Neuroscience and Biobehavioral Reviews 57 (2015) 328–349

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

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

Review

From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics

Thomas Wolfers[a,b,*], Jan K. Buitelaar[c,d], Christian F. Beckmann[b,c,e], Barbara Franke[a,f], Andre F. Marquand[b,g]

[a] Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, The Netherlands
[b] Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University, PO Box 9101, 6500 HB Nijmegen, The Netherlands
[c] Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, The Netherlands
[d] Karakter Child and Adolescent Psychiatry University Centre, Radboud University Medical Centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands
[e] Centre for Functional MRI of the Brain (FMRIB), University of Oxford, Oxford, United Kingdom
[f] Department of Psychiatry, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands
[g] Department of Neuroimaging, Institute of Psychiatry, King's College London, LondonUnited Kingdom

ARTICLE INFO

Article history:
Received 18 May 2015
Received in revised form 29 July 2015
Accepted 2 August 2015
Available online 4 August 2015

Keywords:
Pattern recognition
Magnetic resonance imaging
Psychiatric diagnostics
Psychiatric disorders
Mental disorders
Schizophrenia
Bipolar disorder
Major depressive disorder
Obsessive compulsive disorder
Social anxiety disorder
Post-traumatic stress disorder
Specific phobia
Attention-deficit/hyperactivity disorder
Autism spectrum disorder

ABSTRACT

Psychiatric disorders are increasingly being recognised as having a biological basis, but their diagnosis is made exclusively behaviourally. A promising approach for 'biomarker' discovery has been based on pattern recognition methods applied to neuroimaging data, which could yield clinical utility in future. In this review we survey the literature on pattern recognition for making diagnostic predictions in psychiatric disorders, and evaluate progress made in translating such findings towards clinical application. We evaluate studies on many criteria, including data modalities used, the types of features extracted and algorithm applied. We identify problems common to many studies, such as a relatively small sample size and a primary focus on estimating generalisability within a single study. Furthermore, we highlight challenges that are not widely acknowledged in the field including the importance of accommodating disease prevalence, the necessity of more extensive validation using large carefully acquired samples, the need for methodological innovations to improve accuracy and to discriminate between multiple disorders simultaneously. Finally, we identify specific clinical contexts in which pattern recognition can add value in the short to medium term.

Contents

* Corresponding author at: Donders Centre for Cognitive Neuroimaging, Radboud University, PO Box 9101, 6500 HB Nijmegen, The Netherlands. Tel.: +31 24 3610098.
E-mail address: t.wolfers@donders.ru.nl (T. Wolfers).

http://dx.doi.org/10.1016/j.neubiorev.2015.08.001

1. Introduction

Diagnostic manuals, such as the Diagnostic Statistical Manual (DSM, American Psychiatric Association, 2013) determine criteria, which are the basis for the diagnosis of psychiatric disorders. Over recent years, a rapidly growing number of studies have been published that aim to complement and improve clinical decision making on the basis of biological measures ('biomarkers') derived from different types of data, such as magnetic resonance imaging (MRI) and genetics. For some neurological diseases, as for example Huntington disease, genetic markers can predict the diagnosis with nearly perfect certainty (Macdonald et al., 1993). However, biomarkers that accurately predict disease state remain to be found for psychiatric disorders. Pattern recognition techniques have shown promise for detecting biomarkers from neuroimaging data and hold the potential to combine complementary information across different sources in an efficient way. This is important because psychiatric disorders are unlikely to be linked to one specific biological process but rather multiple factors that act together. Therefore, it is essential to investigate various types of data, which might capture different aspects of biology, and investigate them jointly. Pattern recognition techniques were first applied to MRI data approximately a decade ago, with the goal of separating and thereby classifying psychiatric patients from controls (Davatzikos et al., 2005). However, despite many subsequent efforts those promising results have not, to date, translated beyond research settings.

In this review, we surveyed the literature on pattern recognition for making diagnostic predictions in psychiatric research and evaluate progress made in translating those techniques towards clinical applications. While previous reviews have focussed on a particular machine learning technique (Orrù et al., 2012), a single imaging modality (Castellanos et al., 2013), or a small subset of disorders and diseases (Klöppel et al., 2012), we aim to provide a comprehensive review of all psychiatric disorders. First, we provide a brief introduction to pattern recognition methods in psychiatric neuroimaging. Second, we survey the pattern recognition literature, comparing studies on the data modalities used, the types of features extracted and subsequently selected as well as the algorithm applied. We expose areas that require further investigation and outline problems that are common to many studies. We evaluate the progress made in translating pattern recognition techniques towards clinical domains and identify clinical as well as research applications for which these techniques are most likely to add value. We conclude by pointing out challenges that are not widely appreciated and that the field needs to overcome in order to translate pattern recognition methods to assist decision making in actual clinical practice.

2. Principles of pattern recognition

The first applications of pattern recognition in neuroimaging was reported about twenty years ago (Lautrup et al., 1995; Morch et al., 1997), however it took some more time until this approach got recognised more widely (e.g. Haxby et al., 2001; Haynes and Rees, 2006). Pattern recognition aims to extract regularities in data which can be used to predict outcome measures such as a particular psychiatric diagnosis (Bishop, 2006; Orrù et al., 2012). The outcome predictions, in the context of pattern recognition are usually learned in a supervised way, which means that the algorithm or classifier is provided a set of predefined labels. However, in psychiatry the labels, that e.g. indicate the diagnoses of participants, are often uncertain which makes pattern leaning more difficult. A scheme which allows us to categorise the reviewed articles based on the methods that they applied and which illustrates the different aspects of pattern recognition is presented in Fig. 1. Subsequent figures and tables follow the logic developed in this scheme.

2.1. Feature extraction

A feature (collectively a 'feature set') is any characteristic that can be extracted from the data and that is believed to be informative about the class labels. In neuroimaging, pattern recognition is sometimes referred to as multi-voxel analysis, as voxels are often used directly as features (Haxby, 2012). However, many other kinds of features can be derived, which may vary in their ability to predict the class labels. One way we will categorise the articles reviewed here is according to the type of feature set they employ: voxel, region, or network-derived sets. In a voxel-based feature set all features are extracted on the voxel level. In a region based feature set, features are derived by parcellating brain images into predefined regions, e.g. on the basis of an anatomical or functional brain atlas. In a network based feature set, features are derived by combining voxels across networks; for example, derived from independent components analysis (Beckmann and Smith, 2004). Importantly, this step is distinct from feature selection.

2.2. Feature selection

The number of features in MRI is large and often many features do not contribute substantially to the prediction of the class labels. Therefore, feature selection, which aims to amplify the signal by restricting the prediction algorithm to only informative features, is an important (but not essential) step (Chu et al., 2012; Cuingnet et al., 2011). Many types of feature selection are possible and can in general be categorised into: (1) expert feature selection, which is based on prior knowledge and (2) automatic feature selection consists of for example a feature selection algorithm. Different varieties

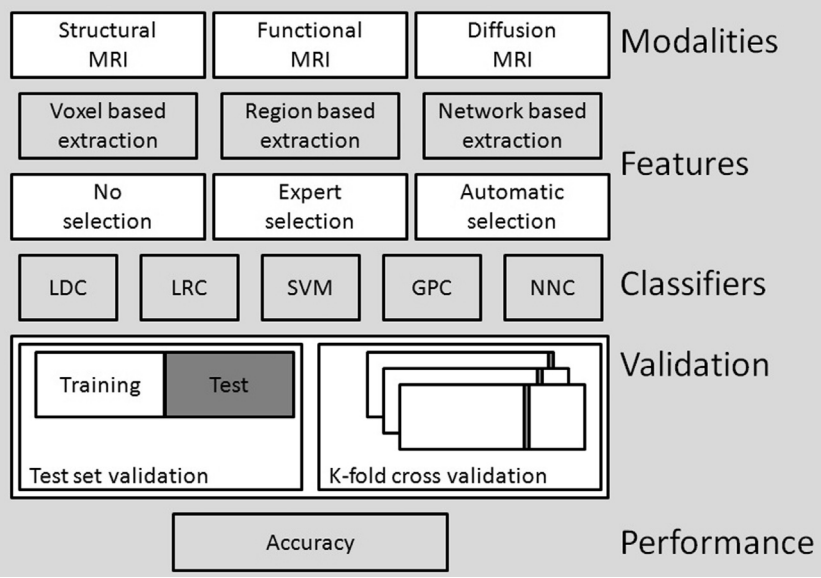

Fig. 1. Depicted is a framework that shows the main components of a pattern recognition pipeline. Data, potentially from different modalities, is processed in such a way that features can be extracted, either on the voxel, regional or network level. The resulting features can subsequently be selected with the goal to enhance the signal to noise ration. This processed data is used to train and test a classifier on their performance in a validation procedure. Different types of classifiers are frequently encountered in particular a linear discriminant classifier (LDC), logistic regression classifier (LRC), support vector machine (SVM), Gaussian process classifier (GPC) and neural network classifiers (NNC).

of those algorithms are used dependent on the type of learning problem and the properties of features (Mwangi et al., 2014). Finally, it is possible to combine those two approaches, for example by selecting a region of interest, which is implicated in a particular disorder and subsequently using an algorithm to favour informative features in this preselected region. A classifier or regression model, which, respectively, learns a rule for the separation of the classes or predicting a quantitative variable, is then applied to the feature set. Furthermore, classifiers often penalise the weight associated with particular features dependent on the constraints of the regularisation applied in the classification algorithm.

2.3. Classifiers

Various classifiers have been applied to neuroimaging datasets. In principle, any type of classification or regression algorithm can be used in pattern recognition ranging from linear or logistic regressions to multilayer neural networks or Gaussian processes (Bishop, 2006). The relatively large number of features and the small number of examples, also referred to as the curse of dimensionality, puts some restrictions to the applicability of algorithms. Each classifier learns a rule, which separates the classes optimally. In principle, classifiers differ with regard to the method determining this rule. In the present review we focus on classifiers dealing with discrete outcome measures such as diagnostic labels, regression methods are thus not included into the review.

A Linear discriminant classifier (LDC), a classical linear model, is used to separate classes by maximising the ratio of between-class to within class variance. A Logistic regression classifier (LRC) is a probabilistic discriminant model that aims to learn an optimal decision rule by modelling the log-odds ratio as a linear combination of predictor variables. Under Gaussian assumptions, LDC and LRC are equivalent (Hastie et al., 2009). Both methods yield probabilistic predictions that a new example corresponds to a particular class and can be transformed into a class label. The support vector machine (SVM) is an algorithm designed for binary classification that maximises the margin between classes in a high dimensional space. Mathematically, the discriminant function is defined by a weight vector orthogonal to the decision boundary, which can be uniquely specified by the samples that lie closest to the decision boundary, referred to as support vectors. The decision boundary represents the rule for classification of new examples. A Gaussian process classifier (GPC) is a probabilistic model and is a Bayesian extension of LRC. In contrast to SVM, the predicted class is augmented by an estimate of the certainty of the prediction. GPCs are best described as a distribution over functions. Based on Bayes' rule the posterior distribution of functions which represent the training data is found in an optimal way. This posterior distribution is used to classify new examples according to the rules of probability. Neural network classifiers (NNC) are a broad class of algorithms that are modelled on biological networks. They consist of a set of artificial neurons that are trained by adjusting the weights connecting them and can be used for a range of pattern recognition tasks including classification. The learned relation of these artificial neurons represents the rule for decision, when new examples are encountered.

2.4. Training and testing of classifiers

The multivariate pattern, a visualisation of the decision rule, is learned on the basis of a training set and validated in a test set. It is essential that the test and training sets are kept independent from one another in order to avoid over-fitting. Over-fitting refers to inflated performance measures as a consequence of testing a trained classifier on previously seen data. During training the classifier learns to predict the labels from the feature set. In case of a relatively simple learning problem, e.g. without iterative feature selection, the trained classifier is subsequently tested on previously unseen data, the test set. Usually, this is repeated using multiple different training and test partitions in a procedure called *k*-fold cross-validation, where *k* denotes the number of data partitions. The special case, where *k* is equal to the number of samples is referred to as leave one out-cross validation (LOO-CV). This procedure is iterative and every example in the sample is left out once for testing; the performance measures, e.g. accuracies, are subsequently averaged across those iterations for the training and test

set. More complex approaches, which include for example a feature selection step, require nested cross validation. This means that the data are partitioned twice. First, in an 'outer' fold, a data partition is excluded for testing. The remaining samples are then repartitioned in an 'inner' cross-validation loop. The outer loop enables an unbiased estimation of generalisability; the inner loop provides an estimate of generalisability independent from the test set that can be used to determine the optimal number of features or to optimise parameters.

2.5. Performance evaluation of classification

The most frequently used performance measures are sensitivity, specificity and accuracy. These measures give an indication of how accurately a classifier can generalise to new cases. In a clinical context, sensitivity refers to the percentage of cases that a classifier correctly predicts the disorder of a participant. In other words, a high sensitivity thus means a high percentage of true positive and a low of false negatives. A high specificity indicates that only a few participants are diagnosed with a disorder while actually being healthy (i.e. a high percentage of true negatives and a low of false positives). The accuracy refers to the total proportion of samples correctly classified. A good practice is to report balanced accuracy measures, which is an average accuracy obtained for each diagnostic label (Brodersen et al., 2010) that is unaffected by potential imbalances between groups. It is important to note that measures like the receiver operating characteristic (ROC) curves are more informative performance measures. The ROC curves provide information on the balance between the true positive rate (sensitivity) and the false positive rate (1-specificity) across a range of decision thresholds. Classifiers are usually evaluated on one of those or a combination of those measures. In practice, these metrics for classifier performance are usually combined with procedures to estimate statistical significance of the pattern recognition model. For classification, a significant deviation from chance level performance is usually estimated using *p*-values derived from permutation testing and in some cases parametric tests may also be appropriate (Stelzer et al., 2013). Unfortunately, the reports on classifier performance differ considerably across studies, which can make comparisons across studies difficult. Therefore, we recommend – at a minimum – reporting sensitivity, specificity, balanced accuracy and the area under the curve as well as the confidence interval[1] for those measures (Klöppel et al., 2009). Positive and negative predictive values are also important for diagnostic studies as they directly quantify the potential utility of the classifier for clinical decision making. The positive predictive value is defined as the number times the classifier correctly predicted participants as having the disorder (positive diagnosis) divided by the total number of positive predictions. The negative predictive value is defined similarly as the number of times the classifier correctly predicted a negative diagnosis divided by the total number of negative predictions. For comparative reasons, we report only accuracy measures in the tables and text, as these measures were provided most consistently across different studies. In case that more than one accuracy measure was reported in a single study, only the maximum for each contrast is shown in the tables. Therefore, accuracies reported in the tables should be understood to represent upper bounds for classification performance. Moreover, most studies do not report confidence intervals for classifier accuracies, which hamper comparison between algorithms and studies.

[1] Calculator of confidence intervals: http://vassarstats.net/clin1.html

2.6. Pattern interpretation

Most linear classifiers permit the weights determining the classification rule to be visualised in the voxel space, which has formed the basis for 'discriminative mapping' of the weight vector (Mourão-Miranda et al., 2005). However, the interpretation with regard to the localisation of an effect within such multivariate patterns is not straightforward (Haufe et al., 2014) and is dependent on the particular classification algorithm employed. An additional problem for non-linear classifiers is that it is usually not possible to exactly map the weights back into the voxel space. Furthermore, weights associated with a specific location can change as a consequence of a reconfiguration of the multivariate pattern due to e.g. feature selection. Therefore, a straight-forward interpretation in terms of effect localisation is often incorrect. For this and other reasons, researchers combined univariate and multivariate approaches to benefit from potentially higher sensitivity of multivariate estimates for detecting spatially distributed effects and a better interpretability of univariate approaches in terms of effect localisation (Hart et al., 2014b). However, in general a multivariate pattern must be considered as a whole and only in some cases allows a more specific interpretation. However, the discriminative mapping approach is often useful to give a qualitative overview of the regional distribution of weights learned by the classifier. This can help, for example, to ensure that the classifier discrimination is not driven by artefactual processes confounded with the class labels (e.g. head motion). In general it is difficult to interpret the overlap of multivariate patterns across different studies. The potential of pattern recognition in psychiatry lies rather in mapping behaviourally diagnostic labels onto biology or to help fractionate disease phenotypes.

3. Pattern recognition for diagnostic predictions of psychiatric disorders

The present review is based on an extensive literature search for research papers applying pattern recognition for making diagnostic predictions of psychiatric disorders. Schizophrenia (SCZ; Davatzikos et al., 2005) was one of the first disorders investigated with pattern recognition, followed by major depressive disorder (MDD; Marquand et al., 2008; Fu et al., 2008), attention-deficit/hyperactivity disorder (ADHD; Zhu et al., 2008), bipolar disorder (BPD; Arribas et al., 2010), autism spectrum disorder (ASD; Ecker et al., 2010b), post-traumatic stress disorder (PTSD; Gong et al., 2011), obsessive compulsive disorder (OCD; Weygandt et al., 2012a), social anxiety disorder (SAD; Liu et al., 2013) and specific phobia (SP; Lueken et al., 2014). A systematic literature search primarily in PubMed was performed and concluded on the 1st of Mai 2015. The search consisted, first, of different terms related to pattern recognition and their abbreviations second, all terms and abbreviations related to MRI third, all names and abbreviations for one of the disorders mentioned above[2]. This search was repeated for all disorders, and their references were checked on missed publications which were included to the review as well. All publications were screened on their relevance for the review. All papers using pattern recognition approach on MRI data in psychiatric diagnostics that reported performance measures of the classification and

[2] Search term: (pattern recognition OR multivariate pattern analysis OR classification OR prediction OR diagnostics OR linear discriminant analysis OR logistic regression analysis OR support vector machine OR Gaussian processes OR neural networks) AND (magnetic resonance imaging OR structural magnetic resonance imaging OR functional magnetic resonance imaging OR diffusion magnetic resonance imaging OR task functional magnetic resonance imaging OR resting state functional magnetic resonance imaging OR diffusion tensor imaging) AND (one of the disorders mentioned).

Table 1
Schizophrenia.

