Schizophrenia Bulletin vol. 39 no. 5 pp. 1105–1114, 2013
doi:10.1093/schbul/sbs095
Advance Access publication September 11, 2012

Distinguishing Prodromal From First-Episode Psychosis Using Neuroanatomical Single-Subject Pattern Recognition

Stefan Borgwardt[*,1,2], Nikolaos Koutsouleris[3], Jacqueline Aston[1], Erich Studerus[1], Renata Smieskova[1,2], Anita Riecher-Rössler[1], and Eva M. Meisenzahl[3]

[1]Department of Psychiatry, University of Basel, Basel, Switzerland; [2]University Hospital Basel, Medical Image Analysis Centre, University of Basel, Basel, Switzerland; [3]Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University, Munich, Germany

*To whom correspondence should be addressed; Department of Psychiatry, University of Basel, Petersgraben 4, CH-4031 Basel, Switzerland; fax: +41 61 265 45 88, e-mail: sborgwardt@uhbs.ch

Background: **The at-risk mental state for psychosis (ARMS) and the first episode of psychosis have been associated with structural brain abnormalities that could aid in the individualized early recognition of psychosis. However, it is unknown whether the development of these brain alterations predates the clinical deterioration of at-risk individuals, or alternatively, whether it parallels the transition to psychosis at the single-subject level.** ***Methods:*** **We evaluated the performance of an magnetic resonance imaging (MRI)-based classification system in classifying disease stages from at-risk individuals with subsequent transition to psychosis (ARMS-T) and patients with first-episode psychosis (FE). Pairwise and multigroup biomarkers were constructed using the structural MRI data of 22 healthy controls (HC), 16 ARMS-T and 23 FE subjects. The performance of these biomarkers was measured in unseen test cases using repeated nested cross-validation.** ***Results:*** **The classification accuracies in the HC vs FE, HC vs ARMS-T, and ARMS-T vs FE analyses were 86.7%, 80.7%, and 80.0%, respectively. The neuroanatomical decision functions underlying these discriminative results particularly involved the frontotemporal, cingulate, cerebellar, and subcortical brain structures.** ***Conclusions:*** **Our findings suggest that structural brain alterations accumulate at the onset of psychosis and occur even before transition to psychosis allowing for the single-subject differentiation of the prodromal and first-episode stages of the disease. Pattern regression techniques facilitate an accurate prediction of these structural brain dynamics at the early stage of psychosis, potentially allowing for the early recognition of individuals at risk of developing psychosis.**

Key words: at-risk mental state/early prediction of psychosis/voxel-based morphometry/multivariate analysis/machine learning/support vector machine

Introduction

Over the past decade, research on the prodromal phase of psychosis has exponentially progressed, allowing for preventive therapeutic interventions in clinical psychiatry.[1] In the light of the severe functional, social, and economic long-term impact of psychosis, psychiatric imaging needs to increasingly shift its research focus to the translation of imaging findings to clinical applications, targeting important stages of the disease course, including transition, remission, and response to preventative interventions.[2] Because it is difficult to predict which subjects with an at-risk mental state (ARMS) will later develop psychosis on the basis of their presenting clinical features, there is a need for objective surrogate markers to identify the individuals at highest risk of developing overt psychosis and those who might benefit most from preventive interventions. Although neuroimaging techniques seem promising for this issue in subjects with an ARMS, inconsistent findings across individual studies prevent applicability of imaging methods to clinical psychiatry. In recent years, a range of neuroimaging techniques showed alterations in brain structure,[3] function,[4] and neurochemistry[5] in the prodromal phase of psychosis.[6] These neuroimaging studies have shown that alterations in brain anatomy and neurophysiology found in established psychosis are also present in people with an ARMS for the disease.[4] Overall, ARMS subjects show qualitatively similar, but less pronounced, structural brain abnormalities than patients with established schizophrenia. Studies comparing ARMS subjects showed reduced gray matter (GM) in prefrontal, temporal, and cingulate regions; insula; and cerebellum in those subjects who develop psychosis.[7] Interestingly, similar areas were also recently found to be related in volume to schizotypal personality, a subclinical schizophrenia spectrum trait.[8] Although there are few structural imaging studies that

© The Author 2012. Published by Oxford University Press on behalf of the Maryland Psychiatric Research Center. All rights reserved.
For permissions, please email: journals.permissions@oup.com

Downloaded from https://academic.oup.com/schizophreniabulletin/article/39/5/1105/1925098 by guest on 16 August 2022

compared ARMS subjects according to clinical outcome, it is still not clear which magnetic resonance imaging (MRI) abnormalities are specific to vulnerability as opposed to a later transition to psychosis. Volumetric reductions in the temporal, cingulate, insular, and prefrontal cortex and in the cerebellum have been specifically associated with the development of psychosis.[9]

However, univariate image analysis methods such as voxel-based morphometry (VBM) detect group differences on the basis of spatially confined image elements like single voxels or clusters, instead of providing information on the complex spatial patterns of neuroanatomical disease characteristics. Thus, VBM is a research tool for structural and functional brain differences[11] between groups of subjects, and not a diagnostic or classification device.[12] In contrast, multivariate pattern recognition methods categorize individual structural brain scans by separation of images taking into account the innate interregional dependencies of different pathologies.[13] In this context, support vector machines (SVMs) emerged as a powerful diagnostic tool to evaluate the categorization of complex, high-dimensional training data and to generalize the learned classification rules to new, unseen individual data.[14,15] In the context of structural MRI, SVM tools successfully identified spatial patterns across brain regions that provide a single-subject, diagnostic separation between different clinical populations eg, in Alzheimer's disease and mild cognitive impairment[16–18]; multiple sclerosis,[19] schizophrenia[15] and presymptomatic Huntington's disease.[20] Therefore, these multivariate neuroimaging tools may promote a potentially accessible and objective way to improve clinical decision making, taking into account the risk of developing psychosis in individuals with an ARMS. The prediction of subsequent disease conversion has been shown by nonlinear MRI-based SVMs operating at the individual level.[15] Recently, we have shown that SVM has the potential to increase the prediction accuracy of established clinical decision in 2 independent ARMS samples from Munich and Basel to over 80%.[15,21] To our knowledge, SVMs have neither been applied to structural MRI data using GM segments for distinguishing individuals with an ARMS with subsequent transition to psychosis (ARMS-T) vs patients with a first-episode psychosis (FE) nor been applied to distinguish FE from healthy controls (HC). This is important because there may be a neurobiological process that parallels the clinical transition process from the prodromal stage to the manifest disease. SVMs can be used to evaluate whether this occurs at the single-subject level. We hypothesized that the neuroanatomical alterations found in the manifest disease are (almost) equally present in the prodrome and make it impossible to distinguish prodromal from FE subjects at the single-subject level using SVM. Secondly, we expected a robust discrimination of FE patients and HC.

