Individualized prediction of psychosis in subjects with an at-risk mental state

Eleni Zarogianni [a,*], Amos J. Storkey [b], Stefan Borgwardt [c], Renata Smieskova [c], Erich Studerus [d], Anita Riecher-Rössler [d], Stephen M. Lawrie [a]

[a] *Division of Psychiatry, School of Clinical Sciences, University of Edinburgh, The Royal Edinburgh Hospital, Morningside Park, UK*
[b] *Institute for Adaptive and Neural Computation, University of Edinburgh, UK*
[c] *Department of Psychiatry (UPK), University of Basel, Switzerland*
[d] *Center for Gender Research and Early Detection, University of Basel Psychiatric Hospital, Switzerland*

ABSTRACT

Early intervention strategies in psychosis would significantly benefit from the identification of reliable prognostic biomarkers. Pattern classification methods have shown the feasibility of an early diagnosis of psychosis onset both in clinical and familial high-risk populations. Here we were interested in replicating our previous classification findings using an independent cohort at clinical high risk for psychosis, drawn from the prospective FePsy (Fruherkennung von Psychosen) study. The same neuroanatomical-based pattern classification pipeline, consisting of a linear Support Vector Machine (SVM) and a Recursive Feature Selection (RFE) achieved 74% accuracy in predicting later onset of psychosis. The discriminative neuroanatomical pattern underlying this finding consisted of many brain areas across all four lobes and the cerebellum. These results provide proof-of-concept that the early diagnosis of psychosis is feasible using neuroanatomical-based pattern recognition.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Over the past 10 years, the target of early intervention strategies has shifted from the early diagnosis of first episode of psychosis (Perkins et al., 2005; Marshall et al., 2005) towards the early identification and treatment of individuals exhibiting early signs of psychosis. Early detection and intervention centres, such as the Personal Assessment and Crisis Evaluation (PACE) in Australia (Yung et al., 1996), the Prevention through Risk Identification, Management and Education (PRIME) in the United States (McGlashan et al., 2003), and the Outreach and Support in South London (OASIS) clinic (Fusar-Poli et al., 2013a, 2013b) have been set up worldwide, aiming to provide case management and provisional treatment for individuals presenting with sub-threshold psychotic symptoms and/or a decline in functioning. Individuals with this clinical presentation are considered at increased risk for developing psychosis and are named as ultra high-risk (UHR) individuals or at an at-risk mental state (ARMS).

The rationale for an ARMS stems from observations that psychosis and schizophrenia begin many years before the emergence of frank, psychotic symptoms, with nonspecific changes, perceptual alterations and often attenuated or transient psychotic disturbances (Riecher-Rössler et al., 2006). Currently, the identification of an ARMS relies on operationalized criteria that are based on a combination of trait and state risk factors (Yung et al., 2008; Yung et al., 1998). Despite the fact that these operationalized tools are highly reliable among raters, their predictive accuracy might be affected by the prevalence of the condition in the groups examined (Fusar-Poli et al., 2015), the heterogeneity in ARMS populations themselves (Fusar-Poli et al., 2016; Nelson et al., 2011) and other factors interplaying with transition risk (such as age, duration of follow-up examination, Fusar-Poli et al., 2012a).

Reliable prognostic biomarkers could complement current early detection tools by providing objective methods to evaluate risk of developing psychosis in ARMS populations. A series of putative biomarkers have been recently identified, suggesting that the ARMS is characterized by abnormalities in the neurocognitive domain (Lencz et al., 2006; Fusar-Poli et al., 2012b; Koutsouleris et al., 2012b) and alterations at the neuroanatomical (Fusar-Poli et al., 2011; Smieskova et al., 2010; Mechelli et al., 2011; Dazzan et al., 2012) and neurofunctional level (Fusar-Poli et al., 2007). Conversion to psychosis in ARMS subjects has been associated with reduced grey matter volume in the prefrontal and temporal cortices and other subcortical brain structures (Smieskova et al., 2010; Borgwardt et al., 2007a; Borgwardt et al., 2007b; Borgwardt et al., 2008; Pantelis et al., 2003; Koutsouleris et al., 2009; Mechelli et al., 2011; Dazzan et al., 2012; Fusar-Poli et al., 2011).

