



Contents lists available at ScienceDirect

Schizophrenia Research

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

Convergence and divergence of neurocognitive patterns in schizophrenia and depression

Sugai Liang^{a,b}, Matthew R.G. Brown^c, Wei Deng^{a,b}, Qiang Wang^a, Xiaohong Ma^a, Mingli Li^a, Xun Hu^e, Michal Juhas^c, Xinmin Li^c, Russell Greiner^d, Andrew J. Greenshaw^c, Tao Li^{a,b,*}

^a Mental Health Centre and Psychiatric Laboratory, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, Sichuan, China

^b Huaxi Brain Research Centre, West China Hospital, Sichuan University, Chengdu, Sichuan, China

^c Department of Psychiatry, University of Alberta, Edmonton, AB, Canada

^d Department of Computing Science, University of Alberta, Edmonton, AB, Canada

^e Huaxi Biobank, West China Hospital, Sichuan University, Chengdu, Sichuan, China

ARTICLE INFO

Article history:

Received 17 November 2016

Received in revised form 28 May 2017

Accepted 3 June 2017

Available online xxxx

Keywords:

Schizophrenia
Major depressive disorder
Endophenotype
Neurocognition
Machine learning

ABSTRACT

Background: Neurocognitive impairments are frequently observed in schizophrenia and major depressive disorder (MDD). However, it remains unclear whether reported neurocognitive abnormalities could objectively identify an individual as having schizophrenia or MDD.

Methods: The current study included 220 first-episode patients with schizophrenia, 110 patients with MDD and 240 demographically matched healthy controls (HC). All participants performed the short version of the Wechsler Adult Intelligence Scale-Revised in China; the immediate and delayed logical memory of the Wechsler Memory Scale-Revised in China; and seven tests from the computerized Cambridge Neurocognitive Test Automated Battery to evaluate neurocognitive performance. The three-class AdaBoost tree-based ensemble algorithm was employed to identify neurocognitive endophenotypes that may distinguish between subjects in the categories of schizophrenia, depression and HC. Hierarchical cluster analysis was applied to further explore the neurocognitive patterns in each group.

Results: The AdaBoost algorithm identified individual's diagnostic class with an average accuracy of 77.73% (80.81% for schizophrenia, 53.49% for depression and 86.21% for HC). The average area under ROC curve was 0.92 (0.96 in schizophrenia, 0.86 in depression and 0.92 in HC). Hierarchical cluster analysis revealed for MDD and schizophrenia, convergent altered neurocognition patterns related to shifting, sustained attention, planning, working memory and visual memory. Divergent neurocognition patterns for MDD and schizophrenia related to motor speed, general intelligence, perceptual sensitivity and reversal learning were identified.

Conclusions: Neurocognitive abnormalities could predict whether the individual has schizophrenia, depression or neither with relatively high accuracy. Additionally, the neurocognitive features showed promise as endophenotypes for discriminating between schizophrenia and depression.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Schizophrenia and major depressive disorder (MDD) are two of the most common psychiatric disorders (Kessler et al., 2005; Walker et al., 2004). Depression is an important co-occurring syndrome in schizophrenia to the extent that approximately 50% of patients with schizophrenia present with comorbid depression (Buckley et al., 2009). Depression in schizophrenia is, however, heterogeneous and the best approaches to its understanding and treatment are based on

appropriate differential diagnosis (Siris, 2000). The proposal for neurocognitive endophenotypes as biomarkers has shed some light on the identification of transdiagnostic processes in these disorders (Bentall et al., 2009). Cognitive abnormalities are widely acknowledged as significant aspects of both schizophrenia and depression. Compared to individuals with MDD, individuals with schizophrenia have more serious cognitive deficits in working memory and selective attention (Egeland et al., 2003a; Egeland et al., 2003b) and MDD with psychotic features is associated with greater levels of cognitive impairment (Busatto, 2013). Both schizophrenia and depression could have significant impairments on working memory, planning and shifting (Barch et al., 2003; Snyder, 2013). Despite these observations, it remains unclear whether these two disorders are associated with distinctly different neurocognitive patterns.

* Corresponding author at: West China Mental Health Centre, West China Hospital, Sichuan University, No. 28th Dianxin Nan Str., Chengdu, Sichuan, China.
E-mail address: litaohx@scu.edu.cn (T. Li).

Unfortunately, conventional statistical group differences might not translate to discovering deviations from normal on a single-subject level and therefore are not sufficient as a significant diagnostic aid. However, machine learning offers a variety of tools that to develop models that may predict the disease status of each individual subject. Previous research has demonstrated that individuals with schizophrenia may be distinguished from healthy subjects with a reasonable classification accuracy based on genetic, neuroimaging or neurocognitive data (Aguilar-Pulido et al., 2010; Lu et al., 2012; Schnack et al., 2014; Shen et al., 2014). Patients with MDD may also be distinguished from healthy subjects using pharmacogenomics or neuroimaging data (Guilloux et al., 2015; Mwangi et al., 2012). One neuroimaging study reported a multi-class classification of schizophrenia versus depression versus healthy with an accuracy rate of 80.9% (Yu et al., 2013). However, those studies did not determine whether multi-class classification methods could use neurocognitive features to distinguish individuals with schizophrenia or depression from healthy controls.

