Classification of multi-site MR images in the presence of heterogeneity using multi-task learning

Qiongmin Ma\(^{a,b,c,*}\), Tianhao Zhang\(^{b}\), Marcus V. Zanetti\(^{d}\), Hui Shen\(^{b}\), Theodore D. Satterthwaite\(^{e}\), Daniel H. Wolf\(^{f}\), Raquel E. Gur\(^{e}\), Yong Fan\(^{b}\), Dewen Hu\(^{a,1}\), Geraldo F. Busatto\(^{d,1}\), Christos Davatzikos\(^{b,1}\)

\(^{a}\) College of Mechatronics and Automation, National University of Defense Technology, Changsha, Hunan 410073, China
\(^{b}\) Center for Biomedical Image Computing and Analytics, and Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, United States
\(^{c}\) Beijing Institute of System Engineering, China
\(^{d}\) Laboratory of Psychiatric Neuroimaging (LIM-21), Department and Institute of Psychiatry, Faculty of Medicine, University of S\ñ?o Paulo, S\ñ?o Paulo, Brazil
\(^{e}\) Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, United States

**ARTICLE INFO**

**Keywords:**
- Multi-task learning
- Multi-site classification
- Imaging heterogeneity
- Sparsity
- Schizophrenia
- MRI

**ABSTRACT**

With the advent of Big Data Imaging Analytics applied to neuroimaging, datasets from multiple sites need to be pooled into larger samples. However, heterogeneity across different scanners, protocols and populations, renders the task of finding underlying disease signatures challenging. The current work investigates the value of multi-task learning in finding disease signatures that generalize across studies and populations. Herein, we present a multi-task learning type of formulation, in which different tasks are from different studies and populations being pooled together. We test this approach in an MRI study of the neuroanatomy of schizophrenia (SCZ) by pooling data from 3 different sites and populations: Philadelphia, Sao Paulo and Tianjin (50 controls and 50 patients from each site), which posed integration challenges due to variability in disease chronicity, treatment exposure, and data collection. Some existing methods are also tested for comparison purposes. Experiments show that classification accuracy of multi-site data outperformed that of single-site data and pooled data using multi-task feature learning, and also outperformed other comparison methods. Several anatomical regions were identified to be common discriminant features across sites. These included prefrontal, superior temporal, insular, anterior cingulate cortex, tempo-limbic and striatal regions consistently implicated in the pathophysiology of schizophrenia, as well as the cerebellum, precuneus, and fusiform, middle temporal, inferior parietal, postcentral, angular, lingual and middle occipital gyri. These results indicate that the proposed multi-task learning method is robust in finding consistent and reliable structural brain abnormalities associated with SCZ across different sites, in the presence of multiple sources of heterogeneity.

1. Introduction

Neuroimaging studies have widely explored the clinical value of machine learning methods for differentiating psychiatric patients from healthy controls at the individual level. In addition to providing individualized indices for diagnostic purposes, machine learning methods may ultimately help identify brain regions affected by disease in subtle ways that can only be elucidated using multi-variate analysis. While most of these neuroimaging studies to date have been performed using single-site datasets, it is essential to integrate multi-site data for two reasons. First, multi-site data provide sufficient statistical power for detecting subtle, but informative patterns of brain structure and function (Brown et al., 2011; Friedman et al., 2006; Schnack et al., 2010), which may be difficult to unravel with the relative small sample sizes usually acquired in single centers (Pearlson, 2009; Segall et al., 2009). Second, large sample sizes enhance sample generalizability by pooling large patient populations with diverse demographic features and clinical characteristics including disease onset, symptom severity, and types and duration of treatment (Brown et al., 2011; Friedman et al., 2006; Glover et al., 2012; Pearson, 2009; Sutton et al., 2008; Van Horn and Toga, 2009). Multi-site studies are therefore becoming increasingly the norm in neuroimaging research (Casey et al., 1998; Van Horn and Toga, ...
Multi-site data reflect a more comprehensive abnormal pattern of disease, and therefore may provide a richer understanding of disease signatures than single-site data. However, two recent studies found that simple pooling of multi-site data did not outperform single-site disease classification. Colby (Colby et al., 2012) and Nielsen (Nielsen et al., 2013) pooled multi-site data and trained a common classifier for all data, to identify attention deficit hyperactivity disorder (ADHD) and autism, respectively. Results in these two studies showed that the pooled dataset exhibited lower accuracy than each single-site datasets. Here we seek a new approach to synergistically integrating multi-site data, by emphasizing two points. First, since datasets collected in multiple imaging centers have a common disorder of interest (e.g. schizophrenia (SCZ), in our experiments herein), the abnormal patterns in each dataset are strongly related and thus, to some extent, may share a common imaging signature. The pattern reproducibility among multi-site data has been repeatedly demonstrated in several multi-site studies on functional MRI (fMRI) (Casey et al., 1998; Costafreda et al., 2007; Gee et al., 2015; Jovicich et al., 2015), morphometric MRI (Cannon et al., 2014; Schnack et al., 2010), and diffusion-tensor imaging (DTI) (Jovicich et al., 2014; Pfefferbaum et al., 2003). Most of these studies arrived at similar conclusions, namely that with appropriate multi-site data collection, different data sites shared highly consistent feature patterns. With this, the site-shared features reveal consistent brain abnormalities in multi-site data, which can lead to a more accurate neurobiological understanding of the psychiatric disorder under investigation. On the other hand, the integration of multiple single-site data is advantageous, the unavoidable presence of site-specific features might decrease the accuracy of a classifier that merely pools data together across studies. Heterogeneity can emanate from multiple sources including scanner differences, differences in image acquisition protocols, or ethnic and treatment differences among participating patient populations (Jovicich et al., 2015; Schnack et al., 2010; Van Horn and Toga, 2009). Given such site-related heterogeneity, multi-site datasets should not simply be merged into larger cohorts for further machine learning investigation (Pan and Yang, 2010). In an attempt to eliminate or reduce the site-specific variability, studies have suggested some scanning protocols, consistent scanner parameters and etc. (Brown et al., 2011; Calhoun and Adali, 2009; Pearlson, 2009; You et al., 2011) in data collection, as well as the utilization of smoothness equalization (Friedman et al., 2006) and independent component analysis (Kim et al., 2009; Meda et al., 2008) in data preprocessing. Despite these efforts, site-specific heterogeneities still exist due to their complex causes (Pearlson, 2009; Segall et al., 2009).

