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

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.

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 (Aguiar-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
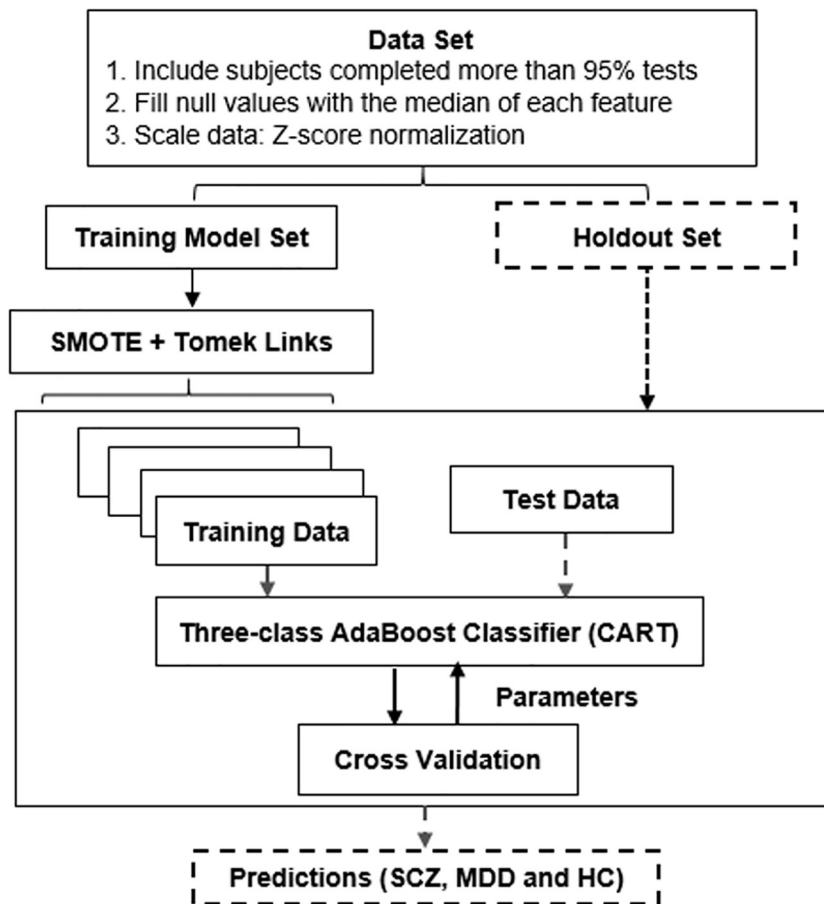
Fig. 1. Data processing diagram.

neurocognitive features. The Gini importance of each feature was computed to identify which neurocognitive features contributed to the discriminative ability of the classifier (Hastie et al., 2005; Ritchie et al., 2014). The formula was based on the work of Louppe et al. (2013).

Here, hierarchical clustering analysis was performed on the original normalized data set (without the SMOTE over-sampling). We selected the 18 neurocognitive features with the highest Gini importance scores, then hierarchically clustered to produce a "neurocognitive hierarchy", generating a dendrogram for each group (see Table S4 & Fig. 4). Each dendrogram illustrates how each cluster is composed by drawing a U-shaped link between a non-singleton cluster and its sub-nodes. The Silhouette Coefficient was used to evaluate clustering, in terms of the clusters' cohesion and separation (in the Supplementary material).

2.5. Performance measures

2.5.1. Accuracy
Predictive accuracy is the performance measure generally associated with machine learning algorithms: accuracy is defined as the fraction of correct predictions.

2.5.2. Precision-Recall curve
Precision-Recall curves are typically used in binary classification to study the output of a classifier. A high value for area under the Precision-Recall curve (AUPRC) represents both high recall and high precision.

2.5.3. Receiver Operating Characteristic (ROC) curve
ROC curves represent the family of best decision boundaries for relative costs of TP and FP. The AUROC is a useful metric for classifier performance as it implicitly considers a range of criterion and prior class probabilities. A classifier associated with an AUROC value greater than 0.8 is considered good.