Methods

Study Design

This MRI study was embedded in the naturalistic, prospective, and multidomain *FePsy*-study on the prediction of psychosis development in ARMS individuals, covering a service area of 200 000 habitants in and around Basel, Switzerland. A more detailed description of the overall study design can be found elsewhere.[22,23] All aspects of the study were reviewed and approved by the institutional ethics committee of the University of Basel, and written informed consent was obtained from each participant before study inclusion.

Participants

Within the prospective *FePsy*-study, ARMS individuals received a structural MRI scan at study inclusion. For screening purposes, we used the Basel Screening Instrument for Psychosis (BSIP),[24] a 46-item instrument based on variables that have been shown to be risk factors or early symptoms of psychosis such as DSM-III-R – "prodromal symptoms," social decline, drug abuse, previous psychiatric disorders, or genetic liability for psychosis. The BSIP enables a reliable identification of vulnerable individuals at risk of developing psychosis using clinical criteria that closely correspond to the well-established ultrahigh-risk definitions of the Personal Assessment and Crisis Evaluation clinic in Melbourne.[24,25] In keeping with previous MRI studies of ARMS cohorts recruited using these high-risk criteria,[9,26] inclusion into the present study required one or more of the following: (a) attenuated psychotic-like symptoms (APS; a score of 2 or 3 on the Brief Psychiatric Rating Scale [BPRS] hallucination item, or 3 or 4 on BPRS items for unusual thought content or suspiciousness—at least several times a week and for more than 1 week duration), (b) brief limited intermittent psychotic symptoms (BLIPS; scores of 4 or above on the hallucination item, or 5 or above on the unusual thought content, suspiciousness or conceptual disorganization items of the BPRS, with each symptom lasting less than 1 week before resolving spontaneously), or (c) a first or second degree relative with a psychotic disorder plus at least 2 further risk factors for or indicators of beginning psychosis according to the BSIP screening instrument. A more detailed description of these ARMS criteria can be found in our previous work.[23] Psychopathology was assessed with the BPRS, in combination with the BSIP and the Scale for the Assessment of Negative Symptoms (SANS).

All subjects were followed-up regularly and were offered supportive counseling and clinical management. Because we are interested in the pure prodromal state of psychosis, we included only those ARMS individuals who subsequently made transition during the follow-up

Downloaded from https://academic.oup.com/schizophreniabulletin/article/39/5/1105/1925098 by guest on 16 August 2022

period. Conversion to frank psychosis was monitored using the criteria described by Yung et al[25]: BPRS scores of 4 or above on the hallucination item; or scores of 5 or above on the unusual thought content, suspiciousness, or conceptual disorganization items. Symptoms had to occur daily and persist for more than 1 week to be deemed a conversion to frank psychosis. Using these definitions, 16 converters (ARMS-T) to psychosis were included.

The first-episode group (n = 23) was defined as subjects who met the operational criteria for FE described by Yung et al[25] as it was done in previous MRI studies of the ARMS.[27] Inclusion required scores of 4 or above on the hallucination item, or 5 or above on the unusual thought content, suspiciousness, or conceptual disorganization items of the BPRS. The symptoms must have occurred at least several times a week and persisted for more than 1 week.

Exclusion criteria were age below 18 years, insufficient knowledge of German, IQ < 70 (measured with the MWT-B), previous psychotic episodes treated with major tranquillizers for more than 3 weeks, a clearly diagnosed brain disease or substance dependency (except for cannabis dependency), or psychotic symptoms within a clearly diagnosed depression, bipolar or borderline personality disorder. 2 out of 16 ARMS-T individuals received low-dose antipsychotic medication prior to MRI scanning. These participants had been administered low doses of atypical antipsychotic medication for behavioral control by the referring psychiatrist or general practitioner (1 participant olanzapine, 1 risperidone) at some time prior to study inclusion, both for less than 3 weeks. A large proportion of FE patients were scanned within 1–3 days of first contact, therefore most of the FE patients (14/23; 61%) were also antipsychotic-naive. Six had been taking antipsychotics for < 1 month and 3 had been taking them for 1–3 months.

Twenty-two HC were recruited from the same geographical area as the ARMS group through local advertisements and were matched to the ARMS sample groupwise for age, gender, handedness, and education level. These individuals had no current psychiatric disorder, no history of psychiatric illness, head trauma, neurological illness, serious medical or surgical illness, substance dependency (except for cannabis and nicotine), and no family history of any psychiatric disorder as assessed by an experienced psychiatrist in a detailed clinical interview.

MRI Data Acquisition

Subjects were scanned using a SIEMENS (Erlangen, Germany) MAGNETOM VISION 1.5T scanner at the University Hospital, Basel. A three-dimensional volumetric spoiled gradient recalled echo sequence generated 176 contiguous, 1-mm-thick sagittal slices. Imaging parameters were time-to-echo, 4 ms; time-to-repetition, 9.7 ms; flip angle, 12°; matrix size, 200 × 256; field of view, 25.6 × 25.6 cm matrix; voxel dimensions, 1.28 × 1 × 1 mm.