References	Sample size	Age Group	Modalities	Features	Classifiers	Validation	Classification	Performance (accuracy) (%)[#]
Yushkevich et al. (2005)	SCZ = 46 C = 46	Adults	Structural MRI	Region-based feature set and automatic feature selection	n.s.	n.s.	SCZ vs. C	71.00
Davatzikos et al. (2005)	SCZ = 69 C = 69	Adults	Structural MRI	Voxel-based feature set	n.s.	LOO-CV	SCZ vs. C	81.10
Shi et al. (2007)	SCZ = 48 C = 35	Adults	Structural MRI	Region-based feature set and automatic feature selection	LDC	LOO-CV	SCZ vs. C	80.00
Fan et al. (2007)	SCZ-Males = 46 C-Males = 41 SCZ-Females = 38 C-Females = 23	Adults	Structural MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	SCZ-Males vs. C SCZ-Females vs. C	90.80 91.80
Kawasaki et al. (2007)	SCZ-Males = 30 C-Males = 30 *Rep. cohort* SCZ-Males = 16 C-Males = 16	Adults	Structural MRI	Voxel-based feature set	LDC	Test-validation	SCZ-Males vs. C	≈80.00
Sun et al. (2009)	SCZ = 36 C = 36	Adults	Structural MRI	Voxel-based feature set	LRC	LOO-CV	SCZ vs. C	86.10
Pohl and Sabuncu (2009)	SCZ = 16 C = 17	Adults	Structural MRI	Region-based feature set	SVM	LOO-CV	SCZ vs. C	90.00
Koutsouleris et al. (2009)	SCZ e.r. = 20 SCZ l.r. = 25 C = 25 *Rep. cohort* SCZ t. = 15 SCZ no t. = 18 C = 17	Adults	Structural MRI	Voxel-based feature set and automatic feature selection	SVM	1/5-CV	SCZ e.r. vs C SCZ l.r. vs C SCZ e.r. vs. SCZ l.r. *Rep.cohort* SCZ t. vs. C SCZ no t. vs. C SCZ t. vs. SCZ no t.	87.00 78.00 82.00 94.00 86.00 82.00
Anderson et al. (2010)	SCZ = 14 C = 6	Adults	Resting state functional MRI	Network-based feature set and automatic feature selection	n.s.	LOO-CV	SCZ vs. C	75.00
Takayanagi et al. (2010)	SCZ = 34 C = 48	Adults	Structural MRI	Region-based feature set	LDC	LOO-CV	SCZ vs. C	75.60
Rathi et al. (2010)	SCZ = 21 C = 20	Adults	Diffusion MRI	Voxel-based feature set	Multiple	LOO-CV	SCZ vs. C	85.50
Shen et al. (2010)	SCZ = 20 C = 32	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	Multiple	LOO-CV	SCZ vs. C	86.50
Yang et al. (2010)	SCZ = 20 C = 20	Adults	Task functional MRI—auditory oddball	Network-based and genetic feature set	SVM	LOO-CV	SCZ vs. C	87.00
Shen et al. (2010)	SCZ = 32 C = 18	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	Multiple	LOO-CV	SCZ vs. C	≈85.00
Ingalhalikar et al. (2010)	SCZ = 27 C = 27 ASD = 25 C = 23	Children; Adults	Diffusion MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	SCZ vs. C ASD vs. C	90.62 89.58
Kasparek et al. (2011)	SCZ = 39 C = 39	Adults	Structural MRI	Voxel-based feature set and automatic feature selection	LDC	LOO-CV	SCZ vs. C	72.00
Karageorgiou et al. (2011)	SCZ = 28 C = 47	Adults	Structural MRI	Region-based feature set and automatic feature selection	LDC	LOO-CV	SCZ vs. C	70.50
Fan et al. (2011)	SCZ = 31 C = 31	Adults	Resting state functional MRI	Network-based feature set and automatic feature selection	SVM	LOO-CV	SCZ vs. C	87.00

Castro et al. (2011)	SCZ = 52 C = 54	Adults	Task functional MRI—auditory oddball	Network/voxel based feature set and automatic feature selection	CKM	LOO-CV	SCZ vs. C	95.00
Ardekani et al. (2011)	SCZ = 50 C = 50	Adults	Diffusion MRI	Voxel-based feature set and automatic feature selection	LDC	Test-validation	SCZ vs. C	96.00
Takayanagi et al. (2011)	SCZ = 52 C = 40	Adults	Structural MRI	Region-based feature set	LDC	n.s.	SCZ vs. C	87.68
Costafreda et al. (2011)	SCZ = 32 BPD = 32 C = 40	Adults	Task functional MRI—verbal fluency	Voxel-based feature set	SVM	LOO-CV	SCZ vs. C BPD vs. C	92.00 79.00
Yoon et al. (2012)	SCZ = 51 C = 51	Adults	Task functional MRI—continous perfomance	Region-based feature set	LDC	LOO-CV	SCZ vs. C	62.00
Greenstein et al. (2012)	SCZ = 98 C = 99	Adolescents	Structural MRI	Region-based feature setand automatic feature selection	n.s.	n.s.	SCZ vs. C	73.70
Bassett et al. (2012)	SCZ = 29 C = 29	Adults	Resting state functional MRI	Region-based feature set	SVM	1/2-CV	SCZ vs. C	75.00
Castellani et al. (2012)	SCZ = 54 C = 54	Adults	Structural MRI	Region-based feature set and expert feature selection	SVM	LOO-CV	SCZ vs. C	75.00
Borgwardt et al. (2012)	SCZ = 23 C = 22	Adults	Structural MRI	Region-based feature set and automatic feature selection	SVM	1/10-CV	SCZ vs. C	86.70
Du et al. (2012)	SCZ = 28 C = 28	Adults	Resting state functional MRI	Network-based feature set and automatic feature selection	LDC	LOO-CV	SCZ vs. C	93.00
Tang et al. (2012)	SCZ = 22 C = 22	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	SCZ vs. C	93.20
Bansal et al. (2012)	SCZ = 36 ADHD = 41 C = 42	Adults	Structural MRI	Region-based feature set and automatic feature selection	n.s.	LOO-CV	SCZ vs. C ADHD vs. C	94.00 91.05
Honorio et al. (2012)	SCZ = 13 C = 15	Adults	Task functional MRI—sensory motor; Oddball; Memory	Region-based feature set and expert feature selection	MVC	LOO-CV	SCZ vs. C	96.40
Venkataraman et al. (2012)	SCZ = 18 C = 18	Adults	Structural MRI; Resting state functional MRI	Region-based feature set and automatic feature selection	NNC	n.s.	SCZ vs. C	75.00
COBRE-sample*	SCZ ≈ 50 C ≈ 50	Adults	Structural MRI; Resting state functional MRI	Feature sets and selections differ	Multiple	Multiple	SCZ vs. C	≈70.00
Nieuwenhuis et al. (2012)	SCZ = 128 C = 111 *Rep. cohort* SCZ = 155 C = 122	Adults	Structural MRI	Voxel-based feature set	SVM	LOO-CV; Test-validation	SCZ vs. C *Rep. cohort*	71.40 70.40
Liu et al. (2012)	SCZ = 24 SCZ-sibl = 25 C = 22	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	SCZ vs. C SCZ vs. SCZ-sibl SCZ-sibl vs. C	80.40 77.60 78.70
Yu et al. (2013)	SCZ = 24 SCZ-sibl = 25 C = 22	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	SCZ vs. SCZ-sibl vs. C	62.00
Zanetti et al. (2013)	SCZ = 62 C = 62	Adults	Structural MRI	Voxel-based feature set	SVM	LOO-CV	SCZ vs. C	73.40

Table 1 (*Continued*)

References	Sample size	Age Group	Modalities	Features	Classifiers	Validation	Classification	Performance (accuracy) (%)[#]
Iwabuchi et al. (2013)	SCZ = 19 C = 20	Adults	Structural MRI	Voxel-based feature set	SVM	LOO-CV	SCZ vs. C	77.00
Hu et al. (2013)	SCZ = 10 C = 10	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	SCZ vs. C	77.50
Fekete et al. (2013)	SCZ = 10 C = 8	Adults	Resting state functional MRI	Region-based feature set	n.s.	LOO-CV	SCZ vs. C	88.00
Arbabshirani et al. (2013)	SCZ = 28 C = 28	Adults	Resting state functional MRI	Network-based feature set and automatic feature selection	Multiple	Test-validation	SCZ vs. C	96.00
Pettersson-Yeo et al. (2013)	u.h.r. SCZ = 19 f.e. SCZ = 19 C = 12	Adults	Structural MRI; Diffusion MRI	Voxel-based feature set	SVM	LOO-CV	u.h.r.SCZ vs. C f.e. SCZ vs. C u.h.r. SCZ vs. f.e. SCZ	68.42 63.16 76.67
MCIC-sample[**]	SCZ ≈ 75 C ≈ 75	Adults	Structural MRI; Functional MRI; Diffusion MRI	Feature sets and selections differ	Multiple	Multiple	SCZ vs. C	≈70.00
Ota et al. (2013)	SCZ-Females = 25 MDD-Females = 25 *Rep. cohort* SCZ-Females = 18 MDD-Females = 16	Adults	Structural MRI Diffusion MRI	Region-based feature set and expert feature selection	LDC	Test-validation	SCZ vs. MDD	≈80.00
Su et al. (2013)	SCZ = 32 C = 32	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	SCZ vs. C	≈80.00
Yu et al. (2013)	SCZ = 32 MDD = 19 C = 38	Adults	Resting state functional MRI	Region-based feature set	SVM	LOO-CV	SCZ vs. MDD&C MDD vs. SCZ&C MDD&SCZ vs. C SCZ vs. MDD vs. C	81.30 84.20 78.90 80.90
Gould et al. (2014)	SCZ = 126 C = 134	Adults	Structural MRI	Voxel-based feature set	SVM	LOO-CV	SCZ vs. C	68.00
Zhu et al. (2014)	SCZ = 10 C = 10	Adults	Structural MRI; Diffusion MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	SCZ vs. C	100.00
Schnack et al. (2014)	SCZ = 66 BPD = 66 C = 66	Adults	Structural MRI	Voxel-based feature set	SVM	LOO-CV	SCZ vs. C BPD vs. C SCZ vs. BPD	75.50 59.00 65.50
Mueller et al. (2015)	SCZ = 31 C = 37	Adults	Resting state functional MRI	Region-based feature set expert based feature selection	SVM	LOO-CV	SCZ vs. C	74.00
Janousova et al. (2015)	SCZ = 49 C = 49	Adults	Structural MRI	Voxel-based feature set and automatic feature selection	LDC	LOO-CV	SCZ vs. C	81.60
Koch et al. (2015)	SCZ = 44 C = 44	Adults	Task functional MRI—monetary reward	Voxel-based feature set and automatic feature selection	SVM	LOO-CV	SCZ vs. C	93.00
Koutsouleris et al. (2015)	SCZ = 158 MDD = 104 *Rep. cohort used for differential diagnosis*	Adults	Structural MRI	Voxel-based feature set and automatic feature selection	SVM	LOO-CV	SCZ vs. MDD	≈76.00
Arribas et al. (2010)	SCZ = 21 BPD = 15 C = 25	Adults	Task functional MRI—Auditory oddball	Voxel-based feature set and automatic feature selection	NNC	LOO-CV Test-validation	SCZ vs. BPD vs. C	71.90

SCZ = schizophrenia; SCZ e.r. = schizophrenia early risk; SCZ l.r. = schizophrenia late risk; SCZ t. = schizophrenia transition; SCZ no t. = schizophrenia no transition; ASD = Autism spectrum disorder; BPD = Bipolar disorder; f.e. SCZ = first episode schizophrenia; u.h.r. SCZ = ultra high risk schizophrenia; SCZ-sibl = Siblings of patients with schizophrenia; MDD = major depressive disorder; C = Controls; SVM = Support vector machine; LRC = Logistic regression classification; LDC = Linear discriminant classification; CKM = Composite kernel machine; MVC = Major vote classifier; NNC = Neural network classifier; LOO-CV = Leave one out-cross validation; n.s. = not specified.

[#] The maximal accuracy reported in the article.

[*] Pattern recognition studies making diagnostic predictions on the COBRE-sample (Overview paper: Calhoun et al. (2012); Anderson and Cohen (2013); Sabuncu and Konukoglu (2014); Sui et al. (2013)).

[**] Pattern recognition studies making diagnostic predictions on the MCIC-sample (Overview paper: Gollub et al. (2013); Sabuncu and Konukoglu (2014)).

Table 2
Bipolar disorder and major depressive disorder.

Ref.	Sample size	Age group	Modalities	Features	Classifiers	Validation	Classification	Performance (accuracy) (%)[#]
Marquand et al. (2008)	MDD = 20 C = 20	Adults	Task functional MRI—N-back	Voxel-based feature set and automatic feature selection	SVM	LOO-CV	MDD vs. C	67.50
Fu et al. (2008)	MDD = 19 C = 19	Adults	Task functional MRI—affective processing	Voxel-based feature set and expert feature selection	SVM	LOO-CV	MDD vs. C	86.00
Costafreda et al. (2009)	MDD = 37 C = 37	Adults	Structural MRI	Voxel-based feature- set and automatic feature selection	SVM	LOO-CV	MDD vs. C	67.60
Craddock et al. (2009)	MDD = 20 C = 20	Adults	Resting state functional MRI	Region-based feature- set and automatic feature selection	SVM	LOO-CV	MDD vs. C	95.00
Hahn et al. (2011)	MDD = 30 C = 30	Adults	Task functional MRI—multiple tasks	Voxel-based feature set	GPC	LOO-CV	MDD vs. C.	83.00
Nouretdinov et al. (2011)	MDD = 19 C = 19	Adults	Task functional MRI - Emotional faces	Voxel-based feature set and automatic feature selection	SVM	LOO-CV	MDD vs. C	86.90
Costafreda et al. (2011)[*]	BPD = 32 SCZ = 32 C = 40	Adults	Task functional MRI—verbal fluency	Voxel-based feature set	SVM	LOO-CV	BPD vs. C SCZ vs. C	79.00 92.00
Gong et al. (2011)	NDD = 23 RDD = 23 C = 23	Adults	Structural MRI	Voxel-based feature set	SVM	LOO-CV	NDD vs. C RDD vs. C NDD vs. RDD	84.65 67.39 69.57
Fang et al. (2012)	MDD = 22 C = 26	Adults	Diffusion MRI; Structural MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	MDD vs. C	91.70
Zeng et al. (2012)	MDD = 24 C = 29	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	MDD vs. C	94.30
Korgaonkar et al. (2012)	MDD = 23 C = 23	Adults	Diffusion MRI Structural MRI	Region-based feature set	LDC	1/10-CV	MDD vs. C	96.00
Mourão-Miranda et al. (2012)	MDD = 18 BPD = 18 C = 18	Adults	Task functional MRI—emotional faces	Voxel-based feature set	GPC	LOO-CV	MDD vs. C BPD vs. C BPD vs. MDD	61.00 64.00 67.00
Bansal et al. (2012)[*]	BPD = 26 C = 40	Adults	Structural MRI	Region-based feature set and automatic feature selection	n.s.	LOO-CV	BP vs C	98.20
Lord et al. (2012)	MDD = 22 C = 22	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	SVM	n.s.	MDD vs. C	99.00
Mwangi et al. (2012)	MDD = 62 C = 62	Adults	Structural MRI	Voxel-based feature set and automatic feature selection	SVM	LOO-CV	MDD vs. C	90.30
Liu et al. (2012)	TRD = 18 TSD = 24 C = 17	Adults	Structural MRI	Voxel-based feature set and automatic feature selection	SVM	LOO-CV	TRD vs. C TSD vs. C TRD vs. TSD	85.70 91.20 82.90
Modinos et al. (2013)	SD = 17 C = 17	Adults	Task functional MRI—emotional images	Voxel-based feature set	SVM	LOO-CV	SD vs. C	77.00

Table 2 (*Continued*)

Ref.	Sample size	Age group	Modalities	Features	Classifiers	Validation	Classification	Performance (accuracy) (%)[#]
Wei et al. (2013)	MDD = 20 C = 20	Adults	Resting state functional MRI	Network-based feature set	SVM	LOO-CV	MDD vs. C	90.00
Ma et al. (2013)	MDD = 29 C = 29	Adults	Resting state functional MRI	Region-based feature set	SVM	LOO-CV	MDD vs. C	90.06
Ota et al. (2013)[*]	Training-set: MDD-Females = 25 SCZ-Females = 25 Test-set: MDD-Females = 18 SCZ-Females = 16	Adults	Structural MRI Diffusion MRI	Region-based feature set and expert feature selection	LDC	Test-validation	SCZ-Females vs. MDD-Females	≈80.00
Grotegerd et al. (2013)	MDD = 10 BPD = 10 C = 10	Adults	Task functional MRI—emotional faces	Voxel-based feature set	SVM; GPC	LOO-CV	BPD vs. C MDD vs. C MDD vs. BPD	80.00 85.00 90.00
Yu et al. (2013)[*]	MDD = 19 SCZ = 32 C = 38	Adults	Resting state functional MRI	Region-based feature set	SVM	LOO-CV	MDD vs. SCZ&C SCZ vs. MDD&C MDD&SCZ vs. C SCZ vs. MDD vs. C	84.20 81.30 78.90 80.90
Cao et al. (2014)	MDD = 39 C = 37	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	MDD vs. C	78.95
Qin et al. (2014)	MDD = 29 C = 30	Adults	Structural MRI Diffusion MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	MDD vs. C	80.05
Guo et al. (2014)	MDD = 36 C = 27	Adults	Resting state fMRI	Region-based feature set and automatic feature selection	NNC	1/3-CV	MDD vs. C	90.05
Schnack et al. (2014)[*]	BPD = 66 SCZ = 66 C = 66	Adults	Structural MRI	Voxel-based feature set	SVM	LOO-CV	BPD vs. C SCZ vs. C SCZ vs. BPD	59.00 75.50 65.50
Serpa et al. (2014)	BPD = 23 MDD = 19 C = 71	Adults	Structural MRI	Region-based feature set	SVM	LOO-CV	BPD vs. C MDD vs. C BPD vs. MDD	66.10 59.60 54.76
Redlich et al. (2014)	*Cohort 1&2* BPD = 29 MDD = 29 C = 29	Adults	Structural MRI	Region-based feature set	SVM; GPC	LOO-CV Test-validation	*First cohort: BPD vs. MDD* *Second cohort: BPD vs. MDD*	79.30 65.50
Koutsouleris et al. (2015)[*]	MDD = 104 SCZ = 158 *Rep. cohort used for differential diagnosis*	Adults	Structural MRI	Voxel-based feature set and automatic feature selection	SVM	LOO-CV	MDD vs. SCZ	≈76.00
Qin et al. (2015)	cMDD = 28 rMDD = 15 C = 30	Adults	Diffusion MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	cMDD&rMDD vs. C cMDD vs. rMDD	100.00 97.67
Rosa et al. (2015)	*Cohort 1* MDD = 19 C = 19 *Cohort 2* MDD,BPD = 30 C = 30	Adults	Task functional MRI—multiple tasks	Region-based feature set	SVM	LOO-CV	MDD vs. C MDD,BPD vs. C	78.95 85.00
Arribas et al. (2010)	BPD = 15 SCZ = 21 C = 25	Adults	Task functional MRI—Auditory oddball	Voxel-based feature set and automatic feature selection	NNC	LOO-CV Test-validation	SCZ vs. BPD vs. C	71.90
Rocha-Rego et al. (2014)	*Cohort 1* BPD = 26 C = 26 *Cohort 2* BPD = 14 C = 14	Adults	Structural MRI	Voxel-based feature set and automatic feature selection	GPC	LOO-CV	*Cohort 1* BPD vs. C *Cohort 2* BPD vs. C	73.00 72.00

BPD = Bipolar disorder; MDD = Major depressive disorder; NDD = Non-refractory depressive disorder; RDD = Refractory depressive disorder; TRD = Treatment-resistant depression; TRS = Treatment-sensitive depression; SCZ = schizophrenia; cMDD = Current major depressive disorde; rMDD = Remitted major depressive disorde; SD = Subclinical depression; C = Controls; SVM = Support vector machine classifier; GPC = Gaussian process classifier; LDC = Linear discriminant classifier; NNC = Neural network classifier; LOO-CV = Leave one out-cross validation; n.s. = not specified.