The above considerations highlight the need for a feature-learning framework in multi-site disease classification that can extract the site-shared features, while also accounting for the site-specific features; this approach generally seeks an overarching signature of disease, whereas it accommodates potential sub-cohort and other differences to be taken into consideration. In recent years, multi-task learning has been successful in learning task-shared and task-specific features simultaneously, which effectively improves generalization compared with traditional machine learning methods. For example, support vector machine (SVM) with single-task learning (Caruana, 1997) learns a distinct feature pattern and finds a maximum margin hyperplane to classify two groups, which extracts information within a single learning task. In contrast, multi-task learning extracts a subset of task-shared features to generate more accurate models on multiple tasks, with the task-specific features learned simultaneously (Marquand et al., 2014). The basic assumption of multi-task learning is that the feature weights of different tasks share similar sparse patterns (Chen et al., 2012), which can be learned by imposing sparsity regularization penalties on the task weight matrix (Kumar and Daume III, 2012). $l_1$-norm and $l_{2,1}$-norm are two commonly used sparsity regulating terms in multi-task learning, which enforces the weight matrix of different tasks to be sparse across all tasks. Particularly, an $l_1$-norm term highlights task-specific features by encouraging the weights of irrelevant features to be very small (Wang et al., 2015), while $l_{2,1}$-norm introduces group sparsity and enforces task-shared features to have larger weights (Watanabe et al., 2014; Yan et al., 2015).

The advantage of multi-task learning makes it suitable for multi-site data learning, considering the site as task, and the site-shared and site-specific features as task-shared and task-specific features. Neuroimaging studies have shown the effectiveness of performing multi-task learning in the brain decoding and disease classification (Marquand et al., 2014; Obozinski et al., 2010; Rao et al., 2013; Wang et al., 2015; Watanabe et al., 2014). Specifically, multi-site fMRI data of ADHD was demonstrated better than single-site classification by learning site-shared and site-specific features using multi-task scheme (Watanabe et al., 2014). In this work, though the multi-task learning scheme successfully extracted site-shared and site-specific features in multi-site data, the form of the objective function was rather complex and specific as it included an $l_{2,1}$-norm group sparsity regularization term and a 6-D spatial structure penalty (generated by the GraphNet, fused Lasso, or the isotropic total variation). In order to make the multi-task learning scheme more simple and applicable, an objective function including $l_1$-norm and $l_{2,1}$-norm penalty terms was used in the current study (Wang et al., 2011; Wang et al., 2015).

Building upon this emerging literature, we aim to distinguish SCZ patients from healthy controls across multiple-site MRI data using multi-task learning. We hypothesized that using multi-task learning framework on multi-site classification would not only have better performance than single-site data classification, but would also identify the abnormalities shared by all sites, and also specific to each site. These site-shared brain structural alterations should be consistent with the previously reported altered regions in SCZ, such as the brain regions involving prefrontal, superior temporal, insular, tempo-limbic regions, among others.

2. Materials and methods

2.1. Participants and MRI acquisitions

MRI data were collected by three academic centers, including locations in the United States (University of Pennsylvania; site A) (Davatzikos et al., 2005), Brazil (University of Sao Paolo; site B) (Schauflberger et al., 2007; Zanetti et al., 2013), and China (Tianjin; site C). From each site, a balanced dataset was obtained with 50 normal controls (NCs) and 50 SCZ patients randomly selected from a larger pool of available subjects. In total, we had 150 NCs and 150 SCZs, which didn’t differ by age and gender significantly (p > 0.05; see Table 1).

All SCZ patients met DSM-IV criteria. Written informed consent was obtained from all participants before MRI scanning. In site A, the 50 SCZ patients had chronic symptoms and were receiving treatment with antipsychotics (mean duration of illness 16.2 ± 12.3 years). In site B, SCZ patients had chronic symptoms and were receiving treatment with antipsychotics (mean duration of illness 16.2 ± 12.3 years). In site C, all SCZ subjects were recruited shortly after they made their first contact with mental health services due to psychotic symptoms, and their duration of illness was 1.0 ± 1.3 years; 31 patients had been on antipsychotic treatment within 3 weeks of MRI, while the remaining 19 patients were free of antipsychotics at the time of MRI scanning. Site C contributed 5 first-episode, never-treated SCZ patients and 45 chronic SCZ patients under antipsychotic treatment (mean duration of illness 10.5 ± 7.2 years).

In site A, the imaging data were acquired using a Siemens Trio 3-T scanner (Siemens Medical Systems, Erlangen Germany), with the following protocol: slice thickness = 1 mm, TE = 3.51 ms, TR = 18.1 ms, flip angle = 9°, acquisition matrix = 240 x 180, and slice number = 160, no gaps, 1-mm isotropic voxels. In site B, the T1 images were acquired using two identical 1.5-T GE Signa scanners (GE Medical Systems, Milwaukee WI, USA) with the following protocol: T1-SPGR...
sequence providing 124 contiguous slices, TE = 5.2 ms, TR = 21.7 ms, flip angle = 20°, FOV = 22 cm, acquisition matrix = 256 × 192. Details of the MRI scanning in site B can be found in Zanetti et al. (2013). In site C, all scans were obtained using a GE 3-T Signa scanner (GE Medical Systems, Milwaukee WI, USA) with the following protocol: slice thickness = 1 mm, TE = 3.2 ms, TR = 8.2 ms, flip angle = 12°, acquisition matrix = 256 × 256, FOV = 25.6 cm.

2.2. Preprocessing

The imaging data were preprocessed using the following steps as previously described (Davatzikos, 1998; Davatzikos et al., 2001; Goldsztal et al., 1998; Shen and Davatzikos, 2003): (1) skull stripping, (2) bias field correction, (3) brain tissue segmentation into gray matter (GM), white matter (WM), cerebrospinal fluid (CSF), (4) spatial registration to Montreal Neurological Institute (MNI) template, (5) generation of the Regional Analysis of Volumes Examined in Normalized Space (RAVENS) (Davatzikos et al., 2001; Goldsztal et al, 1998) maps of GM, WM, and CSF using deformable registration (DRAMMS) (Ou et al., 2011), and (6) spatial smoothing of RAVENS maps using a 6-mm Full Width at Half Maximum (FWHM) Gaussian filter. In our study, we investigated the RAVENS value changes in the GM images, which indicated regional GM volumetric alterations in the brain. Note that the RAVENS method has been demonstrated to provide more accurate and reliable quantification of volumes of different brain regions (Goldsztal et al., 1998). In order to reduce computational requirements, and since SCZ is not known to be associated with very localized volumetric reductions, the GM RAVENS maps were resampled to have a spatial resolution of 3 × 3 × 3 mm³. After applying a GM mask, 66,732 voxel-wise features for each subject were extracted. These images yielded a feature matrix of dimensionality of 100 × 66,732 for each site dataset.