MRI Data Preprocessing

After inspection for artifacts and gross abnormalities, the images were segmented into GM, white matter (WM), and cerebrospinal fluid (CSF) maps in native space using the VBM5 toolbox (http://dbm.neuro.uni-jena.de), an extension of the SPM5 software package (Wellcome Department of Cognitive Neurology, London, UK). Details of this segmentation protocol have been described in our previous work.[28] Then, the estimated tissue maps of each individual were combined into a single labeled volume (CSF: 10, GM: 150, and WM: 250) and registered to the single-subject brain template of the Montreal Neurological Institute, using a well-established, high-dimensional, elastic warping algorithm. The volumetric changes occurring during this normalization process were written out to the registered tissue maps, allowing for a Regional Analysis of Volumes in Normalized Space (RAVENS). Similar to the "modulation" step used in VBM, RAVENS maps low for local comparisons in standard space that are equivalent to volumetric comparisons of the original tissue maps in native space. The individual GM-RAVENS maps were proportionally scaled to the global GM volume computed from the native tissue maps and entered the subsequent multivariate pattern classification analysis.

Multivariate Pattern Classification Analysis

SVMs are multivariate statistical methods that have been increasingly employed for diagnostic purposes in a wide range of biomedical applications because they provide optimal decision rules for classifying individuals. Instead of describing statistical between-group differences, nonlinear classification models that reliably predict the study participants' group membership are used. As customary in predictive analytics, the SVM models were constructed from one set of subjects (the training sample) and applied to a different set of subjects (the test sample), using cross-validation (CV). This process produced an unbiased estimate of the method's expected diagnostic accuracy on new individuals rather than merely fitting this study population. The principles of generating and validating predictive models on separate training and testing samples have been previously described.[15] Based on the LIBSVM software (www.csie.ntu.edu.tw/cjlin/libsvm), our machine learning toolbox NeuroMiner running under MATLAB (R2009b, The MathWorks Inc.) produced compact ensembles of SVMs

Downloaded from https://academic.oup.com/schizophreniabulletin/article/39/5/1105/1925098 by guest on 16 August 2022

that optimally separated single individuals from different groups, while avoiding the danger of overfitting to the peculiarities of the training data. It consisted mainly of three successive steps that were wrapped into a repeated nested CV framework (see online supplementary material):

Neuroanatomical Feature Generation. First, each training sample's GM-RAVENS maps were adjusted for age and gender effects using partial correlations and scaled voxelwise to the range [0,1]. These scaled and adjusted maps entered a recently proposed, multivariate filter method,[29] which automatically determined those sets of voxels that *conjointly* maximize the geometric distance between the training subjects in the HC vs ARMS-T, HC vs FE, and ARMS-T vs FE analyses. This algorithm removed irrelevant/unreliable voxels from the high-dimensional MRI input space that did not contribute to the respective binary classification problem. Then, correlated voxels within the extracted discriminative patterns were projected to a number of uncorrelated principal components (PC) using principal component analysis (PCA). This further reduced the dimensionality of the discriminative patterns to compact sets of neuroanatomical features. The optimum number of PC was determined using CV (see online supplementary material).

SVM Training. These discriminative PC features were projected to a high-dimensional feature space using the radial basis functions in order to account for possible nonlinear relations between the training subjects' neuroanatomical features and their group membership. In this feature space, the SVM found the optimal between-group boundary by maximizing the geometric distance between the neuroanatomically most similar subjects of opposite groups (the "support vectors"). It has been shown that this maximum-margin principle in conjunction with the nonlinear projection generates classification rules that are adaptive to subtle between-group differences and therefore generalize well to unseen individuals.

Classification of Unseen Test Data. The group assignment of unseen test subjects was predicted after applying all training parameters successively to their MRI data, including (a) the adjustment for age and gender effects, (b) the selection of optimally discriminative voxels, (c) the projection of these voxels to PC, and (d) the nonlinear transformation of these neuroanatomical features. Then, for each subject, the 3 trained binary SVM models (HC vs ARMS-T, HC vs FE, and ARMS-T vs FE) determined its geometric position relative to their learned decision boundaries, resulting in 3 decision values and group assignment predictions. We used these decision values to construct a multigroup classifier (HC vs ARMS-T vs FE), where the binary SVM model with the maximum decision value decided about the test subject's group assignment (one-vs-one-max-wins method). Feature generation, model training, and test subject prediction were wrapped into a repeated nested CV framework (see online supplementary material).[30] The main goal of this framework was to *completely separate* the process of estimating the SVMs' prediction performance in a large number of unseen validation samples (outer CV loop) from the process of constructing optimally discriminative SVM models from a large number of training samples (inner CV loop). On the outer CV loop, we performed 10 repetitions of the following CV cycle. First, the order of the subjects was permuted within each group and the entire population was split into 10 nonoverlapping samples. Each of these samples was iteratively held back as validation data, while the nine remaining samples entered the inner CV loop as the training data. At this inner loop, we used 10-fold CV with 10 repetitions to generate ensembles of SVM models. More specifically, for each validation sample at the outer CV level, 100 different training data partitions were created at the inner CV level. In each of these 100 training partitions the most discriminative sets of neuroanatomical features were determined. Each of these sets was used to train a separate SVM model. Then, each of these models predicted the group assignment of the unseen validation subjects on the outer loop. These predictions were averaged across all 100 training partitions to yield an ensemble decision. Finally, for each validation subject, all SVM ensemble decisions were aggregated across those outer training partitions, in which this subject had not been involved in the training process. Majority voting was used to determine the validation subject's *class probability* and thus its final *out-of-training* group assignment.