[#] The maximal accuracy reported in the article.

[*] Reported in Table 1.

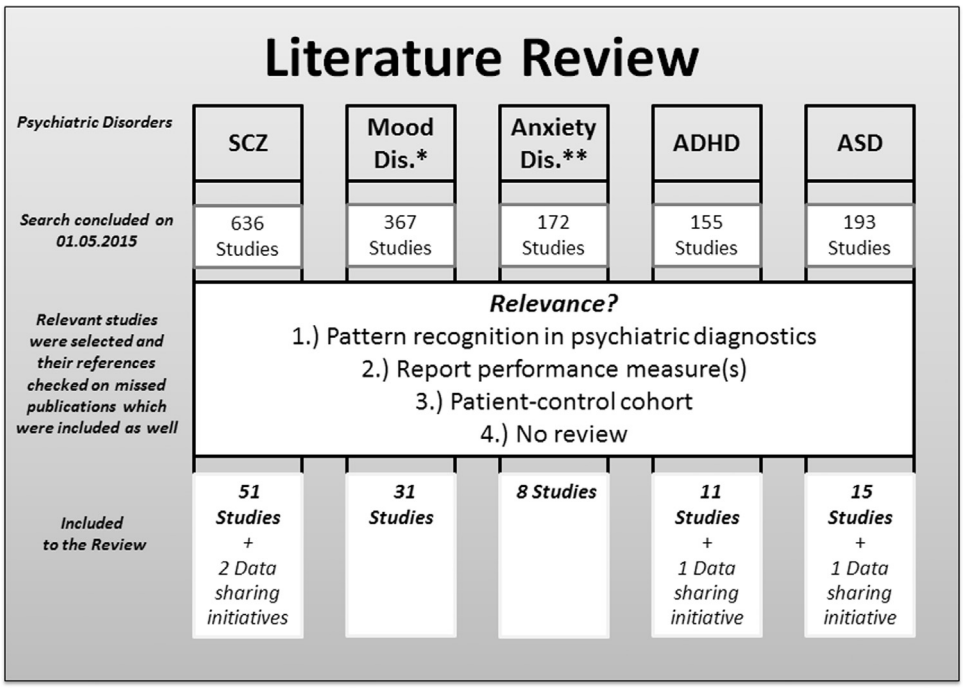

Fig. 2. Depicted is the search procedure for the literature review and the inclusion criteria as well as the number of reviewed publications per disorder. SCZ = schizophrenia; Mood Dis. = mood disorders; * BPD = bipolar disorder, MDD = major depressive disorder; Anx. Dis = anxiety disorders; ** OCD = obsessive compulsive disorder, SAD = social anxiety disorder, PTSD = post-traumatic stress disorder, SP = specific phobia; ADHD = attention-deficit/hyperactivity disorder; ASD = Autism spectrum disorder.

Table 3
Obsessive compulsive, social anxiety, post-traumatic stress disorder and specific phobia.

Ref.	Sample size	Age group	Modalities	Features	Classifiers	Validation	Classification	Performance (accuracy) (%)[#]
Soriano-Mas et al. (2007)	OCD = 72 C = 72 Rep. Cohort OCD = 30 C = 30	Adults	Structural MRI	Neither feature set, mean difference value of OCD and controls	n.s.	Test-validation	OCD vs. C	76.60
Weygandt et al. (2012a)	OCD = 10 C = 10	Adults	Task functional MRI—emotional images	Voxel-based feature set and automatic feature selection	SVM	LOO-CV	OCD vs. C	100.00
Liu et al. (2013)	SAD = 20 C = 20	Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	SAD vs. C	82.50
Shenas et al. (2014)	OCD = 12 C = 12	Adults	Resting state functional MRI; task functional MRI—imagination	Region-based feature set and automatic feature selection	LDC; SVM	LOO-CV	OCD vs. C	74.00
Li et al. (2014)	OCD = 28 C = 28	Adults	Diffusion MRI	Voxel-based feature set	SVM	LOO-CV	OCD vs. C	84.00
Frick et al. (2014)	SAD = 14 C = 12	Adults	Task functional MRI—emotional faces, Structural MRI	Voxel-based feature set expert feature selection	SVM	LOO-CV	SAD vs. C	84.50
Lueken et al. (2014)	SP = 33 DP = 26 C = 37	Adults	Structural MRI	Voxel-based feature set	GPC	LOO-CV	phobics vs. C SP vs. C DP vs. C DP vs. SP	88.00 89.00 89.00 89.00
Gong et al. (2014)	PTSD-Earthquake = 50 noPTSD-Earthquake = 50 C = 40	Adults	Structural MRI	Voxel-based feature set	SVM	LOO-CV	PTSD vs. C noPTSD vs. C noPTSD vs. PTSD	91.00 85.00 76.00

OCD = Obsessive compulsive disorder; SAD = Social anxiety disorder; PTSD = Post traumatic stress disorder; SP = Snake phobics; DP = Dental phobics; C = Controls; SVM = Support vector machine; LDC = Linear discriminant classifier; GPC = Gaussian process classifier; LOO-CV = Leave one out-cross validation.

[#] The maximal accuracy reported in the article.

Table 4
Attention-deficit/hyperactivity disorder.

Ref.	Sample size	Age group	Modalities	Features	Classifiers	Validation	Classification	Performance (accuracy) (%)[#]
Zhu et al. (2008)	ADHD = 12 C = 12	Children; Adolescents	Resting state functional MRI	Voxel-based feature set and automatic feature selection	LDC SVM	LOO-CV	ADHD vs. C	85.00
Wang et al. (2011)	ADHD = 21 C = 28	Adults	Resting state functional MRI	Region-based feature set	SVM	LOO-CV	ADHD vs. C	81.00
Igual et al. (2012)	ADHD = 39 C = 39	Children	Structural MRI	Region-based feature set	SVM	1/5-CV	ADHD vs. C	72.48
Bansal et al. (2012)[*]	ADHD = 41 SCZ = 36 C = 42	Children	Structural MRI	Region-based feature set and automatic feature selection	n.s.	LOO-CV	ADHD vs. C SCZ vs. C	91.05 94.00
ADHD-200-sample[**]	ADHD = 350 C = 554	Children; adolescents	Structural MRI; Resting state functional MRI	Feature sets and selections differ	Multiple	Multiple	ADHD vs. C	≈61.00
Hart et al. (2014a)	ADHD-boys = 30 C-boys = 30	Adolescents	Task functional MRI—stop signal	Voxel-based feature set	GPC	LOO-CV	ADHD vs. C	77.00
Wang et al. (2013)	ADHD = 23 C = 23	Adults	Resting state functional MRI	Voxel-based feature set	SVM	LOO-CV	ADHD vs. C	80.00
Peng et al. (2013)	ADHD = 55 C = 55	Children; adolescents	Structural MRI	Region-based feature set	NNC SVM	LOO-CV	ADHD vs. C	90.18
Lim et al. (2013)	ADHD = 29 ASD = 19 C = 29	Adolescents	Structural MRI	Voxel-based feature set	GPC	LOO-CV	ADHD vs. C ADHD vs. ASD ADHD vs. ASD vs. C	79.30 85.20 68.20
Hart et al. (2014b)	ADHD = 20 C = 20	Adolescents	Task functional MRI—temporal discounting	Voxel-based feature set	GPC	LOO-CV	ADHD vs. C	75.00
Johnston et al. (2014)	ADHD = 34 C = 34	Children; adolescents	Structural MRI	Voxel-based feature set and automatic feature selection	SVM	LOO-CV	ADHD vs. C	93.00
Iannaccone et al. (2015)	ADHD = 18 C = 18	Adolescents	Task functional MRI—flanker task	Voxel-based feature set	SVM	LOO-CV	ADHD vs. C	77.78

ADHD = Participants with attention deficit hyperactivity disorder; C = Control participants; LDC = Linear discriminant classifier; SVM = Support vector machine classifier; GPC = Gaussian process classifier; NNC = Neural network classifier; LOO-CV = Leave one out-cross validation;

[#] The maximal accuracy reported in the article.

[*] Reported in Table 1.

[**] Pattern recognition studies making diagnostic predictions on the ADHD-200 sample (Overview paper: The ADHD Consortium (2012); http://fcon_1000.projects.nitrc.org/indi/adhd200/results.html).

had case-control design were included to the review (Fig. 2). The search procedure was repeated in Google scholar to decrease the likelihood of missing relevant articles. Articles that were based on one of the big data sharing efforts such as the Autism Brain Imaging Data Exchange[3] (ABIDE; Di Martino et al., 2014), ADHD-200 Global Competition[4] (The ADHD Consortium, 2012), Centre for Biomedical Research Excellence[5] (COBRE; Calhoun et al., 2012) or MIND Clinical Imaging Consortium[6] (MCIC; Gollub et al., 2013) were combined and a representative accuracy included to the table. In total we reviewed 118 articles and categorised them on the basis of a scheme developed for this review depicted in Fig. 1, the summary of all articles is presented in Tables 1–5.

3.1. Schizophrenia

One of the first studies on SCZ was published in 2005 (Davatzikos et al., 2005; Table 1). After a short period of time in which no publications were reported a few studies on SCZ using structural MRI derived feature were conducted. Those studies yielded promising accuracies of up to 90%. At time of writing, about fifty one different studies on pattern recognition in Schizophrenia have been published and range in predictive accuracy from 62% to 100% (Table 1). Several of those studies were performed in the context of a task-based functional MRI experiment, with auditory oddball, verbal fluency, working memory and sensory motor tasks (Castro et al., 2011; Costafreda et al., 2011; Honorio et al., 2012; Yang et al., 2010). An important early study predicted SCZ based on structural MRI and used a principal component feature selection algorithm prior to the predictions. An optimal number of principal components were determined based on the overall predictive performance of the algorithm. The study could show that different subcategories of SCZ could be reliably predicted, and three-class classification for this patient group was feasible with a maximal accuracy of 82% (Koutsouleris et al., 2009). An interesting study in this context, investigating the differential diagnostics of, SCZ, BPD and healthy participants could show that verbal fluency let to a reliable diagnostic specificity for SCZ (Costafreda et al., 2011). The first large scale study with a cohort of 128 patients and an equal number of controls as well as a similar sized replication cohort predicted SCZ based on structural MRI-derived multivariate patterns with an accuracy of just above 70%. Their cross validation was about the same as the replication accuracy (Nieuwenhuis et al., 2012). Different data sharing efforts for SCZ cohorts, such as the MCIC and the COBRE, have been initiated. However, the sample numbers were relatively

[3] http://fcon_1000.projects.nitrc.org/indi/abide/
[4] http://fcon_1000.projects.nitrc.org/indi/abide/
[5] http://fcon_1000.projects.nitrc.org/indi/retro/cobre.htmls
[6] http://coins.mrn.org/

Table 5
Autism spectrum disorder.

References	Sample size	Age group	Modalities	Features	Classifiers	Validation	Classification	Performance (accuracy) (%)[#]
Ecker et al. (2010b)	ASD-Males = 22 C-Males = 22	Adults	Structural MRI	Voxel-based feature set	SVM	LOO-CV	ASD vs. C	81.00
Jiao et al. (2010)	ASD = 22 C = 16	Children	Structural MRI	Region-based feature set	multiple	1/10-CV	ASD vs. C	87.00
Ecker et al. (2010a)	ASD = 20 C = 20	Adults	Structural MRI	Region-based feature set	SVM	LOO-CV	ASD vs. C	90.00
Ingalhalikar et al. (2010)[*]	ASD = 25 C = 23 SCH = 27 C = 27	Children; Adults	Diffusion MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	ASD vs. C SCH vs. C	89.58 90.62
Ingalhalikar et al. (2011)	ASD = 45 C = 30	Children	Diffusion MRI	Region-based feature set and automatic feature selection	SVM	LOO-CV	ASD vs. C	80.00
Uddin et al. (2011)	ASD = 24 C = 24	Children; Adolescents	Structural MRI	Voxel-based feature set and automatic feature selection	SVM	LOO-CV	ASD vs. C	92.00
Anderson et al. (2011)	ASD = 40 C = 40 *rep. cohort* ASD = 20 C = 20	Adolescents; Adults	Resting state functional MRI	Region-based feature set and automatic feature selection	n.s.	LOO-CV	ASD vs. C *rep. cohort*	79.00 71.00
Murdaugh et al. (2012)	ASD = 13 C = 14	Adults	Resting state functional MRI	Region-based feature set	LRC	LOO-CV	ASD vs. C	96.00
Uddin et al. (2013)	ASD = 20 C = 20	Children	Structural MRI; Resting state functional MRI	Network-based feature set	LRC	n.s.-CV	ASD vs. C	78.00
Deshpande et al. (2013)	ASD = 15 C = 15	Adults	Task functional MRI—theory of mind; Diffusion MRI	Region-based feature set and automatic feature selection	SVM	1/10-CV	ASD vs. C	95.90
ABIDE-sample[***]	ASD = 539 C = 573	Children; Adolescents; Adults	Structural MRI; Resting state functional MRI	Feature sets and selections differ	multiple	multiple	ASD vs. C	≈68.00
Lim et al. (2013)[**]	ASD = 19 ADHD = 29 C = 29	Adolescents	Structural MRI	Voxel-based feature set	GPC	LOO-CV	ADHD vs. ASD ADHD vs. C ADHD vs. ASD vs. C	85.20 79.30 68.20
Ingalhalikar et al. (2014)	ASD = 57 C = 42	Children	Diffusion MRI	Region-based feature set	LDC	1/5-CV	ASD vs. C	74.00
Wee et al. (2014)	ASD = 58 C = 59	Children; Adolescents	Structural MRI	Region-based feature set	SVM	1/2-CV	ASD vs.C	96.30
Segovia et al. (2014)	ASD = 52 ASD-sibl = 40 C = 40	Adults	Structural MRI	Voxel-based feature set and automatic feature selection	SVM	n.s.-CV	ASD vs. C ASD vs. ASD-sibl	80.00 85.00

ASD = Participants with autism spectrum disorder; SCH = schizophrenia; ADHD = Attention deficit hyperactivity disorder; C = Control participants; ASD-sibl = Siblings of participants with autism spectrum disorder; SVM = Support vector machine classifier; LRC = Logistic regression classifier; GPC = Gaussian process classification; LDC = Linear discriminant classifier; LOO-CV = Leave one out-cross validation; n.s. = not specified.

[#] The maximal accuracy reported in the article.
[*] Reported in Table 1.
[**] Reported in Table 4.
[***] Pattern recognition studies making diagnostic predictions on the ABIDE sample (Overview paper: Di Martino et al. (2014); Zhou et al. (2013); Iidaka (2014); Nielsen et al. (2013); Sabuncu and Konukoglu (2014)).

small compared to other disorders and the predictions in those samples showed similar performance compared to other samples in SCZ (Anderson and Cohen, 2013; Sabuncu and Konukoglu, 2014; Sui et al., 2013). Recently, a study combining diffusion-weighted as well as structural MRI reached perfect accuracies, by building connectivity matrices based on both image modalities (Zhu et al., 2014). However, it is important to note that this study was small, with only ten subjects per condition, and therefore, it is not clear whether those results will generalise well to a different sample.

In comparison to other psychiatric disorders, SCZ is the disorder to which pattern recognition methods have most commonly been applied. Initial results indicate that this disorder can be accurately predicted, although accuracies vary considerable between studies. Primarily, functional and structural MRI modalities were researched and it is difficult to make a distinction between certain MRI modalities, with regard to the performance of a classifier. Large scale studies in SCZ are still missing.

