Neuroscience and Biobehavioral Reviews 36 (2012) 1140–1152

Contents lists available at SciVerse ScienceDirect

Neuroscience and Biobehavioral Reviews

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

Review

Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review

Graziella Orrù[a,1], William Pettersson-Yeo[a,1], Andre F. Marquand[b], Giuseppe Sartori[c], Andrea Mechelli[a,*]

[a] Department of Psychosis Studies, Institute of Psychiatry, King's College London, De Crespigny Park, PO Box 67, London SE5 8AF, UK
[b] Department of Neuroimaging, Centre for Neuroimaging Sciences, Institute of Psychiatry, King's College London, De Crespigny Park, London SE5 8AF, UK
[c] Department of Psychology, University of Padua, Via Venezia, 8, 35100 Padova, Italy

ARTICLE INFO

Article history:
Received 25 July 2011
Received in revised form 16 January 2012
Accepted 20 January 2012

Keywords:
Machine learning
Support Vector Machine
Neuroimaging
Diagnosis
Prediction
Alzheimer's disease
Schizophrenia
Major depression
Bipolar disorder
Autistic spectrum disorder

ABSTRACT

Standard univariate analysis of neuroimaging data has revealed a host of neuroanatomical and functional differences between healthy individuals and patients suffering a wide range of neurological and psychiatric disorders. Significant only at group level however these findings have had limited clinical translation, and recent attention has turned toward alternative forms of analysis, including Support-Vector-Machine (SVM). A type of machine learning, SVM allows categorisation of an individual's previously unseen data into a predefined group using a classification algorithm, developed on a training data set. In recent years, SVM has been successfully applied in the context of disease diagnosis, transition prediction and treatment prognosis, using both structural and functional neuroimaging data. Here we provide a brief overview of the method and review those studies that applied it to the investigation of Alzheimer's disease, schizophrenia, major depression, bipolar disorder, presymptomatic Huntington's disease, Parkinson's disease and autistic spectrum disorder. We conclude by discussing the main theoretical and practical challenges associated with the implementation of this method into the clinic and possible future directions.

Contents

1. Introduction

In the past 20 years the development of Position Emission Tomography (PET) and Magnetic Resonance Imaging (MRI), has allowed the non-invasive investigation of the structure and function of the human brain in health and pathology (Friston,

* Corresponding author. Tel.: +44 020 7848 0833.
E-mail address: a.mechelli@kcl.ac.uk (A. Mechelli).
[1] These two authors have contributed equally to this work.

0149-7634/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neubiorev.2012.01.004

2009). These techniques have been applied to patients with neurological or psychiatric disorders in order to identify possible biomarkers which could be used for early diagnosis, treatment planning and monitoring of disease progression. This has revealed structural and functional alterations in several disorders including, amongst others, mild cognitive impairment, probable dementia of Alzheimer type, major depression, bipolar disorder, schizophrenia and generalised anxiety disorder (Arnone et al., 2011; Davatzikos and Resnick, 2002; Ellison-Wright and Bullmore, 2010; Etkin and Wager, 2007; Smieskova et al., 2010; Zakzanis et al., 2003). To date however, the results of these studies have had minimal clinical impact and despite much interest in the use of brain scans for diagnostic and prognostic purposes, neurologists and psychiatrists are still forced to rely on traditional and often ineffective diagnostic and prognostic tools. One of the reasons for the limited impact of the findings on clinical practice, is that neuroimaging studies have typically reported differences between patients and controls at group level; in contrast, doctors working in a psychiatric or neurological ward have to make clinical decisions about individuals. For neuroimaging to be useful in a clinical setting therefore, one must be able to make inferences at the level of the individual rather than the group.

Over the past few years, there has been growing interest within the neuroimaging community in the use of analytical methods that allow such inference. One such method is supervised machine learning (ML), an area of artificial intelligence concerned with the development of algorithms and techniques able to automatically extract information from the data (Hastie et al., 2001). Relative to traditional methods of analysis based on the general linear model, the advantages of applying supervised ML to neuroimaging data are twofold. Firstly, supervised ML methods allow characterisation at the level of the individual therefore yielding results with a potentially high level of clinical translation. Secondly, as inherently multivariate approaches, supervised ML methods are sensitive to spatially distributed and subtle effects in the brain that would be otherwise undetectable using traditional univariate methods which focus on gross differences at group level.

Support Vector Machine (SVM) is a specific type of supervised ML method that aims to classify data points by maximising the margin between classes in a high-dimensional space (Pereira et al., 2009; Vapnik, 1995). The optimum algorithm is developed through a "training" phase in which training data are used to develop an algorithm able to discriminate between groups previously defined by the operator (e.g. patients vs. controls), and a "testing" phase in which the algorithm is used to blind-predict the group to which a new observation belongs. The initial applications of SVM to neuroimaging data were primarily aimed at decoding the mental states of healthy volunteers (see Haynes and Rees, 2006 for review); Davatzikos et al. (2005) for example successfully demonstrated that it was possible to discriminate between subjects giving truthful and non-truthful responses with an accuracy of 99.3% based solely on discriminative patterns of brain activity. Since then several studies have used SVM to examine the diagnostic and prognostic potential of neuroimaging in a range of neurological and psychiatric disorders and, to date, a number of promising results have been reported (e.g. Klöppel et al., 2008).

In what follows, we begin by providing a brief overview of SVM and its application to neuroimaging data. We then summarise the results of those studies which have examined the diagnostic and prognostic value of structural and functional neuroimaging in neurological and psychiatric disorders before concluding with a discussion regarding the main theoretical and practical challenges associated with actually implementing the use of this methodology into the clinic and possible future directions.

2. Overview of SVM

Within ML, there are two main approaches that one can take: supervised and unsupervised learning. In supervised learning, one seeks to develop a function which maps two or more sets of observations with two or more, operator defined, categories through an iterative procedure which gradually reduces the difference between the predicted and expected result; subsequently, the algorithm can then be used to assign new, previously unseen, data to one of the predefined categories with a given accuracy. In unsupervised learning by contrast, one seeks to determine how the data are organised without the availability a priori information supplied by the operator; with the primary objective being to discover unknown, but potentially useful structure in the data.

One specific form of supervised pattern recognition algorithm is that used for classification, concerned with the automatic discovery of regularities in the data that can be used to classify the data into different predefined categories. Using this approach individuals (represented by their brain scan for example) are referred to as "examples" and the categories to which they might belong, "labels"; the aim is to generate the decision function or "classifier" which most accurately captures the relationship between each example and its respective label (Pereira et al., 2009). Examples of multivariate techniques for pattern recognition include, but are not limited to, artificial neural networks, decision trees, Gaussian process classification and SVM. Of these the most popular technique in the neuroimaging literature is SVM (Vapnik, 2000), which to date has been applied in several studies of neurological and psychiatric disorder, allowing the classification of individual observations (e.g. scans) into distinct groups or classes (e.g. diagnostic categories) based on data in high-dimensional space. The main steps of the method comprise (i) preparing data for input into classifier training, (ii) training and testing the classifier and (iii) evaluating its performance.

2.1. Preparing data for classifier training

Prior to developing a classifier, two preliminary steps are required known as "feature extraction" and "feature selection", respectively. The first of these, feature extraction, involves the transformation of the original data into a set of "features" which can be used as input for SVM. In the case of neuroimaging, this may consist of transforming each three-dimensional image into a column vector of features with each value corresponding to the intensity of a single corresponding voxel. This feature vector thus encodes the pattern of grey or white matter volume (for structural neuroimaging data) or brain activation (for functional neuroimaging data) across the whole brain. Feature selection by comparison involves the selection of a subset of features expected to facilitate learning, in addition to the removal of any remaining features considered to be either of minimal importance, if not wholly redundant, for the purpose of classification. In the context of neuroimaging, this may consist of selecting one or more regions of interest that based on the previous literature are expected to differ between groups or, alternatively, more data-driven approaches such as recursive feature elimination. The rationale for feature selection is threefold. Firstly, it is hoped that by reducing the number of features supplied to classifier, it will be able to more accurately discriminate between classes. Secondly, feature selection can aid interpretation of the predictive model (e.g. localising which brain regions carry discriminative information). Thirdly, for some classification algorithms, removing irrelevant or redundant features reduces computational load and speeds up the learning process. Whilst the use of feature extraction is a prerequisite of SVM, feature selection represents an optional step only used when there are specific features believed not to aid the learning process. For further details

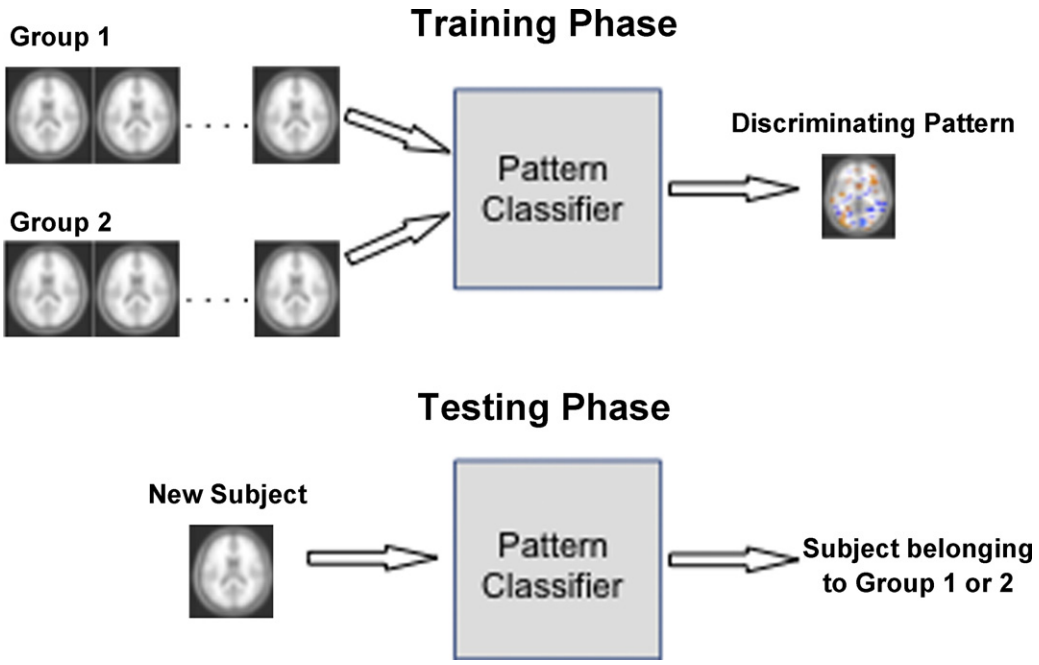

Fig. 1. Training and testing in SVM (top and bottom, respectively). In the training phase, an algorithm is developed which captures the key differences between groups (e.g. patients vs. controls). In the "testing" phase, the algorithm is used to determine the group to which a new observation belongs to.

concerning the issues involved in feature selection see Chu et al. (2011).

