. First, we reviewed a large number of MRI-based brain disorder diagnostic/prognostic studies in schizophrenia, ASD, ADHD, depressive disorder, MCI and Alzheimer's disease. These studies are compared in a number of key aspects such as type of features, classifier and reported accuracies. Next, we formed our opinion on the issues associated with how machine learning is applied in neuroimaging and have suggested solutions that might address these pitfalls. Considering the immense potential of neuroimaging tools for clinical adoption, careful implementation and interpretation of machine learning in neuroimaging is crucial. Machine learning is a relatively new domain for many neuroimaging researchers coming from other fields and therefore pitfalls are unfortunately not rare. We attempt to identify and emphasize some common mistakes that resulted in these shortcomings and biases. At the end, we discuss emerging trends in neuroimaging such as data sharing, multimodal brain imaging and differential diagnosis.

Group difference vs. classification

As pointed out in the Introduction section, many brain disorder studies have shown abnormality in the average sense in one or more brain features in a patient cohort in comparison with a healthy group using statistical tests. The success of such methodology is usually measured by the means of p-values. On the other hand, the goal of single subject prediction is to automatically classify each subject into one of the groups in the study (e.g., healthy vs. patient). The success of classification studies is usually measured by accuracy.

These two problems are very different in essence as they try to address distinct research questions. In general, showing group differences is much easier compared to single subject prediction. To better illustrate the difference between these types of analysis, we show an example in Fig. 1. Suppose there are two groups each with 100 samples (subjects) and we have measurements of one brain feature for each subject. Fig. 1A shows a case where the mean values of the two groups are different as measured by a two-sample t-test. The difference is statistically significant (p-value = 0.001). However, if one tries to classify subjects based on a threshold on this brain feature (the dotted red line placed between the mean of two groups), a weak classification rate of 60.0% will be achieved. The reason for this is the range of values for that specific feature is highly overlapping for the two groups. So, a highly significant group difference does not necessarily translate into a strong classification result. But the opposite is also true, as high classification based on a feature doesn't necessarily mean that group-level mean differences exist. Fig. 1B shows a case where the two-sample t-test on the two groups is not significant (p-value = 0.86) but the classification based on two thresholds (red dotted lines placed between each mode of group 2 and mean of group 1) is very strong (94.5%). In this case, the abnormality is bidirectional, which does not cause significant mean differences but makes it possible to separate the groups with two thresholds (dotted lines). Interestingly, bidirectional abnormalities are observed in neuroimaging studies (Arbabshirani and Calhoun, 2011; Calhoun et al., 2006b). Fig. 1C shows a case where strong group differences and successful classification go hand in hand. The abnormality is one-directional and the mean difference is very significant (p-value < 2e−16). The mean of two groups is so far apart that the values of most of the samples of the two groups do not overlap. Therefore, a strong classification rate of 93.5% is achieved (based on one threshold).

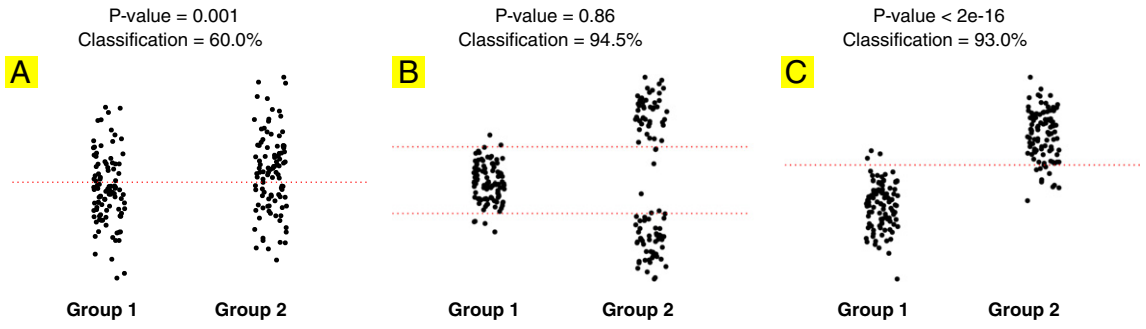

Fig. 1. Comparison of group difference analysis and classification in three different scenarios using toy data. Group difference is analyzed by two-sample t-tests and classification is performed by simple thresholding (red dotted lines). Each group/class has 100 samples. A: Significant group difference (p-value < 0.001) but poor classification (60.0%). B: Insignificant group difference (p-value = 0.865) but high classification accuracy (94.5%). C: Significant group difference (p-value < 2e−16) and high classification accuracy (93.0%). Significant group difference doesn't necessarily cause high classification and vice versa.