In order to ensure that features in each site are within the same level and contribute equally in the process of feature weight learning, the feature matrices were then standardized to zero mean and unit variance and contribute equally in the process of feature weight learning in the multi-task learning problem with the task learning tasks. The columns of $W_r \in R^d$ are the feature weights in each task, and the rows of $W (w^r \in R^d$, where $p = 1, 2, \ldots, d$) hold the weight vectors of each feature across all learning tasks. Our goal is to learn the feature weight matrix $W$ by minimizing the following cost function:

$$
\arg\min \sum_{i=1}^{d} \sum_{j=1}^{d} \max(0, 1 - y_j (w^r_i, x_i)) + \alpha \|W\|_1 + \beta \|W\|_{2,1},
$$

where $y_j \in \{1, -1\}$ stands for NCs and SCZs, respectively. Let $w_r \in R^d$ denotes a feature weight vector for the $i$th learning task, and $W = [w_1, w_2, \ldots, w_d] \in R^{d \times d}$ denotes the feature weight matrix of all tasks. The columns of $W_r \in R^d$ are the feature weights in each task, and the rows of $W (w^r \in R^d$, where $p = 1, 2, \ldots, d$) hold the weight vectors of each feature across all learning tasks.

2.3. Multi-site classification using multi-task learning framework

In our study, the multi-task learning framework was used to learn the site-shared and site-specific features in the three data sites. These selected features were then used to classify the different data sites respectively. The workflow of multi-site learning is shown in Fig. 1. We model the classification between SCZs and NCs of different sites as a multi-task learning problem with $t$ learning tasks (one task for each site). In the $i$th learning task, $X_i = [x_{i1}, x_{i2}, \ldots, x_{in}] \in R^{d \times n}$ is features of training samples, where $d$ is the number of features and $n$ is number of training samples, and their class labels are $Y_i = [y_{i1}, y_{i2}, \ldots, y_{in}] \in R^n$,
greater than the irrelevant features. Though the site-shared and site-specific features were both very important, the site-shared features were considered underlying the pathology of SCZ and more critical in characterizing disease alterations. Thus, to further emphasize the influence of site-shared features in the classification, we increased the weights of site-shared features using the following equation,

$$W' = \gamma W_w$$

where $\gamma \geq 1$.

Note that the $\gamma$ was not set to be a free parameter, which was also tuned using the nested-loop cross-validation in our study.

### 2.4 Classification with top ranked features

The $K$ top ranked features with the new feature weights were used to classify each dataset. The classifier equation is shown as follows.

$$y = \text{sign}(\sum w_i' x_i),$$

where $w_i'$ denotes the new weight value of the $i$th feature and $x_i$ denotes the feature value. Given one testing sample in the $i$th data site, the dot product of the feature weights and the feature values of the $K$ selected features were summed up to get the sample label.

### 2.5 Validation experiments

To evaluate the proposed method, ten comparison experiments were designed in the consideration of three aspects. (1) To show the effectiveness of gathering multi-site data, we reported the classification results of single-site, pooled, and multi-site data, respectively. Single-site classification learns features of each single-site data separately for classification; pooled classification pools the three data sites together as a larger dataset regardless of the site differences, which contributed to a simple binary classification of 150 controls and 150 patients; and multi-site classification uses multi-task learning to learn the site-specific and site-shared features simultaneously, in the three data sites, and combine the two types of features to classify each dataset separately. (2) To show the effectiveness of using multi-task feature learning framework, two widely used feature learning methods including principal component analysis (PCA) and two-sample $t$-test ($ttest2$) were also tested. (3) The proposed classifier based on the learned feature weights was compared to the SVM classifier with a linear kernel. Differences between these two classifiers exist as the SVM focuses on the selected features and calculates new feature weights for classification, while our proposed classifier takes advantage of the feature weights obtained in the feature learning step, and then generate the classification label. Table 2 listed the ten comparison methods in the current study.
In the first three experiments, the same proposed feature learning and classifier framework was applied in the multi-site, single-site and pooled data classification, respectively. In the single-site data classification, for each data site, we used the $l_2$-norm feature selection (Wang et al., 2015), which included the hinge loss function and the $l_1$-norm sparsity regularization term in the cost function. In the pooled data classification, the feature learning cost function was similar to the single-site classification, which considered the pooled data as a single cohort. Note that the feature learning framework applied in the single-site data classification and the pooled data classification was a special case of the multi-task feature learning framework, with the $l_2,1$ sparsity regularization term disabled. The site-specific features were learned in the same way in all three classification experiments, while the site-shared features were only emphasized in multi-site learning, which could show the influence of site-shared features more clearly. The classifier for single-site data and pooled data classification was the same as Eq. (3), however, increasing weights of site-shared features in Eq. (2) was not applied as we did in multi-site learning, since there were no site-shared features in the single-site and pooled data classification.

In other seven experiments, the SVM classifier with a linear kernel was used after different feature learning methods. As shown in Table 2, the 4th to 6th experiments used the proposed multi-task feature learning technique, taking the feature learning method in single-site and pooled data as a special case of multi-site feature learning. In the 7th and 8th experiments, the PCA method was used, while in the 9th and 10th experiments, the test2 method was used in the single-site and pooled data classification, respectively.

The nested loop 5-fold cross-validation was used in the ten experiments to evaluate the classification performance. In specific, the samples were randomly separated into five subsets: one (20% of all samples) was used for testing, and the other four (80% of all samples) were divided further into two parts with the proportion rates of 80% (64% of all samples) and 20% (16% of all samples), respectively. The larger of the two parts was used as a training set, and the smaller part was used as a validation set, as described previously (Koutsouleris et al., 2014). The training samples were used for feature learning and classifier training, while the validation samples were used to tune all parameters with the trained classifier, and the best parameter set which contributed to the best classification accuracy of the validation samples was selected. Finally, the best parameters in the validation step were used to classify the testing samples and obtain the classification result. The whole process was repeated for 20 repeated experiments. The average accuracies of the 20 iterations for each site was considered as the final classification result.

The parameters of the multi-task feature learning framework were tuned by grid searching the regularization factors in the range of $\alpha, \beta \in \{10^{-8}, 10^{-6}, 10^{-4}, 10^{-2}, 10^{-1}\}$, 0, 0.3, 0.5, 1, 5, 10, $10^2, 10^3\}$, while in the special case of using multi-task feature learning in the single-site and pooled data, the $l_2,1$-norm term was disabled and $\beta$ was set to 0. In the proposed classifier in the multi-site classification (the first experiment), the ratio to increase the site-shared feature weights $\gamma$ was tuned in the range of $\{1, 2, \ldots, 100\}$. The number of selected features $K$ ranged from $\{500, \ldots, 20,000\}$, with the step-size of 100. In the experiments with the SVM classifier, the linear kernel was used and the cost value $C$ was set to 1.