2.2. Training and testing

As touched upon above, SVM aims to develop an algorithm that allows the classification of individual observations into distinct classes based on data in high-dimensional space. This involves optimising the parameters of the algorithm using training data and then assessing its generalisation performance using test data (Fig. 1).

In the training phase, a SVM is trained using training data predefined into set groups by the operator in order to estimate a "decision function" or "hyperplane" which best distinguishes between the experimental set groups of interest (e.g. patients and controls). Fig. 2 illustrates a hypothetical pattern recognition problem between patients (red circles) and healthy controls HCs (blue circles) for the simplified case of two variables or voxels. Each axis represents the measurement in one voxel, with each symbol (circle or square) representing a brain scan of a different subject. As represented by the dashed lines in Fig. 2(a), there are a number of possible ways one can feasibly delineate between the data. In SVM, the classifier is trained by maximising the margin of separation between the two groups by using those examples lying closest to the separating plane (and hence the most difficult to classify) as defining points; these are called support vectors. In Fig. 2(b) the optimal classifier is represented by the dashed red line, and the support vectors are represented by circled symbols. This hypothetical example shows a linear classifier, in which each feature affects classification on the basis of its weight alone. One well known linear discriminant algorithm is the Fisher's Linear discriminant Analysis (LDA), which projects all the data points to a one dimensional space with the aim of maximising inter-group separation and minimising intra-class variation. Another important feature of SVM is that it does not require the training and test data to be explicitly represented and instead can be trained solely using a kernel, which is a matrix of pairwise similarities between data points. There are two main advantages of employing kernels in an SVM context: first, the kernel is an efficient representation of the similarity between

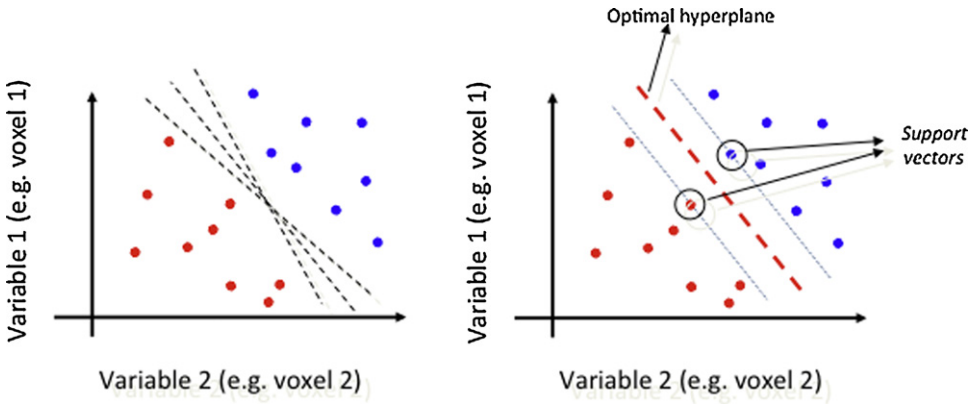

Fig. 2. Hypothetical pattern classification problem between patients (red circles) and controls (blue circles). Each symbol (circle or square) represents the brain scan of a different subject. (a) Using different pattern classification methods, it is possible to obtain several different classifiers that correctly separate the two groups; these are represented by dashed lines. (b) Using SVM, an optimal classifier is obtained as represented by the red dashed line and the support vectors are represented by circled symbols. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

samples. Thus the computational complexity of SVM is governed by the number of samples, rather than the number of features, which is advantageous in high-dimensional settings. Second, kernels can be used to learn non-linear decision boundaries. In this case, a non-linear kernel is used to implicitly map the data from the input space (where no linear separation of the data is possible) to a higher dimensional feature space (where the data can be separated by a linear hyperplane). In theory, non-linear kernels allow the classifier to solve more difficult classification problems than linear kernels but in practice, it is important to note that non-linear kernels do not guarantee more accurate generalisation to unseen samples since it can be difficult to know the optimal mapping in advance. Further, for problems where the number of data dimensions exceeds the number of samples, it is always possible to find a linear decision boundary that perfectly separates the data. Thus, many researchers in neuroimaging prefer linear kernels. For further details of the issues linked with linear and nonlinear methods see Muller et al. (2001) and Schölkopf and Smola (2002).

In the subsequent testing phase, the algorithm is used to predict the group to which a new, previously unseen, observation belongs. For an unbiased estimation of generalisability, it is important that the data used for testing does not overlap with those used for training. To achieve this, one option is to use two completely different sets of data for training and testing respectively; this however increases the total amount of data to be acquired. Alternatively one can use cross-validation, in which the total sample is split into training and testing data several times using different partitions and the resulting accuracies of classification are averaged across repetitions (Lemm et al., 2011). "Leave-one-out" cross-validation for example, involves excluding a single subject from each group and training the classifier using the remaining subjects; the subject pair excluded are then used to test the ability of the classifier to reliably classify new cases; this procedure is repeated for each subject pair in order to obtain a relatively unbiased estimate of generalisability (Hastie et al., 2001).

2.3. Performance evaluation

The performance of an algorithm can be described by its sensitivity, specificity and accuracy. Sensitivity refers to the proportion of true positives correctly identified (e.g. percentage of people with a disease who are identified as being ill). In contrast, specificity refers to the proportion of true negatives correctly identified (e.g. percentage of healthy people identified as not being ill). Finally accuracy represents the overall proportion of correct classifications; when there are an equal number of test subjects in both classes, this is equivalent to the mean of sensitivity and specificity. In the context of predictive classifiers sensitivity refers to the proportion of subjects who, having been predicted to fall ill, did fall ill, whilst specificity refers to the proportion of subjects who, having been predicted to remain stable, did remain stable. The statistical significance of any given classification accuracy can be determined using parametric tests such as the binomial test or by permutation testing. The latter involves repeating the classification procedure with training group labels randomly allocated by the computer multiple times in order to generate a null distribution of accuracies. The number of permutations in this null distribution achieving a higher accuracy than the true labels is then divided by the total number of permutations. This provides an estimate of the significance of the accuracy relative to chance.

3. A review of SVM studies of neurological and psychiatric disorders

In the past few years, an increasing cohort of studies have used SVM or other pattern recognition methods to investigate possible neuroanatomical biomarkers of neurological and psychiatric disorders. For simplicity, these studies can be divided into three main categories: (i) studies which examine the diagnostic value of neuroimaging data by comparing patients and HCs; (ii) studies which examine the potential of neuroimaging data for predicting the onset of a disease by comparing the brain scans (acquired at baseline) of individuals with prodromal symptoms who subsequently did and did not become ill, and (iii) studies which examine the prognostic value of imaging data by comparing the brain scans obtained from patients prior to treatment onset who subsequently did and did not respond. In order to identify suitable publications for inclusion, an online search of the Pubmed, Medline and PsychInfo. databases using the search terms; '("Support Vector Machine") AND ("diagnosis" OR "prognosis" OR "prediction") AND (MRI OR "positron-emission tomography")' was conducted on the 17th November 2011. A total of 88 hits were returned of which we included 31. Those studies investigating non-neurological or non-psychiatric disorders, or those not meeting our strict operating methodological definitions were excluded. In addition to the online search criteria described, we also hand cross-referenced the publication list referenced by each of the studies we included to ensure that no studies of significance were omitted from the review. Using these criteria a total of 50 papers were included comprising 40 investigating diagnostic classification, 15 investigating prognostic classification and 5 investigating prediction of treatment response (with some studies falling into more than one category) (see Tables 1–3). Below we review the main findings of these studies, before discussing the potential limitations of the methodology and possible future directions.

3.1. Diagnostic studies

To date several studies have examined the diagnostic value of imaging data in neurological and psychiatric disorders utilising a range of neuroimaging techniques including structural MRI, functional MRI, diffusion tensor imaging (DTI) and PET (see Table 1).

With respect to neurological disorders, the majority of diagnostic investigations conducted to date have focused on mild cognitive impairment (MCI) and probable dementia of Alzheimer type (PDAT). Characterised by an impairment in cognition extending beyond what one would expect based on age and education alone, MCI is thought to be a transitional stage between normal ageing and PDAT. A devastating and increasingly wide spread illness, early diagnosis of PDAT patients and its preceding prodromal MCI phase is becoming rapidly more important since it potentially facilitates earlier individual treatment intervention with cholinesterase inhibitors, shown to be effective in terms of delaying the course of the illness and improving or stabilising cognitive and behavioural symptoms (Whitehead et al., 2004). Accordingly, given the vastly positive downstream consequences that come from providing such intervention for those in the earliest stages of the illness, there has naturally been much interest in the potential use of brain scans combined with SVM to identify individuals with MCI and PDAT.