The nonlinearity of the decision rules determining the test subjects' group assignment made it difficult to directly visualize each voxel's contribution to the average SVM ensemble decision. Therefore, we first approximated the average neuroanatomical decision boundary used by the binary nonlinear SVM models as described in Koutsouleris et al[15] and then measured each voxel's probability of *reliably* contributing to this discriminative pattern across the entire experiment at the 95% confidence interval. The exact visualization procedure has been detailed in the legend of figure 1. Moreover, a supplementary parcellation analysis (see online supplementary material figure) was conducted in order to measure the distribution of reliably discriminative voxels across the 116 brain regions of the automated anatomical labeling template.[31]

Results

Sociodemographic, Clinical and Global Anatomical Findings

Subsequent converters, HC, and FE patients did not significantly differ with respect to age, gender, educational

Downloaded from https://academic.oup.com/schizophreniabulletin/article/39/5/1105/1925098 by guest on 16 August 2022

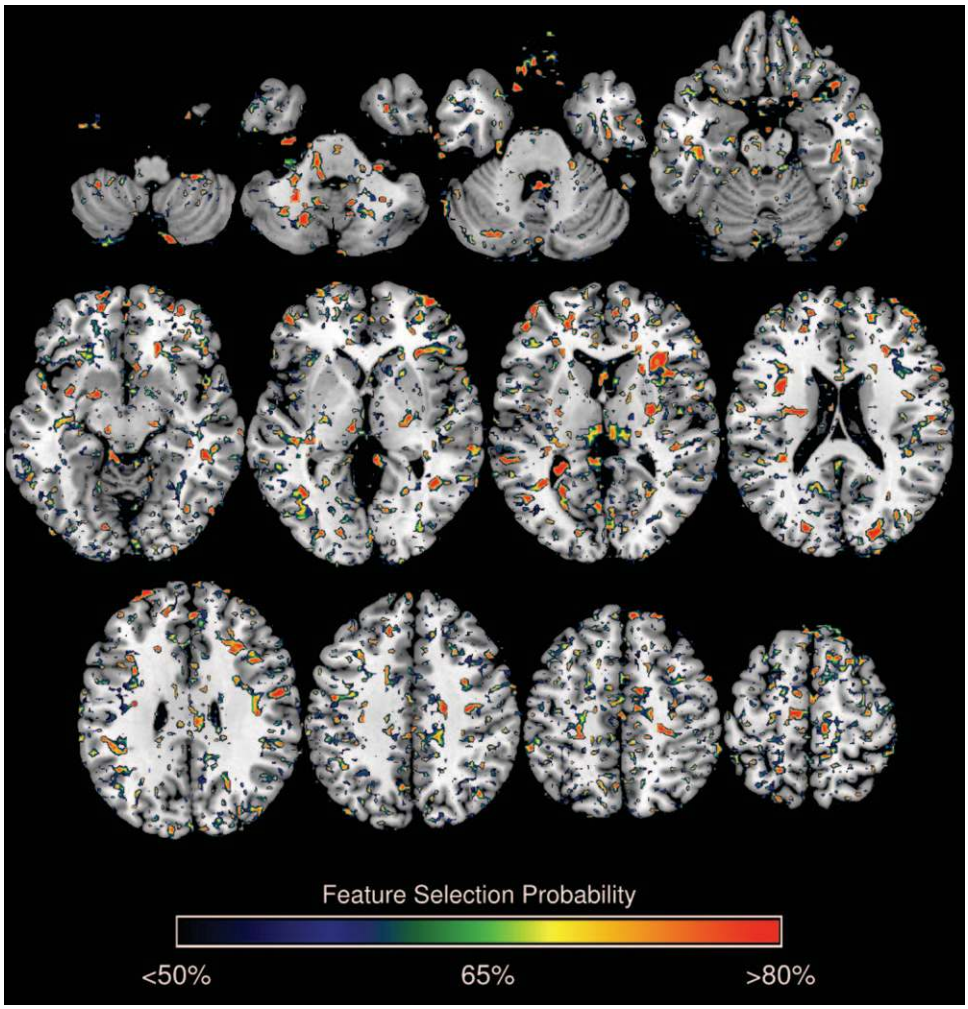

Fig. 1. Voxel probability map (VPM) of reliable contributions to the HC vs ARMS-T decision boundary.The approximation of each voxel's contribution to the average nonlinear classification used to separate HC from ARMS-T subjects was obtained as follows: (1) In principal component analysis space, the average minimum difference vector (*SV* mindiff) across the support vectors of a given SVM model was computed and projected back to voxel space as described previously. This computation was performed for every training sample on the inner cross-validation (CV) loop resulting in 100 *SV* mindiff images for a given training partition on the outer CV loop. (2) The average and standard error volumes of these 100 *SV* mindiff images were computed. (3) For every outer CV partition, the average *SV* mindiff image was binarized, in that voxels with an absolute value greater than their respective standard error were set to one, or to zero otherwise. This thresholding procedure extracted only those voxels that *reliably* contributed to the average neuroanatomical decision boundary of a given outer CV partition at the 95% confidence interval. (4) The obtained binary images were summed across all 100 outer CV partitions and divided by 100, thus forming a single map that specified every voxel's *probability* of reliably contributing to the average neuroanatomical decision boundary across the entire experiment.

level, and global brain volumes (table 1). Furthermore, no significant baseline differences were found between the FE and ARMS-T samples regarding educational level, family history of psychosis, duration of symptoms prior to the MRI examination, BPRS, and SANS (table 1).

SVM Classification Analysis

Classification Performance. Among the 3 binary classification analyses (table 2a), the highest diagnostic performance (balanced accuracy [BAC] = 86.7%) was observed in the HC vs FE comparison, where out of 55 subjects, 2 FE individuals were classified as HC and 3 HC subjects were assigned to the FE group (sensitivity = 87.0%, specificity = 86.4%). A diagnostic performance in the HC vs ARMS-T analysis was BAC = 80.7% because 2 ARMS-T were wrongly assigned to the HC group and 4 HC were classified as ARMS-T (sensitivity = 75%, specificity = 86.%). In the critical ARMS-T vs FE analysis, the BAC was 80.0% with 4 FE subjects being misclassified as ARMS-T and 5 ARMS-T being wrongly labeled as FE (sensitivity = 91.3%, specificity = 68.8%).