3.2. Bipolar disorder and major depressive disorder

Several studies have used pattern recognition to make diagnostic predictions of BPD and MDD (Table 2). The initial results indicate that classification of BPD is challenging but possible (Arribas et al., 2010; Costafreda et al., 2011; Lueken et al., 2014; Redlich et al., 2014; Rocha-Rego et al., 2014). In those studies, structural images provide features, which could reliably be used to distinguish BPD cases from healthy participants. However, the number of studies on BPD is relatively low, and therefore, conclusions can only be tentative (Table 2). In contrast, many studies have employed pattern recognition methods to predict MDD diagnosis. This started

in 2008 with two studies, one using a verbal memory N-back task (Marquand et al., 2008) and the other using a sad facial processing task (Fu et al., 2008), with diagnostic accuracies of 67.5% and 84%, respectively. The first structural MRI and resting state fMRI studies were performed a year later. The accuracies obtained from resting state fMRI-derived features were particularly promising (Costafreda et al., 2009; Craddock et al., 2013). In the following years, the number of studies accelerated rapidly and the accuracies ranged from 67% to 99% (Table 2). However, all studies were performed in relatively small to medium sized samples; the largest study preformed in MDD using pattern recognition consisted of 62 patients and an equal number of control participants. Based on a structural features the two groups could be distinguished with 90% accuracy (Mwangi et al., 2012). Cross-disorder classifications are particularly difficult especially when the disorders that need to be distinguished are similar as the case in BPD and MDD. However, BPD could be distinguished from MDD on the basis of multivariate patterns based on gray matter differences (Redlich et al., 2014). In this study, the findings were validated in an independent cohort and the test-set validation yielded comparable accuracies to the within sample leave one out cross-validation. SCZ could also be distinguished from MDD, with even higher accuracies (Table 2; Yu et al., 2013; Ota et al., 2013; Costafreda et al., 2011). Those studies employed different feature sets, ranging from task based functional to structural and diffusion MRI, indicating that differential diagnostics based on multiple MRI modalities is feasible.

In summary, for MDD and BPD large studies are still missing to a degree that allows definitive conclusions to be drawn. Especially BPD requires further research across different modalities.

3.3. Obsessive compulsive, social anxiety, post-traumatic stress disorder and specific phobia

As shown in Table 3, only few, recent studies have applied pattern recognition to OCD, SAD, PTSD, and SP. In an early study on OCD, using an approach calculating the distance between individual participants and the mean of OCD and control group based on structural MRI derived measures, showed that OCD patients could be distinguished from controls relatively reliable in an independent test-set (Soriano-Mas et al., 2007). This initial study set the stage for subsequent pattern recognition studies some years later. In a task-based functional MRI study, in which stimuli with emotional valence were presented, OCD patients could be classified with perfect accuracy in a very small sample of ten participants per group (Weygandt et al., 2012a). These promising results required further verification in a larger sample to exclude specific study or sample characteristics as cause for the good performance. However, subsequent studies could not substantiate these results to the same degree (Li et al., 2014; Shenas et al., 2014). Two studies on SAD have been published in small samples, reporting accuracies above 80% (Frick et al., 2014; Liu et al., 2013). Multivariate patterns in these studies were derived from different MRI modalities indicating that features relevant for disorder classification can be extracted and analysed across modalities. In general informative features were found to be distributed across widespread brain areas (Frick et al., 2014; Liu et al., 2013) and not easily localisable to brain regions usually associated with anxieties, such as the limbic lobes (Damsa et al., 2009). An additional study showed that structural images can be used as a diagnostic medium for different specific phobias, which could reliably be classified based on grey and white matter densities (Lueken et al., 2014). In a study conducted after an earthquake in Sichuan (China), 50 survivors with and without PTSD were compared to controls using structural imaging. Patients with PTSD could be classified with an accuracy of 91% (Gong et al., 2014). If replicated, such data suggest that severe traumas can alter brain structure to a degree that can be picked up as clinically meaningful multivariate patterns. The highest discriminative weights were found in different areas of the brain in particular in left and right parietal regions.

In summary, it can be determined that the number of studies on obsessive, compulsive and anxiety disorders is too small to determine general conclusions. While first results appear promising, these disorders do require further research especially in larger samples.

3.4. Attention-deficit/hyperactivity disorder

For the neurodevelopmental disorder ADHD, the first report of an application of pattern recognition was published in 2008 describing a sample of 12 children/adolescents and an equal number of controls (Table 4). Resting-state fMRI data was used to classify these two groups with an accuracy of 75 to 80% (Zhu et al., 2008). Four later studies focused on structural MRI and yielded accuracies for the classification of ADHD ranging from about 72% to 93% (Igual et al., 2012; Johnston et al., 2014; Lim et al., 2013; Peng et al., 2013). With the exception of one study, all other used feature selection methods prior to classification. The study with the highest predictive accuracy employed an automated feature selection algorithm (Johnston et al., 2014). The weights associated with areas in the brainstem contributed mostly to the classification. However, these brainstem regions have not been associated with ADHD prior to this finding, and further exploration is needed. Another study on structural MRI data used Gaussian process classification on whole brain grey and white matter and could predict ADHD with an accuracy of 79.3% (Lim et al., 2013). The same type of classifier was also employed in two different task fMRI studies, a stop signal and a temporal discounting task. The accuracies in those studies were 77% and 75%, respectively (Hart et al., 2014a, 2014b), indicating that voxel-based feature-sets can yield good predictive accuracies. While small studies with presumably relatively homogeneous samples have shown promising predictions, the question about generalisability remains. For this reason, a data-sharing project, the ADHD-200 Global Competition, was set-up in 2012 (The ADHD Consortium, 2012), which allowed different groups from all over the world to train their machine-learning algorithms on a legacy multi-site dataset, with the goal to find classifiers suitable for ADHD diagnostics. The study collected existing data on about 350 patients and 554 controls, and multiple classifiers were employed and trained on data ranging from demographic tables to structural and resting-state fMRI. Unfortunately, the classification results were disappointing with accuracies not exceeding 61% on the held out test sample (Castellanos et al., 2013; Sabuncu and Konukoglu, 2014).

In summary, it appears that while small studies performed well in predicting case status, the large ADHD-200 sample seemed to suffer from the high heterogeneity of the disorder and/or of the experimental characteristics of the studies contributing to this sample. No research on pattern recognition of diffusion MRI-derived features has been published up until today, while widespread alterations in white matter have been linked to ADHD (see for meta-analysis van Ewijk et al., 2012).

3.5. Autism spectrum disorder

Autism spectrum disorder, another (male-dominant) neurodevelopmental disorder, has been studied using pattern recognition since 2010, when three studies, one in males (Ecker et al., 2010a), and two in mixed gender samples of about 20 patients and controls (Ecker et al., 2010b; Jiao et al., 2010), were published. These initial results were promising, and showed accuracies ranging from 81% for whole brain structural features to 90% for regional features (Table 5). The first resting state and diffusion MRI studies in ASD

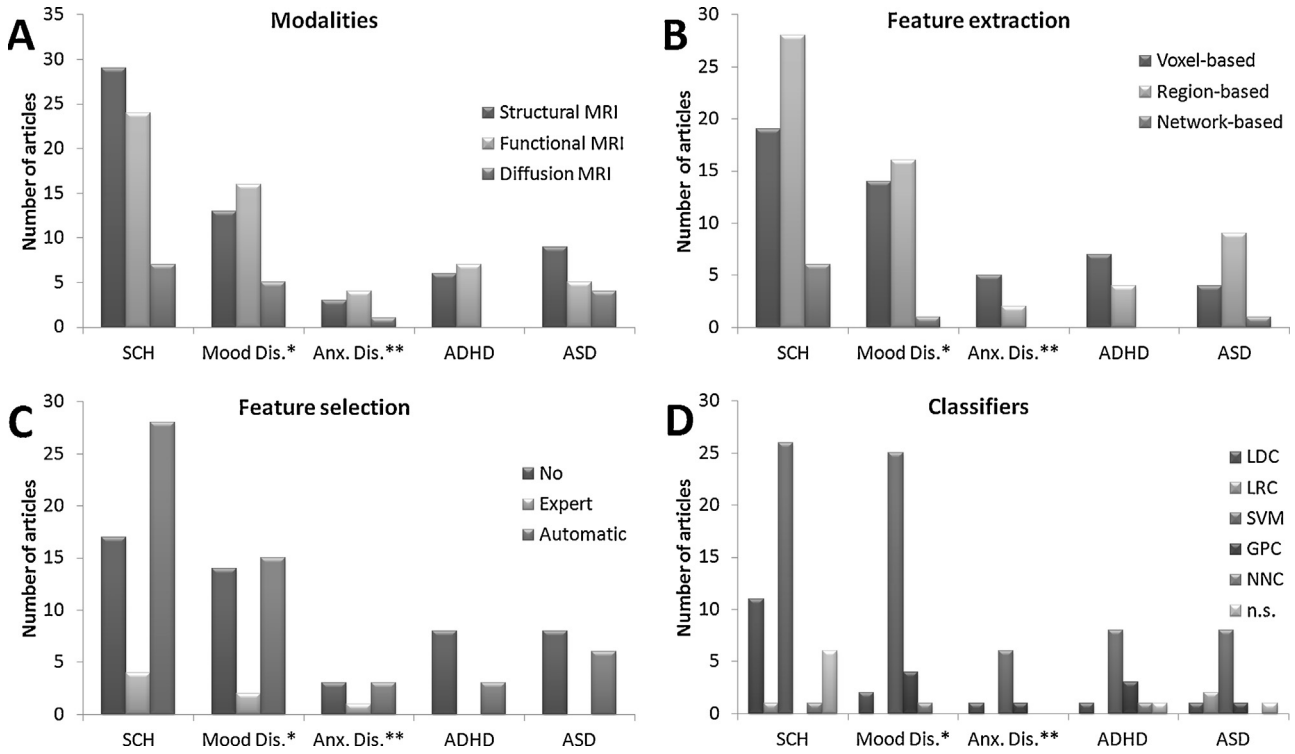

Fig. 3. Visual summary of articles reviewed. (A) Number of articles on different modalities; (B) number of articles of different feature extractions; (C) number of articles on different feature selections; (D) number of articles on different classifiers; SCZ = schizophrenia; mood Dis. = mood disorders; * BPD = bipolar disorder, MDD = major depressive disorder; Anx. Dis = anxiety disorders; ** OCD = obsessive compulsive disorder, SAD = social anxiety disorder, PTSD = post-traumatic stress disorder, SP = specific phobia; ADHD = attention-deficit/hyperactivity disorder; ASD = Autism spectrum disorder.

were analysed with pattern recognition. This provided comparable accuracies, and showed that other modalities than the structural can provide informative features for ASD classification (Anderson et al., 2011; Ingalhalikar et al., 2011). After those initial publications, eight additional studies were published up until December 2014, which extracted features from different modalities. The accuracies in those studies ranged from 70% to 96% (Table 5; Zhou et al., 2014; Segovia et al., 2014; Iidaka, 2014; Ingalhalikar et al., 2014; Uddin et al., 2013; Deshpande et al., 2013; Murdaugh et al., 2012; Uddin et al., 2011). In one of those studies, a searchlight algorithm was applied to structural MRI data, yielding a maximal accuracy of 92% for subsets of voxels (Uddin et al., 2011). In this type of analysis, a small number of voxels in spatial proximity with one another provide the features for prediction. The searchlight moves across the whole feature space and repeats the prediction at every location, potentially allowing a better interpretability with regard to the localisation of an effect. A large data-sharing effort was also established for ASD, the ABIDE consortium (Di Martino et al., 2014), but – similar to the situation for ADHD – the diagnostic predictions in this cohort were disappointing (Iidaka, 2014; Nielsen et al., 2013; Sabuncu and Konukoglu, 2014; Zhou et al., 2014).

In summary most MRI modalities have been researched in ASD, although the number of studies per modality is still relatively low. As in ADHD, clinical and experimental heterogeneity might hinder good classification in ASD in larger (retrospective) sample collections. Importantly, specific subgroups in ASD, such as lower functioning ASD patients and female patients are currently underrepresented in the existing studies.

3.6. Other psychiatric disorders

A few studies have investigated additional psychiatric disorders, especially substance use disorders or eating disorders. In one study it was shown that patients with different types of eating disorders could be distinguished based on their brain response viewing pictures of food and neutral stimuli, indicating that differential diagnosis between eating disorders is feasible (Weygandt et al., 2012b). Drug dependence could also be classified by MR-derived features (Zhang et al., 2005, 2011), in the latter study classification on the basis of frontal regions appeared especially promising.

In summary, few studies have employed pattern recognition techniques for psychiatric disorders other than those described in the previous sections. However, preliminary studies published in eating disorders and drug dependence suggest that pattern recognition methods may be applicable more widely than has been demonstrated to date.

4. Discussion

In the present article we extensively and systematically reviewed the literature applying different pattern recognition methods for neuroimaging to psychiatric disorders. Despite many promising results, the use of pattern recognition for assisting diagnosis of psychiatric disorders is still in its infancy. Considering the complexity of the problem, starting with the diagnostic process of psychiatric disorders to the challenges of MRI, this is perhaps not surprising. Structural and functional MRI have been investigated approximately equally often, but diffusion MRI was relatively neglected, especially in ADHD (Fig. 3A). Features sets were most frequently based on either the regional or voxel level (Fig. 3B), and automatic feature selection was more often applied than expert feature selection (Fig. 3C). While the body of literature on SCZ, MDD, ADHD, and ASD is already relatively extensive, only a few studies have applied pattern recognition to OCD, SAD, PTSD, SP, and BPD. Many studies suffer from common limitations, such as a relatively small sample size and a primary focus on estimating generalisability within a single study. Furthermore, as described

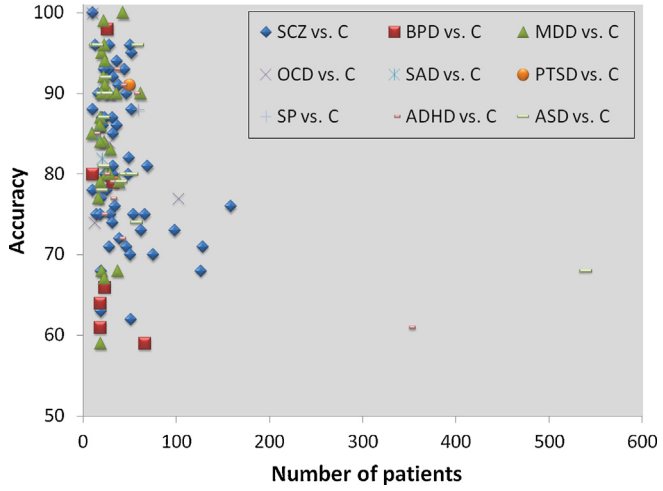

Fig. 4. Scatter plot with the number of patients included in a specific study on the *x*-as and the, respectively, reported maximal accuracy for the patient vs. control prediction on the *y*-as. SCZ = schizophrenia; BPD = bipolar disorder; MDD = major depressive disorder; OCD = obsessive compulsive disorder; SAD = social anxiety disorder; PTSD = post-traumatic stress disorder; SP = specific phobia; ADHD = attention-deficit/hyperactivity disorder; ASD = Autism spectrum disorder; C = controls.

below, the few larger studies that have been performed – mainly based on multi-site retrospective collection of data – tend to have reduced performance in comparison with smaller studies (Fig. 4).

Classifiers used in the various studies included linear discriminant classifier, logistic regression classifier, support vector machines, Gaussian process classifiers, and neural network classifiers differ little in their performance but more with respect to their applicability to large datasets and their computational requirements. Unsurprisingly SVM, which is applicable to a large number of features in combination with just a few examples, often represents the method of choice (Fig. 3D). However, as diagnostic labels in psychiatry are inherently uncertain, predictive algorithms that provide probabilistic predictions, such as GPC, may provide advantages for clinical studies. For example, the probabilistic predictions can be easily recalibrated to accommodate variations in disease prevalence (see below).

Only a few studies have directly compared either different modalities or different feature selection approaches. Especially for less frequently researched disorders, this makes a definite conclusion with regard to the usefulness of specific modalities impossible at the current time. Wherever different modalities have been used, they appear to result in comparable accuracies. Earlier, we introduced the concept of feature selection and made a distinction between expert and automatic selection. Both approaches have been applied in the reviewed articles. Automatic feature selection has been a useful tool to improve classifications in some cases (Craddock et al., 2009; Qin et al., 2014), while in others feature selection did not make a difference (Ecker et al., 2010b). However, extra care needs to be taken when interpreting articles applying feature selection, as there is a danger of over-fitting if the split between training and test sets during feature selection is incomplete and parameter optimisation may be an issue in some studies.

In general, reported accuracies are likely to be optimistic and confidence intervals are often not reported. Only few studies have reported accuracy measures on additional datasets, limiting generalisability of the results. The observation that larger studies tend to have reduced performance in comparison with smaller studies is hard to interpret, and at the same time surprising, as it may be expected that the number of examples should increase predictive performance of classifiers[7]. It could point to a bias in the literature, or simply be due to the fact that large-scale studies are hampered by an increase in heterogeneity within clinical groups. This heterogeneity can stem from many sources: large cohorts are often not as well matched on important demographic and clinical variables and, further, data are often pooled across research centres providing data from different scanners and/or acquisition sequences. This may be particularly problematic for large studies that aggregate legacy datasets that were originally acquired for other purposes. On a positive note, there is increasing evidence from pattern recognition studies of neurological disorders that when data acquisition parameters for structural MRI are carefully controlled across sites, accurate generalisation across sites is feasible (e.g. Abdulkadir et al., 2014; Doyle et al., 2014; Klöppel et al., 2008).

In contrast, data derived from task fMRI, resting state fMRI or diffusion MRI may suffer from more problems when pooled in comparison to structural MRI data, which allows for less degrees of freedom in terms of procedures or protocols of acquisition. Furthermore, larger samples often suffer from missing data. These factors may impair the ability of the classifier to learn an appropriate decision rule and thus reduce accuracy. These considerations highlight the need for well-powered and carefully controlled prospectively acquired samples and for further research into developing acquisition and preprocessing pipelines that are robust to inter-scanner differences.

In the following sections we evaluate the progress made in translating PR techniques towards application in the clinical diagnostics and identify clinical as well as research applications for which these techniques are most likely to add value.