* Corresponding author at: Kennedy Tower, Division of Psychiatry, School of Clinical Sciences, University of Edinburgh, the Royal Edinburgh Hospital, Morningside Park, EH10 5HF, UK.

E-mail address: ezarogia@exseed.ed.ac.uk (E. Zarogianni).

Recently, multivariate pattern recognition approaches, including SVM, have emerged as powerful tools in the indetification of objective neuromanatomical biomarkers by taking into account inter-regional correlations between brain regions and they also have great potential for clinical translation as these methods can make inferences at a single-subject level (Orrù et al., 2012; Zarogianni et al., 2013). These methods may thus provide the means for an individualized risk assessment and prediction of psychosis conversion and possibly deliver increased sensitivity and specificity, both of which are essential for informing individualized prevention care (Phillips et al., 2006).

Previous machine learning studies have shown that a neuroanatomical-based prediction of psychosis is possible at the single-subject level (Koutsouleris et al., 2009; Borgwardt et al., 2013), providing diagnostic accuracy up to 85% (Koutsouleris et al., 2012a; Koutsouleris et al., 2015). Based on our previous study (Zarogianni et al., 2016), we have shown that a linear MRI-based Support Vector Machine (SVM) classifier can predict with significant accuracy the later transition to schizophrenia in a cohort of familial high-risk individuals. Here, we aimed to examine the generalizability of our classification method in predicting transition to psychosis using baseline structural MRI data from an independent cohort of subjects with an ARMS.

2. Materials and methods

2.1. Subjects

Subjects included in this analysis were part of the prospective, early psychosis, FePsy (Fruherkennung von Psychosen) study (Riecher-Rössler et al., 2006; Riecher-Rössler et al., 2007). Individuals potentially in an early stage of psychosis were referred to the 'Early Psychosis Clinic' mainly from the Psychiatric Outpatient Department, at the University Hospital in Basel, Switzerland. Regular information campaigns with general practicioners, private psychiatrists and social workers were organized locally in order to encourage low-threshold referrals from professionals outside the University Hospital. For more information on recruitment strategies please see Riecher-Rössler et al. (2007).

Subjects were initially screened using the Basel Screening Instrument for Psychosis (BSIP, Riecher-Rössler et al., 2008), which is a 46-item checklist based on variables reported as risk factors or predictors of psychosis (Riecher-Rössler et al., 2006, Riecher-Rössler et al., 2007). For psychosis items, the Brief Psychiatric Rating Scale (BPRS; expanded version (Ventura et al., 1993)) for assessing (pre)-psychotic phenomena was incorporated. The BSIP allows the rating of individuals regarding the inclusion/exclusion criteria, corresponding to the PACE criteria (Yung et al., 1998), and has been shown to have a good interrater reliability ($\kappa = 0.67$) for the assessment of the main outcome category "at risk for psychosis" and a high predictability (Riecher-Rössler et al., 2008).

Subjects were identified as at an ARMS if they manifested one (or more) of the following criteria, corresponding to the widely used PACE criteria (Yung et al., 1998): i) presence of attenuated psychotic-like symptoms (APS), ii) brief limited intermittent psychotic symptoms (BLIPS) or iii) a genetic risk of psychosis plus at least 2 further risk factors according to the BSIP checklist. Inclusion because of attenuated psychotic symptoms required that change in mental state had to be present at least several times a week and for >1 week duration (a score of 2 or 3 on the Brief Psy- chiatric Rating Scale (BPRS) hallucination item or 3 or 4 on BPRS items for unusual thought content or suspi- ciousness). Inclusion because of BLIPS 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, with each symptom lasting <1 week before resolving spontaneously.

Additionally, negative symptomatology was assessed using the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1989), in combination with the BSIP. Briefly, the SANS assessment is a well-recognised rating scale that consists of 19 items assessing negative symptoms of psychosis, which are further grouped into five domains (affective flattening, alogia, avolition-apathy, anhedonia-asociality, and inattention).

Exclusion criteria included age below 18 years, insufficient knowledge of German, IQ <70, previous episodes of schizophrenic psychosis (treated with major tranquilizers for >3 weeks), a clearly diagnosed brain disease or substance dependency (except for cannabis dependency), or psychotic symptoms within a clearly diagnosed depression or borderline personality disorder.