In the 3-group classification (table 2b), 21 of the 23 FE patients were correctly assigned to their group, while 9 of the 16 ARMS-T and 7 of the 22 HC subjects were misclassified as FE (sensitivity = 91.3%, specificity = 57.9%, BAC = 74.6%). Of the 16 ARMS-T subjects, 3 were correctly assigned to their group, while 1 HC individual

Downloaded from https://academic.oup.com/schizophreniabulletin/article/39/5/1105/1925098 by guest on 16 August 2022

Table 1. Sociodemographic, Clinical, and Global Anatomical Characteristics of the 3 Study Groups

	Study Groups			
	ARMS-T	FE	HC	*P*
Sociodemographic variables				
N	16	23	22	
Mean age (years) at baseline (SD)	26.4 (6.5)	26.78 (6.5)	23.0 (4.3)	ns
Sex (male)	11 (69%)	17 (74%)	13 (59%)	ns
Handedness (mixed or left)	3 (19%)	5 (22%)	6 (29%)	ns
Educational level				ns
<9 yrs	4 (25%)	12 (52%)	2 (9%)	
9–11 yrs	6 (38%)	8 (34%)	7 (32%)	
12–13 yrs	5 (31%)	1 (4%)	10 (46%)	
>13 yrs	1 (6%)	2 (9%)	3 (14%)	
Clinical variables				
Individuals with a 1° relative with schizophrenia	3 (19%)	4 (17%)	na	ns
Mean BPRS global score at intake (SD)	41.9 (10.6)	52.7 (13.6)	na	ns
Mean SANS at intake (SD)	9.5 (5.4)	10.0 (5.3)	na	ns
Mean duration (months) of symptoms (SD)	42.6 (39.5)[a]	54.9 (74.1)	na	ns
Mean interval (days) between baseline MRI scan and disease transition (SD)	306.3 (318.3)	na	na	
Global anatomical volumes				
Mean global gray matter volume [ml] (SD)	680.5 (57.5)	680.9 (55.9)	692.2 (52.6)	ns
Mean global white matter volume [ml] (SD)	613.0 (79.9)	627.0 (78.9)	615.2 (68.7)	ns
Mean global cerebrospinal fluid volume [ml] (SD)	212.6 (36.8)	215.1 (55.2)	204.8 (30.9)	ns

Note: ARMS-T, at-risk mental state with subsequent transition to psychosis; FE, first-episode psychosis; HC, healty controls; BPRS, Brief Psychiatric Rating Scale; SANS, Scale for the Assessment of Negative Symptoms; SD, standard deviation; na, not applicable; ns, not significant.
[a]Duration of the ARMS symptoms at the inclusion of the study (before transition to psychosis).

Table 2a. Two-Group Classification Performance

Binary Classifiers	TP	TN	FP	FN	Sens (%)	Spec (%)	BAC (%)	FPR (%)	PPV (%)	NPV (%)
HC vs ARMS-T	19	12	4	3	75	86.4	80.7	25	82.7	80
HC vs FE	19	20	3	3	87	86.4	86.7	13	86.4	87
ARMS-T vs FE	11	21	2	5	91.3	68.8	80	8.7	84.6	80.8

Notes: The performance of the binary SVM ensemble classifiers (group "+1" vs group "−1") was evaluated (1) by constructing a binary SVM ensemble from all SVM base learners of a inner cross-validation (CV) partition, in which the respective outer CV test subjects had not been included, (2) by computing the average decision value in each of these binary inner CV ensembles in order to determine the group membership (average decision value > 0 or < 0) of the respective outer CV test subjects, and (3) through majority voting across those binary inner CV loop SVM ensembles, in which the outer CV test subjects had not participated in the training process (see also the Methods section for a detailed explanation of the employed ensemble learning framework). Sensitivity (Sens), specificity (Spec), balanced accuracy (BAC), false positive rate (FPR), positive/negative predictive values (PPV/NPV) as well as positive/negative Likelihood Ratios (LR+/LR−) were calculated from the confusion matrix containing the number of true positives (TP), false negatives (FN), true negatives (TN) and false positives (FP).

was wrongly labeled as ARMS-T (sensitivity = 18.8%, specificity = 97.8%, BAC = 58.3%). Fourteen out of 22 HC individuals were correctly identified by the pattern recognition system, while 4 ARMS-T subjects and 2 FE patients were misclassified as HC (sensitivity = 63.6%, specificity = 894.6%, BAC = 74.1%).

Neuroanatomical Mapping of SVM Decision Functions

In summary, the approximation of the three neuroanatomical SVM decision functions (methodological descriptions in figure 1) revealed that reliable voxels were not confined to single brain regions, but instead were distributed across a broad range of cortical and subcortical areas. Within these distributed patterns shown in figures 1–3, foci of moderate-high probability voxels (>65% probability) were detected particularly in the prefrontal, limbic, basal ganglia, and cerebellar structures. More specifically, the average neuroanatomical decision function of the HC vs ARMS-T ensemble classifier involved high-probability hotspots particularly in the prefrontal cortex, the limbic lobe (e.g. amygdala and olfactory regions), and cerebellum. There were no clusters of contiguous high-probability voxels involved in the average HC vs FE ensemble decision. Reliable high-probability voxels contributing to the average

Downloaded from https://academic.oup.com/schizophreniabulletin/article/39/5/1105/1925098 by guest on 16 August 2022

Table 2b. Three-Group Classification Performance

	SVM Predictions		
Clinical Groups	HC	ARMS-T	FE
HC	14	1	7
ARMS-T	4	3	9
FE	2	0	21
OOT-performance			
TP	14	3	21
TN	33	44	22
FP	6	1	16
FN	8	13	2
Sensitivity (%)	63.6	18.8	91.3
Specificity (%)	84.6	97. 8	57.9
Balanced accuracy (%)	74.1	58.3	74.6
False positive rate (%)	15.4	2.2	42.1
Positive predictive value (%)	70	75	56.8
Negative predictive value (%)	80.5	77.2	91.7

Notes: Multigroup decisions were obtained by (1) constructing a multigroup ensemble classifier for each CV2 data partition using error-correcting output codes (see Methods section and online supplementary material) and (2) computing the final out-of-training (OOT) group membership of a given CV2 test subject through majority voting of all CV2 multigroup ensemble classifiers, in which this test subject had not been part of the training data and thus had not been seen by these classifier ensembles. The OOT classification performance of the multigroup ensembles was then evaluated for 1 group against all other groups. For example, in the HC vs ARMS-T vs FE analysis 21 FE subjects of 23 (sensitivity: 91.3%) were correctly assigned to their group, while 22 of the other 38 (57.9%) subjects were correctly not labeled as FE, resulting in a BAC of (91.3% + 57.9%)/2 = 74.6%.

ARMS-T vs FE ensemble decision mainly mapped to basal ganglia (including thalamus, pallidum, and putamen) and the cerebellum.