The site-shared and site-specific features were obtained in the multi-site classification, which used the proposed multi-task learning framework. After the parameter tuning, the best parameter set of the regularization terms $\alpha, \beta$, the number of selected features $K$, and the site-shared feature increasing ratio $\gamma$ were obtained. In each iteration, the best classification accuracy in each site corresponded to a specific $W$ vector. We first calculated the mean $W$ vectors of the 20 iterations that were contributed to the best accuracy in all three sites respectively, then the mean $W$ vector of sites A, B, and C was obtained. The consensus features with weights that ranked the top 5% of the three sites were shown as the site-shared features, and the remaining features were considered as the site-specific features for each site, respectively. The feature weights of the site-shared and site-specific features were displayed using the Caret package (Van Essen, 2005), in which the clusters comprising of $> 50$ voxels were shown.

3. Results

The classification results of multi-site, single-site, and pooled data were listed in Table 3. Using the proposed multi-task feature learning framework and the SVM classifier based on the learned feature weights, the average classification accuracies of sites A, B, and C, based on the multi-site data learning, were 78.5%, 64.5%, and 84.5%, which outperformed the accuracies of 76%, 56%, and 82.5% on single-site learning, as well as the accuracies of 70%, 57.5%, and 75.3% on the pooled classification.

Using the multi-task feature learning framework and the SVM classifier, the average accuracies in multi-site (67.3%), single-site (69.0%) and pooled (61.7%) data classification were lower than the

| Table 2 | Summary of the comparison methods. |
|---------|------------------------------|
| Method | Description |
| $l_1,1 + l_1$ (MS) | The proposed multi-task classification framework on multi-site data, which contains $l_2,1 + l_1$ regularization terms in the feature learning step, and the classifier based on the learned feature weights. |
| $l_1$ (SS) | Single-task classification on each single site data: the single-task feature learning step uses the $l_1$ regularization term, and the classifier uses the same classifier formula as the multi-task classification framework. |
| $l_1$ (PO) | Using multi-site data, the feature learning step is multi-task learning, and uses SVM classifier with the linear kernel. |
| $l_1,1 + l_1 + SVM$ (MS) | Using multi-site data, the feature learning step is multi-task learning, and uses SVM classifier with the linear kernel. |
| $l_1 + SVM$ (SS) | Use single-task feature learning framework and a SVM classifier to classify each single data site. |
| $l_1 + SVM$ (PO) | Use single-task feature learning framework and a SVM classifier to classify the pooled data. |
| PCA + SVM (SS) | Use principal component analysis (PCA) to learn features in each single site data and a SVM classifier with the linear kernel. |
| PCA + SVM (PO) | Use principal component analysis (PCA) to learn features on the pooled data and a SVM classifier with the linear kernel. |
| test2 + SVM (SS) | Use two-sample t-test (test2) to learn features in each single site data and a SVM classifier with the linear kernel. |
| test2 + SVM (PO) | Using two-sample t-test (test2) to learn features on the pooled data and a SVM classifier with the linear kernel. |

| Table 3 | The average accuracy values of multi-site, single-site and pooling classification with 20 repetitions in 10 experiments. |
|---------|---------------------------------|
| Method | Site A | Site B | Site C | Average of 3 sites |
| $l_1,1 + l_1$ (MS) | 0.785 | 0.645 | 0.845 | 0.758 |
| $l_1$ (SS) | 0.76 | 0.56 | 0.825 | 0.713 |
| $l_1$ (PO) | 0.70 | 0.575 | 0.753 | 0.676 |
| $l_1,1 + l_1 + SVM$ (MS) | 0.713 | 0.648 | 0.668 | 0.673 |
| $l_1 + SVM$ (SS) | 0.728 | 0.598 | 0.825 | 0.690 |
| $l_1 + SVM$ (PO) | 0.610 | 0.573 | 0.668 | 0.617 |
| PCA + SVM (SS) | 0.668 | 0.474 | 0.724 | 0.622 |
| PCA + SVM (PO) | 0.646 | 0.494 | 0.732 | 0.624 |
| test2 + SVM (SS) | 0.562 | 0.468 | 0.680 | 0.570 |
| test2 + SVM (PO) | 0.672 | 0.530 | 0.646 | 0.616 |

MS: multi-site data; SS: single-site data; PO: pooled data.
results using the proposed classifier (75.8%, 71.3%, and 67.6%), respectively. However, the average accuracy of multi-site classification was poorer than the single-site classification using multi-task feature learning framework and SVM classifiers. In the experiments using PCA feature learning method and SVM classifiers, the average results of single-site and pooled data classification were 62.2% and 62.4%. In the experiments using test2 feature learning method, the classification accuracies were the lowest comparing to all other methods, which were 57.0% and 61.6%, respectively in single-site and pooled data classification.

The best average accuracy was 75.8% with our proposed multi-task feature learning classification method. The results indicated that multi-site learning captured more precise feature patterns in revealing the differences between SCZ patients and controls, than using only one single-site data. Simply pooling all data sites together did not improve classification performance compared with using single-site data.

The site-shared and site-specific features with the largest weights were mapped onto a brain image template. As shown in the left panel of Fig. 2, multi-site learning extracted a series of feature patterns as site-shared features, which revealed gray matter abnormalities in a variety of regions, including the fusiform, middle temporal, superior temporal, inferior parietal, postcentral, angular, inferior frontal, middle occipital and lingual gyri, insular, anterior cingulate cortex, precuneus, the dorsolateral prefrontal cortex, and the cerebellum. The most discriminative site-shared regions are all listed in Table 4.

The site-specific features were displayed in the right panel of Fig. 2. The site-specific features of site A included regions of the fusiform, lingual, inferior temporal, middle temporal, superior temporal, middle frontal, superior frontal, superior medial frontal, postcentral, superior parietal, precentral gyri, precuneus, putamen, supplemental motor area, and the cerebellum. The site-specific features of site B included regions of the fusiform, anterior cingulate cortex, posterior cingulate cortex, hippocampus, parahippocampus, angular, inferior frontal, middle frontal, postcentral, precentral, inferior parietal, middle occipital gyri, precuneus, and the cerebellum. As shown in the right panel of Fig. 2, multi-site learning extracted a series of feature patterns as site-shared and site-specific features were shown with cluster size of > 50 voxels obtained by multi-task learning on multi-site schizophrenia classification. The colorbar represents the weight values of features. The warm and hot colors corresponded to negative and positive weight values, respectively.

4. Discussion

We present a method utilizing multi-task l2,1 + l1-norm learning to simultaneously learn the site-shared and site-specific features, for multi-site classification problems. The method was tested to differentiate SCZ patients from healthy controls in three data sites. Compared to single-site and pooled data classification, the multi-site classification performance was significantly higher, which demonstrated that multi-site data can enhance accuracy in identifying brain alterations in SCZ if multi-task learning is applied. The multi-task learning framework thus provides a new way to take advantage of multi-site data to better classify and understand brain diseases.