3. Results

3.1. Classification results

We applied a multi-class AdaBoost tree-based ensemble algorithm onto neurocognitive features in individuals with schizophrenia, individuals with depression and healthy controls (Table S1 in Supplemental information illustrates subject characteristics). Three-class machine learning classification (schizophrenia vs. MDD vs. healthy control) achieved an average accuracy on the holdout data of 77.73% at the individual level (80.81% for patients with schizophrenia, 53.49% for patients with MDD and 86.21% for healthy controls). Table 1 illustrates the confusion matrix for results in the holdout data set. All of these results were based on the optimal parameter setting based on the cross validation.

Table 1
Confusion matrix for results in holdout data set.

Classes		Predicted Subjects		
		MDD	SCZ	HC
Actual subjects	MDD	23	3	17
	SCZ	4	80	15
	HC	9	3	75

The rows of this matrix present the groups of the subjects (ground truth), and the columns present the predictions by the classifier. The cells in each row contain the number of trials in which subjects responded with the category indicated by the column.

The average area under the Precision-Recall curve (AUPRC) of the three groups was 0.88. The AUPRC was 0.96 in the schizophrenia group, 0.63 in the depression group and 0.87 in the healthy control group (Fig. 2).

The average AUROC of the three groups was 0.92. The AUROC was 0.96 in the schizophrenia group, 0.86 in the depression group and 0.92 in the healthy control group (Fig. 3).

3.2. Hierarchical cluster analysis

The dendrogram of the healthy control group revealed a separation of features into two main clusters: general intelligence in WAIS-RC and WMS-RC, and visual-spatial neurocognitive function in CANTAB. It also illustrated that completing the cognitive items tends to require higher and higher neurocognitive function from right to left (Fig. 4(A)).

The dendrogram of the depression group had two main clusters, but the clustering sequence and its distance were different from those of the HC dendrogram. In the depression dendrogram, the hierarchy of the general intelligence cluster was lower than the cluster including cognitive items of CANTAB. This may reflect that individuals with depression could have more deficits in visual-spatial neurocognitive function compared to their general intelligence. The hierarchy of motor speed (BLC_CRL), shifting (IED_CST), sustained attention (RVP_ML), working memory (SWM_MLR, SWM_MTP) and visual memory (DMS_MCLS, PRM_MCLi) in the depression group were different from those in HC group. This may reflect significant neurocognitive impairments in motor speed, shifting, sustained attention, working memory and visual memory in the MDD group. The smaller clustering distance in the MDD group is likely indicative of similarities in the level of difficulty in terms of cognitive abilities required to complete the subtest tasks. These results probably reflected the broad neurocognitive decline in depression, especially in motor speed and shifting (Fig. 4(B)).

The dendrogram of the schizophrenia group had three clusters. The hierarchy of neurocognition in schizophrenia was as follows: Cluster 1: shifting (IED_CST), perceptual sensitivity (RVP_BDP), and general intelligence; Cluster 2: working memory (SWM_MFR), reversal learning (IED_EB9) and shifting (IED_EB8); Cluster 3: planning, visual memory, motor speed and sustained attention. The schizophrenia group had an obviously different neurocognitive hierarchy sequence compared to the healthy control group. Moreover, the greater clustering distance in the schizophrenia group is likely indicative of differences in the level of difficulty in terms of cognitive abilities required to complete the subset tasks (Fig. 4(C)). These results were consistent with the generalized dysfunction of neurocognition in the schizophrenia group, especially in general intelligence, shifting, perceptual sensitivity, reversal learning, working memory, planning and sustained attention.

Individuals with MDD and those with schizophrenia exhibited convergent patterns in altered neurocognition patterns related to shifting, sustained attention, planning, working memory and visual memory. By contrast, individuals with MDD and with schizophrenia, respectively, had different clustering distances and hierarchy orders. Divergent neurocognition patterns of MDD and schizophrenia were related to motor speed, general intelligence, perceptual sensitivity and reversal learning.