In this context a number of studies have been conducted comparing individuals with MCI against HCs with reports of considerable success. Notably, ten separate studies using a range of different modalities including both structural and functional MRI, PET, DTI, and CSF biomarkers, both individually and in combination, have been able to discriminate between individuals with MCI and HCs matched for age and gender with accuracies ranging from 71.09% (Cui et al., 2011b) to 100% (Fan et al., 2008). Moreover these results build on the increasingly established finding that individuals with PDAT can be successfully discriminated from HCs, which to date has been demonstrated by a total of nineteen studies (see Table 1). Once again, the accuracies reported vary with

Table 1
Studies investigating the diagnostic potential of neuroimaging data and SVM.

Author	Comparison	Sample size	Technique	Accuracy (%)
Arimura et al. (2008)	PDAT vs. HC	PDAT = 29 HC = 25	Structural MRI	82.7
Davatzikos et al. (2008)	MCI vs. HC	MCI = 15 HC = 15	Structural MRI	90
Duchesne et al. (2008)	PDAT vs. HC	PDAT = 75 HC = 75	Structural MRI	92
Fan et al. (2008a)	MCI vs. HC	MCI = 15 HC = 15	Structural MRI & PET	100
Ferrarini et al. (2008)	PDAT vs. HC	PDAT = 58 HC = 28	Structural MRI	84
Klöppel et al. (2008a)	PDAT vs. HC	PDAT = 20 HC = 20	Structural MRI	95
	PDAT vs. FTLD	PDAT = 18 FTLD = 19	Structural MRI	89.2
Vemuri et al. (2008)	PDAT vs. HC	PDAT = 190 HC = 190	Structural MRI & genetic data	89.3
Gerardin et al. (2009)	PDAT vs. HC	PDAT = 23 HC = 23	Structural MRI	94
	MCI vs. HC	MCI = 23 HC = 25	Structural MRI	83
Magnin et al. (2009)	PDAT vs. HC	PDAT = 16 HC = 22	Structural MRI	94.5
Haller et al. (2010)	HC vs. MCI	HC = 35 MCI = 67	Structural MRI, DTI FA	91.4
Nho et al. (2010)	PDAT vs. HC	PDAT = 182 HC = 226	Structural MRI	90.5
Oliveira et al. (2010)	PDAT vs. HC	PDAT = 14 HC = 20	Structural MRI	88.2
Plant et al. (2010)	PDAT vs. HC	PDAT = 32 HC = 18	Structural MRI (GM + WM)	90
	MCI vs. HC	MCI = 24 HC = 18	Structural MRI (GM + WM)	97.62
Salas-Gonzalez et al. (2010)	PDAT + MCI vs. HC	PDAT + MCI = 167 MCI = 52	PET	86
	PDAT vs. HC	PDAT = 53 HC = 52	PET	95
	MCI vs. HC	MCI = 114 HC = 53	PET	88
Abdulkadir et al. (2011)	PDAT vs. HC	PDAT = 95 HC = 95	Structural MRI (GM)	84.5
Chen et al. (2011)	PDAT vs. HC	PDAT = 21 HC = 20	Resting-state functional MRI	87
	MCI vs. HC	MCI = 15 HC = 20	Resting-state functional MRI	95
Chincarini et al. (2011)	PDAT vs. HC	PDAT = 144 HC = 189	Structural MRI	Sens. = 89 Spec. = 94
Cui et al. (2011a)	MCI vs. HC	MCI = 79 HC = 204	Structural MRI & DTI	71.09
Dukart et al. (2011)	PDAT vs. FTLD vs. HC	PDAT = 21 FTLD = 14 HC = 13	Structural MRI (GM) & PET	91.7
Graña et al. (2011)	PDAT vs. HC	PDAT = 20 HC = 25	DTI FA	100
		PDAT = 20 HC = 25	DTI MD	98
Hinrichs et al. (2011)	PDAT vs. HC	PDAT = 48 HC = 66	Structural MRI & PET & CSF(tau, amygdaloid-beta 142, p-tau 181P, t-tau, APOE genotype) & NMs	88.2
Zhang et al. (2011a)	PDAT vs. HC	PDAT = 45 HC = 50	Structural MRI & PET & CSF(Aβ42, t-tau and p-tau)	92
	MCI vs. HC	MCI = 91 HC = 50	Structural MRI & PET & CSF(Aβ42, t-tau and p-tau)	80
Zhang et al. (2011b)	PDAT vs. HC	PDAT = 51 HC = 52	Structural MRI & PET & CSF(Aβ42, t-tau and p-tau)	93.2
	MCI vs. HC	MCI = 99 HC = 52	Structural MRI & PET & CSF(Aβ42, t-tau and p-tau)	76.4
Fu et al. (2008)	MD vs. HC	MD = 19 HC = 19	Functional MRI	86
Marquand et al. (2008)	MD vs. HC	MD = 20 HC = 20	Functional MRI	67.5
Costafreda et al. (2009a)	MD vs. HC	MD = 37 HC = 37	Structural MRI	67.6
Hahn et al. (2010)	MD vs. HC	MD = 30 HC = 30	Functional MRI	83
Nouretdinov et al. (2010)	MD vs. HC	MD = 19 HC = 19	Structural and fMRI	76.3

Table 1 (*Continued*)

Author	Comparison	Sample size	Technique	Accuracy (%)
Gong et al. (2011)	RDD vs. HC	RDD = 23 HC = 23	Structural MRI (GM)	67.39
	NDD vs. HC	NDD = 23 HC = 23	Structural MRI (GM)	76.09
	RDD vs. HC	RDD = 23 HC = 23	Structural MRI (WM)	58.70
	NDD vs. HC	NDD = 23 HC = 23	Structural MRI (WM)	84.65
Koutsouleris et al. (2009)	HC vs. ARMS	HC = 17 ARMS-T = 15 ARMS-NT = 18	Structural MRI	82
Davatzikos et al. (2005)	SCH vs. HC	SCH = 69 HC = 79	Structural MRI	81.1
Shen et al. (2010)	SCH vs. HC	SCH = 32	Functional MRI	84.37
Yang et al. (2010)	SCH vs. HC	SCH = 20 HC = 20	Combined fMRI and genetic data	83
Costafreda et al. (2011b)	SCH vs. HC	SCH = 32 HC = 40	Functional MRI	92
	BD vs. HC	BD = 32 HC = 40	Functional MRI	79
Sun et al. (2009)	PS vs. HC	PS = 36 HC = 36	Structural MRI	86.1
Ingalhalikar et al. (2010)	SCH vs. HC	SCH = 27 HC = 37	DTI	90.62
	ASD vs. HC	ASD = 25 HC = 23	DTI	89.58
Ecker et al. (2010a)	HC vs. ASD	HC = 22 ASD = 22	Structural MRI (GM)	81
		HC = 22 ASD = 22	Structural MRI (WM)	68
		HC = 22 ASD = 22	Structural MRI (GM + WM)	77
Ecker et al. (2010b)	HC vs. ASD	HC = 20 ASD = 20	Structural MRI	85
Duchesne et al. (2009)	IPD vs. PPS	IPD = 16 PPS = 16	Structural MRI	90.6
Focke et al. (2011)	IPD vs. PSP	IPD = 21 PSP = 10	Structural MRI (GM)	87.1
		IPD = 21 PSP = 10	Structural MRI (WM)	96.77
	IPD vs. MSA	IPD = 21 MSA = 11	Structural MRI (GM)	71.87
		IPD = 21 MSA = 11	Structural MRI (WM)	n.s.
	IPD vs. HC	IPD = 21 HC = 22	Structural MRI (GM or WM)	n.s.

ASD, autistic spectrum disorders; BD, bipolar disorder; FTLD, frontotemporal lobar degeneration; HC, healthy controls; MCI, mild cognitive impairment; MD, major depression; NDD, non-refractory depressive disorder; PDAT, probable dementia of Alzheimer's type; PS, psychosis; RDD, refractory depressive disorder; SCH, schizophrenia; IPD, Idiopathic Parkinson's Disease; PPS, Parkinson Plus Syndromes; PSP, Progressive Supranuclear Palsy; MSA, Multiple Systems Atrophy; NMs, Neuropsychological Measures; n.s., non-significant result.

figures ranging from 82.7% (Arimura et al., 2008) to a perfect 100% (Graña et al., 2011), with a range of modalities and integrative approaches being represented. At this point it is worth noting that whilst not producing definitively superior results, a number of these reports suggest the integration of modalities may improve the level of accuracy with which groups can be discriminated. Fan et al. (2008a) for example demonstrated that the integration of structural MRI and PET allowed discrimination between individuals with MCI and matched controls with an accuracy of 100% on the basis of hippocampal and medial temporal lobe grey matter which contrasted with their previous work where, based on sMRI alone, they were only able to achieve an accuracy of 87%.

Evidence also exists suggesting that SVM can be used to distinguish not only between HCs and patients with dementia, but between patients with different types of dementia as well. For instance in 2008 Klöppel and colleagues reported that SVM could correctly distinguish between sporadic PDAT and controls with an accuracy of 95% and between sporadic PDAT and frontotemporal lobar degeneration (FLDT) with an accuracy of 89.2%. Furthermore, it was found that the algorithm could be generalised across data sets acquired using different scanners from multiple research centres, a major hurdle for any new diagnostic tool, providing considerable support for the potential role of computer-based diagnosis in everyday clinical practice. Furthermore, this conclusion is strengthened by the observation that the unmanned SVM analysis compared favourably with radiologist accuracy on the same data set. Consistent with these findings Dukart et al. (2011) were able to classify between PDAT, FLDT and control subjects with an accuracy of 91.7% on the basis of structural MRI and PET data combined.