4.1. Reliable identification of multi-site schizophrenia data using multi-task learning

Single-site data contains information about pathophysiology, but may simultaneously include information related to different scanners or imaging protocols, as well as demographic differences that may influence the pattern investigations across individual subject samples. Using single-site data is limited by smaller sample sizes and smaller diversity.
in imaging features, both of which often lead to poor generalization performance of classifiers on new patient populations. Data pooled from multiple sites alleviates both of these problems at the expense of introducing confounding heterogeneity in the imaging data, which renders the pooled analysis challenging. In our study, using the proposed multi-task feature learning scheme, the classification accuracy of using pooled data was no better than using each single site, and, in the proposed multi-task feature learning scheme, the classification accuracy of using the pro-

| Regions                          | Side | BA       | Cluster size (voxels) | MNI coordinates     |
|---------------------------------|------|----------|-----------------------|---------------------|
| Cerebellum Crus I, fusiform gyrus, lingual gyrus | L    | 18,19    | 146                   | −32, −86, −17/ −29, −74, −3 |
| Middle temporal gyrus           | L    | 38,21,22 | 77                    | −44,12, −35/ −59, −14, −8/ −56, −59, −13/ −65, −35, −1 |
| Superior temporal gyrus, middle temporal gyrus, insula | R    | 22,41,13 | 77                    | 52, −41, −10/49, −17,1 |
| Anterior cingulate, superior medial frontal gyrus | L    | 32      | 51                    | −12,34,1/ −12,34,25 |
| Middle occipital gyrus, angular | L/R  | 39,40    | 111                   | 45, −55,20/33, −76,32/ −51,67,39 |
| Insula                          | R    | 13      | 27                    | 33,16,12 |
| Middle frontal gyrus, inferior triangular frontal gyrus | L/R  | 9,10,44,45 | 162                  | −51,26,28/ −33,52,16/- 42,22,13/39,35,24 |
| Postcentral                     | L    | 2       | 18                    | −49, −31,50 |
| Precuneus                       | L/R  | 31      | 25                    | −6, −43,36 |
| Inferior parietal gyrus         | L    | 40      | 15                    | −40, −49,51 |
| Superior parietal gyrus         | L    | 7,5     | 42                    | −28, −51,75 |
| Superior frontal gyrus, middle frontal gyrus | R    | 9,6     | 56                    | 27,47,32/33,41,35,35,8,42 |
| Superior frontal gyrus, middle frontal gyrus | L    | 32,6   | 63                    | −16,23,47/- 34,11,52 |
| dIPFC                           | L/R  | 9,6,32  | 119                   | −16,23,47/- 34,11,52/27,47,32/33,41,35,35,8,42 |

Note: BA = Brodmann area, R = right, L = left.

Table 6
The site-specific gray matter alteration features of brain regions in site B.

| Regions                          | Side | BA       | Cluster size (voxels) | MNI coordinates     |
|---------------------------------|------|----------|-----------------------|---------------------|
| Cerebellum IV, V, fusiform gyrus | R    | 19,36    | 51                    | 22, −35, −20 |
| Hippocampus, parahippocampus, fusiform gyrus | R    | 36      | 75                    | 34, −35, −11 |
| Anterior cingulate cortex       | L    | 32,9,24,10 | 294                  | −3,38,23 |
| Inferior parietal gyrus, middle occipital gyrus, angular | R    | 40,39,7 | 105                   | −46, −49,47 |
| Posterior cingulum              | L    | 65      | 9                    | −9, −43,36 |
| Inferior frontal gyrus          | L    | 13,44,47 | 38                    | −39,22,13 |
| Middle frontal gyrus            | L    | 8,9     | 53                    | −34,35,37 |
| Precuneus                       | L    | 7       | 120                   | −13, −67,37 |
| Postcentral gyrus, precentral gyrus | L/R  | 6,4,3,40 | 180                   | −52, −31,53 |

Note: BA = Brodmann area, R = right, L = left, dIPFC = dorsolateral prefrontal cortex.

SVM classifier only used the sequencing information of the feature weights, which had lost some valuable information of these selected features. Therefore, better results using our proposed classifier showed that the feature weights learned by the multi-task feature learning method were accurate.

4.2. Site-shared and site-specific regions of gray matter alterations in schizophrenia

A pattern of gray matter intensity alterations implicating several brain regions in the SCZ sample was revealed by the site-shared feature analysis of multi-site MRI data in our study. This is consistent with a
vast number of large-scale structural MRI investigations of SCZ, including both meta-analyses of MRI results (Chan et al., 2011; Ellison-Wright et al., 2008; Honea et al., 2005; Palaniyappan et al., 2012; Steen et al., 2006; Vita et al., 2012; Wright et al., 2000), and mega-group comparisons of mean indices combining data from multiple sites (Gupta et al., 2015; Segall et al., 2009; Torres et al., 2016). These previous studies have repeatedly highlighted the prefrontal cortex, lateral temporal cortex, insular, striatum, anterior cingulate cortex, hippocampus and other temporo-limbic structures as key components of the network of brain regions showing volumetric deficits relative to healthy controls from early stages of SCZ onwards. Most of those brain regions were implicated in our study, indicating a high degree of consistency between the spatial distribution of findings in our site-shared feature analysis and the previous MRI literature on SCZ. The brain regions involved in this network are critical to emotion processing and several cognitive domains known to be affected in schizophrenia, including language processing, working memory and other executive functioning operations, episodic memory, attention, self-monitoring, error detection, and attribution of salience to emotionally relevant stimuli, among others (Backasch et al., 2014; Belin et al., 2009; Gabeza and Nyberg, 2000; Chao et al., 1999; Jeong et al., 2009; Kim et al., 2003; Noppeney and Price, 2002; Ojemann et al., 2002; Palaniyappan and Liddle, 2012a; Palaniyappan and Liddle, 2012b; Petkin et al., 2009; Price, 2010; Tranel et al., 1997; Wright et al., 2003). Thus, the findings presented herein provide evidence that our method robustly identifies consistent and reliable structural brain abnormalities associated with SCZ across different sites. Additional brain regions often implicated in the pathophysiology of SCZ were also highlighted in our site-shared analysis, such as the cerebellum (Andreasen and Pierson, 2008). Though the cerebellum is traditionally believed to participate in motor function, neuroimaging studies in both nonhuman primes and humans, have demonstrated that it is also involved in brain cognitive and affective processing (Bernard and Mittal, 2015; Dum and Strick, 2003; Hu et al., 2008; Kelly and Strick, 2003). Several structural and functional neuroimaging studies have shown cerebellar abnormalities in SCZ subjects relative to controls (Bernard and Mittal, 2015; Gupta et al., 2015; Ha et al., 2004; Kim et al., 2014; Mothersill et al., 2016; Rusch et al., 2007; Shen et al., 2010; Shenton et al., 2001; Wagle et al., 2015). Abnormalities in the cerebellum, through its connections to the prefrontal cortex via thalamus, have been proposed as critical to the state of “cognitive dysmetria” put forward by some authors to explain the behavior manifestations and cognitive deficits of schizophrenia (Andreasen et al., 1999; Rusch et al., 2007; Shen et al., 2010). However cerebellar findings should be interpreted cautiously, since segmentation between GM and WM in that region is difficult and prone to errors, due to poor contrast at this imaging resolution. Finally, other brain sites showing volume abnormalities in our study, including the precuneus and lateral parietal cortical regions, fusiform, lingual and middle occipital gyri, have been variably implicated in previous neuroimaging studies of SCZ (Borgwardt et al., 2010; Brunet-Gouet and Decety, 2006; Chou et al., 2014; Davatzikos et al., 2005; Di Rosa et al., 2009; Dong et al., 2006; Faget-Agius et al., 2012; Gaser et al., 1999; Herrmann et al., 2006; Johnston et al., 2005; Lee et al., 2002; Mashal et al., 2014; Nierenberg et al., 2005; Niznikiewicz et al., 2000; Onitsuka et al., 2005; Onitsuka et al., 2003; Quandt et al., 2014; Schultz et al., 2010; Seifert et al., 2008; Shenton et al., 2001; Singh et al., 2014; Takahashi et al., 2011; Tanskanen et al., 2010; Walther et al., 2009; Whalley et al., 2006; Yoon et al., 2006). These brain regions are involved in a variety of cognitive operations of relevance to schizophrenia, including language, visual attention and emotional processing (Cohen et al., 1996; Dong et al., 2006; Herrmann et al., 2006; Johnston et al., 2005; Kaas, 1995; Mashal et al., 2014; Nierenberg et al., 2005; Niznikiewicz et al., 2000; Seifert et al., 2008; Walther et al., 2009; Whalley et al., 2006; Yoon et al., 2006).