4.1. Translating pattern recognition towards clinical utility

The diagnostic process in real clinical populations is more complex than in research settings, where two, usually balanced, groups of well-matched patients and controls are carefully diagnosed before a supervised pattern recognition algorithm is trained. Often, subjects with uncertain diagnoses or comorbidities are excluded. Therefore, considerable work remains for pattern recognition to become applicable in clinical practice. In many cases, the question to be answered will not be related to distinguishing patients from controls; rather, distinction between different disorders in the same population will be needed. In other words, what is required is a differential diagnosis. Furthermore, it is necessary to identify individuals that have comorbidities and may be members of multiple diagnostic classes. There are very few studies in the current literature that tackle these problems. A few studies across different disorders have demonstrated multiclass classification of three or more disorders (Costafreda et al., 2011; Grotegerd et al., 2013; Koutsouleris et al., 2015; Lim et al., 2013; Redlich et al., 2014; Serpa et al., 2014; Yu et al., 2013). We are not aware of any pattern recognition studies that have tackled the issue of comorbidities. Further work in this area is therefore needed.

While most pattern recognition algorithms are trained and tested on well balanced samples, the prevalence of most psychiatric disorders in the general population is five percent or lower. Even if the classification models are not applied at the population level, the relative class frequencies in the training set are often different from the test set or target application domain. Therefore, in such cases, accuracies that are derived from those samples are not immediately representative for predictions in clinical samples. In settings with

[7] Whether or not an increase in sample size improves accuracy depends on how well the data meets the assumptions made by the classifier (e.g. whether the data are independent and identical distributed) and what the Bayes error of the classification problem is.

low prevalence, clinical diagnostic algorithms require very high specificity to prevent an unacceptably high false positive rate (i.e. to prevent too many healthy individuals being erroneously diagnosed). Therefore, it is important to tailor the predictive properties of specific pattern recognition algorithms depended on the setting in which they are used, as they will have to be biased differently in a population as opposed to a high-risk or clinical setting. Here again, the use of probabilistic pattern recognition algorithms is indicated. An important advantage of these algorithms is that they allow the decision threshold to be easily recalibrated to accommodate different diagnostic settings (Bishop, 2006). A proof of concept of this idea for neuroimaging was demonstrated in Hahn et al. (2013), in which an unbalanced test sample was predicted based on a balanced training sample, the classifier was adjusted for class priors yielding better performance.

Another important consideration for diagnostic applications of pattern recognition is the relative expense of MRI relative to other measures. The prospective value of employing an automated approach should be balanced against the costs of data acquisition. This, combined with moderate diagnostic accuracies for most disorders, means that pattern recognition techniques are not likely to be applied directly to making diagnostic predictions in the immediate future. However, pattern recognition may well provide a useful adjunct to clinical decision-making in cases where uncertainty is high or if a decision for an expensive or invasive procedure needs to be taken. Perhaps more importantly than such direct diagnostic applications, we identify several specific applications for which pattern recognition in a diagnostic context is more immediately applicable.

First, pattern recognition techniques can provide a richer endophenotype than behaviour alone, which may ultimately be informative about disease mechanisms, staging and progression. One way that this can be achieved is by means of mapping classifier weights back into brain space. In that way biological pathways that might be implicated in the aetiology of a particular disorder could be more appropriately described. This line of research might eventually yield new insights into the mechanisms of particular disorders and thus to diagnoses based on 'biomarkers' rather than based primarily on clinical interviews. However, care needs to be taken when such weight-maps are interpreted in terms of effect localisation (Haufe et al., 2014).

Second, pattern recognition algorithms that provide estimates of predictive confidence across multiple classes could be used to disentangle comorbidities in patient groups. True multiclass methods (Filippone et al., 2012), provide a prediction for all classes for each subject. Therefore, if a subject has a high predictive confidence for multiple disorders it may indicate that the subject is comorbid for those disorders. In contrast, if the predictive confidence is only high for one of the disorders, this may be considered a relatively definitive diagnostic prediction. Another intriguing possibility for accommodating comorbidities is multi-label classification (Zhang and Zhou, 2014), or multi-task learning (Pan and Yang, 2010) where each data sample has multiple labels associated to it. Multi-task learning has been applied in neuroimaging for predicting multiple clinical variables (e.g. Marquand et al., 2014; Zhang and Sen, 2013) but to our knowledge, it has not been applied to accommodate comorbidities and may therefore be a promising avenue of future research.

Third, pattern recognition might provide an efficient validation procedure for the stratification of patient groups. The rationale here is that, if a subcategorisation reduces the biological heterogeneity of a patient group, prediction of the subgroups should be more accurate.

Fourth, classification based on MRI-derived feature sets might be employed as a 'triage' or screening step before an expert opinion is sought. For example, for certain clinical conditions, 'gold standard' diagnoses require specially trained and thus expensive medical specialists and several hours of interview. This is the case, for example, for the administration of the gold standard diagnostic instruments for ASD, the autism diagnostic interview (Lord et al., 1994) and the autism diagnostic observation schedule (Lord et al., 1989). In such cases, a prior screening of patients using automated techniques might improve cost-efficiency.

Fifth, in disorders that are characterised by dangerous behaviours of the patient for himself or his environment as well as in situation when difficult clinical decisions have to be made, pattern recognition might be used to further specify a clinical indication. Concretely, schizophrenic patients might be classified based on their likelihood of a psychotic episode or whether a prescription of Clozapine should be indicated, an antipsychotic that is often more effective but requires weekly white blood cell counts. In this way, the (still quite considerable) expenses of acquiring MRI could be justified by the added information provided to the clinician, leading to direct adjustment of treatment. In addition to the diagnostic studies reviewed here, pattern recognition has a clear merit in predicting treatment response (e.g. Gong et al., 2011) or naturalistic outcome (e.g. Schmaal et al., 2014). Outcome predictions provide a promising way to improve cost-efficiency, by targeting potential treatments to patients most likely to benefit from them.

Finally, another promising application of pattern recognition in the clinical domain could be sample enrichment in clinical trials. Pattern recognition algorithms could help to select a biologically more homogenous group to reduce the sample size needed to detect an effect of a new medication.

In summary, there are many challenges that need to be addressed to move pattern recognition towards clinical practice. Studies that demonstrate applicability to differential diagnosis and the identification of comorbidities are needed. In the shorter term, pattern recognition could be an effective tool in assisting complex clinical diagnostic process in specific diagnostic settings. Especially for specific diagnostic questions, biological validation of the stratification of diagnostic groups and the identification of biological pathways implicated in a particular disorder, this approach has its merits. Employing pattern recognition for widespread screening of clinical populations is premature, and further research on how to accommodate variations in disease prevalence is needed.

4.2. *Reducing heterogeneity of psychiatric diagnoses*

Psychiatric disorders may be characterised by different etiologic factors across individuals, with different disease causes potentially resulting in the same symptoms. Thus, different features might be meaningful across patients belonging to the same diagnostic group, increasing the difficulty to find a multivariate pattern, which predicts a specific diagnostic category. In addition, there is heterogeneity on the diagnostic level. The diagnostic process for many mental disorders is complex and requires considerable expertise, experience, and time. Despite best efforts, and due to the complexity of the underlying pathology, some examples within a diagnostic group may be mislabelled. Furthermore, patients can be diagnosed with the same disorder despite showing different patterns of symptoms. It is not always clear, whether or not these symptoms are a consequence of similar biological mechanisms. Classification of cases can be impaired if the number of examples with an uncertain or incorrect diagnosis is high.

One way to reduce the uncertainty of diagnostic categories is to stratify groups into smaller, potentially less heterogeneous subgroups, by for example selecting those individuals which have shown a representative pattern of brain abnormalities specific to a certain disorder, indicated by high predictive probabilities across different MRI modalities for this disorder. As exemplified by the reduced diagnostic accuracy reported for larger studies, current

pattern recognition algorithms seem insufficiently able to accommodate such biological (and/or experimental) heterogeneity (see above). It is reasonable to expect that accurate diagnostic biomarkers will be easiest to find for diagnostic labels with a consistent pattern of underlying pathology. One way to achieve this is to stratify diagnostic groups into smaller biologically more meaningful subgroups. The biological validity of these smaller groups could then be estimated by pattern recognition algorithms based on MRI-derived feature sets. This process could generate hypotheses and can be used to evaluate alternative diagnostic stratifications. Related to this point, the misclassifications that a classifier makes can be highly informative about stratifications within clinical groups. To take a simple example, a classifier might predict that the imaging-derived pattern for a given individual resembles one typical for some disorder but the person shows no symptoms. This could indicate a prodromal or at-risk disease phase, but might also mean that this individual has developed compensatory mechanisms that inhibit symptom expression.

4.3. The importance of the choice of features

MRI-derived features are an indirect measure of underlying brain biology. To illustrate, the resolution of a modern MRI scan rarely exceeds 1 cubic millimetre. While abnormalities at this scale already allow clinical predictions based on neuroimaging, some processes in the brain implicated in the aetiology of psychiatric disorders are present in small regional networks, single neurons, or even take place at the molecular level in axon terminals (Glausier et al., 2014). Increasing the spatial resolution of MRI technology might therefore provide us with means to infer discriminative information hidden today, which might improve clinical predictions in the future.

As mentioned earlier, psychiatric disorders are characterised by considerable heterogeneity, and many patients belonging to the same diagnostic group could do so for various reasons. Therefore, it is important to consider a range of different data types, from genetics to information on the social network of a patient. The information contained in those features would need to be combined to interpret the data efficiently. Multi-modal techniques such as 'multi-kernel learning' (Filippone et al., 2012; Hinrichs et al., 2011; Zhang and Sen, 2013) or linked independent component analysis (Francx et al., in preparation), which combine different data types, might help to make sense of those multimodal datasets. Although potential improvements would largely depend on the information contained in such data, these techniques might yield better interpretability and a more complete picture of the aetiology of a disorder.

Feature selection guided by prior knowledge can improve accuracy in disorders, which allow clear enough hypotheses on specific relevant imaging features; this has already been shown for Alzheimer's disease (Chu et al., 2012). For most psychiatric disorders this prior knowledge, if available today at all, is only accessible with considerable uncertainty, as the results with respect to specific imaging features still diverge between studies. Therefore, hypotheses which could guide feature selection are scarce and often not robust enough at the current time. Large scale meta-analyses capturing information on altered brain phenotypes are necessary for individual psychiatric disorders as well as across disorders. In addition, such meta-analyses should also assess different MRI modalities. Recent efforts, like the ENIGMA Working Groups on different brain disorders and the ENIGMA Cross-disorder Working Group, might provide a first step in this direction[8] (e.g. Erp et al., in press; Hibar et al., 2015; Schmaal et al., in press). In addition, recent efforts combining information from voxel-based morphological studies across different disorders into a unified profile of results within and across disorders could be a right step in this direction (Goodkind et al., 2015), and certainly sets the scene for research focussing on unrevealing common and unique factors contributing to psychiatric disorders. It is important that the publications on such data contain specific recommendations for the machine learning community and that prior knowledge on informative features for psychiatric disorders is organised in suitable and accessible ways, so that it can guide expert-based feature selection more effectively. Importantly, those publications require estimates of effect sizes of specific potential features and should not shy away when the overall pattern of results seem inconclusive. As our understanding of the pathology underlying psychiatric disorders improves, this knowledge can increase the accuracy of future pattern recognition models beyond what is currently achievable. Increasing clinical prediction today is (next to improving or adapting pattern recognition to clinical requirements) predominantly a quest for informative features.

5. General conclusion

Pattern recognition has shown promising but mixed results for predicting diagnosis in clinical neuroimaging research. Despite many promising proof of concept results, it is clear that the techniques are still in their infancy, and many challenges remain to be solved before they can be employed in clinical practice. Some of the main challenges stem from the very high aetiological and phenotypic heterogeneity that characterises psychiatric disorders. The studies we have reviewed are mostly employing relatively small samples acquired in single imaging centres, with only a few efforts to simultaneously discriminate different disorders. Clearly, more extensive validation using large, carefully acquired samples is necessary, as is the development of methodological innovations to improve accuracy, to discriminate between multiple disorders simultaneously, and to translate to settings having realistic disease prevalence. In the short to medium term, however, there are specific clinical contexts where pattern recognition can add value; for example as triage tools in cases where the diagnosis of a trained clinician is expensive or access to such resources is limited, to subdivide patient populations into more homogeneous subgroups, and to help fractionate disease phenotypes. In order for the field to move forward, we argue that it is essential to combine clinical and methodological expertise to make optimal use of the rich source of information provided by neuroimaging data. International research consortia bringing together clinicians and researchers across disciplines, like the EU-funded, IMAGEMEND[9] (Frangou et al., in press), PRONIA[10], PSYSCAN[11], and Aggressotype[12] consortia, provide valuable platforms for such interaction. In the future, we anticipate that technological developments in pattern recognition in combination with the acquisition of large, multimodal, and prospectively acquired clinical samples will enable the construction of more accurate models and will move the field closer towards biomarkers that can be used to guide clinical decision making in psychiatric disorders.

[8] http://ENIGMA.ini.usc.edus

[9] www.imagemend.eu

[10] www.pronia.eu

[11] http://ec.europa.eu/research/health/medical-research/brain-research/projects/psyscan_en.html

[12] www.aggressotype.eu

Conflict of interest statement

Barbara Franke has received a speaker fee from Merz. Jan K Buitelaar has been in the past 3 years a consultant to/member of advisory board of/and/or speaker for Janssen Cilag BV, Eli Lilly, Shire, Lundbeck, Roche and Servier. He is not an employee of any of these companies, and not a stock shareholder of any of these companies. He has no other financial or material support, including expert testimony, patents, royalties. The other authors do not report conflicts of interest. None of these, nor any of the funding agencies have had any influence on the content of this manuscript.

Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 602450 (IMAGEMEND). In addition, the study was supported by grants from The Netherlands Organization for Scientific Research (NWO), i.e. a Vici grant to Franke (grant 016-130-669), Brain & Cognition grants 433-09-242 and 056-13-015 to Buitelaar, and the Gravitation Programme Language in Interaction (grant 024.001.006). The research of Franke and Buitelaar also receives funding from the FP7 under grant agreements no. 602805 (AGGRESSOTYPE), no. 603016 (MATRICS) and no. 278948 (TACTICS), from the European Community's Horizon 2020 Programme (H2020/2014-2020) under grant agreements no. 643051 (MiND) and no. 642996 (BRAINVIEW), and from a grant for the ENIGMA Center for world-wide Medicine Imaging and Genomics from the National of Health's BD2K Initiative (grant U54 EB020403). Andre Marquand gratefully acknowledges support from the Language in Interaction consortium, funded by the NWO under the Gravitation Programme.

References

Abdulkadir, A., Peter, J., Brox, T., Ronneberger, O., Klöppel, S., 2014. Voxel-based multi-class classification of AD, MCI, and elderly controls: Blind evaluation on an independent test set. In: Proc MICCAI Workshop Challenge on Computer-Aided Diagnosis of Dementia Based on Structural MRI Data, pp. 8–15.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, fifth ed. American Psychiatric Publishing, Arlington, VA.

Anderson, A., Cohen, M.S., 2013. Decreased small-world functional network connectivity and clustering across resting state networks in schizophrenia: an fMRI classification tutorial. Front. Hum. Neurosci. 7, 1–18, http://dx.doi.org/10.3389/fnhum.2013.00520.

Anderson, A., Dinov, I.D., Sherin, J.E., Quintana, J., Yuille, A.L., Cohen, M.S., 2010. Classification of spatially unaligned fMRI scans. NeuroImage 49, 2509–2519, http://dx.doi.org/10.1016/j.neuroimage.2009.08.036.

Anderson, J.S., Nielsen, J.a., Froehlich, A.L., DuBray, M.B., Druzgal, T.J., Cariello, A.N., Cooperrider, J.R., Zielinski, B.a., Ravichandran, C., Fletcher, P.T., Alexander, A.L., Bigler, E.D., Lange, N., Lainhart, J.E., 2011. Functional connectivity magnetic resonance imaging classification of autism. Brain 134, 3742–3754, http://dx.doi.org/10.1093/brain/awr263.

Arbabshirani, M.R., Kiehl, K.a., Pearlson, G.D., Calhoun, V.D., 2013. Classification of schizophrenia patients based on resting-state functional network connectivity. Front. Neurosci. 7, 1–16, http://dx.doi.org/10.3389/fnins.2013.00133.

Ardekani, B.A., Tabesh, A., Sevy, S., Robinson, D.G., Bilder, R.M., Szeszko, P.R., 2011. Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers. Hum. Brain Mapp. 32, 1–9, http://dx.doi.org/10.1002/hbm.20995.

Arribas, J.I., Calhoun, V.D., Adali, T., 2010. Automatic bayesian classification of healthy controls, bipolar disorder and schizophrenia using intrinsic connectivity maps from fMRI data. IEEE Trans. Biomed. Eng. 57, 2850–2860, http://dx.doi.org/10.1109/TBME.2010.2080679.

Bansal, R., Staib, L.H., Laine, A.F., Hao, X., Xu, D., Liu, J., Weissman, M., Peterson, B.S., Zhan, W., 2012. Anatomical brain images alone can accurately diagnose chronic neuropsychiatric illnesses. PLoS ONE 7, http://dx.doi.org/10.1371/journal.pone.0050698.

Bassett, D.S., Nelson, B.G., Mueller, B.a., Camchong, J., Lim, K.O., 2012. Altered resting state complexity in schizophrenia. NeuroImage 59, 2196–2207, http://dx.doi.org/10.1016/j.neuroimage.2011.10.002.

Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans. Med. Imaging 23, 137–152.

Bishop, C.M., 2006. Pattern Pecognition and Machine Learning, first ed. Springer, http://dx.doi.org/10.1117/1.2819119.

Borgwardt, S., Koutsouleris, N., Aston, J., Studerus, E., Smieskova, R., Riecher-Rössler, A., Meisenzahl, E.M., 2012. Distinguishing prodromal from first-episode psychosis using neuroanatomical single-subject pattern recognition. Schizophr. Bull. 39, 1105–1114, http://dx.doi.org/10.1093/schbul/sbs095.

Brodersen, K.H., Ong, C.S., Stephan, K.E., Buhmann, J.M., 2010. The balanced accuracy and its posterior distribution. In: Proceedings - International Conference on Pattern Recognition, pp. 3121–3124, http://dx.doi.org/10.1109/ICPR.2010.764.

Calhoun, V.D., Sui, J., Kiehl, K., Turner, J., Allen, E., Pearlson, G., 2012. Exploring the psychosis functional connectome: aberrant intrinsic networks in schizophrenia and bipolar disorder. Front. Psychiatry 2, 1–13, http://dx.doi.org/10.3389/fpsyt.2011.00075.