A number of investigations have also investigated PDAT and MCI concurrently, directly discriminating between the two, including a selection of those discussed above. In 2009 for example Gerardin and colleagues compared patients with MCI and PDAT with HCs matched for age, using the hippocampus as a selected region of interest. The authors reported good accuracy of classification both for PDAT vs. controls (accuracy: 94%, sensitivity: 96%, specificity: 92%) and MCI vs. controls (accuracy: 83%, sensitivity: 83%, specificity: 84%, respectively). Rather than using neuroanatomical data, a similar study that compared MCI, PDAT and normal cognitive status also found a similar pattern of results, but in this instance by

relying on a Large-Scale Network analysis based on resting-state functional MRI (Chen et al., 2011).

Taken collectively, the implications of the studies described above suggest that (i) both structural and functional neuroimaging data provide important information for distinguishing between MCI, PDAT and normal ageing, and (ii) the integration of structural and functional data may in some cases improve accuracy of classification. Furthermore, the findings reported by Klöppel et al. (2008) provide support to the fundamental idea that the methodology can be robust and hence generalised across different research centres.

With respect to diagnostic studies of psychiatric disorders, these have to date predominantly focused on major depression and, more recently, psychosis. Costafreda et al. (2009a) for example applied SVM to structural brain scans of patients with major depression and healthy volunteers reporting an accuracy of classification of 67.6%. Interestingly this was driven by a distributed network which included the right subgenual anterior cingulate, medial frontal gyrus, precuneus, hippocampus and thalamus. One limitation of the finding however is that diagnostic accuracy based on structural neuroanatomy may be critically dependent on the specific type of depression being examined. For instance, Gong et al. (2011) reported a different accuracy of classification for patients with refractory (RDD) vs. non-refractory (NDD) depression: specifically diagnostic accuracy with respect to grey and white matter respectively was 67.39% for RDD and 76.09% for NDD, and 70% for RDD and 84.65% for NDD. SVM studies investigating major depression using functional neuroimaging data have also been conducted, including studies by Fu et al. (2008), Marquand et al. (2008) and Hahn et al. (2010). In the first of these, Fu and colleagues reported being successfully able to discriminate between patients with major depression and HCs with an accuracy of 86% based on the functional activity occurring in response to the processing of sad facial expressions. Similarly, Marquand et al. were able to use subject responses to a variable load version of the 'n-back' task which allowed them to discriminate between unipolar depression patients and HCs with a significant accuracy of 67.5%. Most recently, Hahn and colleagues used a decision tree approach in order to integrate neuroimaging data with multiple, symptom-related neural processes. Consequently, they reported the ability to discriminate between depressed patients and HCs using the integrated classifier with 83% accuracy which was greater than any of the single best classifiers generated.

In the context of psychosis, the first study to use SVM to examine the diagnostic value of structural MRI in schizophrenia was published by Davatzikos et al. (2005), who was able to discriminate between schizophrenia (SZ) and HCs with an accuracy of 81.1%. Notably, discrimination was based not only on frontotemporal regions but also on parts of the occipital cortex, an area not traditionally implicated in schizophrenia – demonstrating SVM's ability to detect subtle and distributed brain alterations. More recently, Shen et al. (2010) used resting-state functional connectivity to compare patients with schizophrenia and HCs, on the basis that poor functional integration is thought to be one of the possible biomarkers of the disorder (Pettersson-Yeo et al., 2011). The authors report being able to identify schizophrenia patients and HCs with a 93.75% sensitivity and 75.0% specificity, resulting in overall accuracy of 84.37%. Interestingly the authors also report that the classification was predominantly driven by group differences in the functional integration between the cerebellum and prefrontal regions. Preliminary evidence also exists suggesting that diagnostic accuracy of neuroimaging data may vary as a function of the type of psychosis, as shown by Costafreda et al. (2011a,b). The authors examined the relative diagnostic value of functional MRI in schizophrenia and bipolar disorder using a phonological verbal fluency task which is known to tap into cognitive processes affected in both disorders. Higher diagnostic accuracy was found for schizophrenia (92%) than bipolar disorder (79%) possibly reflecting the fact that, although both disorders are associated with altered activation in prefrontal, anterior and posterior cingulate and striatal regions, the magnitude of this dysfunction was more pronounced in schizophrenia. Similarly, Yang et al. (2010) examined whether diagnostic accuracy could be improved by integrating different types of data within the same algorithm. Though basically consistent with earlier work (Vemuri et al., 2008) their own efforts proved little however with the integration of genetic and functional MRI data yielding only a slight improvement in diagnostic accuracy (87%) relative to the use of genetic data or functional MRI data alone (74% and 82%, respectively). Whilst the above studies were performed with chronic schizophrenia patients, more recent efforts have investigated those in the very earliest stages of the illness. For example Sun et al. (2009) used a cortical pattern matching method to compare HCs to patients with recent-onset psychosis. Significantly patients showed lower grey matter density, particularly in prefrontal, cingulate and lateral temporal brain regions relative to the controls, and the pattern classification analysis using a leave-one-out cross-validation was able to discriminate between the two groups with an accuracy of 86.1%. Further consistent with this finding, Koutsouleris and colleagues reported being able to discriminate between HCs and those with an At-Risk Mental State (ARMS) based on grey matter volume with 82% accuracy (Koutsouleris et al., 2009).

Neuroimaging and SVM have also been combined to examine autistic spectrum disorders (ASD). Most notably, Ecker et al. (2010a) was able to distinguish between adults with ASD and HCs based on both grey matter (GM) and white matter (WM) volume with a classification accuracy of 81% in the former and 68% in the latter. Furthermore the distance between the individual observations and the hyperplane which separated the two groups was greater for those with greater symptom severity indicating that this was the main factor driving the classifier as opposed to other potentially confounding variables. A follow up study (Ecker et al., 2010b) from the same research group replicated this first finding reportedly being able to discriminate between individuals with ASD and HCs, using structural MRI, with an accuracy of 85%. Further consistent with these findings, Ingalhalikar and colleagues using DTI, a wholly different approach, were able to distinguish between children with ASD and HC children with 89.58% accuracy (Ingalhalikar et al., 2010).

Focke et al. (2011) also used SVM to distinguish between those suffering Idiopathic Parkinsonian Syndrome (IPS) and two similar afflictions it is frequently misdiagnosed as, namely, Progressive Supranuclear Palsy (PSP) and Multiple Systems Atrophy (MSA) (Parkinsonian variant), with accuracies of 87.1% and 71.87% using grey matter images, and 96.77% using white matter images, respectively.

In summary, the growing breadth of studies published to date, collectively suggest that using structural and/or functional neuroimaging data as input to SVM has significant potential as a diagnostic aid for major neurological and psychiatric illnesses.

3.2. Studies on prediction of disease onset

In contrast to the diagnostic classifiers reviewed so far, the second category of studies reviewed here examined the potential of neuroimaging data for predicting the onset of a disease by comparing the brain scans of individuals with possible prodromal signs at baseline who subsequently did and did not become ill by a predefined follow-up date (see Table 2).

Although approximately 40–60% of those suffering from MCI develop AD within 4–6 years, a large proportion of those diagnosed with the impairment will remain cognitively stable over time in their impaired state, or even return to a normal functional level

Table 2
Studies investigating the potential of imaging data and SVM for predicting conversion to illness.

Author	Comparison	Sample size	Technique	Accuracy (%)
Fan et al. (2008b)	PDAT vs. HC	PDAT = 66 HC = 56	Structural MRI	94.3
	MCI vs. HC	MCI = 88 HC = 56	Structural MRI	81.8
	PDAT vs. MCI	PDAT = 66 MCI = 88	Structural MRI	74.3
Davatzikos et al. (2010)	MCI-C vs. MCI-NC	MCI-C = 69 MCI-NC = 170	Structural MRI, SPARE-AD and t-tau and Aβ42	60.8
Haller et al. (2010)	HC vs. MCI	HC = 35 MCI = 67	Structural MRI, DTI FA	91.4
	Stable MCI vs. progressive MCI	Stable MCI = 40 Progressive MCI = 27	Structural MRI, DTI FA	98.4
Nho et al. (2010)	MCI vs. PDAT	MCI = 267 PDAT = 182	Structural MRI (FreeSurfer VBM)	72.3
Plant et al. (2010)	MCI-C vs. MCI-NC	MCI-C = 13 MCI-NC = 9	Structural MRI (GM + WM)	95.83
Chincarini et al. (2011)	MCI-C vs. MCI-NC	MCI-C = 136 MCI-NC = 166	Structural MRI	Sens. = 72 Spec. = 65
	MCI-C vs. HC	MCI-C = 136 HC = 189	Structural MRI	Sens. = 89 Spec. = 80
Costafreda et al. (2011a)	MCI vs. PDAT	MCI = 103 PDAT = 71	MRI, 3D hippocampal morphology	80
Cui et al. (2011b)	MCI-C vs. MCI-NC	MCI-C = 56 MCI-NC = 87	Structural MRI & PET & CSF& NMs	67.13
Zhang et al. (2011a)	MCI-C vs. MCI-NC	MCI-C = 43 MCI-NC = 48	Structural MRI & PET & CSF(Aβ42, t-tau and p-tau)	65.4
Zhang et al. (2011b)	MCI-C vs. MCI-NC	MCI-C = 43 MCI-NC = 56	Structural MRI & PET & CSF(Aβ42, t-tau and p-tau)	Sens. = 91.5 Spec. = 73.4
Koutsouleris et al. (2009)	ARMS-NT vs. ARMS-T	ARMS-NT = 15 ARMS-T = 18	Structural MRI	82
Koutsouleris et al. (2011)	ARMS-T vs. ARMS-NT	ARMS-T = 16 ARMS-NT = 21	Structural MRI	84.2
	ARMS-T vs. HC	ARMS-T = 16 HC = 22	Structural MRI	92.3
	ARMS-NT vs. HC	ARMS-NT = 21 HC = 22	Structural MRI	66.9
Mourao-Miranda et al. (2011)	PS-CP vs. PS-EP	PS-CP = 28 PS-EP = 28	Structural MRI	70
	PS-CP vs. HC	PS-CP = 28 HC = 28	Structural MRI	67
	PS-EP vs. HC	PS-EP = 28 HC = 28	Structural MRI	n.s.
Klöppel et al. (2008b)	PSH vs. HC	PSH = 25 HC = 20	DTI FA	82.2
Klöppel et al. (2009)	PSH vs. HC	PSH(<10%) = 32 HC = 32	Structural MRI (GM)	n.s.
		PSH(10–33%) = 32 HC = 32	Structural MRI (GM)	n.s.
		PSH(>33%) = 96 HC = 95	Structural MRI (GM)	69