Findings of structural brain abnormalities in SCZ may be influenced by clinical variables not specifically related to the pathophysiology of the disorder, such as the use of antipsychotic medications. Several recent studies have shown that the continued use of antipsychotics may be significantly associated with some of the neuroanatomical changes detected in MRI studies of SCZ, even after taking into account the influence of other potential moderators such as illness severity and substance abuse (Fusar-Poli et al., 2013; Ho et al., 2011; Torres et al., 2016). In the present study, participants with SCZ from site B were all in their first psychotic episode and had a relatively shorter duration of illness in comparison to the samples contributed by sites A and C (both of which included a substantial proportion of chronic, medicated SCZ patients). Thus the overall SCZ sample from site B was less exposed to antipsychotic treatment in comparison to the other two sites. The selection of first-episode SCZ patients in site B may help to explain the more modest diagnostic accuracy figures obtained in the single-site analysis for that center in comparison to the two other sites, given the fact that medicated SCZ patients in chronic disease stages display more widespread brain volumetric abnormalities relative to healthy controls than first-episode SCZ patients (Meisenzahl et al., 2008; Torres et al., 2016). Since the three imaging sites contributed each an equal number of SCZ patients and controls to the present investigation, disease chronicity and antipsychotic exposure should be taken as clinical variables that increased inter-site variability in our study, instead of representing unifying features present in all sites. Therefore, the results of the site-shared analysis presented herein can be more safely assumed to reflect the commonality of SCZ neuropathology across the three sites, rather than the potential influence of confounding clinical variables ubiquitously present in the three centers. Disentangling the specific neurobiological features underlying psychiatric disorders from clinically relevant confounders is challenging in neuroimaging research, given the highly prevalent influence of factors such as use of nicotine and other drugs of abuse, cardiovascular risk factors, and exposure to pharmacological treatment, all of which can affect brain structure (Weinberger and Rudenscu, 2016).

The site-specific regions of the three sites revealed in our study consisted of two parts. The larger part of the site-specific features was
located in the same regions of the site-shared features, though the coordinates were slightly different. These regions involved most of the site-shared features, including the fusiform, lingual, postcentral, precuneus, prefrontal gyri and the cerebellum. The findings further implied that these site-shared and also the site-specific regions revealed the abnormal feature patterns of the SCZ accurately, and multi-task feature learning framework provided a promising method in detecting these features. However, though these site-specific features were located in the same regions with the site-shared features, the feature distributed in adjacent areas. This was reasonable due to the data heterogeneity resulted from each site. Another part of the site-specific regions included some site-specific regions found exclusively in each site, such as putamen and supplemental motor area in site A, hippocampus, parahippocampus, posterior cingulate cortex in site B, and supramarginal gyrus in site C.

Studies have shown that supplemental motor area dysfunction was related to motor disturbances in SCZ patients (Schröder et al., 1995). The putamen was proposed to be related to the treatment response (Buchsbaum et al., 2003; Li et al., 2012; Mitelman et al., 2009), and also reflected the deficits of auditory verbal hallucinations of SCZ patients (Cui et al., 2016). The gyrus of hippocampus and parahippocampus were reduced (Razi et al., 1999), and hippocampus was demonstrated to be associated with the positive symptoms of SCZ patients (Duan et al., 2015). Meanwhile, the posterior cingulate cortex has recently been implicated in the pathophysiology of schizophrenia (Newell et al., 2006). These site-specific regions may provide potential evidences on interpreting the underlying multisite data heterogeneity introduced by different scanners, subjects, treatment exposure, or disease symptoms or etc.

4.3. Limitations and future directions

Although the classification result of multi-site MRI data using multi-task learning was encouraging, there are limitations in our work. First, due to the limited sample sizes, small site numbers, and the specific psychiatric disease type and data modality in our study, the effectiveness of using mult-task learning scheme on multi-site classification needs to be further validated in larger datasets. Moreover, an important challenge in multi-task learning is that when new data is analyzed, the multi-task learner would need to be recalculated. This would obviously limit clinical utility. One solution is to apply our approach to obtain a set of $t$ discriminants, each associated with one of our training sites. It is possible that as new data and sites are added to our analyses, one would not need to solve the multi-task learning problem again, but rather could apply all these $t$ vectors and pool the classification results via weighted consensus approaches.

5. Conclusion

This work proposes a multi-task learning framework for high-dimensional classification, which simultaneously learns the site-specific and site-shared features in multi-site MRI data of brain morphology. This method was tested on multi-site anatomical images of SCZ patients. The classification results of using multi-task learning outperformed single-task learning and the pooled data. The most significant site-shared features revealed brain structural alterations in SCZ consistent with those previously reported in the literature. In summary, multi-task learning provides a promising technique to learn the site-shared features in multi-site data study without neglecting site-specific features, which gives a new angle on integrating multiple single-site data together for big data studies.