Study inclusion started in March 2000 and continued until February 2004. Individuals were followed up at monthly intervals during the first year, at 3-month intervals during the second and third year and annually thereafter until transition to psychosis was established or until the end of the follow-up period (2007). In general, all ARMS subjects were followed up for over 4 years during which they were also offered supportive counselling and clinical management.

In total, 37 ARMS individuals were recruited. Thirty of the 37 ARMS individuals never received antipsychotic medication while 7 participants were administered low doses of antipsychotic medication for behavioural control (2 participants on olanzapine, 2 Chlorprothixene and 3 risperidone) prior to study inclusion, all for <3 weeks.

Transition to psychosis was operationally defined by meeting PACE criteria (Yung et al., 1998): 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 >1 week to be deemed a conversion to frank psychosis.

Follow-up information for 2 ARMS subjects was not available. In this regard, 16 of the 35 ARMS individuals with retained follow-up information made a transition to psychosis (ARMS-T) and 19 did not convert (ARMS-NT, see Table 1).

2.2. Image acquisition and preprocessing

Subjects were scanned using a Siemens (Erlangen, Germany) Magnetom Vision 1.5 T scanner at the University Hospital Basel. Head movement was minimized by foam padding and velcro straps across the forehead and chin. A three-dimensional volumetric spoiled gradient recalled echo sequence generated 176 contiguous, 1 mm thick sagittal slices. Imaging parameters were: time-to-echo, 4 msec; time-to-repetition, 9.7 msec; flip angle, 12; matrix size, 200×256; field of view, 25.6 $\times$ 25.6 cm matrix; voxel dimensions, $1.28 \times 1 \times 1$ mm.

Standard VBM procedures (Ashburner and Friston, 2000) were followed using SPM5 (http://www.fil.ion.ucl.ac.uk/spm/). Study-specific templates and customized prior probability maps were constructed using data from all study groupds in order to represent the entire study population and therefore minimise bias for spatial normalization. The scans were normalized to the generic SPM T1 template using 12-point linear affine transformation to minimise the residual sum of squares differences between the images and the template. A study-specific T1 template was created from the mean image calculated from all the normalized T1 images and smoothed at 8-mm full-width at half maximum (FWHM). To generate study-specific brain tissue a priori maps, the normalized images were segmented and then mean images for the normalized GM, WM and CSF segments were produced and finally smoothed at 8-mm full-width at half maximum (FWHM).

Then, the baseline T1 scans entered the same pre-processing pipeline, described previously (Zarogianni et al., 2016). Briefly, T1 brain scans were segmented in native space into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) and study-specific images for GM, WM and CSF, after which the SPM brain extraction function returned a tissue mask for each scan. These masks were, then, applied to the original T1 images to remove non-brain tissue. T1 brain images were, then, spatially normalized to the study-specific T1 template, using a 12-parameter linear affine transformation. Bilinear interpolation was used to resample the normalized images and write MNI-

Table 1
Socio-demographic and clinical information of the ARMS-T and ARMS-NT study groups.

	Study groups		p
	ARMS-T	ARMS-NT	
Socio-demographic variables			
N	16	19	
Mean age at baseline, y (sd)	26.8 (6.5)	23.9 (6.2)	ns[a]
Sex (male), n (%)	11 (69)	10 (52)	ns[b]
Educational level			ns[b]
<9 y, n (%)	3	8	
9–11 y, n (%)	7	7	
12–13 y, n (%)	5	2	
<13 y, n (%)	1	2	
Mean verbal IQ (MWT-B) (sd)	109.6 (12.6)	107.3 (15.4)	ns[a]
Cannabis use at baseline			ns[b]
None	10	11	
Rarely	1	1	
Several times/month	0	2	
Several times/week	4	0	
Daily	1	5	
Antipsychotics before entry, n (%)	6 (37.5)	1 (5)	<0.05[b]
Antidepressants at baseline, n (%)	7 (44)	5 (26)	ns[b]
Family history			ns[b]
No relative	15	16	
One 1rst degree	1	2	
One 2nd degree	0	1	
Clinical variables			
Mean BPRS total score at baseline (sd)	42.3 (10.6)	35.7 (7.1)	ns[c]
Mean SANS global score at baseline (sd)	9.75 (5.8)	7.7 (4.2)	ns[c]
Mean interval between MRI and disease onset, d (sd)	306.3 (318.3)	na	

