Journal of Neuroscience Methods 235 (2014) 1–9

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Computational Neuroscience

Creating multimodal predictors using missing data: Classifying and subtyping autism spectrum disorder

Madhura Ingalhalikar[a,*], William A. Parker[a], Luke Bloy[b], Timothy P.L. Roberts[b], Ragini Verma[a]

[a] Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA
[b] Lurie Family Foundations MEG Imaging Center, Department of Radiology, Children's Hospital of Philadelphia, Philadelphia, PA, USA

HIGHLIGHTS

- Novel classifiers for multimodality imaging data that can handle missing data.
- Solves the real life issue of incomplete studies/missing data in clinical studies.
- Shows clear distinction in classifying autism spectrum disorder from normal controls as well as in stratifying Autism based on language impairment based on MEG and DTI features.
- Ranks and specifies the distinctive features involved in the classification for further investigations.

ARTICLE INFO

Article history:
Received 8 January 2014
Received in revised form 28 May 2014
Accepted 23 June 2014
Available online 28 June 2014

Keywords:
Magnetoencephalography
Diffusion tensor imaging
Autism spectrum disorder
Pattern classification
Language impairment
Missing data

ABSTRACT

Background: Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by wide range of symptoms and severity including domains such as language impairment (LI). This study aims to create a quantifiable marker of ASD and a stratification marker for LI using multimodality imaging data that can handle missing data by including subjects that fail to complete all the aspects of a multimodality imaging study, obviating the need to remove subjects with incomplete data, as is done by conventional methods.

Methods: An ensemble of classifiers with several subsets of complete data is employed. The outputs from such subset classifiers are fused using a weighted aggregation giving an aggregate probabilistic score for each subject. Such fusion classifiers are created to obtain a marker for ASD and to stratify LI using three categories of features, two extracted from separate auditory tasks using magnetoencephalography (MEG) and the third extracted from diffusion tensor imaging (DTI).

Results: A clear distinction between ASD and neurotypical controls (5-fold accuracy of 83.3% and testing accuracy of 87%) and between ASD/+LI and ASD/−LI (5-fold accuracy of 70.1% and testing accuracy of 61.1%) was obtained. One of the MEG features, mismatch field (MMF) latency contributed the most to group discrimination, followed by DTI features from superior temporal white matter and superior longitudinal fasciculus as determined by feature ranking.

Comparison with existing methods: Higher classification accuracy was achieved in comparison with single modality classifiers.

Conclusion: This methodology can be readily applied in large studies where high percentage of missing data is expected.

* Corresponding author at: Section of Biomedical Image Analysis, Department of Radiology, 3600 Market St., Suite 380, Philadelphia, PA 19104, USA. Tel.: +1 215 662 7099.
E-mail addresses: Madhura.Ingalhalikar@uphs.upenn.edu (M. Ingalhalikar), William.Parker@uphs.upenn.edu (W.A. Parker), Luke.Bloy@uphs.upenn.edu (L. Bloy), robertstim@email.chop.edu (T.P.L. Roberts), Ragini.Verma@uphs.upenn.edu (R. Verma).

http://dx.doi.org/10.1016/j.jneumeth.2014.06.030
0165-0270/

1. Introduction

Novel techniques for creating imaging based biomarkers for ASD that can be clinically useful in complementing diagnostic and other cognitive markers, have gained substantial importance in the past few years (Ecker et al., 2010a,b; Ecker et al., 2010a,b; Lange et al., 2010; Bosl et al., 2011; Ingalhalikar et al., 2011; Tsiaras et al., 2011). Beyond diagnosis, which is made by expert and experienced clinicians, such markers have important additional potential utility in clinical assessment of ASD by providing a quantifiable score for each subject, reflecting the extent of pathology over the wide etiological spectrum. Furthermore, such classifiers have the capability to identify key anatomic/functional substrates and circuitry elucidating the neuropathology. However, these methods classify the subject only in two categories: ASD or typically developing (TD), therefore not taking into account the fact that ASD in general consists of a broad set with multiple etiologies captured in the same diagnostic category. In other words, classifiers to this date are not developed for stratifying the heterogeneous ASD population into more homogenous subgroups. For example, language abilities in ASD are highly variable with difficulties ranging from mild to severe impairment in social communication with a subset of individuals having characteristic language impairment (ASD/+LI) demonstrated via delay or absence of spoken language (Kjelgaard and Tager-Flusberg, 2001).