Discussion

To our best knowledge, this is the first SVM analysis exploring GM pattern in pure prodromal subjects (ARMS-T) in relation to FE. The present investigation expanded our previous findings in the same cohort of a clinically defined at-risk population (ARMS) and HC subjects[28] in that our fully automated classification system reliably identified ARMS-T and FE individuals using only their MRI scans acquired at study inclusion. Interestingly, we found that the highest diagnostic performance was observed in the HC vs FE comparison with a sensitivity of 87% and a specificity of 86.4%. This underlines the ability of this multivariate image analysis tool to translate imaging findings in the clinical field targeting diagnosis, remission, and response to preventative interventions. The high discriminative power of FE and HC confirms categorization of individual brain scans by separation of images from different groups, taking into account the interregional dependencies of different pathologies. Interestingly, these classification rates compare to those observed in studies of other illnesses such as in dementia[16–18] or MS[19] attempting the prediction of illness conversions.

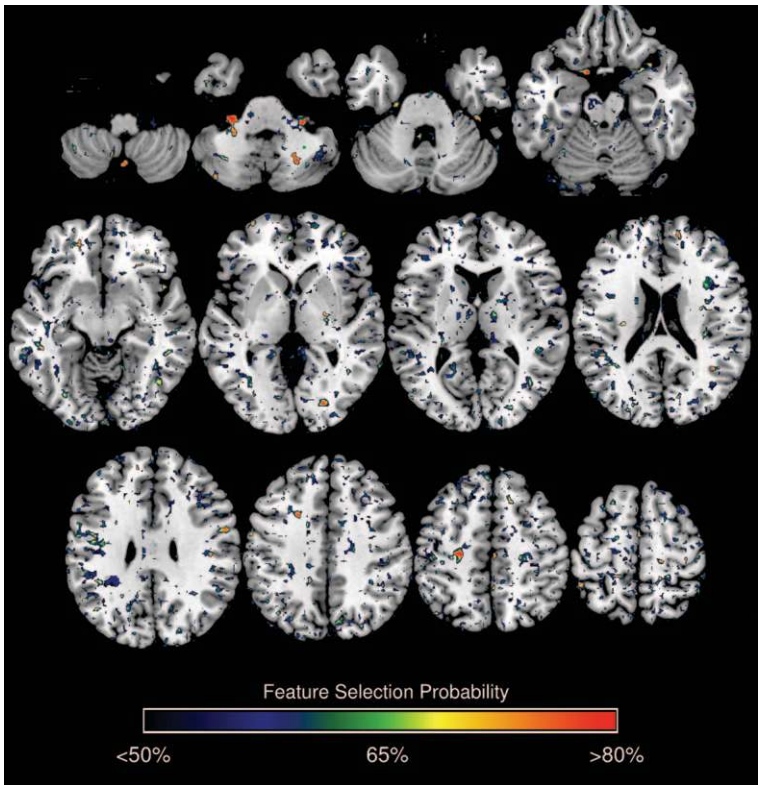

Fig. 2. Voxel probability map (VPM) of reliable contributions to the HC vs FE decision boundary. See legend of figure 1.

Downloaded from https://academic.oup.com/schizophreniabulletin/article/39/5/1105/1925098 by guest on 16 August 2022

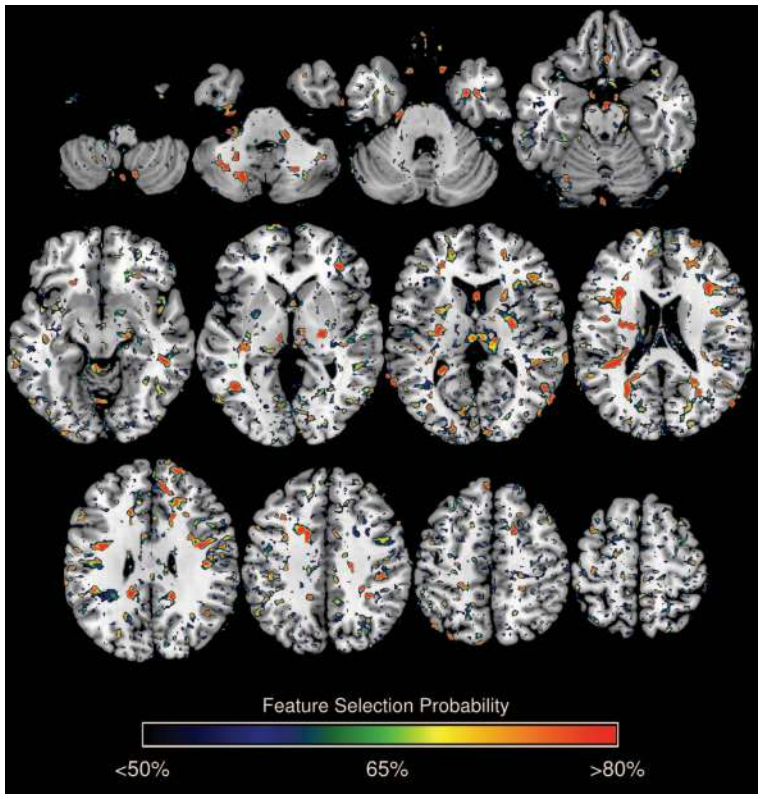

Fig. 3. Voxel probability map (VPM) of reliable contributions to the ARMS-T vs FE decision boundary. See legend of figure 1.

The diagnostic performance in the HC vs ARMS-T analysis was similar to the HC vs FE. In contrast to the high classification performance observed in the ARMS-T vs FE comparison, we found that the ARMS-T subjects were frequently misclassified as FE in the 3-group analysis. In this regard, it is well known that SVMs were primarily developed to solve binary classification problems rather than multigroup classification tasks. Similarly, our feature extraction method operated on binary classification tasks (see Methods section). Taken together, our machine learning pipeline was primarily devised to detect pairwise discriminative patterns, which may have limited its ability to detect robust neuroanatomical signatures separating the study population at the multigroup level. Nevertheless, the main purpose of the current analysis was to test the hypothesis whether neuroanatomical abnormalities differentiate the pure prodromal vs the first-episode stage of psychosis at the single-subject level. In this context, the BAC of 80% found in the ARMS-T vs FE comparison supports that an active neurobiological disease process parallels the clinical transition from the prodromal to the manifest stage of psychosis.