Acknowledgement

This work was supported by NIH Grants (R01 MH112070); the National Science Foundation of China (61420106001, 91420302, 61773391, 61722313); and Fok Ying Tung Education Foundation (161057).

References

Andreasen, N.C., Pierson, R., 2008. The role of the cerebellum in schizophrenia. Biol. Psychiatry 64, 81–88.
Andreasen, N.C., Nopoulos, P., O’Leary, D.S., Miller, D.D., Wassink, T., Flau, M., 1999. Defining the phenotype of schizophrenia: cognitive dysmetria and its neural mechanisms. Biol. Psychiatry 46, 908–920.
Azazi, S., Sra, S., 2014. Towards an optimal stochastic alternating direction method of multipliers. In: Proceedings of the 31st International Conference on Machine Learning (ICML).
Backsch, B., Sommer, J., Klohn-Saghatiolaif, F., Muller, M.J., Kircher, T.T., Leube, D.T., 2014. Disconnectivity of the inferior frontal gyrus: implications for an impaired self-other distinction in patients with schizophrenia. Psychiatry Res. 223, 202–209.
Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature 403, 309–312.
Bernard, J.A., Mittal, V.A., 2015. Dysfunctional activation of the cerebellum in schizophrenia: a functional neuroimaging meta-analysis. Clin. Psychol. Sci. 3, 545–566.
Borgwardt, S.J., Piccioni, M.M., Ettinger, U., Troupoupolou, T., Murray, R., McGuire, P.K., 2010. Regional gray matter volume in monzygotic twins discordant and discordant for schizophrenia. Biol. Psychiatry 67, 956–964.
Brown, G.G., Mathalon, D.H., Stern, H., Ford, J., Mueller, R., Greve, D.N., McCarthy, G., Voyvodic, J., Glover, G., Ogryzlo, P., de Jonge, P., de Weerth, F., Turner, J.A., Thompson, W.K., Potkin, S.G., Function Brain Biomedical Informatics Research Network. 2011. Multisite reliability of cognitive BOLD data. Neuroimage 54, 2163–2175.
Brunet-Gour, E., Decety, J., 2006. Social brain dysfunctions in schizophrenia: a review of neuroimaging studies. Psychiatry Res. 148, 75–92.
Buchsbaum, M.S., Shiabudden, L., Brickman, A.M., Miozzo, R., Prikril, R., Shaw, R., Davis, K., 2003. Caudate and putamen volumes in good and poor outcome patients with schizophrenia. Schizophr. Res. 64, 53–62.
Cabeza, R., Nyberg, L., 2000. Imaging cognition II: an empirical review of 275 PET and fMRI studies. J. Cogn. Neurosci. 12, 1–47.
Calloway, V.D., Adali, T., 2009. Weighted consensus approaches.
Chao, L.L., Haxby, J.V., Martin, A., 1999. Attribute-based neural substrates in temporal lobe of normal and patient populations. J. Cogn. Neurosci. 11, 475–492.
Chou, P.H., Koike, S., Nishimura, Y., Kawasaki, S., Satomura, Y., Kinoshita, A., Takizawa, R., Kasai, K., 2014. Distinct effects of duration of untreated psychosis on brain cortical activity in different treatment phases of schizophrenia: a multi-channel near-infrared spectroscopy study. Prog. Neuro-Psychopharmacol Biol. Psychiatry 49, 289–241.
Chou, P.H., Koike, S., Nishimura, Y., Kawasaki, S., Satomura, Y., Kinoshita, A., Takizawa, R., Kasai, K., 2014. Distinct effects of duration of untreated psychosis on brain cortical activity in different treatment phases of schizophrenia: a multi-channel near-infrared spectroscopy study. Prog. Neuro-Psychopharmacol Biol. Psychiatry 49, 63–69.
Cohen, M., Koslyn, S., Breiter, H., DiGirolamo, G., Thompson, W., Anderson, A., Brookheimer, S., Rosen, B., Belliveau, J., 1996. Changes in cortical activity during mental rotation. Brain 119, 89–100.
Colby, J.B., Rudie, J.D., Brown, J.A.,月圆, D.K., Cohen, M.S., Sheline, Z., 2012. Insights into multimodal imaging classification of AD/HD. Front. Syst. Neurosci. 6, 59.
Costafreda, S.G., Brammer, M.J., Veness, R.Z., Mourao, M., Portela, L.A., de Carvalho, C.C., Giampietro, V.P., Amaro Jr., E., 2007. Multisite fMRI reproducibility of a motor task using identical MR systems. J. Magn. Reson. Imaging 26, 1122–1126.
Cui, L.B., Liu, K., Li, C., Wang, L.X., Guo, F., Tian, P., Wu, Y.J., Guo, L., Liu, W.M., Xi, Y.B., Wang, H.N., Yin, H., 2016. Putamen-related regional and network functional deficits in first-episode schizophrenia with auditory verbal hallucinations. Schizophr. Res. 173, 12–22.
Davatzikos, C., 1998. Mapping image data to stereotaxic spaces: applications to brain mapping. Hum. Brain Mapp. 6, 334–354.
Davatzikos, C., Genc, A., Xu, D., Rensick, S.M., 2001. Voxel-based morphometry using the RAVENTS maps: methods and validation using simulated longitudinal atrophy. Neuroimage 14, 1361–1369.
Davatzikos, C., Shen, D., Gur, R.C., Wu, X., Liu, D., Fan, Y., Hughtt, P., Turetsky, B.I., Gur, R.E., 2005. Whole-brain morphometric study of schizophrenia revealing a spatially complex set of focal abnormalities. Arch. Gen. Psychiatry 62, 1218–1227.
Di Rosa, E., Crow, T.J., Walker, M.A., Black, G., Chance, S.A., 2009. Reduced neuron density, enlarged minicolumn spacing and altered ageing effects in fusiform cortex in schizophrenia. Psychiatry Res. 166, 102–115.
neuroimaging: advantages, problems, and challenges. Schizophr. Bull. 35, 1–2.
Pfeiferbaum, A., Adalsteinsson, E., Sullivan, E.V., 2003. Replicability of diffusion tensor imaging measurements of fractional anisotropy and trace in brain. J. Magn. Reson. Imaging 18, 427–433.
Potkin, S.G., Turner, J.A., Brown, G.G., McCarthy, G., Greve, D.N., Glover, G.H., Manoach, D.S., Belger, A., Diaz, M., Wibbe, C.G., Ford, J.M., Mathalon, D.H., Gollub, R., Lauriello, J., O’Leary, D., van Erp, T.G., Toga, A.W., Preda, A., Lim, K.O., FBIRN, 2009. Working memory and DLPCF inefficiency in schizophrenia: the FBIRN study. Schizophr. Bull. 35, 19–31.
Price, C.J., 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann. N.Y. Acad. Sci. 1191, 62–88.