4. Discussion

To our knowledge, this is the first study on multi-class (three-class) machine learning classification of individuals with schizophrenia, depression and healthy controls on the basis of neurocognitive tests. In this study, the AdaBoost tree-based ensemble classifier was trained to perform the three-class classification task using 65 features including general intelligence and executive function. It also confirmed that motor speed, shifting, general intelligence, perceptual sensitivity, reversal learning, sustained attention, working memory and planning
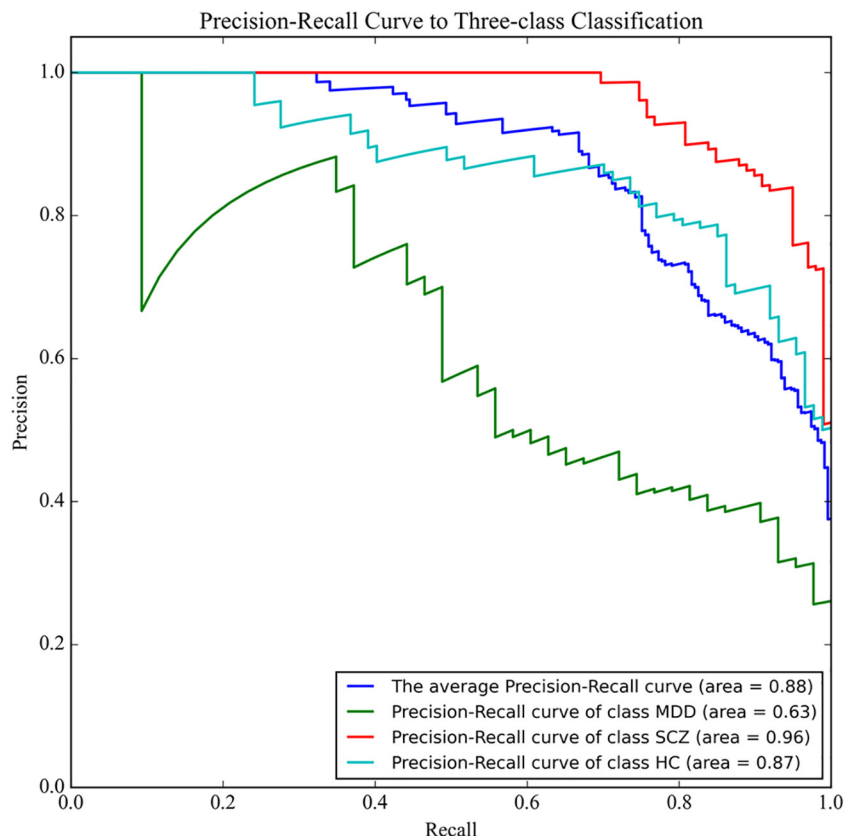

Fig. 2. The Precision-Recall curve. Average precision was calculated from prediction scores which corresponded to the area under the precision-recall curve.
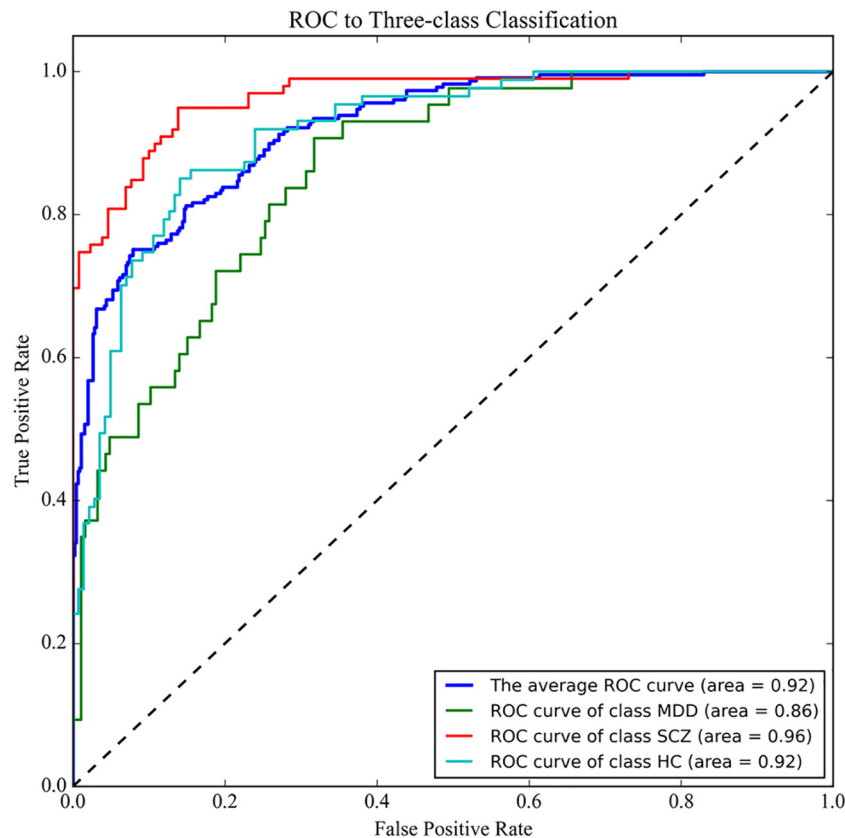
Fig. 3. The ROC curve. Average ROC was calculated from prediction scores which corresponded to the area under the ROC curve.