The brain regions most relevant for the present discriminative multivariate classification were in line with previously shown GM abnormalities[22,27,33–35] we observed in the temporal, limbic, and prefrontal cortex within the ARMS group and in the temporoinsular cortex and cerebellum within the FE subjects (for review of specific regions).[32,22,27,33–35] A recently published voxel-based meta-analysis of antipsychotic-naive ARMS and FE patients indicated similar consistent GM abnormalities in the patient groups as compared with controls.[10] In line with the above findings, a recent meta-analysis of functional imaging studies confirmed abnormal neural activity in schizophrenic patients during auditory hallucinations.[36] In terms of predicting clinical outcome and sustaining clinical applications of psychiatric imaging, there is further evidence from functional and neurochemical ARMS studies that the extent of abnormality at baseline is predictive of subsequent conversion to psychosis.[9]

Limitations of this study are as follows. First, our results may have been influenced by the heterogeneity of ARMS subgroups such as patients with APS or BLIPS.[1] To account for this, we controlled for the potential effects of covariates such as age, gender, and education, but other factors like substance abuse and cognitive functioning[23,37] could have played a confounding role. Although only a minor number of patients received antipsychotic treatment, we cannot rule out a modifying effect on neuroanatomical patterns, in particular, in contrasts including FE patients. Although there is clear evidence demonstrating the impact of antipsychotic medication on symptoms, its impact on neuroanatomical patterns of GM is widely unclear.[38] However, in a recent meta-analysis, we have selectively included studies in antipsychotic-naive subjects to control for the confounding effect of antipsychotics[10] because there is evidence indicating that chronic antipsychotic treatment can influence

Downloaded from https://academic.oup.com/schizophreniabulletin/article/39/5/1105/1925098 by guest on 16 August 2022

GM in subjects with established psychosis[39] and in the early phases of psychosis, the temporal and prefrontal cortex.[38] As in previous neuroimaging studies in ARMS subjects using SVM methods, our group sizes were modest; we can therefore not exclude the existence of group differences because of limited statistical power. It is important to validate our findings in larger, multicentre data sets.

Conclusions

This study provides further evidence that pattern recognition methods may indeed have the potential to delineate neuroanatomical phenotypes that constitute disease signatures beyond the level of coarse between-group differences. Psychosis onset in patients with FE is characterized by known temporoinsular, anterior cingulate, and cerebellar GM reductions; differences in neuroanatomical patterns were evident between FE and HC. To reveal the diagnostic specificity of SVM methods, other clinical samples at high risk for developing neuropsychiatric conditions as bipolar disorder, major depression, or borderline personality disorder are needed. To ultimately help treating patients, future single-subject imaging studies will also benefit from multimodal image analyses[6] and are needed to link basic research and clinical outcomes.[40]

Supplementary Material

Supplementary material is available at http://schizophreniabulletin.oxfordjournals.org.

Funding

Swiss National Science Foundation (3200-057216.99, 3200-0572216.99, PBBSB-106936, and 3232BO-119382); and the Nora van Meeuwen-Haefliger Stiftung, Basel (CH).

Acknowledgments

We are particularly grateful for Dr. Reinhold Bader's support in integrating the VBM5 and HAMMER/RAVENS and SVM algorithms into the batch system of the Linux Super-computing Cluster for the Munich and Bavarian Universities. We also would like to thank Martina Klemm, Doris Blaser, and Claudine Pfister for their help in preparing and editing this manuscript. The Authors have declared that there are no conflicts of interest in relation to the subject of this study.

References

1. Fusar-Poli P, Borgwardt S, Bechdolf A, *et al.* The psychosis high risk state: a comprehensive state of the art review. *Arch Gen Psychiatry*. In press.
2. Fusar-Poli P, Yung AR. Should attenuated psychosis syndrome be included in DSM-5? *Lancet*. 2012;379:591–592.
3. Mechelli A, Riecher-Rössler A, Meisenzahl EM, *et al.* Neuroanatomical abnormalities that predate the onset of psychosis: a multicenter study. *Arch Gen Psychiatry*. 2011;68:489–495.
4. Fusar-Poli P, Perez J, Broome M, *et al.* Neurofunctional correlates of vulnerability to psychosis: a systematic review and meta-analysis. *Neurosci Biobehav Rev*. 2007;31:465–484.
5. Howes OD, Montgomery AJ, Asselin MC, Murray RM, Grasby PM, McGuire PK. Molecular imaging studies of the striatal dopaminergic system in psychosis and predictions for the prodromal phase of psychosis. *Br J Psychiatry Suppl*. 2007;51:s13–s18.
6. Fusar-Poli P, McGuire P, Borgwardt S. Mapping prodromal psychosis: a critical review of neuroimaging studies. *Eur Psychiatry*. 2012;27:181–191.
7. Fusar-Poli P, Borgwardt S, Crescini A, *et al.* Neuroanatomy of vulnerability to psychosis: a voxel-based meta-analysis. *Neurosci Biobehav Rev*. 2011;35:1175–1185.
8. Ettinger U, Williams SC, Meisenzahl EM, Moller HJ, Kumari V, Koutsouleris N. Association between brain structure and psychometric schizotypy in healthy individuals [published online ahead of print October 24, 2011]. *World J Biol Psychiatry*.
9. Smieskova R, Fusar-Poli P, Allen P, *et al.* Neuroimaging predictors of transition to psychosis–a systematic review and meta-analysis. *Neurosci Biobehav Rev*. 2010;34:1207–1222.
10. Fusar-Poli P, Radua J, McGuire P, Borgwardt S. Neuroanatomical maps of psychosis onset: voxel-wise meta-analysis of antipsychotic-naive VBM studies [published online ahead of print November 17, 2011]. *Schizophr Bull*. doi:10.1093/schbul/sbr134
11. Henley SM, Ridgway GR, Scahill RI, *et al.*; EHDN Imaging Working Group. Pitfalls in the use of voxel-based morphometry as a biomarker: examples from huntington disease. *AJNR Am J Neuroradiol*. 2010;31:711–719.
12. Friston KJ, Ashburner J. Generative and recognition models for neuroanatomy. *Neuroimage*. 2004;23:21–24.
13. Klöppel S, Stonnington CM, Chu C, *et al.* Automatic classification of MR scans in Alzheimer's disease. *Brain*. 2008;131:681–689.
14. Noble WS. What is a support vector machine? *Nat Biotechnol*. 2006;24:1565–1567.
15. Koutsouleris N, Meisenzahl EM, Davatzikos C, *et al.* Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Arch Gen Psychiatry*. 2009;66:700–712.
16. Kawasaki Y, Suzuki M, Kherif F, *et al.* Multivariate voxel-based morphometry successfully differentiates schizophrenia patients from healthy controls. *Neuroimage*. 2007;34:235–242.
17. Teipel SJ, Born C, Ewers M, *et al.* Multivariate deformation-based analysis of brain atrophy to predict Alzheimer's disease in mild cognitive impairment. *Neuroimage*. 2007;38:13–24.
18. Davatzikos C, Fan Y, Wu X, Shen D, Resnick SM. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. *Neurobiol Aging*. 2008;29:514–523.
19. Bendfeldt K, Klöppel S, Nichols TE, *et al.* Multivariate pattern classification of gray matter pathology in multiple sclerosis. *Neuroimage*. 2012;60:400–408.
20. Klöppel S, Chu C, Tan GC, *et al.*; PREDICT-HD Investigators of the Huntington Study Group. Automatic