Classification techniques have been mainly based on single imaging modality or a single measure derived from advanced modalities like diffusion tensor imaging (DTI), magnetoencephalography (MEG), electroencephalography (EEG) or functional MRI (fMRI) that cannot individually provide a comprehensive brain-level characterization of ASD or even a symptomatic subgroup such as LI. This necessitates the creation of composite multivariate imaging profiles representative of the underlying pathology, by combining multiple "weak" effects. Multimodality classifiers can therefore aid in exploring multivariate dimensions of pathology patterns and provide a rich multiparametric marker with increased accuracy. However, multivariate population profiles are challenging to create, especially in ASD populations, due to implicit disorder and development induced heterogeneity. A few recent studies have attempted to perform such analysis on other neuro-psychiatric disorders. For example, a simple technique of combining features from different modalities was implemented by Wang et al. (2012), while more advanced methods were employed that included multi-parametric analysis using multi-kernel learning (Zhang et al., 2011) and tensor factorization (Batmanghelich et al., 2011).

However, in ASD, young children are often unable to complete the full scanning protocol or may undergo severely confounding motion in the scanner, plaguing these clinical studies with incomplete data thus reducing the sample size for any traditional and/or learning paradigm. This limits the effectiveness of the multimodal approach since the probability of a subject being excluded increases with the number of modalities utilized in the classifier. Furthermore, if the missing data is associated with a severity of pathology, which is widely true in ASD, the learnt classifier is not representative of the more severe forms of pathology that includes e.g. ASD/+LI. In this work, we address these issues by creating markers from partial data, by building an ensemble of classifiers based on different modalities, combining information from subjects with missing data, so that the variation in the population including LI is learned with maximal data utilization.

Imaging based diagnostic and prognostic classifiers for ASD have been constructed over features that include the volume and other structural features captured from T1 MRI images (Ecker et al., 2010a,b; Ecker et al., 2010a,b; Uddin et al., 2011) using high-dimensional multivariate learning algorithms like support vector machines (SVMs) and searchlight techniques. Structural white matter changes have been captured via DTI based classifiers resulting in high accuracy in characterizing the subject over the patient-control spectrum. For example, the study by Lange et al. implemented features derived from DTI only from superior temporal gyrus and the temporal pole (Lange et al., 2010) and achieved ∼94% classification accuracy. In contrast to the study by Lange et al., another recent study employed whole brain DTI features that not only could classify the patients from controls but also aided in understanding the anatomical changes occurring in ASD patients (Ingalhalikar et al., 2011). Further, functional differences between ASD and the controls were captured via fMRI where in domains such as social interaction or language tasks, ASDs displayed atypical recruitment of brain regions (Critchley et al., 2000; Perkins et al., 2010). Studies have also shown atypical changes occurring during the resting state (Paakki et al., 2010; Weng et al., 2010). Such resting state differences were quantified on a subject by subject basis via multivariate pattern classifiers (Anderson et al., 2011). Non-MRI functional modalities like MEG and EEG have also been employed in analysis of ASD. These modalities offer high temporal resolution and are of interest as ASD symptoms are increasingly thought to be due to a disruption in the excitatory/inhibitory balance of neural activity (Hughes, 2008). Classifiers that are highly predictive of being a control or ASD patient were created based on MEG recordings (Tsiaras et al., 2011) and EEG features (Bosl et al., 2011). However, all these aforementioned classification studies have concentrated on information from single modality, neither tackling the incomplete data issues nor endeavoring to stratify the ASD population into homogenous subgroups.

This work aims at classifying and subgrouping ASD using spatio-temporal multimodal imaging data whilst addressing the practical issues of small sample size, population heterogeneity and missing data that is representative of clinical studies. In our work, we use temporal characteristics derived from MEG based auditory tasks (Roberts et al., 2010, 2011) and spatial anisotropy and diffusivity measures from DTI that are associated with language impairment (LI) in ASD. Our classifiers not only aid the ASD diagnosis but also stratify the heterogeneous ASD population based on language impairment (LI). Thus we employ classifiers that can (i) discriminate each subject as ASD or TD, as well as classify the subjects with ASD on LI spectrum (ii) include all the subjects even with partial missing data (iii) assign abnormality scores to a subject that is more representative of the underlying pathology, as no data has been discarded and (iv) identify the optimal combination of features and of the modalities that can best describe the pathological underpinnings.

2. Methods

2.1. Dataset and pre-processing

Subjects with ASD were recruited from the Regional Autism Center of the Children's Hospital of Philadelphia and the Neuropsychiatry program of the Department of Psychiatry of the University of Pennsylvania School of Medicine. All children screened for inclusion in the ASD sample had a prior ASD diagnosis made by an expert clinician based on extensive clinical interview, documentation of DSM-IV criteria for ASD, and use of various ASD diagnostic tools, such as the Childhood Autism Rating Scale and the Autism Diagnostic Observation Schedule (ADOS). To evaluate presence of LI, all subjects were tested using the Clinical Evaluation of Language Fundamentals-Fourth Edition (CELF-4) (Semel et al., 2003). The ASD group with language impairment (ASD/+LI) was comprised of subjects with a CELF-4 core language score below the 16th percentile.