(as endophenotypic features) may be used to distinguish among individuals with schizophrenia or depression relative to healthy subjects.

In this study, although classification for the depression group (i.e. depression vs. non-depression) had relatively lower precision and recall, the AUROC was higher than 0.80. This inconsistency between the Precision-Recall curve and the AUROC was probably due to the unbalanced sample size in these three groups (Davis and Goadrich, 2006).

We also computed hierarchical clustering to explore further the neurocognitive patterns in schizophrenia, depression and healthy control. Compared with healthy controls, individuals with depression and with schizophrenia exhibited altered neurocognitive patterns in motor speed, sustained attention, working memory and planning. This may account for overlap in symptoms such as psychomotor slowing, difficulty concentrating and remembering details, and a decline in working memory and learning seen in these two disorders. Consistent with previous research, the major depression was associated with broad impairments on neuropsychological measures of executive function including shifting, processing speed, working memory, planning and problem-solving (Snyder, 2013). The abnormality of motor speed in depression group could be tightly linked to psychomotor retardation (Albus et al., 1996; White et al., 1997). The impairment in shifting and inhibition could result in difficulty with respect to regulating negative information, associated with rumination in major depression (Snyder, 2013). Schizophrenia is associated with generalized neurocognitive dysfunction and more serious deficits in general intelligence, perceptual sensitivity, reversal learning, working memory, selective attention relative to major depression (Albus et al., 1996; Blanchard and Neale, 1994; Egeland et al., 2003a; Egeland et al., 2003b; Mussgay and Hertwig, 1990). The results of the current study suggest that schizophrenia and depression are not only associated with convergent cognitive deficits in shifting, sustained attention, planning, working memory and visual memory, but also potential divergent neurocognitive patterns. The latter are manifest as differences in motor speed, general intelligence,

perceptual sensitivity and reversal learning associated with these two disorders.

Numerous neuroimaging studies have supported the involvement of brain regions significantly related to emotion processing and to models of psychotic symptoms in schizophrenia and depression. These include the hippocampus, insula, prefrontal cortex and inferior parietal cortex (Bernstein et al., 2016; Busatto, 2013). A considerable degree of overlap in the regional pattern of brain abnormalities across these two diagnostic categories is notable. Recently, genetic research also provided some evidence related to the pathogenesis of both schizophrenia and depression, including Retinoic acid-inducible or induced gene 1, the α-1C subunit of the L-type voltage-gated calcium channel and immune genes (Bufalino et al., 2013; Fillman et al., 2013; Green et al., 2010; Haybaeck et al., 2015). Convergent findings in neuroimaging and genetics of depression and schizophrenia provide a plausible biological basis for similar neurocognitive impairments observed in previous studies and the current one.

Connections associated with the prefrontal cortex and the affective network have been reported to be differentially influenced by schizophrenia and depression (Yu et al., 2013). Additionally, bilaterally reduced claustral volumes could relate to sensory processing impairments in schizophrenia, in contrast to a possible association with disturbance in salience in major depression (Bernstein et al., 2016). Individuals with schizophrenia or depression could also have characteristic differences in their cortical responses to dynamic affective stimuli, potentially related to pathology-specific problems in social cognition (Regenbogen et al., 2015). In accord with this overall pattern of findings, the current study also revealed differences in neurocognitive patterns associated with schizophrenia and depression, providing further evidence for different mechanisms underlying pathology in these two disorders.