ARMS-T: at-risk mental state individuals that later developed psychosis; ARMS-NT: at-risk mental state subjects that did not make a transition. BPRS: the Brief Psychiatric Rating Scale; SANS: the Scale for the Assessment of Negative Symptoms. Verbal IQ Mehrfach-Wortwahl-Test-B.

[a] Student's *t*-test.
[b] Fisher's exact test.
[c] Mann-Whitney *U*-test.

normalized images into the stereotactic space at a $1 \times 1 \times 1$ mm voxel resolution. Normalized images were again segmented using study-specific a priori templates and spatially normalized segments for GM, WM and CFS were returned. Finally, the spatially normalized, segmented images were smoothed with an 8 mm full-width at maximum (FWHM) isotropic Gaussian kernel.

2.3. Multivariate pattern classification analysis

2.3.1. Support vector machine

The SVM is a multivariate pattern recognition technique that has been widely used in neuroimaging-based studies because it can provide optimal decision rules for classification. Here, a linear SVM classifier was used for the classification task because it allows the straightforward extraction of the corresponding discrimination map. A detailed description of the SVM was given in our previous work (Zarogianni et al., 2016).

2.3.2. Feature extraction

Prior to SVM, all smoothed and normalized GM maps were mapped to the Automated Anatomical Labeling (AAL) brain atlas (Tzourio-Mazoyer et al., 2002) and GM density volumes corresponding to the 116 brain regions of the template were returned. Features in the training set were also scaled to the [0 1] template before applying the same normalization template to the testing set.

2.3.3. Recursive feature elimination

To identify the most significant features in the classification task and simultaneously increase classification performance, the recursive feature elimination (RFE) method (Guyon et al., 2002) was embedded in a nested leave-one-out cross-validation (LOO-CV), as described in Fig. 1

and previously (Zarogianni et al., 2016). For a more detailed description of the RFE methodology, please see the Supplementary material.

2.3.4. Permutation testing

Permutation testing was performed in order to derive a p value for the accuracy of our classifier. We permuted the class labels 1000 times (randomly assigning patient and control labels to the training subjects) and repeated the entire nested LOO-CV procedure. We then calculated the number of times in which the specificity (percentage of true negative) and sensitivity (percentage of true positive) for the permuted labels were higher than those obtained for the real labels. Dividing this number by 1000 we derived a p value for the classification accuracies.

2.3.5. Discrimination map

A discrimination map was again generated based on the weight coefficients of the features that were selected by the RFE method (Fig. 2). The discrimination map consists of brain regions that according to the RFE methodology are the most distinctive in the classification task and provides a spatial representation of the decision function in that every feature contributes with a certain weight to this function (or hyperplane). The SVM weight vector is a linear combination or weighted average of the support vectors and defines the decision boundary. The weight vector is therefore a spatial representation of the decision boundary. Every feature contributes with a certain weight to the decision boundary or classification function. Given a positive and a negative class ($+1 = $ ARMS-T; $-1 = $ ARMS-NT group), a positive weight means the weighted average in that region was higher for the ARMS-T group,
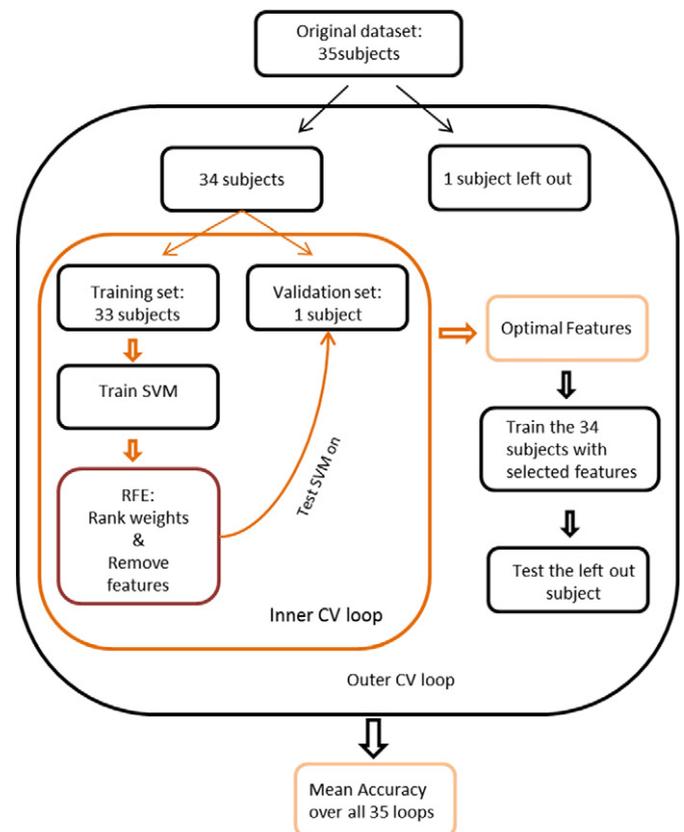