ARMS-NT, individuals with an at-risk mental state who did not convert to psychosis; ARMS-T, individuals with an at-risk mental state who converted to psychosis; HC, healthy controls; MCI, mild cognitive impairment; MCI-C, MCI individuals who converted to PDAT; MCI-NC, MCI individuals who did not convert to PDAT; PDAT, probable dementia of Alzheimer's type; PSH (<10%, 10–33%, >33%), presymptomatic Huntington's disease (which has a less than 10%, 10–33% more than 33% chance, respectively, of clinically manifesting within 5 years); PS-CP/EP/, continuous psychosis/episodic psychosis; FA, fractional anisotropy; n.s., non-significant result.

(Gauthier et al., 2006; Petersen, 2004). To date however, no reliable method exists to predict who will, and will not, develop dementia amongst those with MCI, and a number of neurological studies have used SVM in an effort to rectify this fact. In one such case Fan et al. (2008b) examined the value of structural MRI for predicting cognitive decline in patients with MCI. The authors reported that (i) some but not all MCI patients have significant and extensive brain atrophy and following on from this (ii) the extent of brain atrophy predicted cognitive decline over the following year with an accuracy of 87%. Whilst Fan and colleagues performed a whole-brain analysis a similar study conducted by Costafreda et al. (2011a,b) focused on the hippocampus as an a priori region of interest. Using a fully automated prognostic procedure based on 3D hippocampal morphology, the team were able to predict those MCI patients who would convert to PDAT within 1 year with an accuracy of 80%, and furthermore, reported that those subjects who were to develop dementia showed significantly more rapid cognitive deterioration.

Based on the notion that the integration of different types of data may improve accuracy of classification, Davatzikos et al. (2010) approached the issue of predicting MCI conversion to PDAT incorporating multiple data modalities. Specifically, they examined whether the joint evaluation of structural MRI and cerebrospinal fluid (CSF) biomarkers (t-tau and $A\beta_{42}$) could predict MCI to PDAT conversion within the relatively short period of 12 months. Significantly, the authors reported that those who did convert were characterised by grey and white matter atrophy in the temporal lobe, the posterior cingulate/precuneus, and the insula combined with specific AD-like CSF biomarkers allowing them to be classified with a statistically significant accuracy of 60.8%. By comparison, in those who did not convert, no clear combined pattern of atrophy and AD-like CSF biomarkers could be discerned. Although not

exceptionally high, it is worth noting that the accuracy detailed represented the average of an 84.2% sensitivity and a 50% specificity the latter of which might be explained by the possibility that some MCI patients categorised as non-converters at the 12-month period might in fact be converters who developed AD after the 12-month window, thus interfering with the current specificity score. This interpretation is consistent with the results of another recent study (Nho et al., 2010), which examined the potential of structural MRI for predicting MCI to PDAT conversion using follow-up periods of 1, 2 and 3 years. The authors developed an algorithm by comparing patients with PDAT and HCs, and tested whether the same algorithm could also be used to predict conversion from MCI to PDAT. It was found that the classifier trained to distinguish between PDAT and HCs predicted conversion from MCI to PDAT with an accuracy of 65.0% after a follow-up of 1 year, and 72.3% at 2- and 3-year follow-up. These findings are further supported by the results from five more investigations, that together used both single and multi-integrative modality approaches, all of whom reported being able to successfully discriminate between those MCI subjects who would, and would not, transition to PDAT in a given time frame with a high level of accuracy (Chincarini et al., 2011; Cui et al., 2011a; Plant et al., 2010; Zhang et al., 2011a,b) (see Table 2).

Whilst the above studies used structural MRI, Haller et al. (2010) attempted to predict cognitive decline in MCI at 1-year follow-up using diffusion tensor imaging (DTI), which provides information on the integrity of white matter tracts. Notably, their SVM-based analysis of DTI data yielded a classification accuracy of 98.4%, discriminating between stable MCI patients and those who would suffer progressive decline. Thus, the assessment of white matter integrity at baseline appears to allow highly accurate prediction of subsequent cognitive decline, even with a relatively short follow-up period. Taken collectively, the results of the many studies discussed so far are consistent with the idea that structural neuroanatomy contains important information for predicting progression to dementia amongst those with MCI. This idea is further supported by data reported by Klöppel and colleagues who in 2008 showed that individuals genetically identified as having pre-symptomatic Huntington's Disease (PSH) could be accurately distinguished from HCs based on DTI neuroimaging data with an accuracy of 82.2% (Klöppel et al., 2008). A result the same team expanded on in 2009 by accurately discriminating between HCs and those PSH subjects who had a greater than 33% chance of manifesting clinical symptoms within 5 years, with an accuracy of 69% (Klöppel et al., 2009).

In comparison, relatively fewer studies have investigated the potential of neuroimaging for predicting onset of psychiatric disorders. Of those that have Koutsouleris and colleagues in 2009 examined whether structural MRI could be used for predicting transition from the so-called ARMS to full blown-psychosis. Individuals with an ARMS have increased genetic vulnerability to psychosis and/or express sub-threshold clinical symptoms, both in conjunction with a decline in social and cognitive functioning. It has been estimated that one third of them will transition into established psychosis within two years of diagnosis, with the remaining two thirds either remaining stable or improving (Cannon et al., 2007), and there is currently growing evidence that early clinical intervention may significantly reduce the risk of an individual developing schizophrenia (McGorry et al., 2002). To date however, based on purely clinical grounds, it has not been possible to distinguish between those ARMS subjects who will and will not go on to make the transition. In an effort to counter this Koutsouleris et al. (2009) acquired structural MRI data from a total of 45 individuals with an ARMS and followed them up clinically for a period of 4 years. The team reported that using the baseline scans, SVM-based analysis of grey matter allowed prediction of subsequent conversion to psychosis with a significant accuracy of 82%. In addition to the diagnostic aid perspective, the data also adds support to the conjecture that grey matter abnormalities diffuse throughout the whole brain may contain important information for predicting individual clinical outcomes, rather than just specifically selected a priori regions or networks. Koutsouleris and colleagues further expanded this result, accurately discriminating between those who did, and did not, transition from the ARMS to full blown psychosis with an accuracy of 84.2% (Koutsouleris et al., 2011). Focusing further along the psychosis timeline, one study has also reported the ability to discriminate between subjects who, having had a first episode of psychosis, would then go on to suffer either 'continuous psychosis' (PS-CP) or alternatively, a more 'episodic psychosis' (PS-EP). Specifically, Mourao-Miranda et al. (2011) were able to retrospectively distinguish PS-CP from PS-EP and PS-CP from HCs with accuracies of 70% and 67%, respectively. Notably however, they were unable to significantly classify PS-EP from HCs (see Table 2).

3.3. Studies on prediction of treatment response

Whilst both diagnostic and predictive tools are of essential importance, a third vital strand requiring investigation is the ability to predict patient response to clinical treatments. The third category of studies reviewed here therefore examined the prognostic value of imaging data by comparing the brain scans obtained from patients who subsequently did, and did not, respond to treatment (Table 3).

To date the vast majority of those studies have focused on response to treatment in individuals suffering major depression. This issue is of great clinical relevance since response to treatment tends to vary from one individual to another in most, if not all, neurological and psychiatric conditions, for example, nearly one third of patients with major depression show no improvement following standard antidepressant treatment (Stimpson et al., 2002). One such study that applied SVM to structural MRI data in order to predict response to antidepressant medication (Costafreda et al., 2009a) found that grey matter allowed prediction of treatment response with a sensitivity of 88.9% and a specificity of 88.9%, resulting in an overall accuracy of 88.9%. One limitation of the study however was the fact that the total sample comprised of only 18 patients thereby reducing the generalisability of the findings. More recently, Gong et al. (2011) in the same study referenced above, attempted to replicate this result using by acquiring the structural MRI of a larger sample of 46 patients who had all received standard antidepressant medication, and also attempted to extend the findings by considering the predictive value of both grey and white matter. Significantly SVM applied to grey and white matter correctly predicted clinical outcome at 12 weeks follow-up with an accuracy of 69.57% and 65.22% respectively. Two alternative studies also investigating treatment response in major depression applied SVM to functional MRI data instead. Firstly, Fu and colleagues examined whether brain activation during the processing of sad facial expressions could predict subsequent response to antidepressant medication, and reported that they could with an accuracy of 68.5%. Secondly, Costafreda et al. (2009b) investigated whether brain activation in response to sad facial expressions predicted subsequent treatment response but this time in relation to cognitive behavioural therapy. Likewise, the team were able to classify with a substantial accuracy of 78.5%. A major limitation of both studies however was the small size of their samples (8 and 7 pairs, respectively), so much so that the classification accuracies detailed could not be deemed significant.

As noted, whilst the majority of psychiatric treatment response studies have looked at depression there is one investigation that examined treatment response in those with schizophrenia. Khodayari-Rostamabad et al. (2010) investigated whether pre-treatment electroencephalography (EEG) data could predict

Table 3
Studies investigating the potential of imaging data and SVM for predicting response to treatment.

Author	Disorder	Treatment	Sample size	Technique	Accuracy (%)
Costafreda et al. (2009a)	MD	Fluoxetine	Responders = 9 Non-responders = 9	Structural MRI (GM)	88.9
Costafreda et al. (2009b)	MD	CBT	Responders = 7 Non-responders = 7	Functional MRI	78.5
Fu et al. (2008)	MD	Antidepressant	Responders = 8 Non-responders = 8	Functional MRI	68.5
Khodayari-Rostamabad et al. (2010)	SCH	Clozapine	Responders = 12 Non-responders = 11	EEG	85
Gong et al. (2011)	MD	Antidepressant (imipramine equivalents)	Responders = 23 Non-responders = 23	Structural MRI (GM)	69.57
	MD	Antidepressant (imipramine equivalents)	Responders = 23 Non-responders = 23	Structural MRI (WM)	65.22

CBT, cognitive behavioural therapy; SCH, schizophrenia.

response to clozapine in adults suffering from chronic schizophrenia, using a supervised ML method known as kernelised *partial least squares regression* (KPLSR) procedure (Rosipal and Kramer, 2006). Significantly, a comparison between responders and non-responders yielded an accuracy of 85%, providing preliminary support for the concept that brain activity is as useful an indicator for predicting treatment outcome as brain structure, and also for the use of EEG methodology specifically.

4. Discussion

Neuroimaging studies have revealed structural and functional alterations in several neurological and psychiatric disorders, yet the impact of these findings on clinical practice has been very limited and at the present time there are no objective, biological, markers which can be used to make early diagnosis, prognosis and prediction of clinical outcome in these disorders. SVM is an analytical method which differs from traditional analytical techniques in that (i) it allows the characterisation at the level of the individual, and therefore has high translational potential in a clinical setting, and (ii) it is inherently multivariate, and therefore is sensitive to spatially distributed and subtle effects in the brain. In this review, we have presented a brief overview of SVM followed by a summary of the studies which have used this method to examine the diagnostic and prognostic value of neuroimaging in neurological and psychiatric disorders.

Collectively, the studies published so far support the idea that structural and functional neuroimaging could be used to inform diagnosis of neurological and psychiatric disorders, particularly Alzheimer's disease, depression, schizophrenia, bipolar disorder, Parkinson's disease, presymptomatic Huntington's disease and autistic spectrum disorder. It should be noted however, that diagnostic accuracy was below 100% with the exception of two studies (Fan et al., 2008; Graña et al., 2011). Any translational implementation should take into account the fact that the cost of erroneously misclassifying someone ill as healthy may be higher than the cost of misclassifying someone healthy as ill; thus an algorithm which provides excellent sensitivity but only good specificity may be preferred to one which provides excellent specificity but only good sensitivity. It should also be noted that the application of SVM to neuroimaging data can only reach the same level of diagnostic accuracy as traditional methods of clinical assessment. This is because the development of a diagnostic algorithm is based on the distinction between patients and controls in the training data, which in turn relies on traditional clinical assessment. Thus, it would be an expression of logical confusion to expect that the application of SVM to neuroimaging data would allow better diagnostic classification than traditional methods of clinical assessment. Nevertheless, the diagnostic application of SVM to neuroimaging data could be very useful in a forensic setting for example, as an objective means of reducing controversy in evaluations of mental insanity and minimising errors in detecting malingering (Sartori et al., 2011).

The studies published so far also provide proof of concept that it may be possible to combine SVM with neuroimaging to predict conversion to PDAT or schizophrenia from those in the respective prodromal phases. This would constitute a significant clinical advancement, since at present there are no reliable methods to identify those at greatest risk of becoming ill, and therefore at greatest need of early clinical intervention, within the MCI and ARMS populations. However, it should be noticed that there have been only few studies on conversion from MCI to PDAT and even fewer studies that focussed on conversion from ARMS to schizophrenia; furthermore, most of these studies had a relatively small sample size (see Table 2). Reports from a handful of the most recent studies however strongly suggest that the results from such single centre investigations with a relatively small sample size may well be proved reliable. Specifically, Abdulkadir et al. (2011), Chincarini et al. (2011), Cui et al. (2011b), Hinrichs et al. (2011), and Zhang et al. (2011a,b) all used data collected by the ADNI protocol in which images are gathered from many different centres using a variety of different scanner types and scanning protocols. In each case, the groups were able to develop classifiers that could successfully discriminate between PDAT and HC subjects demonstrating the ability of SVM to generalise across different centres. Further to this, two of the five studies (Chincarini et al., 2011; Cui et al., 2011b) provided further validation by applying the classifier generated from PDAT vs. HC subjects to MCI-C and MCI-NC subjects. In each case the classifier was able to accurately distinguish the two groups with sensitivity 72%/sensitivity 65% and accuracy 67.13% respectively. These results support the notion that SVM could lead to the development of a translational tool for predicting conversion to illness; however this will require the investigation of larger samples and the careful characterisation of the impact of demographic variables such as age, gender and ethnicity which may influence the performance of a classifier.

The studies published so far also indicate that the application of SVM to neuroimaging data may allow prediction of treatment response at the individual level. This could help clinicians make more effective clinical decision in the early stage of the disorder, resulting in fewer ineffective trials and higher remission rates. However, it should be noticed that none of these studies were able to dissociate between treatment response and general remission, because treatment was administered to all patients. Clearly the inclusion of a placebo group in order to dissociate between treatment response and general remission would be unethical, however future studies could compare patient groups receiving different treatments. Also most of the studies were based on small samples (see Table 3) and there have been no attempts to test the generalisability of the results across different research centres. The development of translational tools for predicting conversion to

illness will therefore require the investigation of much larger samples, and a better understanding of the effects of age, gender and ethnicity.

5. Current challenges and future directions

SVM and other supervised ML methods are increasingly being applied to several real world problems, not only in medical research but also in bioinformatics, natural language processing, telecommunications, finance and forensic sciences. Although the application of these methods to neuroimaging data has yielded promising results, there are still significant theoretical and practical challenges for the translational implementation of the results in neurology and psychiatry. Firstly, most studies have been carried out on only a small number of participants, and it is hence not possible to draw definite conclusions on the diagnostic and prognostic value of neuroimaging at individual level. Secondly, MRI and other neuroimaging techniques are still only available to a small minority of the people who each year develop neurological and psychiatric disorders worldwide; whilst their availability is increasing in industrialised countries, this is likely to remain a long-lasting problem in third-world countries. Thirdly, the application of SVM or other supervised ML methods to neuroimaging data involves a series of analytical steps which require technical expertise and resources beyond the capabilities of most clinical units. Fourthly, a clinician treating a new patient with a serious disorder such as schizophrenia has to make a prompt decision on treatment, whereas the application of SVM to neuroimaging data can take several days to complete depending on the pre-processing of the images; it would be impractical and potentially harmful to the patient to delay a clinical decision until the results of such analysis became available. Fifth and finally, SVM would not be suitable for those subjects identified as having a gross neuroanatomical abnormality that is comorbid to their neurological or psychiatric illness. This is because by effectively introducing noise into the data such an abnormality would have the potential to malignly alter the given accuracy if generating a classifier for the first time, or produce a false classification if using a decision function that's already been generated.

There are also several outstanding questions of interest which could be addressed by future studies. For instance, it is often assumed that neuroimaging would allow more accurate diagnostic and prognostic assessment than demographic, clinical and cognitive information, but no previous studies have examined this. It would therefore be of great interest to examine the relative diagnostic and prognostic value of neuroimaging, demographic, clinical and cognitive in the main neurological and psychiatric disorder. Another question of interest is to what extent an algorithm developed in one research centre can be generalised to individuals examined in other research centres. As the many studies to use data from the ADNI database successfully discriminating between patient and control group images, produced from a range of centres and scanners, have shown there is strong evidence to suggest that in principle, a SVM classifier generated in one centre can indeed be generalised for use on data from a different centre. Another question of interest is whether the integration of different types of data within the same SVM model could provide a means of improving accuracy of classification and how such integration can be best implemented technically (see Yang et al., 2010; Hinrichs et al., 2011; Zhang et al., 2011 for three recent examples). Though the optimal method of integration will be continuously improved, it is likely to depend on several variables such as the type of data being combined, the analytical technique used to combine them and the disease under investigation. The results where multiple modalities have been integrated with positive results provide tentative evidence that integrative methods can be not only highly accurate, but most importantly, viable also. Whilst SVM allows categorical classification of each observation (e.g. responders vs. non-responders), there are also probabilistic ML methods that provide an estimate of the probability that a given observation belongs to each category (e.g. 80% responder, 20% non-responders) and therefore aim to quantify the uncertainty in each prediction. Two of the most promising probabilistic classification methods include Gaussian processes (Rasmussen and Williams, 2006) and relevance vector machines (Tipping, 2001), both of which are now beginning to show applications in neuroimaging (e.g. Marquand et al., 2010; Phillips et al., 2011). Another promising development is the application of regression methods to neuroimaging data, which aim to predict a continuous outcome (e.g. symptom severity) instead of a categorical class label. Indeed, pattern regression methods are now being applied to neuroimaging data in both health (Franke et al., 2010) and disease (Stonnington et al., 2010). Probabilistic classification and regression methods could both be particularly useful in a clinical setting, where clinicians have to make a balanced treatment decision after considering potential risks and benefits.

The difficult ethical implications that could potentially arise from the predictive nature of SVM must also be highlighted. For example, any future development of the methodology that could result in an individual being informed of an illness likely to occur at 'some stage' in the future, has profound psychological ramifications on the individual in question. Whilst beyond the scope of the current article, it is an issue that requires further discussion and substantial consideration.

In short, there is increasing interest in the application of supervised ML methods to neuroimaging data to aid diagnosis and prognosis in neurological and psychiatric disorders. The key advantage of these methods over traditional analytical techniques, is that they allow inferences to be made at the level of the individual and therefore could be used to inform treatment decisions in individual cases. Whilst there are significant theoretical and practical challenges to the translational implementation of this approach, the results of the studies published so far are encouraging and may provide the first steps towards the development of computer-based diagnostic and prognostic tools in neurology and psychiatry.

Acknowledgements

William Pettersson-Yeo is supported by a PhD studentship from the Medical Research Council (MRC). Andre Marquand gratefully acknowledges support from King's College Annual Fund and the King's College London Centre of Excellence in Medical Engineering. Andrea Mechelli is supported by an Investigator Award from the National Alliance for Research on Schizophrenia and Depression (NARSAD) Independent Award.

References

Abdulkadir, A., Mortamet, B., Vemuri, P., Jack, C.R., Krueger, G., Klöppel, S., Initiative, A.s.D.N., 2011. Effects of hardware heterogeneity on the performance of SVM Alzheimer's disease classifier. Neuroimage 58, 785–792.

Arimura, H., Yoshiura, T., Kumazawa, S., Tanaka, K., Koga, H., Mihara, F., Honda, H., Sakai, S., Toyofuku, F., Higashida, Y., 2008. Automated method for identification of patients with Alzheimer's disease based on three-dimensional MR images. Acad. Radiol. 15, 274–284.

Arnone, D., McIntosh, A.M., Ebmeier, K.P., Munafò, M.R., Anderson, I.M., 2011. Magnetic resonance imaging studies in unipolar depression: sdystematic review and meta-regression analyses. Eur. Neuropsychopharmacol. [Epub ahead of print].

Cannon, T.D., Cornblatt, B., McGorry, P., 2007. The empirical status of the ultra-risk (prodromal) research paradigm. Schizophr. Bull. 33, 661–664.

Chen, G., Ward, B.D., Xie, C., Li, W., Wu, Z., Jones, J.L., Franczak, M., Antuono, P., Li, S.-J., 2011. Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. Radiology 259 (1), 213–221.

Chincarini, A., Bosco, P., Calvini, P., Gemme, G., Esposito, M., Olivieri, C., Rei, L., Squarcia, S., Rodriguez, G., Bellotti, R., Cerello, P., De Mitri, I., Retico, A., Nobili, F.,

Initiative, A.s.D.N., 2011. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. Neuroimage 58, 469–480.
Chu, C., Hsu, A.-L., Chou, K.-H., Bandettini, P., Lin, C.-P., Initiative, f.t.A.s.D.N., 2011. Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. Neuroimage.
Costafreda, S.G., Chu, C., Ashburner, J., Fu, C.H., 2009a. Prognostic and diagnostic potential of the structural neuroanatomy of depression. PLoS One 4 (7), e6353.
Costafreda, S.G., Khanna, A., Mourao-Miranda, J., Fu, C.H.Y., 2009b. Neural correlates of sad faces predict clinical remission to cognitive behavioural therapy in depression. Neuroreport 20, 637–641.
Costafreda, S.G., Dinov, I.D., Tu, Z., Shi, Y., Liu, C.-Y., Kloszewska, I., Mecocci, P., Soininen, H., Tsolaki, M., Vellas, B., Wahlund, L.-O., Spenger, C., Toga, A.W., Lovestone, S., Simmons, A., 2011a. Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment. Neuroimage 56 (1), 212–219.
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., 2011b. Pattern of neural responses to verbal fluency shows diagnostic specificity for schizophrenia and bipolar disorder. BMC Psychiatr. 11, 18.
Cui, Y., Wen, W., Lipnicki, D.M., Beg, M.F., Jin, J.S., Luo, S., Zhu, W., Kochan, N.A., Reppermund, S., Zhuang, L., Raamana, P.R., Liu, T., Trollor, J.N., Wang, L., Brodaty, H., Sachdev, P.S., 2011a. Automated detection of amnestic mild cognitive impairment in community-dwelling elderly adults: a combined spatial atrophy and white matter alteration approach. Neuroimage.
Cui, Y., Liu, B., Luo, S., Zhen, X., Fan, M., Liu, T., Zhu, W., Park, M., Jiang, T., Jin, J.S., Initiative, A., 2011b. s.D.N.Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. PLoS One 6, e21896.
Davatzikos, C., Resnick, S.M., 2002. Degenerative age changes in white matter connectivity visualized in vivo using magnetic resonance imaging. Cereb. Cortex 12 (7), 767–771.
Davatzikos, C., Ruparel, K., Fan, Y., Shen, D.G., Acharyya, M., Loughead, J.W., Gur, R.C., Langleben, D.D., 2005. Classifyng spatial patterns of brain activity with machine learning methods: application to lie detection. Neuroimage 28 (3), 663–668.
Davatzikos, C., Fan, Y., Wu, X., Shen, D., Resnick, S.M., 2008. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiol. Aging 29 (4), 514–523.
Davatzikos, C., Bhatt, P., Shaw, L.M., Batmanghelich, K.N., Trojanowski, J.Q., 2010. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. Neurobiol. Aging [Epub ahead of print].
Duchesne, S., Caroli, A., Geroldi, C., Barillot, C., Frisoni, G.B., Collins, D.L., 2008. MRI-based automated computer classification of probable AD versus normal controls. IEEE Trans. Med. Imaging 27, 509–520.
Duchesne, S., Rolland, Y., Vérin, M., 2009. Automated computer differential classification in Parkinsonian Syndromes via pattern analysis on MRI. Acad. Radiol. 16, 61–70.
Dukart, J., Mueller, K., Horstmann, A., Barthel, H., Möller, H.E., Villringer, A., Sabri, O., Schroeter, M.L., 2011. Combined evaluation of FDG-PET and MRI improves detection and differentiation of dementia. PLoS One 6, e18111.
Ecker, C., Rocha-Rego, V., Johnston, P., Mourao-Miranda, J., Brammer, M.J., Murphy, C., Murphy, D., the, M.R.C., Consortium, A.I.M.S., 2010a. Investigating the predictive value of whole-brain structural MR scans in autism: a pattern classification approach. Neuroimage 49, 44–56.
Ecker, C., Marquand, A., Maurao-Miranda, J., Johnston, P., Daly, E.M., Brammer, M.J., Maltezos, S., Murphy, C.M., Robertson, D., Williams, S.C., Murphy, G.M., 2010b. Describing the brain in autism in five dimension-magnetic resonance Imaging-Assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. J. Neurosci. 30 (32), 10612–10623.
Ellison-Wright, I., Bullmore, E., 2010. Anatomy of bipolar disorder and schizophrenia: a meta-analysis. Schizophr. Res. 117 (1), 1–12.
Etkin, A., Wager, T.D., 2007. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am. J. Psychiatr. 164 (10), 1476–1488.
Fan, Y., Resnick, S.M., Wu, X., Davatzikos, C., 2008a. Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study. Neuroimage 41, 277–285.
Fan, Y., Batmanghelich, N., Clark, C.M., Davatzikos, C., 2008b. Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. Neuroimage 39, 1731–1743.
Ferrarini, L., Palm, W.M., Olofsen, H., van der Landen, R., van Buchem, M.A., Reiber, J.H.C., Admiraal-Behloul, F., 2008. Ventricular shape biomarkers for Alzheimer's disease in clinical MR images. Magn. Reson. Med. 59, 260–267.
Focke, N.K., Helms, G., Scheewe, S., Pantel, P.M., Bachmann, C.G., Dechent, P., Ebentheuer, J., Mohr, A., Paulus, W., Trenkwalder, C., 2011. Individual voxel-based subtype prediction can differentiate progressive supranuclear palsy from idiopathic Parkinson syndrome and healthy controls. Hum. Brain Mapp. 32, 1905–1915.
Franke, K., Ziegler, G., Klöppel, S., Gaser, C., 2010. Alzheimer's disease neuroimaging initiative estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. Neuroimage 50 (3), 883–892.
Friston, K.J., 2009. Modalities, modes, and models in functional neuroimaging. Science 326 (5951), 399–403.
Fu, C.H., 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. Psychiatr. 63, 656–662.
Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R.C., Ritchie, K., Broich, K., Belleville, S., Brodaty, H., Bennett, D., Chertkow, H., Cummings, J.L., de Leon, M., Feldman, H., Ganguli, M., Hampel, H., Scheltens, P., Tierney, M.C., Whitehouse, P., Winblad, B., 2006. Mild cognitive impairment. Lancet 367 (9518), 1262–1270.
Gerardin, E., Chetelat, G., Chupin, M., Cuingnet, R., Desgranges, B., Kim, H.-S., Niethammer, M., Dubois, B., Lehericy, S., Garnero, L., Eustache, F., Colliot, O., The Alzhemer's Disease Neuroimaging Inititive, 2009. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. Neuroimage 47, 1486–2476.
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.
Graña, M., Termenon, M., Savio, A., Gonzalez-Pinto, A., Echeveste, J., Pérez, J.M., Besga, A., 2011. Computer aided diagnosis system for alzheimer disease using brain diffusion tensor imaging features selected by Pearson's correlation. Neurosci. Lett. 502, 225–229.
Hahn, T., Marquand, F.A., Ehlis, A.C., Dresler, T., Kittel-Schneider, S., Jarczok, A.T., Lesch, K.P., Jakob, M.P., Mourao-Miranda, J., Brammer, J.M., Fallgatter, J.A., 2010. Integrating Neurobiological Markers of Depression. Arch Gen Psychiatry.
Haller, S., Nguyen, D., Rodriguez, C., Emch, J., Gold, G., Bartsch, A., Lovblad, K.O., Giannakopoulos, P., 2010. Invidividual prediction of cognitive decline in mild cognitive impairment using support vector machine-based analysis of diffisuion tensor imaging. J. Alzheimer's Dis. 22 (1), 315–327.
Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer-Verlag, New York, NY.
Haynes, D., Rees, G., 2006. Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. 7, 523–534.
Hinrichs, C., Singh, V., Xu, G., Johnson, S.C., 2011. The ADNI: predicitve markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population. Neuroimage 55, 574–589.
Ingalhalikar, M., Kanterakis, S., Gur, R., Roberts, T.P.L., Verma, R., 2010. DTI based diagnostic prediction of a disease via pattern classification. Med. Image Comput. Comput. Assist. Interv. 13, 558–565.
Khodayari-Rostamabad, A., Hasey, G.M., MacCrimmon, D.J., Reilly, J.P., de Bruin, H., 2010. A pilot study to determine whether machine learning methodologies using pre-treatment electroencephalography can predict the symptomatic response to clozapine therapy. Clin. Neuropsychol. 121, 1998–2006.
Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Rachael, I.S., Jonathan, D., Fox, R.N.C., Ashburner, J., Frackowiak, R.S.J., 2008a. Automatic classification of MR scans in Alzheimer's disease. Brain 131 (3), 681–689.
Klöppel, S., Draganski, B., Golding, C.V., Chu, C., Nagy, Z., Cook, P.A., Hicks, S.L., Kennard, C., Alexander, D.C., Parker, G.J.M., Tabrizi, S.J., Frackowiak, R.S.J., 2008b. White matter connections reflect changes in voluntary-guided saccades in pre-symptomatic Huntington's disease. Brain 131, 196–204.
Klöppel, S., Chu, C., Tan, G.C., Draganski, B., Johnson, H., Paulsen, J.S., Kienzle, W., Tabrizi, S.J., Ashburner, J., Frackowiak, R.S.J., Group, P.-H.I.o.t.H.S., 2009. Automatic detection of preclinical neurodegeneration: presymptomatic Huntington disease. Neurology 72, 426–431.
Koutsouleris, N., Meisenzahl, E.V., 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. Psychiatr. 66 (7), 700–712.
Koutsouleris, N., Borgwardt, S., Meisenzahl, E.M., Bottlender, R., Möller, H.-J., Riecher-Rössler, A., 2011. Disease prediction in the at-risk mental state for psychosis using neuroanatomical biomarkers: results from the FePsy study. Schizophr. Bull.
Lemm, S., Blankertz, B., Dickhaus, T., Muller, K.-R., 2011. Introduction to machine learning for brain imaging. Neuroimage 56 (2), 387–399.
Magnin, B., Mesrob, L., Kinkingnéhun, S., Pélégrini-Issac, M., Colliot, O., Sarazin, M., Dubois, B., Lehéricy, S., Benali, H., 2009. Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology 51, 73–83.
Marquand, A.F., Mourao-Miranda, J., Brammer, M.J., Cleare, A.J., Fu, C.H., 2008. Neuroanatomy of verbal working memory as a diagnostic biomarker for depression. Neuroreport 19 (15), 1507–1511.
Marquand, A., Howard, M., Brammer, M., Chu, C., Coen, S., Mourao-Miranda, J., 2010. Quantitative prediction of subjective pain intensity from whole-brain fMRI data using Gaussian processes. Neuroimage 49, 2178–2189.
McGorry, P.D., Yung, A.R., Phillips, L.J., Yuen, H.P., Francey, S., Cosgrave, E.M., Dominic, Germano, D., Bravin, J., McDonald, T., Blair, A., Adlard, S., Jackson, H., 2002. Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms. Arch. Gen. Psychiatr. 59, 921–928.
Mourao-Miranda, J., Reinders, A.A.T.S., Rocha-Rego, V., Lappin, J., Rondina, J., Morgan, C., Morgan, K.D., Fearon, P., Jones, P.B., Doody, G.A., Murray, R.M., Kapur, S., Dazzan, P., 2011. Individualized prediction of illness course at the first psychotic episode: a support vector machine MRI study. Psychol. Med., 1–11.
Muller, K.R., Mika, S., Ratsch, G., Tsuda, K., Scholkopf, B., 2001. An introduction to kernel-based learning algorithms. IEEE Neural Netw. 12 (2), 181–201.
Nho, K., Shen, L., Kim, S., Risacher, S.L., West, J.D., Foroud, T., Jack, C.R., Weiner, M.W., Saykin, A.J., 2010. Automatic prediction of conversion from mild cognitive impairment to probable Alzheimer's disease using structural magnetic resonance imaging. AMIA Annu. Symp. Proc., 542–546.