Quarantelli, M., Palladino, O., Prinster, A., Schiavone, V., Carotenuto, B., Brunetti, A., Marsili, A., Castello, M., Musettola, G., Salvatore, M., de Bartolomeis, A., 2014. Patients with poor response to antipsychotics have a more severe pattern of frontal atrophy: a voxel-based morphometry study of treatment resistance in schizophrenia. Biomed. Res. Int. 2014, 325052.
Rao, N.S., Cox, C.R., Nowak, R.D., Rogers, T.T., 2013. Sparse overlapping sets Lasso for multitask learning and its application to fMRI analysis. In: Conference on Neural Information Processing Systems. 2013.
Razi, K., Greene, K.P., Sakuma, M., Ge, S., Kushner, M., Delisi, L.E., 1999. Reduction of the parahippocampal gyrus and the hippocampus in patients with chronic schizophrenia. Br. J. Psychiatry 174, 512–519.
Ruch, N., Spointi, I., Wilke, M., Bria, P., Di Paola, M., Di Iulio, F., Evans, A., Murray, R.M., Kahn, R.S., Hulshoff Pol, H.E., 2010. Mapping reliability in multicenter MRI: voxel-based morphometry and cortical thickness. Hum. Brain Mapp. 31, 1967–1982.
Schroder, J., Wenz, F., Schad, L.R., Baudendistel, K., Knopp, M.V., 1995. Sensorimotor cortex and supplementary motor areas changes in schizophrenia. A study with functional magnetic resonance imaging. Br. J. Psychiatry 167, 197–201.
Schulz, C.C., Koch, K., Wagner, G., Roebel, M., Nenadic, I., Gaser, C., Schachtzabel, C., Schnack, H.G., van Haren, N.E., Brouwer, R.M., van Baal, G.C., Picchioni, M., Rusch, N., Wilke, M., Bria, P., Di Paola, M., Di Iulio, F., Martinotti, G., Sutton, B.P., Goh, J., Hebrank, A., Welsh, R.C., Chee, M.W., Park, D.C., 2011. DRAMMS: deformable registration investigation and validation of intersite fMRI studies using the same imaging hardware. J. Magn. Reson. Imaging 28, 21–28.
Takahashi, T., Zhou, S.Y., Nakamura, K., Tanino, R., Furuschi, A., Kido, M., Kawasaki, Y., Noguchi, K., Seto, H., Kurachi, M., Suzuki, M., 2011. A follow-up MRI study of the fusiform gyrus and middle and inferior temporal gyr in schizophrenia spectrum. Prog. Neuro-Psychopharmacol. Biol. Psychiatry 35, 1957–1964.
Tang, L., Chee, Y., Maloney, T.C., Caré, M.M., Holland, S.K., Lu, L.J., 2013. Combined analysis of sMRI and fMRI imaging data provides accurate disease markers for hearing impairment. NeuroImage Clin. 3, 416–428.
Tanskanen, P., Ridler, K., Murray, G.K., Haapela, M., Veijola, J.M., Jaakelainen, E., Miettunen, J., Jones, P.B., Bullmore, E.T., Iohannides, M.K., 2010. Morphometric brain abnormalities in schizophrenia in a population-based sample: relationship to duration of illness. Schizophr. Bull. 36, 766–777.
Torres, U.S., Duran, F.L., Schaufelberger, M.S., Crippa, J.A., Loua, M.R., Sallet, P.C., Kanegusuku, C.Y., Elks, H., Gattar, W.F., Barstot, D.P., Zauder, A.W., Hallak, J.E., Leite, C.C., Castro, C.C., Santos, A.C., Murray, R.M., Busatto, G.F., 2016. Patterns of regional gray matter loss at different stages of schizophrenia: a multisite, cross-sectional VBM study in first-episode and chronic illness. NeuroImage Clin. 12, 1–15.
Tranel, D., Damasio, H., Damasio, A.R., 1997. A neural basis for the retrieval of conceptual knowledge. Neuropsychologia 35, 1319–1327.
Van Essen, D.C., 2005. A population-averaged, landmark- and surface-based (PALS) atlas of human cerebral cortex. NeuroImage 28, 635–662.
Van Horn, J.D., Toga, A.W., 2009. Multisite neuroimaging trials. Curr. Opin. Neurol. 22, 370–378.
Vita, A., De Peri, L., Deste, G., Sacchetti, E., 2012. Progressive loss of cortical gray matter in schizophrenia: a meta-analysis and meta-regression of longitudinal MRI studies. Transl. Psychiatry 2, e190.
Wagemol, D., Knowlton, B.J., Cohen, J.R., Bookheimer, S.Y., Bilder, R.M., Fernandez, V.G., Asarnow, R.F., 2015. Cognitive correlates of gray matter abnormalities in adolescent siblings of patients with childhood-onset schizophrenia. Schizophr. Res. 161, 345–356.
Waltz, S., Federups, A., Horn, H., Bianchi, P., Wiest, R., Wirth, M., Strik, W., Muller, T.J., 2009. Encoding deficit during face processing within the right fusiform face area in schizophrenia. Psychiatry Res. 172, 184–191.
Wang, H., Nie, F., Huang, H., Risher, S., Ding, C., Suykin, A.J., Shen, L. (Sparse multitask regression and feature selection to identify brain imaging predictors for memory performance). In: 2011. p 557–562.
Wang, X., Zhang, T., Chaim, T., Zanetti, M., Davatzikos, C., 2015. Classification of MRI under the presence of disease heterogeneity using multi-task learning: application to bipolar disorder. In: Proceeding of the 18th Annual International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI).
Watanabe, N., Kesler, S., Scott, C., Spriada, C., 2014. Multisite Disease Classification with Second International Workshop on Sparsity Techniques in Medical Imaging. Boston.
Weinberger, D.R., Radelicus, E., 2016. Finding the elusive psychiatric “lesion” with 21st-century neuroanatomy: a note of caution. Am. J. Psychiatry 173, 27–33.
Whalley, H.C., Simonton, E., Moorhead, W., McIntosh, A., Marshall, I., Ebmeier, K.P., Owens, D.G., Goddard, N.H., Johnstone, E.C., Lavrie, S.M., 2006. Functional imaging as a predictor of schizophrenia. Biol. Psychiatry 60, 454–462.
Wright, I.C., Rabe-Hesketh, S., Woodruff, P.W., David, A.S., Murray, R.M., Bullmore, E.T., 2000. Meta-analysis of regional brain volumes in schizophrenia. Am. J. Psychiatry 157, 16–25.
Wright, T.M., Phelpes, K.A., Allison, T., McKeown, M.J., McCarthy, G., 2003. Prefrontal and subcortical involvement in a population-based sample of first-episode schizophrenia. Prog. Neuro-Psychopharmacol. Biol. Psychiatry 34, 116–125.