Although the classification results of the present study are encouraging, possible limitations should be considered. Firstly, the current study
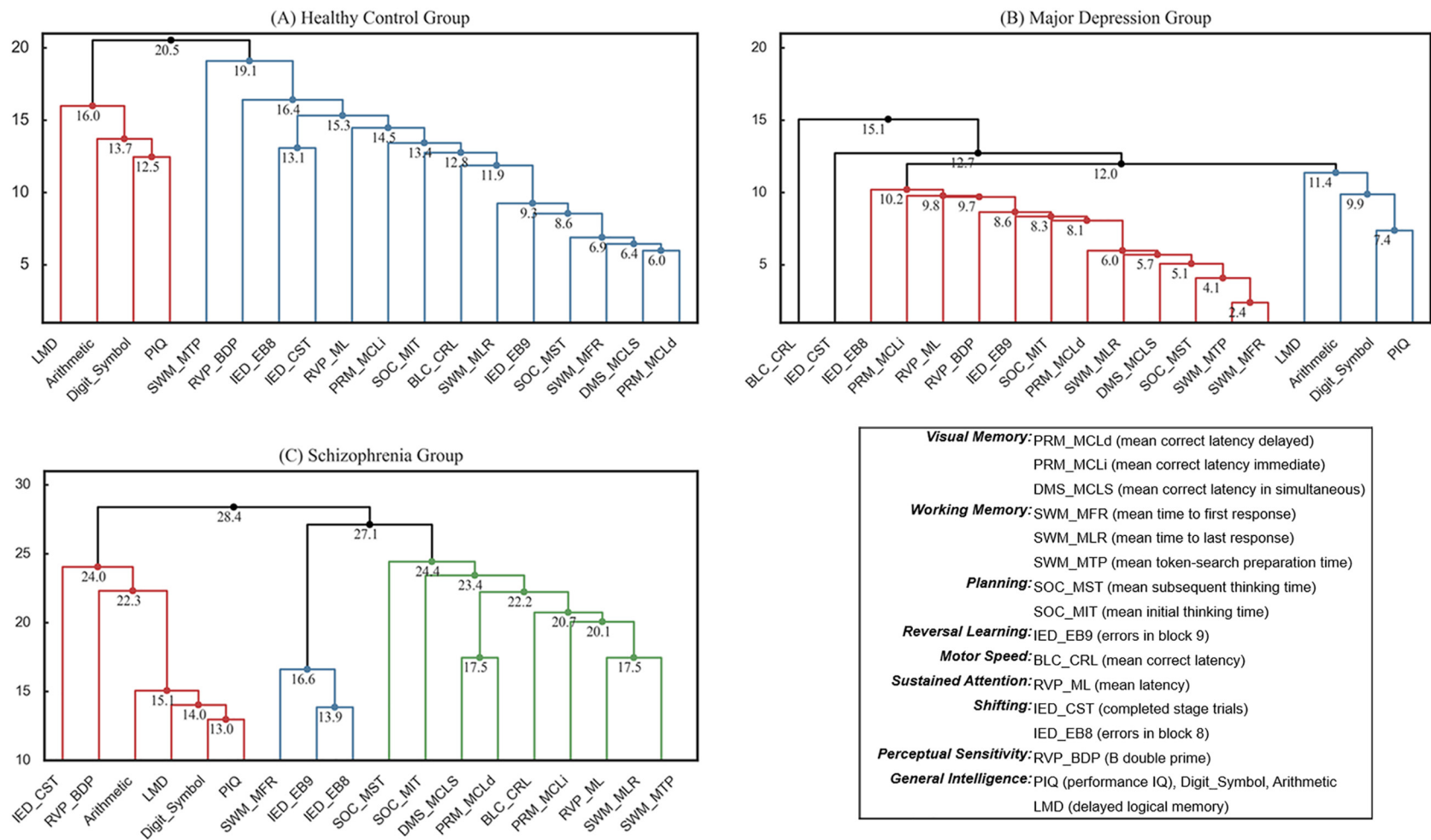
Fig. 4. The dendrograms of hierarchy clustering analysis in schizophrenia, depression and control. The X-axis represents the feature. The Y-axis of the dendrogram represents the dissimilarity or Euclidean distance between neurocognitive features.