Nouretdinov, I., Costafreda, S.G., Gammerman, A., Chervonenkis, A., Vovk, V., Vapnik, V., Fu, C.H.Y., 2010. Machine learning classification with confidence: application of transductive conformal predictors to MRI-based diagnostic and prognostic markers in depression. NeuroImage.

Oliveira, P.P.d.M., Nitrini, R., Busatto, G., Buchpiguel, C., Sato, J.R., Amaro, E., 2010. Use of SVM methods with surface-based cortical and volumetric subcortical measurements to detect Alzheimer's disease. J. Alzheimers Dis. 19, 1263–1272.

Pettersson-Yeo, W., Allen, P., Benetti, S., McGuire, P., Mechelli, A., 2011. Dysconnectivity in schizophrenia: where are we now? Neurosci. Biobehav. Rev. 35 (5), 1110–1124.

Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fMRI: A tutorial overview. Neuroimage 45 (Suppl. 1), S199–S209 (Review).

Petersen, R.C., 2004. Mild cognitive impairment as a diagnostic entity. J. Intern. Med. 256 (3), 183–194.

Phillips, C., Bruno, M., Maquet, P., Boly, M., Noirhomme, Q., Schnakers, C., Vanhaudenhuyse, A., Bonjean, M., Hustinx, R., Moonen, G., Luxen, A., Laureys, S., 2011. "Relevance vector machine" consciousness classifier applied to cerebral metabolism of vegetative and locked-in patients. Neuroimage 15 (56), 797–808, 2.

Plant, C., Teipel, S.J., Oswald, A., Böhm, C., Meindl, T., Mourao-Miranda, J., Bokde, A.W., Hampel, H., Ewers, M., 2010. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. Neuroimage 50, 162–174.

Rasmussen, C., Williams, C.K.I., 2006. Gaussian Processes for Machine Learning. The MIT Press, Cambridge, MA.

Rosipal, R., Kramer, N.N., 2006. Overview and recent advances in partial least squares. In: Saunders, C., Grobelnik, M., Gunn, S., Shawe-Taylor, J. (Eds.), Subspace, Latent Structure and Feature Selection Techniques. Springer, Lecture Notes in Computer Science, pp. 34–51.

Salas-Gonzalez, D., Górriz, J.M., Ramírez, J., Illán, I.A., López, M., Segovia, F., Chaves, R., Padilla, P., Puntonet, C.G., 2010. Feature selection using factor analysis for Alzheimer's diagnosis using 18F-FDG PET images. Med. Phys. 37, 6084–6095.

Sartori, G., Pellegrini, S., Mechelli, A., 2011. Forensic neurosciences: from basic research to applications and pitfalls. Curr. Opin. Neurol. 24 (4), 371–377.

Schölkopf, B., Smola, A.J., 2002. Learning with kernels: support vector machines, regularization, optimization, and beyond.

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 (4), 3110–3121.

Smieskova, R., Fusar-Poli, P., et al., 2010. Neuroimaging predictors of transition to psychosis—a systematic review and meta-analysis. Neurosci. Biobehav. Rev. 34 (8), 1207–1222.

Stimpson, N., Agrawal, N., Lewis, G., 2002. Randomised controlled trials investigating pharmacological and psychological interventions for treatment-refractory depression. Systematic review. Br. J. Psychiatr. 181, 284–294.

Stonnington, C.M., Chu, C., Kloppel, S., Jack Jr., C.R., Ashburner, J., Frackowiak, R.S.J., 2010. Predicting clinical scores from magnetic resonance scans in Alzheimer disease. Neuroimage 51, 1405–1413.

Sun, D., van Erp, T.G.M., Thompso, P.M., Bearden, C.E., Daley, M., Kushan, L., 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. Psychiatr. 66 (11), 1055–1060.

Tipping, M., 2001. Sparse Bayesian learning and the relevance vector machine. J. Mach. Learn. Res. 1, 211–244.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. Springer.

Vapnik, V.N., 2000. The Nature of Statistical Learning Theory. Springer, NY.

Vemuri, P., Gunter, J.L., Senjem, M.L., Whitwell, J.L., Kantarci, K., Knopman, D.S., Boeve, B.F., Petersen, R.C., Jack Jr., C.R., 2008. Alzheimer's disease diagnosis in individual subjects using structural MR image: validation studies. Neuroimage 39 (3), 1186–1197.

Whitehead, A., Perdomo, C., Pratt, R.D., Birks, J., Wilcock, G.K., Evans, J.G., 2004. Donepezil for the symptomatic treatment of patients with mild to moderate Alzheimer's disease: a meta-analysis of individual patient data from randomised controlled trials. Int. J. Geriatr. Psychiatr. 19, 624–633.

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. Neurosci. 4, 1–9 (article 192).

Zakzanis, K.K., Graham, S.J., et al., 2003. A meta-analysis of structural and functional brain imaging in dementia of the Alzheimer's type: a neuroimaging profile. Neuropsychol. Rev. 13 (1), 1–18.

Zhang, D., Shen, D., Initiative, T.A., 2011a. s.D.N. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. Neuroimage.

Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., Initiative, A.s.D.N., 2011b. Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage 55, 856–867.