Cao, L., Guo, S., Xue, Z., Hu, Y., Liu, H., Mwansisya, T.E., Pu, W., Yang, B., Liu, C., Feng, J., Chen, E.Y.H., Liu, Z., 2014. Aberrant functional connectivity for diagnosis of major depressive disorder: a discriminant analysis. Psychiatry Clin. Neurosci. 68, 110–119, http://dx.doi.org/10.1111/pcn.12106.

Castellani, U., Rossato, E., Murino, V., Bellani, M., Rambaldelli, G., Perlini, C., Tomelleri, L., Tansella, M., Brambilla, P., 2012. Classification of schizophrenia using feature-based morphometry. J. Neural Transm. 119, 395–404, http://dx.doi.org/10.1007/s00702-011-0693-7.

Castellanos, F.X., Di Martino, A., Craddock, R.C., Mehta, A.D., Milham, M.P., 2013. Clinical applications of the functional connectome. NeuroImage 80, 527–540, http://dx.doi.org/10.1016/j.neuroimage.2013.04.083.

Castro, E., Martínez-Ramón, M., Pearlson, G., Sui, J., Calhoun, V.D., 2011. Characterization of groups using composite kernels and multi-source fMRI analysis data: application to schizophrenia. NeuroImage 58, 526–536, http://dx.doi.org/10.1016/j.neuroimage.2011.06.044.

Chu, C., Hsu, A.-L., Chou, K.-H., Bandettini, P., Lin, C., 2012. Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. NeuroImage 60, 59–70, http://dx.doi.org/10.1016/j.neuroimage.2011.11.066.

Costafreda, S.G., Chu, C., Ashburner, J., Fu, C.H.Y., 2009. Prognostic and diagnostic potential of the structural neuroanatomy of depression. PLoS ONE 4, http://dx.doi.org/10.1371/journal.pone.0006353.

Costafreda, S.G., Fu, C.H.Y., Picchioni, M., Toulopoulou, T., McDonald, C., Kravariti, E., Walshe, M., Prata, D., Murray, R.M., McGuire, P.K., 2011. Pattern of neural responses to verbal fluency shows diagnostic specificity for schizophrenia and bipolar disorder. BMC Psychiatry 11, 2–10, http://dx.doi.org/10.1186/1471-244X-11-18.

Craddock, R.C., Holtzheimer, P.E., Hu, X., Mayberg, H.S., 2013. Disease state prediction form resting state functional connectivity. Magn. Reson. Med. 62, 1619–1628, http://dx.doi.org/10.1002/mrm.22159.Disease.

Craddock, R.C., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S., 2009. Disease state prediction from resting state functional connectivity. Magn. Reson. Med. 62, 1619–1628, http://dx.doi.org/10.1002/mrm.22159.

Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M.-O., Chupin, M., Benali, H., Colliot, O., 2011. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. NeuroImage 56, 766–781, http://dx.doi.org/10.1016/j.neuroimage.2010.06.013.

Damsa, C., Kosel, M., Moussally, J., 2009. Current status of brain imaging in anxiety disorders. Curr. Opin. Psychiatry 22, 96–110, http://dx.doi.org/10.1097/YCO.0b013e328319bd10.

Davatzikos, C., Shen, D., Gur, R.C., Wu, X., Liu, D., Fan, Y., Hughett, 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, http://dx.doi.org/10.1001/archpsyc.62.11.1218.

Deshpande, G., Libero, L.E., Sreenivasan, K.R., Deshpande, H.D., Kana, R.K., 2013. Identification of neural connectivity signatures of autism using machine learning. Front. Hum. Neurosci. 7, 670, http://dx.doi.org/10.3389/fnhum.2013.00670.

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., Deen, B., Delmonte, S., Dinstein, I., Ertl-Wagner, B., Fair, D.a., Gallagher, L., Kennedy, D.P., Keown, C.L., Keysers, C., Lainhart, J.E., Lord, C., Luna, B., Menon, V., Minshew, N.J., Monk, C.S., Mueller, S., Müller, R., Nebel, M.B., Nigg, J.T., O'Hearn, K., Pelphrey, K.a., Peltier, S.J., Rudie, J.D., Sunaert, S., Thioux, M., Tyszka, J.M., Uddin, L.Q., Verhoeven, J.S., Wenderoth, N., Wiggins, J.L., Mostofsky, S.H., Milham, M.P., 2014. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol. Psychiatry 19, 659–667, http://dx.doi.org/10.1038/mp.2013.78.

Doyle, O.M., Westman, E., Marquand, A.F., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Soininen, H., Lovestone, S., Williams, S.C.R., Simmons, A., 2014. Predicting progression of Alzheimer's disease using ordinal regression. PLoS ONE 9, e105542, http://dx.doi.org/10.1371/journal.pone.0105542.

Du, W., Calhoun, V.D., Li, H., Ma, S., Eichele, T., Kiehl, K.a., Pearlson, G.D., Adali, T., 2012. High classification accuracy for schizophrenia with rest and task FMRI data. Front. Hum. Neurosci. 6, 145, http://dx.doi.org/10.3389/fnhum.2012.00145.

Ecker, C., Marquand, A., Mourão-Miranda, J., Johnston, P., Daly, E.M., Brammer, M.J., Maltezos, S., Murphy, C.M., Robertson, D., Williams, S.C., Murphy, D.G.M., 2010a. Describing the brain in autism in five dimensions—magnetic resonance

imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. J. Neurosci. 30, 10612–10623, http://dx.doi.org/10.1523/JNEUROSCI. 5413-09.2010.
Ecker, C., Rocha-Rego, V., Johnston, P., Mourao-Miranda, J., Marquand, A., Daly, E.M., Brammer, M.J., Murphy, C., Murphy, D.G., 2010b. Investigating the predictive value of whole-brain structural MR scans in autism: a pattern classification approach. NeuroImage 49, 44–56, http://dx.doi.org/10.1016/j.neuroimage.2009.08.024.
Erp, T., Hibar, D., Rasmussen, J., Glahn, D., Pearlson, G., Andreassen, O., Agartz, I., Westlye, L., Haukvik, U., Dale, A., Melle, I., Hartberg, C., Gruber, O., Kraemer, B., Zilles, D., Donohoe, G., Kelly, S., McDonald, C., Morris, D., Cannon, D., Corvin, A., Machielsen, M., Koenders, L., Haan, L., Veltman, D., Satterthwaite, T., Wolf, D., Gur, R., Gur, R., Potkin, S., Mathalon, D., Mueller, B., Preda, A., Macciardi, F., Ehrlich, S., Walton, E., Hass, J., Calhoun, V., Bockholt, H., Sponheim, S., Shoemaker, J., Haren, N., Pol, H., Ophoff, R., Kahn, R., Roiz-Santiañez, R., Crespo-Facorro, B., Wang, L., Alpert, K., Jönsson, E., Dimitrova, R., Bois, C., Whalley, H., McIntosh, A., Lawrie, S., Hashimoto, R., Thompson, P., Turner, J., 2015. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. Mol. Psychiatry (in press).
Fan, Y., Liu, Y., Wu, H., Hao, Y., Liu, H., Liu, Z., Jiang, T., 2011. Discriminant analysis of functional connectivity patterns on Grassmann manifold. NeuroImage 56, 2058–2067, http://dx.doi.org/10.1016/j.neuroimage.2011.03.051.
Fan, Y., Shen, D., Gur, R.C., Gur, R.E., Davatzikos, C., 2007. COMPARE: classification of morphological patterns using adaptive regional elements. IEEE Trans. Med. Imaging 26, 93–105.
Fang, P., Zeng, L.-L., Shen, H., Wang, L., Li, B., Liu, L., Hu, D., 2012. Increased cortical-limbic anatomical network connectivity in major depression revealed by diffusion tensor imaging. PLoS ONE 7, http://dx.doi.org/10.1371/journal.pone.0045972.
Fekete, T., Wilf, M., Rubin, D., Edelman, S., Malach, R., Mujica-Parodi, L.R., 2013. Combining classification with fMRI-derived complex network measures for potential neurodiagnostics. PLoS ONE 8, http://dx.doi.org/10.1371/journal.pone.0062867.
Filippone, M., Marquand, a., Blain, F., Williams, C.R.V., Mourão-Miranda, S.C.R., Girolami, J.M., 2012. Probabilistic prediction of neurological disorders with a statistical assessment of neuroimaging data modalities. Ann. Appl. Stat. 6, 1883–1905, http://dx.doi.org/10.1214/12-AOAS562.
Frangou, S., Schwarz, E., Meyer-Lindenberg, A., Imaging Genetics for Mental Disorders (IMAGEMEND). World Psychiatry, in press.
Frick, A., Gingnell, M., Marquand, A.F., Howner, K., Fischer, H., Kristiansson, M., Williams, S.C.R., Fredrikson, M., Furmark, T., 2014. Classifying social anxiety disorder using multivoxel pattern analyses of brain function and structure. Behav. Brain Res. 259, 330–335, http://dx.doi.org/10.1016/j.bbr.2013.11.003.
Fu, C.H.Y., Mourao-Miranda, J., Costafreda, S.G., Khanna, A., Marquand, A.F., Williams, S.C.R., Brammer, M.J., 2008. Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. Biol. Psychiatry 63, 656–662, http://dx.doi.org/10.1016/j.biopsych.2007.08.020.
Glausier, J.R., Fish, K.N., Lewis, D.a., 2014. Altered parvalbumin basket cell inputs in the dorsolateral prefrontal cortex of schizophrenia subjects. Mol. Psychiatry 19, 30–36, http://dx.doi.org/10.1038/mp.2013.152.
Gollub, R.L., Shoemaker, J.M., King, M.D., White, T., Ehrlich, S., Sponheim, S.R., Clark, V.P., Turner, J.a., Mueller, B.a., Magnotta, V., O'Leary, D., Ho, B.C., Brauns, S., Manoach, D.S., Seidman, L., Bustillo, J.R., Lauriello, J., Bockholt, J., Lim, K.O., Rosen, B.R., Schulz, S.C., Calhoun, V.D., Andreasen, N.C., 2013. The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. Neuroinformatics 11, 367–388, http://dx.doi.org/10.1007/s12021-013-9184-3.
Gong, Q., Li, L., Tognin, S., Wu, Q., Lui, S., Huang, X., Marquand, A.F., 2014. Using structural neuroanatomy to identify trauma survivors with and without post-traumatic stress disorder at the individual level. Psychol. Med., 195–203, http://dx.doi.org/10.1017/S0033291713000561.
Gong, Q., Wu, Q., Scarpazza, C., Lui, S., Jia, Z., Marquand, A., Huang, X., McGuire, P., Mechelli, A., 2011. Prognostic prediction of therapeutic response in depression using high-field MR imaging. NeuroImage 55, 1497–1503, http://dx.doi.org/10.1016/j.neuroimage.2010.11.079.
Goodkind, M., Eickhoff, S.B., Oathes, D.J., Jiang, Y., Chang, A., Jones-hagata, L.B., Ortega, B.N., Zaiko, Y.V., Roach, E.L., Korgaonkar, M.S., Grieve, S.M., Galatzer-levy, I., Fox, P.T., Etkin, A., 2015. Identification of a common neurobiological substrate for mental illness. JAMA Psychiatry 4, 305–315, http://dx.doi.org/10.1001/jamapsychiatry.2014.2206.
Gould, I.C., Shepherd, A.M., Laurens, K.R., Cairns, M.J., Carr, V.J., Green, M.J., 2014. Multivariate neuroanatomical classification of cognitive subtypes in schizophrenia: a support vector machine learning approach. NeuroImage Clin. 6, 229–236, http://dx.doi.org/10.1016/j.nicl.2014.09.009.
Greenstein, D., Malley, J.D., Weisinger, B., Clasen, L., Gogtay, N., 2012. Using multivariate machine learning methods and structural MRI to classify childhood onset schizophrenia and healthy controls. Front. Psychiatry 3, 1–11, http://dx.doi.org/10.3389/fpsyt.2012.00053.
Grotegerd, D., Suslow, T., Bauer, J., Ohrmann, P., Arolt, V., Stuhrmann, A., Heindel, W., Kugel, H., Dannlowski, U., 2013. Discriminating unipolar and bipolar depression by means of fMRI and pattern classification: a pilot study. Eur. Arch. Psychiatry Clin. Neurosci. 263, 119–131, http://dx.doi.org/10.1007/s00406-012-0329-4.
Guo, H., Cheng, C., Cao, X., Xiang, J., Chen, J., Zhang, K., 2014. Resting-state functional connectivity abnormalities in first-onset unmedicated depression. Neural Regen. Res. 9, 153–163, http://dx.doi.org/10.4103/1673-5374.125344.
Hahn, T., Marquand, A.F., Ehlis, A.-C., Dresler, T., Kittel-Schneider, S., Jarczok, T.a., Lesch, K.-P., Jakob, P.M., Mourao-Miranda, J., Brammer, M.J., Fallgatter, A.J., 2011. Integrating neurobiological markers of depression. Arch. Gen. Psychiatry 68, 361–368, http://dx.doi.org/10.1001/archgenpsychiatry.2010.178.
Hahn, T., Marquand, A.F., Plichta, M.M., Ehlis, A.-C., Schecklmann, M.W., Dresler, T., Jarczok, T.a., Eirich, E., Leonhard, C., Reif, A., Lesch, K.-P., Brammer, M.J., Mourao-Miranda, J., Fallgatter, A.J., 2013. A novel approach to probabilistic biomarker-based classification using functional near-infrared spectroscopy. Hum. Brain Mapp. 34, 1102–1114, http://dx.doi.org/10.1002/hbm.21497.
Hart, H., Chantiluke, K., Cubillo, A.I., Smith, A.B., Simmons, A., Brammer, M.J., Marquand, A.F., Rubia, K., 2014a. Pattern classification of response inhibition in ADHD: toward the development of neurobiological markers for ADHD. Hum. Brain Mapp. 35, 3083–3094, http://dx.doi.org/10.1002/hbm.22386.
Hart, H., Marquand, A.F., Smith, A., Cubillo, A., Simmons, A., Brammer, M., Rubia, K., 2014b. Predictive neurofunctional markers of attention-deficit/hyperactivity disorder based on pattern classification of temporal processing. J. Am. Acad. Child Adolesc. Psychiatry 53, 569–578, http://dx.doi.org/10.1016/j.jaac.2013.12.024.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning, second ed. Springer, http://dx.doi.org/10.1007/b94608.
Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., Bießmann, F., 2014. On the interpretation of weight vectors of linear models in multivariate neuroimaging. NeuroImage 87, 96–110, http://dx.doi.org/10.1016/j.neuroimage.2013.10.067.
Haxby, J.V., 2012. Multivariate pattern analysis of fMRI: the early beginnings. NeuroImage 62, 852–855, http://dx.doi.org/10.1016/j.neuroimage.2012.03.016.
Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293, 2425–2430, http://dx.doi.org/10.1126/science.1063736.
Haynes, J.-D., Rees, G., 2006. Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. 7, 523–534, http://dx.doi.org/10.1038/nrn1931.
Hibar, D.P., Stein, J.L., Renteria, M.E., Arias-Vasquez, A., Desrivières, S., Jahanshad, N., Toro, R., Wittfeld, K., Abramovic, L., Andersson, M., Aribisala, B.S., Armstrong, N.J., Bernard, M., Bohlken, M.M., Boks, M.P., Bralten, J., Brown, A.A., Chakravarty, M.M., Chen, Q., Ching, C.R.K., Cuellar-Partida, G., den Braber, A., Giddaluru, S., Goldman, A.L., Grimm, O., Guadalupe, T., Hass, J., Woldehawariat, G., Holmes, A.J., Hoogman, M., Janowitz, D., Jia, T., Kim, S., Klein, M., Kraemer, B., Lee, P.H., Olde Loohuis, L.M., Luciano, M., Macare, C., Mather, K.A., Mattheisen, M., Milaneschi, Y., Nho, K., Papmeyer, M., Ramasamy, A., Risacher, S.L., Roiz-Santiañez, R., Rose, E.J., Salami, A., Sämann, P.G., Schmaal, L., Schork, A.J., Shin, J., Strike, L.T., Teumer, A., van Donkelaar, M.M.J., van Eijk, K.R., Walters, R.K., Westlye, L.T., Whelan, C.D., Winkler, A.M., Zwiers, M.P., Alhusaini, S., Athanasiu, L., Ehrlich, S., Hakobjan, M.M.H., Hartberg, C.B., Haukvik, U.K., Heister, A.J.G.A.M., Hoehn, D., Kasperaviciute, D., Liewald, D.C.M., Lopez, L.M., Makkinje, R.R.R., Matarin, M., Naber, M.A.M., Reese McKay, D., Needham, M., Nugent, A.C., Pütz, B., Royle, N.A., Shen, L., Sprooten, E., Trabzuni, D., van der Marel, S.S.L., van Hulzen, K.J.E., Walton, E., Wolf, C., Almasy, L., Ames, D., Arepalli, S., Assareh, A.A., Bastin, M.E., Brodaty, H., Bulayeva, K.B., Carless, M.A., Cichon, S., Corvin, A., Curran, J.E., Czisch, M., de Zubicaray, G.I., Dillman, A., Duggirala, R., Dyer, T.D., Erk, S., Fedko, I.O., Ferrucci, L., Foroud, T.M., Fox, P.T., Fukunaga, M., Gibbs, J.R., Göring, H.H.H., Green, R.C., Guelfi, S., Hansell, N.K., Hartman, C.A., Hegenscheid, K., Heinz, A., Hernandez, D.G., Heslenfeld, D.J., Hoekstra, P.J., Holsboer, F., Homuth, G., Hottenga, J.-J., Ikeda, M., Jack Jr, C.R., Jenkinson, M., Johnson, R., Kanai, R., Keil, M., Kent Jr, J.W., Kochunov, P., Kwok, J.B., Lawrie, S.M., Liu, X., Longo, D.L., McMahon, K.L., Meisenzahl, E., Melle, I., Mohnke, S., Montgomery, G.W., Mostert, J.C., Mühleisen, T.W., Nalls, M.A., Nichols, T.E., Nilsson, L.G., Nöthen, M.M., Ohi, K., Olvera, R.L., Perez-Iglesias, R., Pike, G.B., Potkin, S.G., Reinvang, I., Reppermund, S., Rietschel, M., Romanczuk-Seiferth, N., Rosen, G.D., Rujescu, D., Schnell, K., Schofield, P.R., Smith, C., Steen, V.M., Sussmann, J.E., Thalamuthu, A., Toga, A.W., Traynor, B.J., Troncoso, J., Turner, J.A., Valdés Hernández, M.C., van 't Ent, D., van der Brug, M., van der Wee, N.J.A., van Tol, M.-J., Veltman, D.J., Wassink, T.H., Westman, E., Zielke, R.H., Zonderman, A.B., Ashbrook, D.G., Hager, R., Lu, L., McMahon, F.J., Morris, D.W., Williams, R.W., Brunner, H.G., Buckner, R.L., Buitelaar, J.K., Cahn, W., Calhoun, V.D., Cavalleri, G.L., Crespo-Facorro, B., Dale, A.M., Davies, G.E., Delanty, N., Depondt, C., Djurovic, S., Drevets, W.C., Espeseth, T., Gollub, R.L., Ho, B.-C., Hoffmann, W., Hosten, N., Kahn, R.S., Le Hellard, S., Meyer-Lindenberg, A., Müller-Myhsok, B., Nauck, M., Nyberg, L., Pandolfo, M., Penninx, B.W.J.H., Roffman, J.L., Sisodiya, S.M., Smoller, J.W., van Bokhoven, H., van Haren, N.E.M., Völzke, H., Walter, H., Weiner, M.W., Wen, W., White, T., Agartz, I., Andreassen, OA., Blangero, J., Boomsma, D.I., Brouwer, R.M., Cannon, D.M., Cookson, M.R., de Geus, E.J.C., Deary, I.J., Donohoe, G., Fernández, G., Fisher, S.E., Francks, C., Glahn, D.C., Grabe, H.J., Gruber, O., Hardy, J., Hashimoto, R., Hulshoff Pol, H.E., Jönsson, E.G., Kloszewska, I., Lovestone, S., Mattay, V.S., Mecocci, P., McDonald, C., McIntosh, A.M., Ophoff, R.A., Paus, T., Pausova, Z., Ryten, M., Sachdev, P.S., Saykin, A.J., Simmons, A., Singleton, A., Soininen, H., Wardlaw, J.M., Weale, M.E., Weinberger, D.R., Adams, H.H.H., Launer, L.J., Seiler, S., Schmidt, R., Chauhan, G., Satizabal, C.L., Becker, J.T., Yanek, L., van der Lee, S.J., Ebling, M., Fischl, B., Longstreth Jr, W.T., Greve, D., Schmidt, H., Nyquist, P., Vinke, L.N., van Duijn, C.M., Xue, L., Mazoyer, B., Bis, J.C., Gudnason, V., Seshadri, S., Ikram, M.A., The Alzheimer's Disease Neuroimaging Initiative, The CHARGE

Consortium, EPIGEN, IMAGEN, SYS, Martin, N.G., Wright, M.J., Schumann, G., Franke, B., Thompson, P.M., Medland, S.E., 2015. Common genetic variants influence human subcortical brain structures. Nature 520 (7546), 224–229.
Hinrichs, C., Singh, V., Xu, G., Johnson, S.C., 2011. Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population. NeuroImage 55, 574–589, http://dx.doi.org/10.1016/j.neuroimage.2010.10.081.
Honorio, J., Tomasi, D., Goldstein, R.Z., Leung, H.-C., Samaras, D., 2012. Can a single brain region predict a disorder? IEEE Trans. Med. Imaging 31, 2062–2072, http://dx.doi.org/10.1109/TMI. 2012.2206047.
Hu, X., Zhu, D., Lv, P., Li, K., Han, J., Wang, L., Shen, D., Guo, L., Liu, T., 2013. Fine-granularity functional interaction signatures for characterization of brain conditions. Neuroinformatics 11, 301–317, http://dx.doi.org/10.1007/s12021-013-9177-2.
Iannaccone, R., Hauser, T.U., Ball, J., Brandeis, D., Walitza, S., Brem, S., 2015. Classifying adolescent attention-deficit/hyperactivity disorder (ADHD) based on functional and structural imaging. Eur. Child Adolesc. Psychiatry, http://dx.doi.org/10.1007/s00787-015-0678-4.
Igual, L., Soliva, J.C., Escalera, S., Gimeno, R., Vilarroya, O., Radeva, P., 2012. Automatic brain caudate nuclei segmentation and classification in diagnostic of attention-deficit/hyperactivity disorder. Comput. Med. Imaging Graph. 36, 591–600, http://dx.doi.org/10.1016/j.compmedimag.2012.08.002.
Iidaka, T., 2014. Resting state functional magnetic resonance imaging and neural network classified autism and control. Cortex 63, 55–67, http://dx.doi.org/10.1016/j.cortex.2014.08.011.
Ingalhalikar, M., Kanterakis, S., Gur, R., Roberts, T.P.L., Verma, R., 2010. DTI based diagnostic prediction of a disease via pattern classification. In: Jiang, T. (Ed.), Medical Image Computing and Computer-Assisted Intervention Part 1. Springer, Berlin, pp. 558–565.
Ingalhalikar, M., Parker, D., Bloy, L., Roberts, T.P.L., Verma, R., 2011. Diffusion based abnormality markers of pathology: toward learned diagnostic prediction of ASD. NeuroImage 57, 918–927, http://dx.doi.org/10.1016/j.neuroimage.2011.05.023.
Ingalhalikar, M., Parker, W.a., Bloy, L., Roberts, T.P.L., Verma, R., 2014. Creating multimodal predictors using missing data: classifying and subtyping autism spectrum disorder. J. Neurosci. Methods 235, 1–9, http://dx.doi.org/10.1016/j.jneumeth.2014.06.030.
Iwabuchi, S.J., Liddle, P.F., Palaniyappan, L., 2013. Clinical utility of machine-learning approaches in schizophrenia: improving diagnostic confidence for translational neuroimaging. Front. Psychiatry 4, 1–9, http://dx.doi.org/10.3389/fpsyt.2013.00095.
Janousova, E., Schwarz, D., Kasparek, T., 2015. Combining various types of classifiers and features extracted from magnetic resonance imaging data in schizophrenia recognition. Psychiatry Res. Neuroimaging, 1–13, http://dx.doi.org/10.1016/j.pscychresns.2015.03.004.
Jiao, Y., Chen, R., Ke, X., Chu, K., Lu, Z., Herskovits, E.H., 2010. Predictive models of autism spectrum disorder based on brain regional cortical thickness. NeuroImage 50, 589–599, http://dx.doi.org/10.1016/j.neuroimage.2009.12.047.
Johnston, B.A., Mwangi, B., Matthews, K., Coghill, D., Konrad, K., Steele, J.D., 2014. Brainstem abnormalities in attention deficit hyperactivity disorder support high accuracy individual diagnostic classification. Hum. Brain Mapp. 35, 5179–5189, http://dx.doi.org/10.1002/hbm.22542.
Karageorgiou, E., Schulz, S.C., Gollub, R.L., Andreasen, N.C., Ho, B.C., Lauriello, J., Calhoun, V.D., Bockholt, H.J., Sponheim, S.R., Georgopoulos, A.P., 2011. Neuropsychological testing and structural magnetic resonance imaging as diagnostic biomarkers early in the course of schizophrenia and related psychoses. Neuroinformatics 9, 321–333, http://dx.doi.org/10.1007/s12021-010-9094-6.
Kasparek, T., Thomaz, C.E., Sato, J.R., Schwarz, D., Janousova, E., Marecek, R., Prikryl, R., Vanicek, J., Fujita, A., Ceskova, E., 2011. Maximum-uncertainty linear discrimination analysis of first-episode schizophrenia subjects. Psychiatry Res. Neuroimaging 191, 174–181, http://dx.doi.org/10.1016/j.pscychresns.2010.09.016.
Kawasaki, Y., Suzuki, M., Kherif, F., Takahashi, T., Zhou, S.Y., Nakamura, K., Matsui, M., Sumiyoshi, T., Seto, H., Kurachi, M., 2007. Multivariate voxel-based morphometry successfully differentiates schizophrenia patients from healthy controls. NeuroImage 34, 235–242, http://dx.doi.org/10.1016/j.neuroimage.2006.08.018.
Klöppel, S., Abdulkadir, A., Jack Jr., C.R., Koutsouleris, N., Mourão-Miranda, J., Vemuri, P., 2012. Diagnostic neuroimaging across diseases. NeuroImage 61, 457–463, http://dx.doi.org/10.1016/j.neuroimage.2011.11.002.
Klöppel, S., Stonnington, C.M., Barnes, J., Chen, F., Chu, C., Good, C.D., Mader, I., Mitchell, L.A., Patel, A.C., Roberts, C.C., Fox, N.C., Jack, C.R., Ashburner, J., Frackowiak, R.S.J., 2008. Accuracy of dementia diagnosis—a direct comparison between radiologists and a computerized method. Brain 131, 2969–2974, http://dx.doi.org/10.1093/brain/awn239.
Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox, N.C., Ashburner, J., Frackowiak, R.S.J., 2009. A plea for confidence intervals and consideration of generalizability in diagnostic studies. Brain 132, 2008–2010, http://dx.doi.org/10.1093/brain/awn090.
Koch, S.P., Hägele, C., Haynes, J.-D., Heinz, A., Schlagenhauf, F., Sterzer, P., 2015. Diagnostic classification of Schizophrenia patients on the basis of regional reward-related fMRI signal patterns. PLoS ONE 10, e0119089, http://dx.doi.org/10.1371/journal.pone.0119089.
Korgaonkar, M.S., Cooper, N.J., Williams, L.M., Grieve, S.M., 2012. Mapping inter-regional connectivity of the entire cortex to characterize major depressive disorder: a whole-brain diffusion tensor imaging tractography study. NeuroReport 23, 566–571, http://dx.doi.org/10.1097/WNR.0b013e3283546264.
Koutsouleris, N., Meisenzahl, E.M., Borgwardt, S., Riecher-Rossler, A., Frodl, T., Kambeitz, J., Kohler, Y., Falkai, P., Moller, H.-J., Reiser, M., Davatzikos, C., 2015. Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. Brain, 1–15, http://dx.doi.org/10.1093/brain/awv111.
Koutsouleris, N., Meisenzahl, E.M., Davatzikos, C., Bottlender, R., Frodl, T., Scheuerecker, J., Schmitt, G., Zetzsche, T., Decker, P., Reiser, M., Möller, H.-J., Gaser, C., 2009. Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. Arch. Gen. Psychiatry 66, 700–712, http://dx.doi.org/10.1001/archgenpsychiatry.2009.62.
Lautrup, B., Hansen, L.K., Law, I., Mørch, N., Svarer, C., Strother, S.C., 1995. Massive weight sharing: a cure for extremely ill-posed problems. In: Herman, H.J., et al. (Eds.), Supercomputing in Brain Research: From Tomography to Neural Networks. World Scientific Pub. Corp., pp. 137–148.
Li, F., Huang, X., Tang, W., Yang, Y., Li, B., Kemp, G.J., Mechelli, A., Gong, Q., 2014. Multivariate pattern analysis of DTI reveals differential white matter in individuals with obsessive–compulsive disorder. Hum. Brain Mapp. 35, 2643–2651, http://dx.doi.org/10.1002/hbm.22357.
Lim, L., Marquand, A., Cubillo, A.a., Smith, A.B., Chantiluke, K., Simmons, A., Mehta, M., Rubia, K., 2013. Disorder-specific predictive classification of adolescents with attention deficit hyperactivity disorder (ADHD) relative to autism using structural magnetic resonance imaging. PLoS ONE 8, e63660, http://dx.doi.org/10.1371/journal.pone.0063660.
Liu, F., Guo, W., Fouche, J.-P., Wang, Y., Wang, W., Ding, J., Zeng, L., Qiu, C., Gong, Q., Zhang, W., Chen, H., 2013. Multivariate classification of social anxiety disorder using whole brain functional connectivity. Brain Struct. Funct., http://dx.doi.org/10.1007/s00429-013-0641-4.
Liu, M., Zeng, L.-L., Shen, H., Liu, Z., Hu, D., 2012. Potential risk for healthy siblings to develop schizophrenia: evidence from pattern classification with whole-brain connectivity. NeuroReport 23, 265–269, http://dx.doi.org/10.1097/WNR.0b013e32834f60a5.
Lord, A., Horn, D., Breakspear, M., Walter, M., 2012. Changes in community structure of resting state functional connectivity in unipolar depression. PLoS ONE 7, http://dx.doi.org/10.1371/journal.pone.0041282.
Lord, C., Rutter, M., Couteur, a.L., 1994. Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J. Autism Dev. Disord. 24, 659–685, http://dx.doi.org/10.1007/BF02172145.
Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., Schopler, E., 1989. Austism diagnostic observation schedule: a standardized observation of communicative and social behavior. J. Autism Dev. Disord. 19, 185–212, http://dx.doi.org/10.1007/BF02211841.
Lueken, U., Hilbert, K., Wittchen, H.-U., Reif, A., Hahn, T., 2014. Diagnostic classification of specific phobia subtypes using structural MRI data: a machine-learning approach. J. Neural Transm., http://dx.doi.org/10.1007/s00702-014-1272-5.
Ma, Q., Zeng, L.-L., Shen, H., Liu, L., Hu, D., 2013. Altered cerebellar-cerebral resting-state functional connectivity reliably identifies major depressive disorder. Brain Res., 86–94, http://dx.doi.org/10.1016/j.brainres.2012.12.002.
Macdonald, M.E., Ambrose, C.M., Duyao, M.P., Myers, R.H., Lin, C., Srinidhi, L., Barnes, G., Taylor, S.A., James, M., Groot, N., Macfarlane, H., Jenkins, B., Anderson, M.A., Wexler, N.S., Gusella, J.F., Bates, G.P., Baxendale, S., Hummerich, H., Kirby, S., North, M., Youngman, S., Mott, R., Zehetner, G., Sedlacek, Z., Poustka, A., Frischauf, A.M., Lehrach, H., Buckler, A.J., Church, D., Doucettestamm, L., Odonovan, M.C., Ribaramirez, L., Shah, M., Stanton, V.P., Strobel, S.A., Draths, K.M., Wales, J.L., Dervan, P., Housman, D.E., Altherr, M., Shiang, R., Thompson, L., Fielder, T., Wasmuth, J.J., Tagle, D., Valdes, J., Elmer, L., Allard, M., Castilla, L., Swaroop, M., Blanchard, K., Collins, F.S., Snell, R., Holloway, T., Gillespie, K., Datson, N., Shaw, D., Harper, P.S., 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntingtons-disease chromosomes. Cell 72, 971–983, http://dx.doi.org/10.1016/0092-8674(93)90585-E.
Marquand, A.F., Brammer, M., Williams, S.C.R., Doyle, O.M., 2014. Bayesian multi-task learning for decoding multi-subject neuroimaging data. NeuroImage 92, 298–311, http://dx.doi.org/10.1016/j.neuroimage.2014.02.008.
Marquand, A.F., Mourão-Miranda, J., Brammer, M.J., Cleare, A.J., Fu, C.H.Y., 2008. Neuroanatomy of verbal working memory as a diagnostic biomarker for depression. NeuroReport 19, 1507–1511, http://dx.doi.org/10.1097/WNR.0b013e328310425e.
Modinos, G., Mechelli, A., Pettersson-Yeo, W., Allen, P., McGuire, P., Aleman, A., 2013. Pattern classification of brain activation during emotional processing in subclinical depression: psychosis proneness as potential confounding factor. Peer J., http://dx.doi.org/10.7717/peerj.42.
Morch, N., Hansen, L.K., Strother, S.C., Svarer, C., Rottenberg, D.A., Lautrup, B., Savoy, R., Paulson, O.B., 1997. Nonlinear versus linear models in functional neuroimaging: learning curves and generalization crossover. In: Proceedings of the 15th International Conference on Information Processing in Medical Imaging.

Mourão-Miranda, J., Almeida, J.R.C., Hassel, S., de Oliveira, L., Versace, A., Marquand, A.F., Sato, J.R., Brammer, M., Phillips, M.L., 2012. Pattern recognition analyses of brain activation elicited by happy and neutral faces in unipolar and bipolar depression. Bipolar Disord. 14, 451–460, http://dx.doi.org/10.1111/j.1399-5618.2012.01019.x.

Mourão-Miranda, J., Bokde, a., Born, L.W., Hampel, C., Stetter, H.M., 2005. Classifying brain states and determining the discriminating activation patterns: support vector machine on functional MRI data. NeuroImage 28, 980–995, http://dx.doi.org/10.1016/j.neuroimage.2005.06.070.

Mueller, S., Wang, D., Pan, R., Holt, D.J., Liu, H., 2015. Abnormalities in hemispheric specialization of caudate nucleus connectivity in schizophrenia. JAMA Psychiatry, http://dx.doi.org/10.1001/jamapsychiatry.2014.3176.

Murdaugh, D.L., Shinkareva, S.V., Deshpande, H.R., Wang, J., Pennick, M.R., Kana, R.K., 2012. Differential deactivation during mentalizing and classification of autism based on default mode network connectivity. PLoS ONE 7, e50064, http://dx.doi.org/10.1371/journal.pone.0050064.

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, 1508–1521, http://dx.doi.org/10.1093/brain/aws084.

Mwangi, B., Tian, T.S., Soares, J.C., 2014. A review of feature reduction techniques in neuroimaging. Neuroinformatics 12, 229–244, http://dx.doi.org/10.1007/s12021-013-9204-3.

Nielsen, J.a., Zielinski, B.a., Fletcher, P.T., Alexander, A.L., Lange, N., Bigler, E.D., Lainhart, J.E., Anderson, J.S., 2013. Multisite functional connectivity MRI classification of autism: ABIDE results. Front. Hum. Neurosci. 7, 1–12, http://dx.doi.org/10.3389/fnhum.2013.00599.

Nieuwenhuis, M., van Haren, N.E.M., Hulshoff Pol, H.E., Cahn, W., Kahn, R.S., Schnack, H.G., 2012. Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. NeuroImage 61, 606–612, http://dx.doi.org/10.1016/j.neuroimage.2012.03.079.