applied machine learning analysis only to neurocognitive data. The ability to determine the specific clinical and neuropsychological profiles associated with each disorder may be improved by the addition of data related to genetic, neurobiological and environmental variables (Nolen-Hoeksema and Watkins, 2011). Secondly, distinguishing psychotic and affective symptoms remains a dilemma of psychiatric classification. In the future studies, it will be fruitful to evaluate both psychotic and affective symptoms in individuals with schizophrenia and depression, respectively, in an attempt to build a differential diagnosis paradigm for comorbid schizophrenia and depression symptoms. Thirdly, although some of the relapsed depressive patients had not taken antidepressants in the preceding three months or more, it is hard to evaluate the significance of previous medication effects. Fourthly, the comparisons of inter-group dendrograms as the results of hierarchical clustering analysis were not based on the evidence of statistical significance, more relative algorithms were urgently demanded.

In conclusion, we trained a data-driven multi-class classifier that used an individual's neurocognitive features, as potential endophenotypes could potentially be used to predict classification of that individual as meeting criteria for MDD, schizophrenia, or as a healthy control. Our results support the proposal that neurocognitive features could potentially be used to reveal some convergent and divergent neurocognitive patterns of symptomatology for schizophrenia and depression, respectively. Moving forward with the development of a clear hierarchy of neurocognitive deficiencies in schizophrenia and major depression could improve diagnostic decision making and may prove beneficial for longitudinal monitoring of therapeutic advances.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.schres.2017.06.004.

References

Aguiar-Pulido, V., Seoane, J.A., Rabunal, J.R., Dorado, J., Pazos, A., Munteanu, C.R., 2010. Machine learning techniques for single nucleotide polymorphism—disease classification models in schizophrenia. Molecules 15 (7), 4875–4889.

Albus, M., Hubmann, W., Wahlheim, C., Sobizack, N., Franz, U., Mohr, F., 1996. Contrasts in neuropsychological test profile between patients with first-episode schizophrenia and first-episode affective disorders. Acta Psychiatr. Scand. 94 (2), 87–93.

Barch, D.M., Sheline, Y.I., Csernansky, J.G., Snyder, A.Z., 2003. Working memory and prefrontal cortex dysfunction: specificity to schizophrenia compared with major depression. Biol. Psychiatry 53 (5), 376–384.

Bentall, R.P., Rowse, G., Shryane, N., Kinderman, P., Howard, R., Blackwood, N., Moore, R., Corcoran, R., 2009. The cognitive and affective structure of paranoid delusions: a transdiagnostic investigation of patients with schizophrenia spectrum disorders and depression. Arch. Gen. Psychiatry 66 (3), 236–247.

Bernstein, H.G., Ortmann, A., Dobrowolny, H., Steiner, J., Brisch, R., Gos, T., Bogerts, B., 2016. Bilaterally reduced claustral volumes in schizophrenia and major depressive disorder: a morphometric postmortem study. Eur. Arch. Psychiatry Clin. Neurosci. 266 (1), 25–33.

Blanchard, J.J., Neale, J.M., 1994. The neuropsychological signature of schizophrenia: generalized or differential deficit? Am. J. Psychiatry 151 (1), 40–48.

Buckley, P.F., Miller, B.J., Lehrer, D.S., Castle, D.J., 2009. Psychiatric comorbidities and schizophrenia. Schizophr. Bull. 35 (2), 383–402.

Bufalino, C., Hepgul, N., Aguglia, E., Pariante, C.M., 2013. The role of immune genes in the association between depression and inflammation: a review of recent clinical studies. Brain Behav. Immun. 31, 31–47.

Busatto, G.F., 2013. Structural and functional neuroimaging studies in major depressive disorder with psychotic features: a critical review. Schizophr. Bull. 39 (4), 776–786.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res. 321–357.

Davis, J., Goadrich, M., 2006. The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd International Conference on Machine Learning. ACM, pp. 233–240.

Egeland, J., Rund, B.R., Sundet, K., Landro, N.I., Asbjornsen, A., Lund, A., Roness, A., Stordal, K.I., Hugdahl, K., 2003a. Attention profile in schizophrenia compared with depression: differential effects of processing speed, selective attention and vigilance. Acta Psychiatr. Scand. 108 (4), 276–284.

Egeland, J., Sundet, K., Rund, B.R., Asbjornsen, A., Hugdahl, K., Landro, N.I., Lund, A., Roness, A., Stordal, K.I., 2003b. Sensitivity and specificity of memory dysfunction in schizophrenia: a comparison with major depression. J. Clin. Exp. Neuropsychol. 25 (1), 79–93.

Fillman, S.G., Cloonan, N., Catts, V.S., Miller, L.C., Wong, J., McCrossin, T., Cairns, M., Weickert, C.S., 2013. Increased inflammatory markers identified in the dorsolateral prefrontal cortex of individuals with schizophrenia. Mol. Psychiatry 18 (2), 206–214.

Freund, Y., Schapire, R.E., 1995. A decision-theoretic generalization of on-line learning and an application to boosting. Computational Learning Theory. Springer, pp. 23–37.

Gong, Y., 1989. Wechsler Memory Scale–Revised in China. Hunan Medical College, Changsha, Hunan/China.

Gong, Y., 1992. Wechsler Adult Intelligence Scale-Revised in China Version. Hunan Medical College, Changsha, Hunan/China.

Green, E.K., Grozeva, D., Jones, I., Jones, L., Kirov, G., Caesar, S., Gordon-Smith, K., Fraser, C., Forty, L., Russell, E., Hamshere, M.L., Moskvina, V., Nikolov, I., Farmer, A., McGuffin, P., Holmans, P.A., Owen, M.J., O'Donovan, M.C., Craddock, N., 2010. The bipolar disorder risk allele at CACNA1C also confers risk of recurrent major depression and of schizophrenia. Mol. Psychiatry 15 (10), 1016–1022.

Guilloux, J.P., Bassi, S., Ding, Y., Walsh, C., Turecki, G., Tseng, G., Cyranowski, J.M., Sibille, E., 2015. Testing the predictive value of peripheral gene expression for nonremission following citalopram treatment for major depression. Neuropsychopharmacology 40 (3), 701–710.

Haring, L., Muursepp, A., Mottus, R., Ilves, P., Koch, K., Uppin, K., Tarnovskaja, J., Maron, E., Zharkovsky, A., Vasar, E., Vasar, V., 2016. Cortical thickness and surface area correlates with cognitive dysfunction among first-episode psychosis patients. Psychol. Med. 46 (10), 2145–2155.

Hastie, T., Tibshirani, R., Friedman, J., Franklin, J., 2005. The elements of statistical learning: data mining, inference and prediction. Math. Intell. 27 (2), 83–85.

Haybaeck, J., Postruznik, M., Miller, C.L., Dulay, J.R., Llenos, I.C., Weis, S., 2015. Increased expression of retinoic acid-induced gene 1 in the dorsolateral prefrontal cortex in schizophrenia, bipolar disorder, and major depression. Neuropsychiatr. Dis. Treat. 11, 279–289.

Johnston, B.A., Coghill, D., Matthews, K., Steele, J.D., 2015. Predicting methylphenidate response in attention deficit hyperactivity disorder: a preliminary study. J. Psychopharmacol. 29 (1), 24–30.

Kessler, R.C., Berglund, P., Demler, O., Jin, R., Merikangas, K.R., Walters, E.E., 2005. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch. Gen. Psychiatry 62 (6), 593–602.

Levaux, M.N., Potvin, S., Sepehry, A.A., Sablier, J., Mendrek, A., Stip, E., 2007. Computerized assessment of cognition in schizophrenia: promises and pitfalls of CANTAB. Eur. Psychiatry 22 (2), 104–115.

Louppe, G., Wehenkel, L., Sutera, A., Geurts, P., 2013. Understanding variable importances in forests of randomized trees. Adv. Neural. Inf. 431–439.

Lu, A.T., Bakker, S., Janson, E., Cichon, S., Cantor, R.M., Ophoff, R.A., 2012. Prediction of serotonin transporter promoter polymorphism genotypes from single nucleotide polymorphism arrays using machine learning methods. Psychiatr. Genet. 22 (4), 182–188.

Mani, I., Zhang, I., 2003. kNN approach to unbalanced data distributions: a case study involving information extraction. Proceedings of Workshop on Learning from Imbalanced Datasets.

Mussgay, L., Hertwig, R., 1990. Signal detection indices in schizophrenics on a visual, auditory, and bimodal Continuous Performance Test. Schizophr. Res. 3 (5–6), 303–310.

Mwangi, B., Ebmeier, K.P., Matthews, K., Steele, J.D., 2012. Multi-centre diagnostic classification of individual structural neuroimaging scans from patients with major depressive disorder. Brain 135 (Pt. 5), 1508–1521.

Nolen-Hoeksema, S., Watkins, E.R., 2011. A heuristic for developing transdiagnostic models of psychopathology: explaining multifinality and divergent trajectories. Perspect. Psychol. Sci. 6 (6), 589–609.

Rahman, M.M., Davis, D., 2013. Addressing the class imbalance problem in medical datasets. Int. J. Mach. Learn. Comput. 3 (2), 224.

Regenbogen, C., Kellermann, T., Seubert, J., Schneider, D.A., Gur, R.E., Derntl, B., Schneider, F., Habel, U., 2015. Neural responses to dynamic multimodal stimuli and pathology-specific impairments of social cognition in schizophrenia and depression. Br. J. Psychiatry 206 (3), 198–205.

Ritchie, G.R., Dunham, I., Zeggini, E., Flicek, P., 2014. Functional annotation of noncoding sequence variants. Nat. Methods 11 (3), 294–296.

Robbins, T.W., James, M., Owen, A.M., Sahakian, B.J., Lawrence, A.D., McInnes, L., Rabbitt, P.M., 1998. A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: implications for theories of executive functioning and cognitive aging. J. Int. Neuropsychol. Soc. 4 (5), 474–490.

Sahakian, B., Owen, A., 1992. Computerized assessment in neuropsychiatry using CANTAB: discussion paper. J. R. Soc. Med. 85 (7), 399.

Schnack, H.G., Nieuwenhuis, M., van Haren, N.E., Abramovic, L., Scheewe, T.W., Brouwer, R.M., Hulshoff Pol, H.E., Kahn, R.S., 2014. Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. NeuroImage 84, 299–306.

Shen, C., Popescu, F.C., Hahn, E., Ta, T.T., Dettling, M., Neuhaus, A.H., 2014. Neurocognitive pattern analysis reveals classificatory hierarchy of attention deficits in schizophrenia. Schizophr. Bull. 40 (4), 878–885.

Siris, S.G., 2000. Depression in schizophrenia: perspective in the era of "atypical" antipsychotic agents. Am. J. Psychiatry 157 (9), 1379–1389.

Snyder, H.R., 2013. Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. Psychol. Bull. 139 (1), 81–132.

Walker, E., Kestler, L., Bollini, A., Hochman, K.M., 2004. Schizophrenia: etiology and course. Annu. Rev. Psychol. 55, 401–430.

White, D.A., Myerson, J., Hale, S., 1997. How cognitive is psychomotor slowing in depression? Evidence from a meta-analysis. Aging Neuropsychol. Cogn. 4 (3), 166–174.

Wu, M.J., Passos, I.C., Bauer, I.E., Lavagnino, L., Cao, B., Zunta-Soares, G.B., Kapczinski, F., Mwangi, B., Soares, J.C., 2016. Individualized identification of euthymic bipolar disorder using the Cambridge Neuropsychological Test Automated Battery (CANTAB) and machine learning. J. Affect. Disord. 192, 219–225.

Yang, X., Ma, X., Huang, B., Sun, G., Zhao, L., Lin, D., Deng, W., Li, T., 2015. Gray matter volume abnormalities were associated with sustained attention in unmedicated major depression. Compr. Psychiatry 63, 71–79.

Yu, Y., Shen, H., Zeng, L.L., Ma, Q., Hu, D., 2013. Convergent and divergent functional connectivity patterns in schizophrenia and depression. PLoS One 8 (7), e68250.

Zhu, J., Zou, H., Rosset, S., Hastie, T., 2009. Multi-class adaboost. Stat. Interface 2 (3), 349–360.