To our knowledge, no studies have investigated whether multi-class machine learning classification methods can find patterns in neurocognitive features that can distinguish individuals with schizophrenia versus depression, versus healthy controls. The purpose of current study was: (1) to classify each individual into one of three categories - schizophrenia, depression or healthy control. Classification of individuals into various disease categories may improve clinical treatment by enabling more effective screening, diagnosis and monitoring of disease trajectory. (2) To examine neurocognitive features to develop further the concept of a neurocognitive hierarchy in the heterogeneous neuropsychological profile of the illness.

2. Methods

2.1. Subjects

This study recruited 570 participants, including 220 first-episode patients with schizophrenia, 110 patients with major depressive disorder and 240 healthy controls. Table S1 summarizes the demographic and clinical characteristics of the subjects. Patients were recruited at the Mental Health Center of West China Hospital, Sichuan University. Healthy controls were recruited by advertisements in local communities. All groups were matched for age, gender and education level. All subjects were right-handed Han Chinese between the ages of 16 and 50 years. Ethical approval for this study was granted by the ethics committee of the West China Hospital, Sichuan University, in accord with the Declaration of Helsinki.

2.2. Neuropsychological assessments

Level of intelligence was estimated at the initial assessment of both patients and healthy controls using the short version of Wechsler Adult Intelligence Scale – Revised in China (WAIS-RC) (Gong, 1992). The seven subtests of WAIS-RC included information, arithmetic, digital symbol, digital span test, block design, picture completion, and similarities.

Both immediate and delayed logical memory were evaluated with the Wechsler Memory Scale–Revised in China (WMS-RC) (Gong, 1989). Lower raw scores represent poorer neuropsychological performance.

The computerized Cambridge Neurocognitive Test Automated Battery (CANTAB - <http://www.cambridgecognition.com>), which comprises visuo-spatial tasks, is sensitive to cognitive impairments in psychiatric disorders (Sahakian and Owen, 1992). Seven CANTAB tests are recognized as sensitive to frontal (including frontostriatal, frontotemporal and frontoparietal), cingulate and temporal brain functions. The variables of CANTAB are also considered as predictive for psychosocial functioning in individuals with schizophrenia and other mental disorders (Johnston et al., 2015; Levaux et al., 2007). The

CANTAB tests included the Big Circle/Little Circle (BLC), the Rapid Visual Information Processing (RVP), the Delayed Matching to Sample (DMS), the Pattern Recognition Memory (PRM), the Spatial Working Memory (SWM), the Stockings of Cambridge (SOC) and the Intra/extra Dimensional Set Shift (IED). Perceptual sensitivity was also assessed through the principles of Signal Detection Theory (SDT) in DMS and RVP (Yang et al., 2015). Variables of interest across tasks included reaction time, accuracy, errors, trials completed and strategy (Haring et al., 2016; Robbins et al., 1998; Wu et al., 2016). These neurocognitive tasks and measurements are briefly described in Table S2 and Table S3. The characterization for each subject was based on 65 features.

2.3. Machine learning analysis

The overall approach used for machine learning analysis involved the following steps: (1) The data was first cleaned, and each feature was normalized using Z-scores. (2) The data was randomly divided: 60% as the training model set and the remaining 40% for testing as hold-out data set. (3) In the training model set, the SMOTE + Tomek links method was applied to help balance the classes, then the three-class AdaBoost algorithm was approached to learn a classifier. (4) The performance of this classifier was evaluated on holdout dataset. The diagram was described in Fig. 1. All analyses were performed on Python 2.7.10 (<https://www.python.org>), scikit-learn 0.17.0 (<http://scikit-learn.org/stable/>), and SciPy (<http://scipy.org/>).

2.3.1. SMOTE + Tomek

Class-imbalance issues become very pronounced in multi-class classification approaches, as the minority class is more likely to be misclassified than the majority class (Rahman and Davis, 2013). Synthetic minority over-sampling technique (SMOTE) is often used to address this problem (Chawla et al., 2002). In this study, the participant sample was partially unbalanced (220 in schizophrenia group, 110 in MDD group and 240 in control group). To address this issue, we applied the SMOTE + Tomek links approach (<https://github.com/fmfn/UnbalancedDataset>) in the training model set. This method constructs additional “synthesized” instances of the minority class, to make the training model set more balanced, based on k-Nearest Neighbor algorithm, here using Euclidean distance and $k = 5$ (Mani and Zhang, 2003). The fraction of the number of MDD group elements to generate was selected as ratio = 1. Before and after the SMOTE + Tomek method, the averages of age and education level were computed to assess distribution of the data.

2.3.2. AdaBoost tree-based ensemble algorithm

AdaBoost is a meta-estimator that tries to produce a strong classifier by combining several weak classifiers (Freund and Schapire, 1995). In this study, we used the multi-class AdaBoost-SAMME (Stagewise Additive Modeling) algorithm with Classification and Regression Trees (CART) as the base learner (Zhu et al., 2009). To train each individual CART classifier, we used the Gini impurity to measure the quality of splits, and set the maximum depth to 5 and the minimum samples per leaf to 15. The number of estimators (CART classifier) was set to 250.

2.3.3. Cross-validation and model grid-search

We used stratified 5-fold cross-validation on the training model set to determine the optimal parameter values, considering each possible combination of parameter values: for CART: maximum depth {3, 4, 5, 6, 7} and minimum samples per leaf {5, 10, 15, 20, 25, 30}; and for AdaBoost classifier: the tree estimator values {50, 100, 150, 200, 250, 300}.

2.4. Hierarchical cluster analysis

Hierarchical cluster analysis was performed with average linkage and the Euclidean distance to reveal close relationships among the

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات