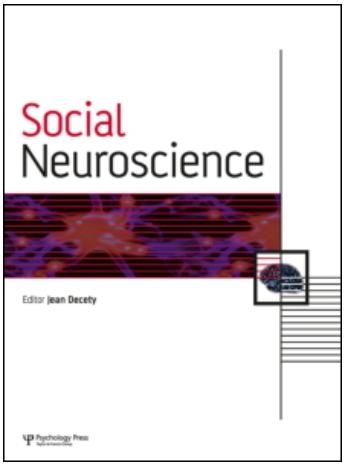

Social Neuroscience

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t741771143

Identification of psychopathic individuals using pattern classification of MRI images

João R. Sato[ab]; Ricardo de Oliveira-Souza[c]; Carlos E. Thomaz[d]; Rodrigo Basílio[c]; Ivanei E. Bramati[c]; Edson Amaro Jr[b]; Fernanda Tovar-Moll[ce]; Robert D. Hare[f]; Jorge Moll[c]

[a] Center for Mathematics, Computation, and Cognition, Universidade Federal do ABC, Santo André, Brazil [b] NIF/LIM44, Departamento de Radiologia da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil [c] Cognitive and Behavioral Neuroscience Unit, D'Or Institute for Research and Education, Rio de Janeiro, Brazil [d] Department of Electrical Engineering, Centro Universitário da FEI, São Bernardo do Campo, Brazil [e] Biomedical Sciences Institute, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil [f] Department of Psychology, University of British Columbia, and Darkstone Research Group, Vancouver, Canada

First published on: 16 May 2011

To cite this Article Sato, João R. , de Oliveira-Souza, Ricardo , Thomaz, Carlos E. , Basílio, Rodrigo , Bramati, Ivanei E. , Amaro Jr, Edson , Tovar-Moll, Fernanda , Hare, Robert D. and Moll, Jorge(2011) 'Identification of psychopathic individuals using pattern classification of MRI images', Social Neuroscience,, First published on: 16 May 2011 (iFirst)
To link to this Article: DOI: 10.1080/17470919.2011.562687
URL: http://dx.doi.org/10.1080/17470919.2011.562687

SOCIAL NEUROSCIENCE, 2011, i*First*, 1–13

Identification of psychopathic individuals using pattern classification of MRI images

João R. Sato[1,3], Ricardo de Oliveira-Souza[2], Carlos E. Thomaz[4], Rodrigo Basílio[2], Ivanei E. Bramati[2], Edson Amaro Jr[3], Fernanda Tovar-Moll[2,5], Robert D. Hare[6], and Jorge Moll[2]

[1]Center for Mathematics, Computation, and Cognition, Universidade Federal do ABC, Santo André, Brazil
[2]Cognitive and Behavioral Neuroscience Unit, D'Or Institute for Research and Education, Rio de Janeiro, Brazil
[3]NIF/LIM44, Departamento de Radiologia da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil
[4]Department of Electrical Engineering, Centro Universitário da FEI, São Bernardo do Campo, Brazil
[5]Biomedical Sciences Institute, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[6]Department of Psychology, University of British Columbia, and Darkstone Research Group, Vancouver, Canada

Background: Psychopathy is a disorder of personality characterized by severe impairments of social conduct, emotional experience, and interpersonal behavior. Psychopaths consistently violate social norms and bring considerable financial, emotional, or physical harm to others and to society as a whole. Recent developments in analysis methods of magnetic resonance imaging (MRI), such as voxel-based-morphometry (VBM), have become major tools to understand the anatomical correlates of this disorder. Nevertheless, the identification of psychopathy by neuroimaging or other neurobiological tools (e.g., genetic testing) remains elusive.
Methods/Principal findings: The main aim of this study was to develop an approach to distinguish psychopaths from healthy controls, based on the integration between pattern recognition methods and gray matter quantification. We employed support vector machines (SVM) and maximum uncertainty linear discrimination analysis (MLDA), with a feature-selection algorithm. Imaging data from 15 healthy controls and 15 psychopathic individuals (7 women in each group) were analyzed with SPM2 and the optimized VBM preprocessing routines. Participants were scanned with a 1.5 Tesla MRI system. Both SVM and MLDA achieved an overall leave-one-out accuracy of 80%, but SVM mapping was sparser than using MLDA. The superior temporal sulcus/gyrus (bilaterally) was identified as a region containing the most relevant information to separate the two groups.
Conclusion/significance: These results indicate that gray matter quantitative measures contain robust information to predict high psychopathy scores in individual subjects. The methods employed herein might prove useful as an adjunct to the established clinical and neuropsychological measures in patient screening and diagnostic accuracy.

Keywords: Psychopathy; Antisocial; Voxel-based morphometry; Moral; Machine learning.

Correspondence should be addressed to: João Ricardo Sato, Center of Mathematics, Computation, and Cognition, Universidade Federal do ABC, Avenida Atlântica, 420, Valparaíso, CEP: 09060-000, Santo André, SP, Brazil. E-mail: joao.sato@ufabc.edu.br or to Jorge Moll, IDOR, R. Diniz Cordeiro 30, 3 andar, CEP:22281-100, Rio de Janeiro, RJ, Brazil. E-mail: jorge.moll@idor.org

This research was supported by FAPERJ (Pronex and INNT grants), FAPESP, CAPES, and CNPq, as well as by intramural grants from IDOR, Brazil.

www.psypress.com/socialneuroscience DOI: 10.1080/17470919.2011.562687

Psychopathy, the first personality disorder recognized in psychiatry (Millon, Simonsen, Davis, & Birket-Smith, 2002), is defined by a cluster of interpersonal, affective, lifestyle, and antisocial traits and behaviors that include grandiosity, egocentricity, deceptiveness, lack of empathy or remorse, irresponsibility, impulsivity, and a tendency to violate social norms (Hare & Neumann, 2005). Psychopathy can be assessed in forensic settings by the Psychopathy Checklist–Revised (PCL–R) and in nonforensic contexts by the Psychopathy Checklist: Screening Version (PCL–SV), each supported by extensive evidence for their reliability and validity (Hare, 2006). As a standalone instrument for assessing psychopathy in civil psychiatric and community populations (de Oliveira-Souza, Ignácio, Moll, & Hare, 2008a; Guy & Douglas, 2006), the PCL–SV is strongly related to the PCL–R both conceptually and empirically (Cooke, Michie, Hart, & Hare, 1999). The widespread adoption of the PCL scales as a common metric for psychopathy has led to a dramatic increase in theoretical and empirical work, paving the way for research on its neurobiological substrates (Glenn & Raine, 2008). In particular, recent advances in magnetic resonance imaging (MRI) acquisition and processing, such as the methods for quantitative and automated assessment of brain structure (Good et al., 2001) have opened up new possibilities. Converging evidence indicates that the core features of the psychopathic personality are related to discrete volumetric changes in a set of frontotemporal and subcortical brain regions (De Brito et al., 2009; de Oliveira-Souza et al., 2008b; Müller et al., 2008; Tiihonen et al., 2008) that underlie moral cognition and behavior (Moll et al., 2005). These recent, quantitative, voxel-based studies have confirmed and extended the findings of earlier ones, which employed radioisotope and volumetric MRI manual tracing techniques (e.g., Soderstrom et al., 2002; Yang et al., 2005). Important progress has also been made on the more proximate causes of psychopathy; evidence suggests that the brain differences in psychopathic individuals are neurodevelopmental in nature, and arise from genetic and environmental factors (e.g., physical abuse) and their interactions (Bezdjian et al., 2011; Gao et al., 2010). A detailed account of these aspects is outside the scope of the current paper, and can be found in recent authoritative reviews (Gao, Glenn, Schug, Yang, & Raine, 2009).

Exciting as these neuroanatomical findings may be, their clinical utility remains elusive. Imaging studies on patient populations generally provide results on a group level, but their use for diagnostic classification of individual patients has yet to be established. The development of reliable and specific neuroanatomical biomarkers for psychiatric disorders——including psychopathy—would obviously be important. Typically, group studies using brain imaging rely on massive, voxel-by-voxel application of the general linear model (GLM) (Friston et al., 1995) or its particular cases, such as *t*-tests and ANOVA. This model is useful to provide statistical hypothesis testing for group differences or linear associations among variables. In neuroimaging, GLM is applied independently to all intracranial voxels, providing *p* values for group comparisons at each voxel. These *p* values are corrected for multiple comparisons, using the false discovery rate (FDR) (Benjamini & Hochberg, 1995) or random fields theory (RFT) (Worsley, 1995), and then compared to a prespecified significance level. This procedure is the core of univariate statistical analysis for brain mapping, by far the most frequently used approach. However, univariate analysis may not be the most suitable approach in clinical neuroimaging for two reasons: (1) the brain is organized in several highly structured networks and univariate, voxel-by-voxel analysis does not take into account this property, treating brain voxels independently; and (2) statistically significant differences do not necessarily mean cognitive or clinically relevant differences.

The interconnected structure of the brain implies that regions may influence one another, both structurally and functionally, and thus multivariate approaches may be more suitable than univariate ones (Lukic, Wernick, & Strother, 2002). Furthermore, groups or conditions may be characterized by the topology or changes of these relationships (Sato et al., 2008a). In addition, statistical hypothesis testing for group differences is an inferential procedure that compares parameters between two populations based on measures calculated over the samples. Therefore, assuming that an adequate test is applied, any difference in a parameter of interest between two populations will be detected, provided the sample sizes are large enough. Thus, complementary to statistical tests, one major point of concern is whether or not differences between groups can be used to allocate each subject to a particular group based on *a priori* defined individual variables (e.g., symptom clusters, diagnostic categories, genetic markers). This is, in fact, an ultimate challenge of clinical diagnosis research, with vast implications for diagnosis and treatment. In the realm of imaging techniques, functional and structural MRI findings have been increasingly employed as "intermediate phenotypes" or "endophenotypes," capturing at a meso- or macroscopic level features that reflect complex and often subtle factors, including environmental,

genetic, and epigenetic ones (Meyer-Lindenberg & Weinberger, 2006). The potential of this approach is reflected by the rapidly growing number of studies using quantitative techniques such as diffusion tensor imaging and voxel-based morphometry (VBM) in this context (Bertisch et al., 2010; Bradley et al., 2009; Camchong, Lim, Sponheim, & Macdonald, 2009; Honea et al., 2008).

Statistical learning or pattern recognition methods have become attractive approaches in computer-aided diagnosis, mostly because they can be applied in a multivariate fashion and they provide classification rules for predicting the group membership of a *new* subject. In the last decade, multivariate pattern recognition providing a joint analysis of all voxels was proposed for neuroimaging analysis (Fan, Shen, & Davatzikos, 2005; Golland, Grimson, Shenton, & Kikinis, 2000; Golland et al., 2002; Lao et al., 2004; Lukic et al., 2002; Sato et al., 2008b, 2009; Thomaz et al., 2007a, 2007b). Nevertheless, only a few studies have explored the predictive power of these approaches in neurological or neuropsychiatric disorders. Emblem et al. (2008) applied support vector machines (SVM) to predict glioma grades, using perfusion images. Gerardin et al. (2009) used hippocampal shape to classify Alzheimer's disease, mild cognitive impairment, and controls with an accuracy of 94%. A classification method for primary progressive aphasia was developed by Wilson et al. (2009), showing good accuracy and generalization power. Ecker et al. (2010) evaluated the predictive power of SVM for whole-brain structural images in autism (gray matter VBM), and also demonstrated good discriminative power between patients and normal controls (specificity of 86% and sensitivity of 88%). Davatzikos, Bhatt, Shaw, Batmanghelich, & Trojanowski, (2010) have shown that it is possible to predict the conversion of mild cognitive impaired patients to Alzheimer's disease, based on a combined analysis between VBM and spatial patterns of abnormalities. Recently, Koutsouleris et al. (2010) studied the prediction of vulnerability and transition to psychosis by using support vector regression.

Despite the clinical and societal relevance of psychopathy, pattern classification of neuroimaging data has not yet been employed in this condition. In this paper, we evaluate the application of pattern recognition methods to gray-matter images, focusing on distinguishing individuals with psychopathy from normal controls. Two classification methods were investigated: SVM and maximum uncertainty linear discriminant analysis (MLDA). Furthermore, we introduced an approach for feature selection to improve classification rates, a useful tool for general subject/group classification when dealing with relatively small samples of brain imaging data. The main goal of the present study was to investigate whether classifiers could discriminate between subject groups by multivariate pattern analysis of whole-brain gray matter voxels, which is essentially distinct from attempting to map which voxels show statistical differences between subject samples. From previous studies (De Brito et al., 2009; de Oliveira-Souza et al., 2008b; Moll, Zahn, de Oliveira-Souza, Krueger, & Grafman, 2005; Müller et al., 2008; Tiihonen et al., 2008), we had, nonetheless, some *a priori* expectations about which brain regions would likely shelter discriminant voxels. These regions included the anterior and ventral sectors of the prefrontal cortex and the superior temporal sulcus region, which were consistently activated in several functional MRI studies on control subjects (see Moll et al., 2005, for a review), and were found to be structurally abnormal in psychopaths (de Oliveira-Souza et al., 2008b; Müller et al., 2008; Tiihonen et al., 2008).

MATERIAL AND METHODS

All participants provided written informed consent before entering the study, which was approved by the D'Or Institutional Review Board (Rio de Janeiro, Brazil). The 15 patients (8 men, 7 women) who agreed to undergo MRI scanning were part of a larger group of 50 patients with neurological and/or neuropsychiatric disorders who were brought to consultation by relatives or acquaintances for a variety of emotional and behavioral problems (de Oliveira-Souza et al., 2008b). Each patient fulfilled the DSM-IV adult criteria for antisocial personality disorder (American Psychiatric Association, 1994) and was assessed with the PCL–SV as the primary measure of interest (Hart, Cox, & Hare, 1995). Their occupational history was erratic and unstable. They lived in the community, but eventually came to medical attention due to chronic and recurrent misbehaviors, which did not result in criminal prosecution. The control group included 15 normal volunteers matched on gender, age, and education, and without a history of neurological or psychiatric disorders or serious misconduct. Further details on participants' characteristics and behavioral results can be found elsewhere (de Oliveira-Souza et al., 2008b). Briefly, there were no significant differences between groups in gender, age, education, handedness, global cognitive status, and executive performance (Table 1).

TABLE 1
Demographic and neuropsychological information of the sample evaluated in the current study

	Controls	*Patients*
Education (years)	11 ± 2	11 ± 2
Age (years)	32 ± 13	32 ± 14
MMSE (0–30)	28.9 ± 1.1	28.7 ± 1.9
Handedness (R/L)	13/2	14/1
PCL–SV (0–24)	0.4 ± 1.0	17.8 ± 3.8
WCST		
Categories completed (0–6)	5.4 ± 1.4	5.0 ± 1.6
Perseverative errors (0–127)	14 ± 12	18 ± 11
Set failures (0–22)	0.80 ± 1.1	1.2 ± 1.0

Notes: MMSE: Mini Mental Status Examination; PCL–SV: Psychopathy Checklist, Screening Version; WCST: Wisconsin Card Sorting Test. Patients and controls were matched in all scores, except for the PCL–SV (t-test, $p < .05$).

Image acquisition

MRI scans were acquired at the Department of Radiology at Barra D'Or Hospital, using a 1.5 Tesla MR System (Siemens Medical Systems, Erlangen, Germany), with a standard quadrature head coil. For each volunteer, a high-resolution, T1-weighted, 3D structural volume was obtained (MPRAGE pulse sequence, TR = 9.7 ms, TE = 4 ms, TI = 300 ms, flip angle = 12°, field of view = 256 mm, slice thickness = 1.25 mm, matrix size = 256 × 256, 128 sagittal slices, in-plane resolution of 1 mm × 1 mm).

Image preprocessing

Automated preprocessing of structural images was carried out using the package SPM2 (Wellcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm). The optimized VBM protocol was used (for details on the preprocessing steps, see the Methods section in de Oliveira-Souza et al. (2008b). For the purpose of the present investigation, the normalized, smoothed (Full width at half maximum = 12 mm) unmodulated gray matter (GM) images (corresponding to GM "concentration," or GMC) were employed for all subsequent pattern classification analyses. The statistical steps of VBM processing and the resulting topographical maps previously reported in (de Oliveira-Souza et al., 2008b) were disregarded for the purposes of the present study, and did not affect the statistical inferences of reported herein.

SVM and maximum uncertainty linear discrimination

The typical task of the statistical learning (or classification) methods is to use the features provided by the previous stages in this work the preprocessed VBM gray matter values to assign the object of interest to a specific group or class. This assignment can be done directly, using risk minimization-based approaches such as SVM (Vapnik, 1998), or indirectly as in the spectral multivariate analysis of the data with the Linear discriminant analysis (LDA)-based approaches (Devijver & Kittler, 1982; Fukunaga, 1990). Both linear kernel SVM and LDA are discriminant methods that seek to find a classification boundary that separates data into different groups with maximum precision. Recent studies have suggested that these two approaches can be successfully applied to neuroimaging data sets (Mourão-Miranda, Bokde, Born, Hampel, & Stetter, 2005; Sato et al., 2008b; Thomaz et al., 2007b). This is because, by being linear methods, they allow the quantification of discriminative information contained at each predictor variable (voxel), which can be directly obtained from the separating hyperplane coefficients. There are, however, important differences between these two approaches on extracting and classifying discriminating information from data.

The primary purpose of SVM is to maximize the width of the margin between two distinct sample classes (Vapnik, 1998). Given a training set that consists of N pairs of $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$, where x_i denotes the K-dimensional training observations and $y_i \in \{-1, +1\}$ the corresponding classification labels, the SVM method seeks to find the hyperplane defined by

$$f(x) = (x \cdot w) + b = 0,$$

which separates positive and negative observations with the maximum margin. It can be shown that the solution vector w_{svm} (hyperplane coefficients) is defined in terms of a linear combination of the training observations; that is,

$$w_{svm} = \sum_{i=1}^{N} \alpha_i y_i x_i,$$

where α_i are non-negative coefficients obtained by solving a quadratic optimization problem with linear inequality constraints. Those training observations x_i with non-zero α_i lie on the boundary of the margin and are called support vectors (Vapnik, 1998). In the

present study, only the linear kernel was applied, and the cost parameter was set to 1. The description of the SVM solution does not make any assumption about the distribution of the data, focusing on the observations that lie close to the opposite class; that is, on the observations that most count for classification (Hastie, Tibshirani, & Friedman, 2001).

The LDA solution, on the other hand, is a spectral matrix analysis of the data and is based on the assumption that each class can be represented by its distribution of data; that is, the corresponding mean vector (or class prototype) and covariance matrix (or spread of the sample group) (Hastie et al., 2001). In other words, LDA depends on all of the data, even points far away from the separating hyperplane; its main objective is to find a projection matrix W_{lda} that maximizes Fisher's criterion (Fukunaga, 1990):

$$\frac{\left|W^T S_b W\right|}{W^T S_w W},$$

where S_b and S_w are respectively the between- and within-class scatter matrices. Fisher's criterion is maximized when the projection matrix W_{lda} is composed of the eigenvectors of $S_w^{-1}S_b$ with at most number of classes – 1 non-zero eigenvalues (Devijver & Kittler, 1982). In the case of a two-class problem, the LDA projection matrix is, in fact, the leading eigenvector w_{lda} of $S_w^{-1}S_b$, assuming that S_w is invertible. However, in limited sample and high dimensional problems, such as the one under investigation, S_w is either singular or mathematically unstable, and the standard LDA cannot be used for the classification task. To avoid these critical issues, we have calculated the leading eigenvector w_{lda} (hyperplane coefficients) by a MLDA that considers the issue of stabilizing the S_w estimate with a multiple of the identity matrix (Thomaz, Kitani, & Gillies, 2006).

Classification and brain mapping

As described in the previous section, SVM and MLDA are classifiers based on finding a discriminative hyperplane, where the decision to which group a subject belongs is achieved by projecting the gray matter maps onto this hyperplane. For each voxel, there is an associated coefficient defining the discriminative hyperplane, which is fully specified by the set of coefficients of all voxels. The absolute value of this coefficient is a measure of how the related voxel predicts the grouping of the subjects; that is, it is an index of the amount of discriminative information contained in this brain region.

A well-known obstacle to the application of classification methods in neuroimaging is the huge dimensionality of the data. In most VBM studies, the number of subjects is in the order of tens, while the number of voxels is in the order of hundreds of thousands. Since we are interested in using the gray matter values at each voxel to predict the class of a subject, this means that we have thousands of variables to predict the class of tens of subjects. This obstacle, known as the "curse of dimensionality," may lead to overfitting; that is, the classifiers are excellent for predicting the subjects used to define the discriminative hyperplane, but they may perform poorly in predicting the class for a new subject. In other words, if a subject is used to train the classifier, the classifier would provide an accurate prediction for this particular subject, but the generalization power for a new individual is not guaranteed. In keeping with this limitation, all accuracies presented in this paper will be based on a leave-one-subject-out cross-validation procedure. This approach consists in leaving one subject out of the training set, training the classifier with the remaining subjects, and then evaluating the generalization power of the classifier by testing the excluded subject.

Despite the fact that both SVM and MLDA are pattern recognition approaches developed to deal with high-dimensional data, the inclusion of confounding variables with little relevant information, (i.e., noise) as predictors may lead to low accuracy rates. A feature (in this case, a voxel) selection step may be useful to improve the accuracy rates. The elimination of voxels that did not contain discriminative information to differentiate the groups may also be useful for brain mapping, since the relevant voxels are identified. Thus, the feature selection step is important to avoid overfitting and is also suitable for brain mapping.

In this paper, image analyses and processing were carried out in the following steps (Figure 1):

1. Process the data with the VBM pipeline to obtain unmodulated gray matter concentration maps (GMC) for each subject. This pipeline includes image intensity normalization, spatial normalization, and spatial smoothing and segmentation for cerebrospinal fluid, and gray and white matter.
2. Mask the volumes for considering only intracranial and gray matter high-probability voxels. This step is important to select only voxels with relevant and interpretative gray matter coefficients.
3. Leave one subject out of the sample.
4. Build a feature matrix X, where the columns correspond to voxels (K features), and each row contains the data of each subject (N individuals).

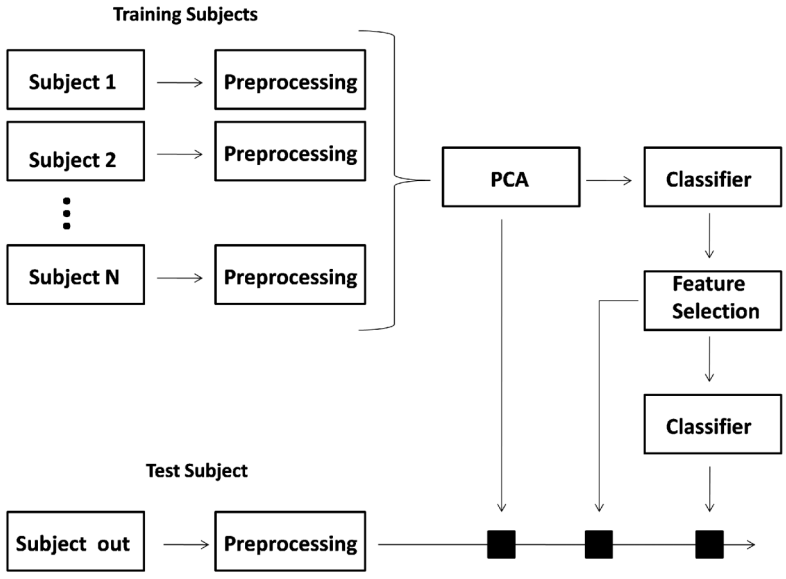

Figure 1. Diagram illustrating the steps for leave-one-out classification and feature selection implemented in this study. The subject left out in each iteration of the leave-one-out procedure is used to evaluate the prediction accuracy. PCA, principal component analysis.

5. Normalize each column of matrix X to have mean equal to zero and variance equal to one, obtaining a feature matrix Z.
6. Train the classifier (SVM or MLDA), using the feature matrix Z and the label (groups) of the subjects included in this matrix.
7. Rank each voxel (feature) by its level of discriminative information measured by the absolute coefficients in hyperplane vector w. Apply the feature selection, keeping the subset (Q%) of more informative voxels set. Note that this step also produces a mask of voxels containing the discriminant information.
8. Retrain the classifiers for this subset of "most informative" voxels.
9. Process the data of the subject-out for normalization (step 4) and feature selection (step 8), and predict his or her group, using the classifier trained at Step 8.
10. Return to Step 3 until all subjects have been processed.

It is important to emphasize that in this procedure, the feature selection is applied for each subject-out. This implies that the voxels used by the classifiers in each leave-one-out may differ. Thus, we propose a brain-mapping strategy based on the proportion of overlap of informative voxels masks (obtained at step 7) between all leave-one-out loops. Note that, in practice, the solution will be sparse, since only a small number of voxels contain discriminative information. Furthermore, the overlap proportion at each voxel allows the evaluation of the procedure's robustness. In this study, these steps were carried out for different values of Q (0.01%, 0.05%, 0.10%, and 1%), in order to evaluate the performance of the nested subset of voxels. In addition, it is important to emphasize that is not necessary to define regions of interest (ROI) *a priori*, and that the automated identification of relevant voxels is not based on mass-univariate statistical tests; instead, relevant voxels are identified in a multivariate fashion from the voxel maps obtained in step 7. The most discriminant regions were based on feature ranking. Furthermore, the discrimination maps were used solely to make sure that the results were not driven by spurious signals (e.g., border effects, CSF), but from brain regions relevant to social cognition and behavior.

RESULTS

The estimated classification rates for $Q\% = \{0.01\%, 0.05\%, 0.10\%, \text{and } 1\%\}$ were {70%, 80%, 73%, 60%} and {67%, 77%, 80%, 70%} for SVM and MLDA, respectively. Note that the classifiers achieved an overall accuracy of 80%, but to provide this rate, MLDA uses the 0.1% highest discriminative features, while SVM is more parsimonious, requiring only 0.05%. Figure 2 highlights the brain regions containing the discriminative information used by the classifiers, after the feature selection step. These maps point out that the unmodulated gray matter coefficient at the left superior temporal gyrus/sulcus (STG/STS)

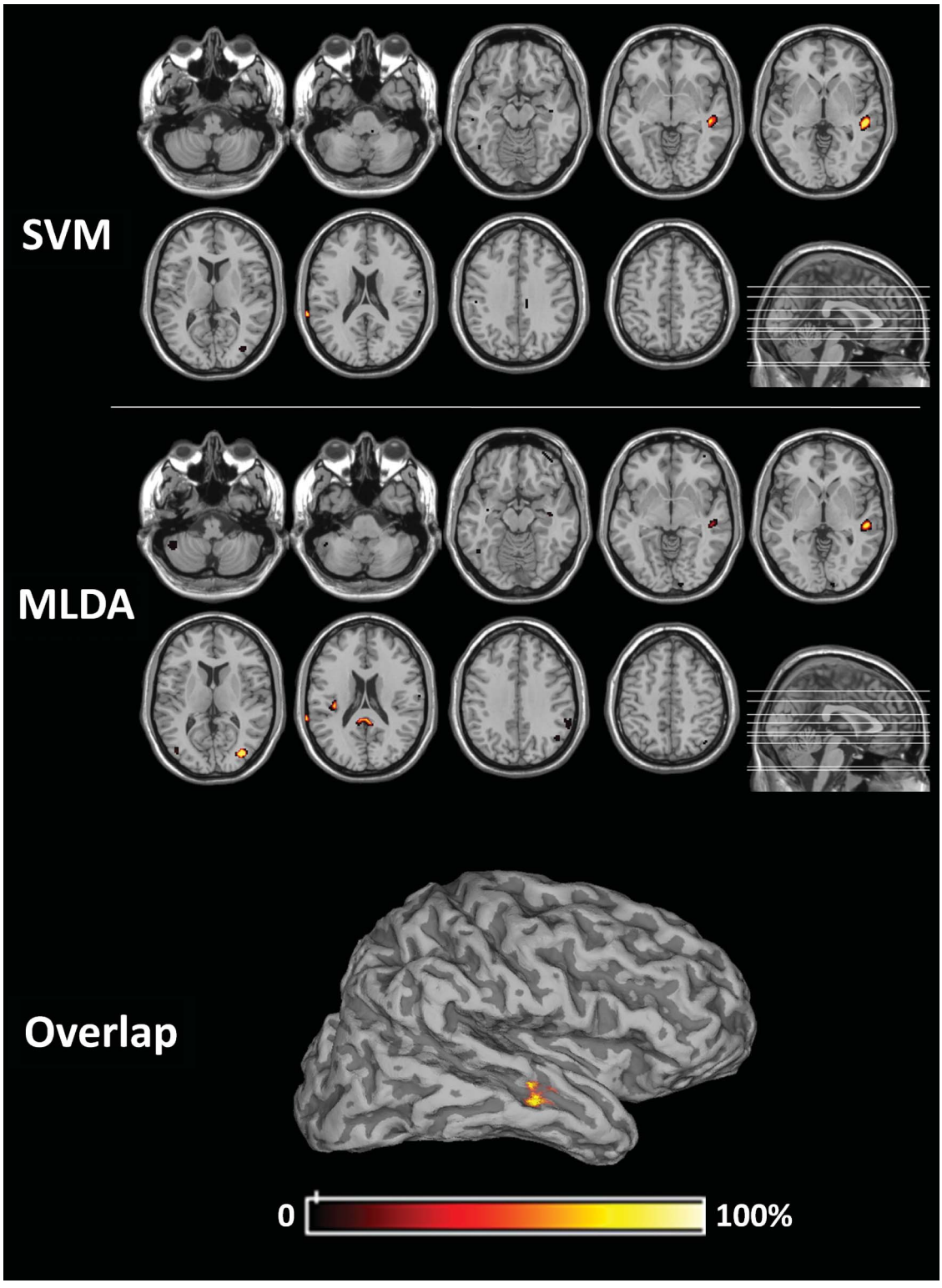

Figure 2. Brain mapping of regions containing the information used by the classifiers to discriminate high PCL–SV patients from controls. The color scale describes the proportion of leave-one-out interactions that selected the respective voxel at feature selection step (consistency).

TABLE 2
Statistical information at local maxima from discriminative clusters with overlap proportion across subjects greater than 70%

	X	*Y*	*Z*	*Side*	*Overlap*	*Size (voxels)*	*BA*	*Area*
SVM	48	−28	−4	R	100%	121	22	Superior temporal sulcus
	−68	−38	18	L	100%	21	22	Superior temporal gyrus
	6	−40	22	R	100%	143	26	Isthmus of cingulate gyrus
MLDA	36	−82	8	R	100%	96	19	Occipital peristriate cortex
	52	−28	−2	R	100%	88	22	Superior temporal sulcus

is an important predictor for both classifiers. On the other hand, since MLDA uses more discriminant regions, it also indicates that the coefficients at the right STG/STS, left occipital cortex, and posterior cingulate gyrus contain relevant information to predict the classes. The statistical information about the clusters of discriminative regions is presented in Table 2.

For each classifier, the leave-one-out projections of each subject onto the discriminative hyperplane space are shown in Figure 3. This figure indicates that the variability within each group is approximately the same at this space. Subject 13 stood out as a "control group" outlier for both classifiers, since his projection was far from the decision boundary and at the control group space set. Finally, note that the misclassified subjects are not exactly the same for the two classifiers, but most of them fall close to the decision boundary.

In addition, in order to explore the sensitivity and specificity of MLDA and SVM, Figure 4 describes the Receiver-operator-characteristic (ROC) curves built by ranking the leave-one-out decision values of each classifier. Note that in some points of the curve the SVM may achieve a specificity and sensitivity of 80% and 86.7%, and MLDA achieved 86.7% and 80%, respectively.

DISCUSSION

The study of the neural basis of psychopathy is of great relevance to the understanding of this severe disorder. This is reflected in the recent surge of electrophysiological, structural, and functional MRI investigations (Blair, Peschardt, Budhani, Mitchell, Pine, & 2006; Fullam, McKie, & Dolan, 2009; Glenn, Raine, & Schug, 2009; Glenn, Raine, Yaralian, & Yang, 2010; Kiehl, Bates, Laurens, Hare, & Liddle, 2006; Müller et al., 2008; Rilling et al., 2007; Veit et al., 2010). In this study, we explored the applicability of multivariate machine learning techniques to psychopathy diagnosis based on gray matter indexes resulting from VBM data preprocessing. Importantly, no group statistical inferences were made by VBM procedures; instead, linear SVM and MLDA classifiers were used to discriminate between patients and controls based only on their GM images ("GM concentration"). In addition, we showed that these approaches may also provide anatomical information on the brain mapping of regions containing relevant information used to discriminate between the two classes.

From the neurobehavioral and anatomical perspective, the STS region identified by SVM and MLDA has figured prominently in several studies of social cognition and emotion, being implicated in social feature representation (e.g., emotional faces and body posture), intentionality inferences, empathy, and moral sentiments such as guilt, compassion, and embarrassment (Decety, Chaminade, Grèzes, & Meltzoff, 2002; de Gelder, 2006; Grèzes, Pichon, & de Gelder, 2007; Kalbe et al., 2009; Leibenluft, Gobbini, Harrison, Haxby, & 2004; Materna, Dicke, & Thier, 2008). Although acquired damage to this region has not so far been associated with the development of severe antisocial behaviors, it has been suggested that this region may critically work in concert with other frontal, temporal, and subcortical regions to enable complex social and emotional abilities, such as interpersonal feelings. Despite these converging lines of evidence, the problem of why the STS showed up more prominently in the present study, and the neuroanatomical implications of this finding for the diagnosis of psychopathy, are matters of careful analyses, as discussed below.

The goals of conventional brain-mapping methods, such as independent samples group comparisons of means, using VBM or ROI volumetry and pattern classification methods, are essentially distinct. In the former, the aim is to reveal statistical differences between groups on a voxel or ROI basis, independently of how well this finding may help categorize a given subject in a diagnostic group. The pattern recognition methods focus less on how much information a given region carries individually, and more

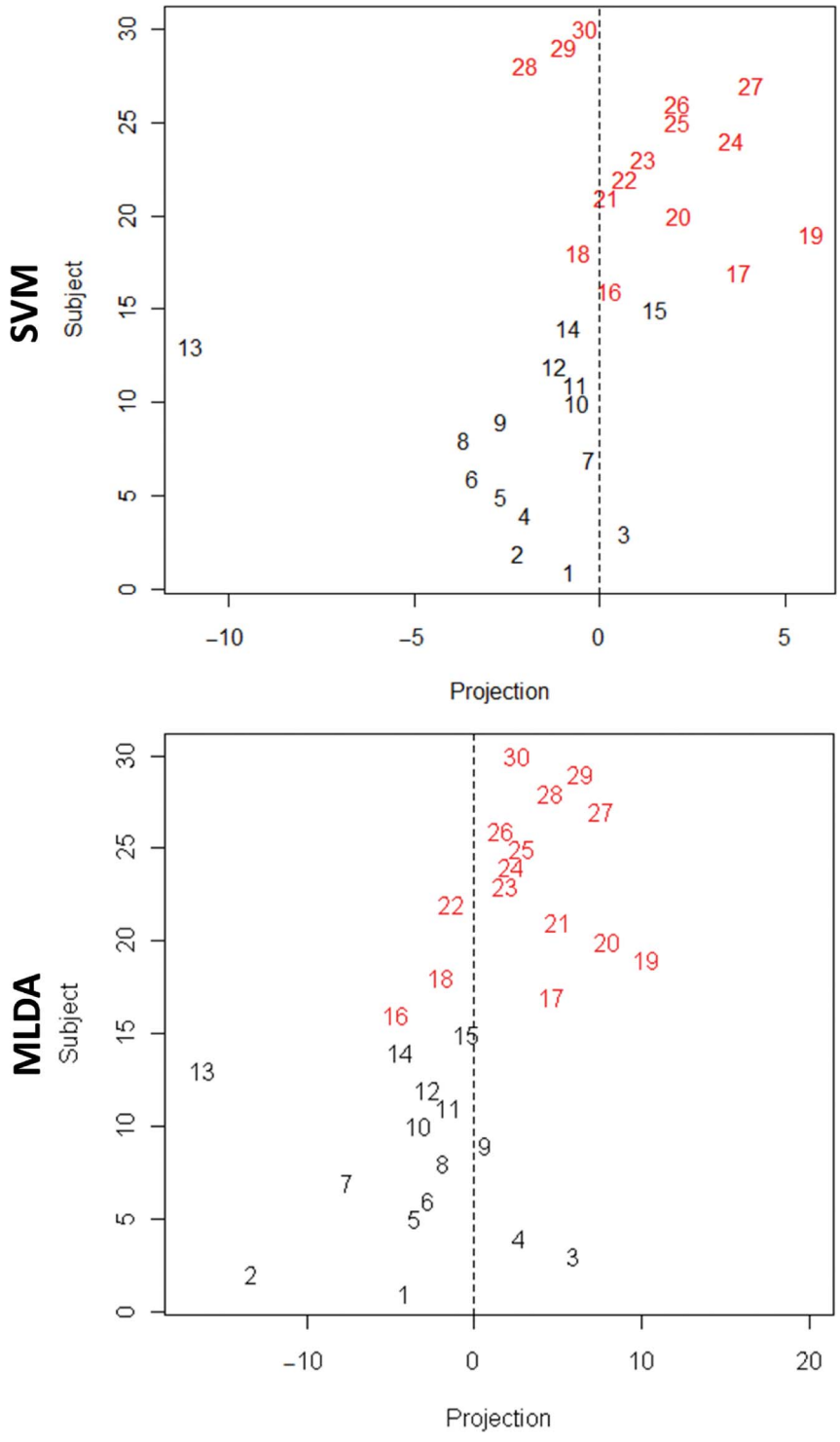

Figure 3. Leave-one-out projections of subjects onto discriminative hyperplane. The decision boundary is at zero in the *x*-axis. The black and red points describe the control and patient groups, respectively.

on how well a combination of different brain regions contributes to the correct classification of a given individual into diagnostic categories. The present results therefore show that the gray matter concentration in the right STS, and to a lesser degree in a few other frontoparieto-occipital regions (including the left

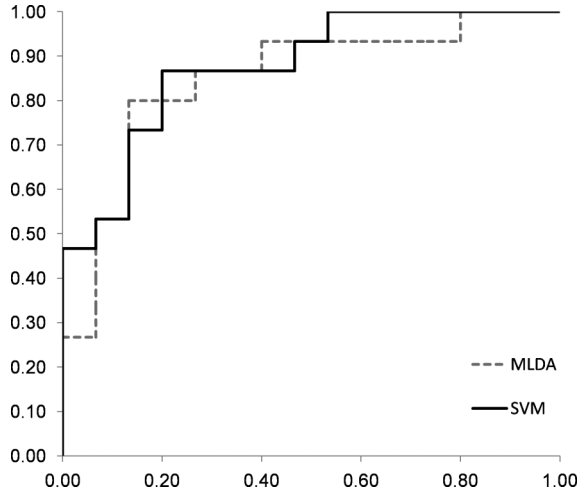

Figure 4. ROC curves for MLDA and SVM obtained by ranking the leave-one-out decision values of each classifier.

STS), when classified according to SVM and MLDA, can reach moderate-to-high accuracies in discriminating patients with high PCL– SV from normal individuals. Thus, the present results emphasize that classification methods are promising in predicting diagnostic categories based on subtle differences in brain structure.

Another advantage of pattern classification methods over more classical approaches is that they are not intrinsically dependent on contiguous spatial relationships of voxels. For example, the influence of anatomical gyral variations is less critical for pattern classification than for methods dependent on contiguous variations (i.e., clustering, Gaussian fields theory). It follows that these techniques can be useful in alterations involving subtle changes spread across the cortical mantle, regardless of spatial distribution. Thus, the alterations can be uniformly distributed, grouped in clusters, or any combination of these—and all could be picked by pattern classification methods.

Although, the potential clinical applicability of this method is straightforward, a number of technical shortcomings and ethical implications must be addressed. Potentially, pattern classification of brain data can increase the diagnostic accuracy of psychopathy, and even help to determine pathological subtypes—but the practical value of this method has yet to be established. One possibility would be to use the method, in conjunction with other neurobehavioral tests, as a supplement to clinical rating tools, such as PCL–R or PCL–SV. Similarly, it is possible that the addition of pattern classifications of brain data will add to the utility of current procedures used to assess the risk of crime and violence, and to select treatment options for psychopathy. Medical diagnostic imaging in general might benefit from pattern classification methods, since interpretation of single-case images is a challenging task, exacerbated by technological advances that may lead to increased diagnostic sensitivity at the expense of specificity.

The small sample size is a limitation of the current study. Leave-one-out and feature selection were implemented in order to avoid results that arise from overfitting. However, leave-one-out accuracy rates were statistically greater than chance ($p < .05$, for binomial distribution with probability of success equal to 50%). It should also be emphasized that discriminating subtle cortical changes in individuals with high levels of psychopathy from healthy controls is less challenging than discriminating psychopathy from other psychiatric conditions associated with a variety of cortical abnormalities, such as borderline personality, drug abuse, attention deficit-hyperactivity disorder, and other neuropsychiatric disorders (Georgopoulos et al., 2010; Nardo et al., 2010; Schaufelberger et al., 2007; Schlaepfer et al., 2006; Zhu et al., 2005). Critically, reduced gray matter volume in the STS region, which had the greatest discriminative power in this study, is found not only in psychopathy (de Oliveira-Souza et al., 2008b; Müller et al., 2008), but also in schizotypal and borderline personality disorders (Goldstein et al., 2009). Moreover, abnormalities in several frontotemporo-limbic regions are implicated in both psychopathy and other psychiatric disorders (Benetti et al., 2010; Brunner et al., 2010; Soloff, Nutche, Goradia, & Diwadkar, 2008; Zou et al., 2010). Clearly, much more research is needed to determine the ability of classification algorithms to discriminate among different conditions, especially given the fact that patients may show multiaxial patterns of psychopathology. Future studies should investigate how the anatomical distributions of cortical anomalies relate to the diagnostic accuracies in discriminating among disorders, not only in offender and patient samples but also in various other settings, such as the corporate world.

Another caveat when interpreting the anatomical results emerging from pattern classification, as well as from other anatomical and functional imaging data, relates to the fact that the pathogenesis of psychopathy are multifactorial, emerging from a range of neurological, genetic, and environmental ingredients (Weber, Habel, Amunts, & Schneider, 2008). Combining data from genetic biomarkers, neuropsychology, and multimodal imaging may be extremely helpful in furthering our understanding of complex disorders such as psychopathy. It is likely that classifiers using these

different types of data will prove to be even more useful and powerful than imaging data alone.

In summary, the results of the current study suggest that it is possible to discriminate patients with psychopathy from healthy controls, with relatively high sensitivity and specificity rates, solely by measuring the gray matter MRI features. Future studies may extend the present approach by employing multivariate approaches to predict the psychopathy scores in larger patient samples. In addition, pattern classification of functional MRI experiments addressing moral judgments and feelings in psychopathy (Glenn et al., 2009; Moll et al., 2002, 2007; Veit et al., 2010) may reveal how functional impairments of specific frontotemporolimbic networks can facilitate the identification of individuals with this severe disorder.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Benetti, S., McCrory, E., Arulanantham, S., De Sanctis, T., McGuire, P., & Mechelli, A. (2010). Attachment style, affective loss and gray matter volume: A voxel-based morphometry study. *Human Brain Mapping*, *1*(10), 1482–1489.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*, 289–300.

Bertisch, H., Li, D., Hoptman, M. J., & Delisi, L. E. (2010). Heritability estimates for cognitive factors and brain white matter integrity as markers of schizophrenia. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *153B*(4), 885–894.

Bezdjian, S., Raine, A., Baker, L. A., & Lynam, D. R. (2011). Psychopathic personality in children: Genetic and environmental contributions. *Psychological Medicine*, *41*, 589–600.

Blair, R. J. R., Peschardt, K. S., Budhani, S., Mitchell, D. G. V., & Pine, D. S. (2006). The development of psychopathy. *Journal of Child Psychology and Psychiatry*, *47*, 262–276.

Bradley, D., Whelan, R., Walsh, R., Reilly, R. B., Hutchinson, S., Molloy, F., et al. (2009). Temporal discrimination threshold: VBM evidence for an endophenotype in adult onset primary torsion dystonia. *Brain*, *132*(9), 2327–2335.

Brunner, R., Henze, R., Parzer, P., Kramer, J., Feigl, N., Lutz, K., et al. (2010). Reduced prefrontal and orbitofrontal gray matter in female adolescents with borderline personality disorder: Is it disorder specific? *NeuroImage*, *49*, 114–120.

Camchong, J., Lim, K. O., Sponheim, S. R., & Macdonald, A. W. (2009). Frontal white matter integrity as an endophenotype for schizophrenia: Diffusion tensor imaging in monozygotic twins and patients' nonpsychotic relatives. *Frontiers in Human Neuroscience*, *3*, 35.

Cooke, D. J., Michie, C., Hart, S. D., & Hare, R. D. (1999). Evaluating the screening version of the Hare Psychopathy Checklist–Revised (PCL:SV): An item response theory analysis. *Psychological Assessment*, *11*(1), 3–13.

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. *Neurobiology of Aging*. Advance online publication. DOI: 10.1016/j.neurobiolaging.2010.05.023

De Brito, S. A., Mechelli, A., Wilke, M., Laurens, K. R., Jones, A. P., Barker, G. J., et al. (2009). Size matters: Increased grey matter in boys with conduct problems and callous-unemotional traits. *Brain*, *132*(4), 843–852.

Decety, J., Chaminade, T., Grèzes, J., & Meltzoff, A. N. (2002). A PET exploration of the neural mechanisms involved in reciprocal imitation. *NeuroImage*, *15*(1), 265–272.

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, *7*(3), 242–249.

de Oliveira-Souza, R., Ignácio, F. A., Moll, J., & Hare, R. D. (2008a). Psychopathy in a civil psychiatric outpatient sample. *Criminal Justice and Behavior, 35*(4), 427–437.

de Oliveira-Souza, R., Hare, R.D., Bramati, I. E., Garrido, G. J., Ignácio, F. A., Tovar-Moll, F., et al. (2008b). Psychopathy as a disorder of the moral brain: Fronto-temporo-limbic gray matter reductions demonstrated by voxel-based morphometry. *NeuroImage*, *40*(3), 1202–1213.

Devijver, P. A., & Kittler, J. (1982). *Pattern classification: A statistical approach*. Englewood Cliffs, NJ: Prentice-Hall.

Ecker, C., Rocha-Rego, V., Johnston, P., Mourão-Miranda, J., Marquand, A., Daly, E. M., et al. (2010). Investigating the predictive value of whole-brain structural MR scans in autism: A pattern classification approach. *NeuroImage*, *49*, 44–56.

Emblem, K. E., Zoellner, F. G., Tennoe, B., Nedregaard, B., Nome, T., Due-Tonnessen, P., et al. (2008). Predictive modeling in glioma grading from MR perfusion images using support vector machines. *Magnetic Resonance in Medicine*, *60*, 945–952.

Fan, Y., Shen, D., & Davatzikos, C. (2005). Classification of structural images via high-dimensional image warping, robust feature extraction, and SVM. *Medical Image Computing and Computer-Assisted Intervention*, *8*, 1–8.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. B., Frith, C., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*, 189–210.

Fukunaga, K. (1990). *Introduction to statistical pattern recognition* (2nd ed.). Boston, MA: Academic Press.

Fullam, R. S., McKie, S., & Dolan, M. C. (2009). Psychopathic traits and deception: Functional magnetic resonance imaging study. *British Journal of Psychiatry*, *194*, 229–235.

Gao, Y., Glenn, A. L., Schug, R. A., Yang, Y., & Raine, A. (2009). The neurobiology of psychopathy: A neurodevelopmental perspective. *Canadian Journal of Psychiatry (Revue Canadienne de Psychiatrie)*, *54*(12), 813–823.

Gao, Y., Raine, A., Chan, F., Venables, P. H., & Mednick, S. A. (2010). Early maternal and paternal bonding, childhood physical abuse and adult psychopathic personality. *Psychological Medicine*, *40*(6), 1007–1016.

Georgopoulos, A. P., Tan, H. M., Lewis, S. M., Leuthold, A. C., Winskowski, A. M., Lynch, J. K., et al. (2010). The synchronous neural interactions test as a functional neuromarker for post-traumatic stress disorder (PTSD): A robust classification method based on the bootstrap. *Journal of Neural Engineering*, *7*, 16011.

Gerardin, E., Chételat, G., Chupin, M., Cuingnet, R., Desgranges, B., Kim, H. S., et al. (2009). Alzheimer's disease neuroimaging initiative. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *NeuroImage*, *47*, 1476–1486.

Glenn, A. L., & Raine, A. (2008). The neurobiology of psychopathy. *Psychiatric Clinics of North America*, *31*, 463–475.

Glenn, A. L., Raine, A., & Schug, R. A. (2009). The neural correlates of moral decision-making in psychopathy. *Molecular Psychiatry*, *14*, 5–6.

Glenn, A. L., Raine, A., Yaralian, P. S., & Yang, Y. (2010). Increased volume of the striatum in psychopathic individuals. *Biological Psychiatry*, *67*, 52–58.

Goldstein, K. E., Hazlett, E. A., New, A. S., Haznedar, M. M., Newmark, R. E., Zelmanova, Y., et al. (2009). Smaller superior temporal gyrus volume specificity in schizotypal personality disorder. *Schizophrenia Research*, *112*, 14–23.

Golland, P., Fischl, B., Spiridon, M., Kanwisher, N., Buckner, R. L., Shenton, M. E., et al. (2002). Discriminative analysis for image-based studies. In *Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention*, LNCS 2488, pp. 508–515, Tokyo, Japan, September 25–28.

Golland, P., Grimson, W., Shenton, M., & Kikinis, R. (2000). Small sample size learning for shape analysis of anatomical structures. In *Proceedings of the 3rd International Conference on Medical Image Computing and Computer-Assisted Intervention*, LNCS 1935, pp. 72–82, Pittsburgh, PA, USA, October 11–14.

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., & Frackowiak, R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*, *14*(1), 21–36.

Grèzes, J., Pichon, S., & de Gelder, B. (2007). Perceiving fear in dynamic body expressions. *NeuroImage*, *35*(2), 959–967.

Guy, L. S., & Douglas, K. S. (2006). Examining the utility of the PCL–SV as a screening measure using competing factor models of psychopathy. *Psychological Assessment*, *18*(2), 225–230.

Hare, R., & Neumann, C. (2005). Structural models of psychopathy. *Current Psychiatry Reports*, *7*(1), 57–64.

Hare, R. D. (2006). Psychopathy: A clinical and forensic overview. *Psychiatric Clinics of North America*, *29*(3), 709–724.

Hart, S., Cox, D., & Hare, R. (1995). *Manual for the Psychopathy Checklist: Screening Version (PCL:SV)*. Toronto, Canada: Multi-Health Systems.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2001). *The elements of statistical learning*: *Data mining, inference, and prediction*. New York, NY: Springer.

Honea, R. A., Meyer-Lindenberg, A., Hobbs, K. B., Pezawas, L., Mattay, V. S., Egan, M. F., et al. (2008). Is gray matter volume an intermediate phenotype for schizophrenia? A voxel-based morphometry study of patients with schizophrenia and their healthy siblings. *Biological Psychiatry*, *63*(5), 465–474.

Kalbe, E., Nowak, D. A., Brand, M., Kessler, J., Dafotakis, M., Bangard, C., et al. (2010). Dissociating cognitive from affective theory of mind: A TMS study. *Cortex*, *46*(6), 769–780.

Kalbe, E., Schlegel, M., Sack, A. T., Nowak, D. A., Dafotakis, M., Bangard, C., et al. (2010). Dissociating cognitive from affective theory of mind: A TMS study. *Cortex, 46*(6), 769–780.

Kiehl, K. A., Bates, A. T., Laurens, K. R., Hare, R. D., & Liddle, P. F. (2006). Brain potentials implicate temporal lobe abnormalities in criminal psychopaths. *Journal of Abnormal Psychology*, *115*, 443–453.

Koutsouleris, N., Gaser, C., Bottlender, R., Davatzikos, C., Decker, P., Jäger, M., et al. (2010). Use of neuroanatomical pattern regression to predict the structural brain dynamics of vulnerability and transition to psychosis. *Schizophrenia Research, 123*(2–3), 175–187.