Nouretdinov, I., Costafreda, S.G., Gammerman, A., Chervonenkis, A., Vovk, V., Vapnik, V., Fu, C.H.Y., 2011. Machine learning classification with confidence: application of transductive conformal predictors to MRI-based diagnostic and prognostic markers in depression. NeuroImage 56, 809–813, http://dx.doi.org/10.1016/j.neuroimage.2010.05.023.

Orrù, G., Pettersson-Yeo, W., Marquand, A.F., Sartori, G., Mechelli, A., 2012. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. Neurosci. Biobehav. Rev. 36, 1140–1152, http://dx.doi.org/10.1016/j.neubiorev.2012.01.004.

Ota, M., Ishikawa, M., Sato, N., Hori, H., Sasayama, D., Hattori, K., Teraishi, T., Noda, T., Obu, S., Nakata, Y., Higuchi, T., Kunugi, H., 2013. Discrimination between schizophrenia and major depressive disorder by magnetic resonance imaging of the female brain. J. Psychiatr. Res. 47, 1383–1388, http://dx.doi.org/10.1016/j.jpsychires.2013.06.010.

Pan, S.J., Yang, Q., 2010. A survey on transfer learning. IEEE Trans. Knowl. Data Eng. 22, 1345–1357.

Peng, X., Lin, P., Zhang, T., Wang, J., 2013. Extreme learning machine-based classification of ADHD using brain structural MRI data. PLoS ONE 8, http://dx.doi.org/10.1371/journal.pone.0079476.

Pettersson-Yeo, W., Benetti, S., Marquand, a.F., Dell'acqua, F., Williams, S.C.R., Allen, P., Prata, D., McGuire, P., Mechelli, A., 2013. Using genetic, cognitive and multi-modal neuroimaging data to identify ultra-high-risk and first-episode psychosis at the individual level. Psychol. Med. 43, 2547–2562, http://dx.doi.org/10.1017/S003329171300024X.

Pohl, K.M., Sabuncu, M.R., 2009. A unified framework for MR based disease classification. Inf. Process. Med. Imaging 21, 300–313.

Qin, J., Wei, M., Liu, H., Chen, J., Yan, R., Hua, L., Zhao, K., Yao, Z., Lu, Q., 2014. Abnormal hubs of white matter networks in the frontal-parieto circuit contribute to depression discrimination via pattern classification. Magn. Reson. Imaging 32, 1314–1320, http://dx.doi.org/10.1016/j.mri.2014.08.037.

Qin, J., Wei, M., Liu, H., Chen, J., Yan, R., Yao, Z., Lu, Q., 2015. Altered anatomical patterns of depression in relation to antidepressant treatment: evidence from a pattern recognition analysis on the topological organization of brain networks. J. Affect. Disord. 180, 129–137, http://dx.doi.org/10.1016/j.jad.2015.03.059.

Rathi, Y., Malcolm, J., Michailovich, O., Goldstein, J., Seidman, L., McCarley, R.W., Westin, C.-F., Shenton, M.E., 2010. Biomarkers for identifying first-episode schizophrenia patients using diffusion weighted imaging. In: Jiang, T. (Ed.), Medical Image Computing and Computer-Assisted Intervention Part 1. Springer, Berlin, pp. 657–665, http://dx.doi.org/10.1007/978-3-642-15705-9_80.

Redlich, R., Almeida, J.J.R., Grotegerd, D., Opel, N., Kugel, H., Heindel, W., Arolt, V., Phillips, M.L., Dannlowski, U., 2014. Brain morphometric biomarkers distinguishing unipolar and bipolar depression: a voxel-based morphometry-pattern classification approach. JAMA Psychiatry 71, 1222–1230, http://dx.doi.org/10.1001/jamapsychiatry.2014.1100.

Rocha-Rego, V., Jogia, J., Marquand, a.F., Mourao-Miranda, J., Simmons, A., Frangou, S., 2014. Examination of the predictive value of structural magnetic resonance scans in bipolar disorder: a pattern classification approach. Psychol. Med. 44, 519–532, http://dx.doi.org/10.1017/S0033291713001013.

Rosa, M.J., Portugal, L., Hahn, T., Fallgatter, A.J., Garrido, M.I., Shawe-Taylor, J., Mourao-Miranda, J., 2015. Sparse network-based models for patient classification using fMRI. NeuroImage 105, 493–506, http://dx.doi.org/10.1016/j.neuroimage.2014.11.021.

Sabuncu, M.R., Konukoglu, E., 2014. Clinical prediction from structural brain MRI scans: a large-scale empirical study. Neuroinformatics 13, 31–46, http://dx.doi.org/10.1007/s12021-014-9238-1.

Schmaal, L., Marquand, A.F., Rhebergen, D., van Tol, M.-J., Henricus, G.R., van der Wee, N.J.a., Veltman, D.J., Penninx, B.W.J.H., 2014. Predicting the naturalistic course of major depressive disorder using clinical and multimodal neuroimaging information: a multivariate pattern recognition study. Biol. Psychiatry 1, 1–9, http://dx.doi.org/10.1016/j.biopsych.2014.11.018.

Schmaal, L., Veltman, D., Erp, T., Sämann, P., Frodl, T., Jahanshad, N., Loehrer, E., Tiemeier, H., Hofman, A., Niessen, W., Vernooij, M., Ikram, M., Wittfeld, K., Grabe, H., Block, A., Hegenscheid, K., Völzke, H., Hoehn, D., Czisch, M., Lagopoulos, J., Hatton, S., Hickie, I., Goya-Maldonado, R., Krämer, B., Gruber, O., Couvy-Duchesne, B., Rentería, M., Strike, L., Mills, N., Zubicaray, G., McMahon, K., Medland, S., Martin, N., Gillespie, N., Wright, M., Hall, G., MacQueen, G., Frey, E., Carballedo, A., Velzen, L., Tol, M., Wee, N., Veer, I., Walter, H., Schnell, K., Schramm, E., Normann, C., Schoepf, D., Konrad, C., Zurowski, B., Nickson, T., McIntosh, A., Papmeyer, M., Whalley, H., Sussmann, J., Godlewska, B., Cowen, P., Fischer, F., Rose, M., Penninx, B., Thompson, P., Hibar, D., 2015. Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. Mol. Psychiatry (in press).

Schnack, H.G., Nieuwenhuis, M., van Haren, N.E.M., 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, http://dx.doi.org/10.1016/j.neuroimage.2013.08.053.

Segovia, F., Holt, R., Spencer, M., Górriz, J.M., Ramírez, J., Puntonet, C.G., Phillips, C., Chura, L., Baron-Cohen, S., Suckling, J., 2014. Identifying endophenotypes of autism: a multivariate approach. Front. Comput. Neurosci. 8, 1–8, http://dx.doi.org/10.3389/fncom.2014.00060.

Serpa, M.H., Ou, Y., Schaufelberger, M.S., Doshi, J., Ferreira, L.K., Machado-Vieira, R., Menezes, P.R., Scazufca, M., Davatzikos, C., Busatto, G.F., Zanetti, M.V., 2014. Neuroanatomical classification in a population-based sample of psychotic major depression and bipolar I disorder with 1 year of diagnostic stability. Biomed. Res. Int., http://dx.doi.org/10.1155/2014/706157.

Shen, H., Wang, L., Liu, Y., Hu, D., 2010. Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI. NeuroImage 49, 3110–3121, http://dx.doi.org/10.1016/j.neuroimage.2009.11.011.

Shenas, S.K., Halici, U., Metehan, C., 2014. A comparative analysis of functional connectivity data in resting and task-related conditions of the brain for disease signature of OCD. In: Conf Proc IEEE Eng Med Biol Soc, pp. 978–981.

Shi, F., Liu, Y., Jiang, T., Zhou, Y., Zhu, W., Jiang, J., Liu, H., Liu, Z., 2007. Regional homogeneity and anatomical parcellation for fMRI image classification: application to schizophrenia and normal controls. Med. Image Comput. Comput. Assist. Interv. 10, 136–143.

Soriano-Mas, C., Pujol, J., Alonso, P., Cardoner, N., Menchón, J.M., Harrison, B.J., Deus, J., Vallejo, J., Gaser, C., 2007. Identifying patients with obsessive–compulsive disorder using whole-brain anatomy. NeuroImage 35, 1028–1037, http://dx.doi.org/10.1016/j.neuroimage.2007.01.011.

Stelzer, J., Chen, Y., Turner, R., 2013. Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. NeuroImage 65, 69–82, http://dx.doi.org/10.1016/j.neuroimage.2012.09.063.

Su, L., Wang, L., Shen, H., Feng, G., Hu, D., 2013. Discriminative analysis of non-linear brain connectivity in schizophrenia: an fMRI Study. Front. Hum. Neurosci. 7, 702, http://dx.doi.org/10.3389/fnhum.2013.00702.

Sui, J., He, H., Yu, Q., Chen, J., Rogers, J., Pearlson, G.D., Mayer, A., Bustillo, J., Canive, J., Calhoun, V.D., 2013. Combination of resting state fMRI, DTI, and sMRI data to discriminate schizophrenia by N-way MCCA + jICA. Front. Hum. Neurosci. 7, 1–14, http://dx.doi.org/10.3389/fnhum.2013.00235.

Sun, D., Erp, T.G.M.V., Thompson, P.M., Bearden, C.E., Daley, M., Hardt, M.E., Nuechterlein, K.H., Toga, A.W., Cannon, T.D., 2009. Elucidating a magnetic resonance imaging-based neuroanatomic biomarker for psychosis: classification analysis using probabilistic brain atlas and machine learning algorithms. Biol. Psychiatry 66, 1055–1060, http://dx.doi.org/10.1016/j.biopsych.2009.07.019.Elucidating.

Takayanagi, Y., Kawasaki, Y., Nakamura, K., Takahashi, T., Orikabe, L., Toyoda, E., Mozue, Y., Sato, Y., Itokawa, M., Yamasue, H., Kasai, K., Kurachi, M., Okazaki, Y., Matsushita, M., Suzuki, M., 2010. Differentiation of first-episode schizophrenia patients from healthy controls using ROI-based multiple structural brain variables. Prog. Neuropsychopharmacol. Biol. Psychiatry 34, 10–17, http://dx.doi.org/10.1016/j.pnpbp.2009.09.004.

Takayanagi, Y., Takahashi, T., Orikabe, L., Mozue, Y., Kawasaki, Y., Nakamura, K., Sato, Y., Itokawa, M., Yamasue, H., Kasai, K., Kurachi, M., Okazaki, Y., Suzuki, M., 2011. Classification of first-episode schizophrenia patients and healthy subjects by automated MRI measures of regional brain volume and cortical thickness. PLoS ONE 6, http://dx.doi.org/10.1371/journal.pone.0021047.

Tang, Y., Wang, L., Cao, F., Tan, L., 2012. Identify schizophrenia using resting-state functional connectivity: an exploratory research and analysis. Biomed. Eng. Online 11, 1–16, http://dx.doi.org/10.1186/1475-925X-11-50.

The ADHD Consortium, 2012. The ADHD-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. Front. Syst. Neurosci. 6, http://dx.doi.org/10.3389/fnsys.2012.00062.

Uddin, L.Q., Menon, V., Young, C.B., Ryali, S., Chen, T., Khouzam, A., Minshew, N.J., Hardan, A.Y., 2011. Multivariate searchlight classification of structural

magnetic resonance imaging in children and adolescents with autism. Biol. Psychiatry 70, 833–841, http://dx.doi.org/10.1016/j.biopsych.2011.07.014.
Uddin, L.Q., Supekar, K., Lynch, C.J., Khouzam, A., Phillips, J., Feinstein, C., Ryali, S., Menon, V., 2013. Salience network-based classification and prediction of symptom severity in children with autism. JAMA Psychiatry 70, 869–879, http://dx.doi.org/10.1001/jamapsychiatry.2013.104.
van Ewijk, H., Heslenfeld, D.J., Zwiers, M.P., Buitelaar, J.K., Oosterlaan, J., 2012. Diffusion tensor imaging in attention deficit/hyperactivity disorder: a systematic review and meta-analysis. Neurosci. Biobehav. Rev. 36, 1093–1106, http://dx.doi.org/10.1016/j.neubiorev.2012.01.003.
Venkataraman, A., Whitford, T.J., Westin, C.-F., Golland, P., Kubicki, M., 2012. Whole brain resting state functional connectivity abnormalities in schizophrenia. Schizophr. Res. 139, 7–12, http://dx.doi.org/10.1016/j.schres.2012.04.021.
Wang, X., Jiao, Y., Lu, Z.,2011. Discriminative analysis of resting-state brain functional connectivity patterns of attention-deficit hyperactivity disorder using kernel principal component analysis. In: Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD). IEEE, pp. 1938–1941, http://dx.doi.org/10.1109/FSKD.2011.6019911.
Wang, X., Jiao, Y., Tang, T., Wang, H., Lu, Z., 2013. Altered regional homogeneity patterns in adults with attention-deficit hyperactivity disorder. Eur. J. Radiol. 82, 1552–1557, http://dx.doi.org/10.1016/j.ejrad.2013.04.009.
Wee, C.-Y., Wang, L., Shi, F., Yap, P.-T., Shen, D., 2014. Diagnosis of autism spectrum disorders using regional and interregional morphological features. Hum. Brain Mapp. 35, 3414–3430, http://dx.doi.org/10.1002/hbm.22411.
Wei, M., Qin, J., Yan, R., Li, H., Yao, Z., Lu, Q., 2013. Identifying major depressive disorder using Hurst exponent of resting-state brain networks. Psychiatry Res. 214, 306–312, http://dx.doi.org/10.1016/j.pscychresns.2013.09.008.
Weygandt, M., Blecker, C.R., Schäfer, A., Hackmack, K., Haynes, J.-D., Vaitl, D., Stark, R., Schienle, A., 2012a. fMRI pattern recognition in obsessive–compulsive disorder. NeuroImage 60, 1186–1193, http://dx.doi.org/10.1016/j.neuroimage.2012.01.064.
Weygandt, M., Schaefer, A., Schienle, A., Haynes, J.D., 2012b. Diagnosing different binge-eating disorders based on reward-related brain activation patterns. Hum. Brain Mapp. 33, 2135–2146, http://dx.doi.org/10.1002/hbm.21345.
Yang, H., Liu, J., Sui, J., Pearlson, G., Calhoun, V.D., 2010. A hybrid machine learning method for fusing fMRI and genetic data: combining both improves classification of schizophrenia. Front. Hum. Neurosci. 4, http://dx.doi.org/10.3389/fnhum.2010.00192.
Yoon, J.H., Nguyen, D.V., McVay, L.M., Deramo, P., Minzenberg, M.J., Ragland, J.D., Niendham, T., Solomon, M., Carter, C.S., 2012. Automated classification of fMRI during cognitive control identifies more severely disorganized subjects with schizophrenia. Schizophr. Res. 135, 28–33, http://dx.doi.org/10.1016/j.schres.2012.01.001.
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, http://dx.doi.org/10.1371/journal.pone.0068250.
Yushkevich, P., Dubb, A., Xie, Z., Gur, R., Gur, R., Gee, J., 2005. Regional structural characterization of the brain of schizophrenia patients. Acad. Radiol. 12, 1250–1261, http://dx.doi.org/10.1016/j.acra.2005.06.014.
Zanetti, M.V., Schaufelberger, M.S., Doshi, J., Ou, Y., Ferreira, L.K., Menezes, P.R., Scazufca, M., Davatzikos, C., Busatto, G.F., 2013. Neuroanatomical pattern classification in a population-based sample of first-episode schizophrenia. Prog. Neuropsychopharmacol. Biol. Psychiatry 43, 116–125, http://dx.doi.org/10.1016/j.pnpbp.2012.12.005.
Zeng, L.-L., Shen, H., Liu, L., Wang, L., Li, B., Fang, P., Zhou, Z., Li, Y., Hu, D., 2012. Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. Brain 135, 1498–1507, http://dx.doi.org/10.1093/brain/aws059.
Zhang, D., Sen, D., 2013. Multi modal multi task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. NeuroImage 59, 895–907, http://dx.doi.org/10.1016/j.neuroimage.2011.09.069.Multi-Modal.
Zhang, L.Z.L., Samaras, D., Tomasi, D., Volkow, N., Goldstein, R., 2005. Machine learning for clinical diagnosis from functional magnetic resonance imaging. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), http://dx.doi.org/10.1109/CVPR.2005.219, 1.
Zhang, M.L., Zhou, Z.H., 2014. A review on multi-label learning algorithms. IEEE Trans. Knowl. Data Eng. 26, 1819–1837, http://dx.doi.org/10.1109/TKDE. 2013.39.
Zhang, Y., Tian, J., Yuan, K., Liu, P., Zhuo, L., Qin, W., Zhao, L., Liu, J., Von Deneen, K.M., Klahr, N.J., Gold, M.S., Liu, Y., 2011. Distinct resting-state brain activities in heroin-dependent individuals. Brain Res. 1402, 46–53, http://dx.doi.org/10.1016/j.brainres.2011.05.054.
Zhou, J., Liu, J., Narayan, V.a., Ye, J., 2013. Modeling disease progression via multi-task learning. NeuroImage 78, 233–248, http://dx.doi.org/10.1016/j.neuroimage.2013.03.073.
Zhou, Y., Yu, F., Duong, T., 2014. Multiparametric MRI characterization and prediction in autism spectrum disorder using graph theory and machine learning. PLoS ONE 9, http://dx.doi.org/10.1371/journal.pone.0090405.
Zhu, C.-Z., Zang, Y.-F., Cao, Q.-J., Yan, C.-G., He, Y., Jiang, T.-Z., Sui, M.-Q., Wang, Y.-F., 2008. Fisher discriminative analysis of resting-state brain function for attention-deficit/hyperactivity disorder. NeuroImage 40, 110–120, http://dx.doi.org/10.1016/j.neuroimage.2007.11.029.
Zhu, D., Shen, D., Jiang, X., Liu, T., 2014. Connectomics signature for characterization of mild cognitive impairment and schizophrenia. In: Proc IEEE IntSymp Biomed Imaging, pp. 325–328, http://dx.doi.org/10.1109/ISBI.2014.6867874.CONNECTOMICS.