Downloaded from https://academic.oup.com/schizophreniabulletin/article/39/5/1105/1925098 by guest on 16 August 2022

detection of preclinical neurodegeneration: presymptomatic Huntington disease. *Neurology*. 2009;72:426–431.

21. Koutsouleris N, Borgwardt S, Meisenzahl EM, Bottlender R, Moller HJ, Riecher-Rössler A. Disease prediction in the at-risk mental state for psychosis using neuroanatomical biomarkers: results from the FePsy study [published online ahead of print November 10, 2011]. *Schizophr Bull*. doi:10.1093/schbul/sbr145
22. Borgwardt SJ, McGuire PK, Aston J, *et al.* Reductions in frontal, temporal and parietal volume associated with the onset of psychosis. *Schizophr Res*. 2008;106:108–114.
23. Riecher-Rössler A, Pflueger MO, Aston J, *et al.* Efficacy of using cognitive status in predicting psychosis: a 7-year follow-up. *Biol Psychiatry*. 2009;66:1023–1030.
24. Riecher-Rössler A, Aston J, Ventura J, *et al.* [The Basel Screening Instrument for Psychosis (BSIP): development, structure, reliability and validity]. *Fortschr Neurol Psychiatr*. 2008;76:207–216.
25. Yung AR, Phillips LJ, McGorry PD, *et al.* Prediction of psychosis. A step towards indicated prevention of schizophrenia. *Br J Psychiatry Suppl*. 1998;172:14–20.
26. Fusar-Poli P, Bonoldi I, Yung AR, *et al.* Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. *Arch Gen Psychiatry*. 2012;69:220–229.
27. Pantelis C, Velakoulis D, McGorry PD, *et al.* Neuroanatomical abnormalities before and after onset of psychosis: a cross-sectional and longitudinal MRI comparison. *Lancet*. 2003;361:281–288.
28. Koutsouleris N, Davatzikos C, Bottlender R, *et al.* Early recognition and disease prediction in the at-risk mental states for psychosis using neurocognitive pattern classification. *Schizophr Bull*. 2011.
29. Sun Y, Todorovic S, Goodison S. Local-learning-based feature selection for high-dimensional data analysis. *IEEE Trans Pattern Anal Mach Intell*. 2010;32:1610–1626.
30. Filzmoser P, Liebmann B, Varmuza K. Repeated double cross validation. *J Chemometr*. 2009;23:160–171.
31. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 2002;15:273–289.
32. Borgwardt S, McGuire P, Fusar-Poli P. Gray matters!–mapping the transition to psychosis. *Schizophr Res*. 2011;133:63–67.
33. Meisenzahl EM, Koutsouleris N, Bottlender R, *et al.* Structural brain alterations at different stages of schizophrenia: a voxel-based morphometric study. *Schizophr Res*. 2008;104:44–60.
34. Chan RC, Di X, McAlonan GM, Gong QY. Brain anatomical abnormalities in high-risk individuals, first-episode, and chronic schizophrenia: an activation likelihood estimation meta-analysis of illness progression. *Schizophr Bull*. 2011;37:177–188.
35. Takahashi T, Wood SJ, Yung AR, *et al.* Progressive gray matter reduction of the superior temporal gyrus during transition to psychosis. *Arch Gen Psychiatry*. 2009;66:366–376.
36. Jardri R, Pouchet A, Pins D, Thomas P. Cortical activations during auditory verbal hallucinations in schizophrenia: a coordinate-based meta-analysis. *Am J Psychiatry*. 2011;168:73–81.
37. Fusar-Poli P, Deste G, Smieskova R, *et al.* Cognitive functioning in prodromal psychosis: a meta-analysis. *Arch Gen Psychiatry*. 2012;69(6):562–571.
38. Smieskova R, Fusar-Poli P, Allen P, *et al.* The effects of antipsychotics on the brain: what have we learnt from structural imaging of schizophrenia?–a systematic review. *Curr Pharm Des*. 2009;15:2535–2549.
39. Ho BC, Andreasen NC, Ziebell S, Pierson R, Magnotta V. Long-term antipsychotic treatment and brain volumes: a longitudinal study of first-episode schizophrenia. *Arch Gen Psychiatry*. 2011;68:128–137.
40. Borgwardt S, Fusar-Poli P. Third-generation neuroimaging in early schizophrenia: translating research evidence into clinical utility. *Br J Psychiatry*. 2012;200:270–272.

Downloaded from https://academic.oup.com/schizophreniabulletin/article/39/5/1105/1925098 by guest on 16 August 2022