Lao, Z., Shen, D., Xue, Z., Karacali, B., Resnick, S., & Davatzikos, C. (2004). Morphological classication of brains via high-dimensional shape transformations and machine learning methods. *NeuroImage*, *21*, 46–57.

Leibenluft, E., Gobbini, M. I., Harrison, T., & Haxby, J. V. (2004). Mothers' neural activation in response to pictures of their children and other children. *Biological Psychiatry*, *56*(4), 225–232.

Lukic, A. S., Wernick, M. N., & Strother, S. C. (2002). An evaluation of methods for detecting brain activations from functional neuroimages. *Artificial Intelligence in Medicine*, *25*, 69–88.

Materna, S., Dicke, P. W., & Thier, P. (2008). The posterior superior temporal sulcus is involved in social communication not specific for the eyes. *Neuropsychologia*, *46*(11), 2759–2765.

Meyer-Lindenberg, A., & Weinberger, D. R. (2006). Intermediate phenotypes and genetic mechanisms of psychiatric disorders [Review]. *Nature Reviews Neuroscience*, *7*(10), 818–827.

Millon, T., Simonsen, E., Davis, R. D., & Birket-Smith, M. (2002). *Psychopathy: Antisocial, criminal, and violent behavior*. New York, NY: Guilford Press.

Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourao-Miranda, J., Andreiuolo, P. A., et al. (2002). The neural correlates of moral sensitivity: A functional magnetic resonance imaging investigation of basic and moral emotions. *Journal of Neuroscience*, *22*, 2730–2736.

Moll, J., de Oliveira-Souza, R., Garrido, G. J., Bramati, I. E., Caparelli-Daquer, E. M. A., Paiva, M. L. M. F., et al. (2007). The self as a moral agent: Linking the neural bases of social agency and moral sensitivity. *Social Neuroscience*, *2*, 336–352.

Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience*, *6*(10), 799–809.

Mourão-Miranda, J., Bokde, A. L.W., Born, C., Hampel, H., & Stetter, S. (2005). Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data. *NeuroImage*, *28*(4), 980–995.

Müller, J. L., Gänssbauer, S., Sommer, M., Döhnel, K., Weber, T., Schmidt-Wilcke, T., et al. (2008). Gray matter changes in right superior temporal gyrus in criminal psychopaths. Evidence from voxel-based morphometry. *Psychiatry Research*, *163*, 213–222.

Nardo, D., Högberg, G., Looi, J. C. L., Larsson, S., Hällström, T., & Pagani, M. (2010). Gray matter density in limbic and paralimbic cortices is associated with trauma load and EMDR outcome in PTSD patients. *Journal of Psychiatric Research*, *44*(7),477–785.

Rilling, J. K., Glenn, A. L., Jairam, M. R., Pagnoni, G., Goldsmith, D. R., Elfenbein, H. A., et al. (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biological Psychiatry*, *61*, 1260–1271.

Sato, J. R., da Graça Morais Martin, M., Fujita, A., Mourão-Miranda, J., Brammer, M. J., & Amaro, E. Jr. (2009). An fMRI normative database for connectivity networks using one-class support vector machines. *Human Brain Mapping*, *30*, 1068–1076.

Sato, J. R., Mourão-Miranda, J., Morais Martin Mda, G., Amaro, E. Jr., Morettin, P. A., & Brammer, M. J. (2008a). The impact of functional connectivity changes on support vector machines mapping of fMRI data. *Journal of Neuroscience Methods*, *172*, 94–104.

Sato, J. R., Thomaz, C. E., Cardoso, E. F., Fujita, A., Martin Mda, G., & Amaro, E. Jr. (2008b). Hyperplane navigation: A method to set individual scores in fMRI group datasets. *Neuroimage*, *42*, 1473–1480.

Schaufelberger, M. S., Duran, F. L. S., Lappin, J. M., Scazufca, M., Amaro, E., Leite, C. C., et al. (2007). Gray matter abnormalities in Brazilians with first-episode psychosis. *British Journal of Psychiatry, Suppl 51*, s117–s122.

Schlaepfer, T. E., Lancaster, E., Heidbreder, R., Strain, E. C., Kosel, M., Fisch, H., et al. (2006). Decreased frontal white-matter volume in chronic substance abuse. *International Journal of Neuropsychopharmacology*, *9*, 147–153.

Soderstrom, H., Hultin, L., Tullberg, M., Wikkelso, C., Ekholm, S., & Forsman, A. (2002). Reduced frontotemporal perfusion in psychopathic personality. *Psychiatry Research*, *114*(2), 81–94.

Soloff, P., Nutche, J., Goradia, D., & Diwadkar, V. (2008). Structural brain abnormalities in borderline personality disorder: A voxel-based morphometry study. *Psychiatry Research*, *164*, 223–236.

Thomaz, C. E., Boardman, J. P., Counsell, S., Hill, D. L. G., Hajnal, J. V., Edwards, A. D., et al. (2007a). A multivariate statistical analysis of the developing human brain in preterm infants. *Image and Vision Computing*, *25*(6), 981–994.

Thomaz, C. E., Duran, F. L. S., Busatto, G. F., Gillies, D. F., & Rueckert, D. (2007b). Multivariate statistical differences of MRI samples of the human brain. *Journal of Mathematical Imaging and Vision*, *29*(2–3),95–106.

Thomaz, C. E., Kitani, E. C., & Gillies, D. F. (2006). A maximum uncertainty LDA-based approach for limited sample size problems – with application to face recognition. *Journal of the Brazilian Computer Society*, *12*(2), 7–18.

Tiihonen, J., Rossi, R., Laakso, M. P., Hodgins, S., Testa, C., Perez, J., et al. (2008). Brain anatomy of persistent violent offenders: More rather than less. *Psychiatry Research*, *163*(3), 201–212.

Vapnik, V. N. (1998). *Statistical learning theory*. New York, NY: Wiley.

Veit, R., Lotze, M., Sewing, S., Missenhardt, H., Gaber, T., & Birbaumer, N. (2010). Aberrant social and cerebral responding in a competitive reaction time paradigm in criminal psychopaths. *NeuroImage*, *49*, 3365–3372.

Weber, S., Habel, U., Amunts, K., & Schneider, F. (2008). Structural brain abnormalities in psychopaths—-a review. *Behavioral Sciences & the Law*, *26*, 7–28.

Williams, K. M., Paulhus, D. L., & Hare, R. D. (2007). Capturing the four-factor structure of psychopathy in college students via self-report. *Journal of Personality Assessment*, *88*(2), 205–219.

Wilson, S. M., Ogar, J. M., Laluz, V., Growdon, M., Jang, J., Glenn, S., et al. (2009). Automated MRI-based classification of primary progressive aphasia variants. *NeuroImage*, *47*, 1558–1567.

Worsley, K. (1995). Estimating the number of peaks in a random field using the Hadwiger characteristic of excursion sets with applications to medical images. *Annals of Statistics*, *23*, 640–669.

Zhu, C. Z., Zang, Y. F., Liang, M., Tian, L. X., He, Y., Li, X. B., et al. (2005). Discriminative analysis of brain function at resting-state for attention-deficit/hyperactivity disorder. *Medical Image Computing and Computer-Assisted Intervention*, *8*, 468–475.

Zou, K., Deng, W., Li T., Zhang, B., Jiang, L., Huang, C., et al. (2010). Changes of brain morphometry in first-episode, drug-naïve, non-late-life adult patients with major depression: An optimized voxel-based morphometry study. *Biological Psychiatry*, *67*, 186–188.

Yang, Y., Raine, A., Lencz, T., Bihrle, S., LaCasse, L., & Colletti, P. (2005). Volume reduction in prefrontal gray matter in unsuccessful criminal psychopaths. *Biological Psychiatry*, *57*(10), 1103–1108.