Fig. 1. Representation of the nested LOO-CV SMV-RFE method. We employed a nested LOO-CV where we repeatedly excluded one subject to comprise the testing set and the remaining subjects were again repeatedly repartitioned in an internal validation loop where one subject was left out for validation and the rest formed the internal training group. In this loop, RFE was repeatedly performed and the mean accuracy on the validation group at each elimination level was recorded until all features were removed. The feature set that produced the maximum accuracy on the validation set was selected and applied to the testing set of the outer testing loop. Finally, mean accuracy was calculated across all outer CV loops.
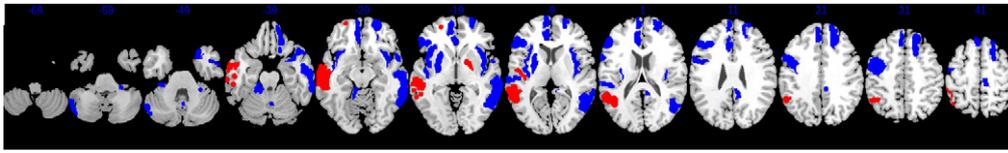
Fig. 2. Discrimination maps for the classification of ARMS-T vs ARMS-NT. The colours represent the weight of each feature in the classification function (the red scale represents positive weights and the blue scale represents negative weights). The SVM weight vector is a linear combination or weighted average of the support vectors and defines the decision boundary. The weight vector is therefore a spatial representation of the decision boundary. Every feature contributes with a certain weight to the decision boundary or classification function. Given a positive and a negative class ($+1$ = ARMS-T; -1 = ARMS-NT group), a positive weight means the weighted average in that region was higher for the ARMS-T group, and a negative weight means the weighted average was higher for ARMS-NT group. Note: features correspond to GM volume measures in the AAL-defined brain regions, and not voxels.

and a negative weight means the weighted average was higher for ARMS-NT group. Since the SVM classifier is multivariate by nature, it should be noted that all brain regions constituting the decision function contribute to the classification.

3. Results

3.1. Socio-demographic and clinical findings

The rate of conversion to psychosis was 45.7% in this ARMS sample of 35 subjects. The mean interval between the baseline scan and disease conversion was 306 days (median: 263, range: 25–1137 days). There were no significant differences between converters and non-converters to psychosis with regards to age, gender, educational level, verbal IQ, cannabis use at study entry, baseline global BPRS and SANS scores (Table 1). However, there was a significant difference in the two groups in the use of antipsychotics, with 6 ARMS-T subjects and 1 AMRS-NT having taken neuroleptics some time before study inclusion (Table 1).

3.2. SVM classification analysis

The application of our LOO-CV SVM-RFE methodology to baseline structural MRI data of the ARMS groups achieved 74% accuracy in predicting conversion to psychosis (Table 2). Six out of 16 subjects in the ARMS-T group were wrongly classified as ARMS—NT, while only 3 out of 19 subjects in the ARMS-NT group were incorrectly labeled as ARMS-T (sensitivity/specificity: 63%/84%; PPV/NPV: 77%/73%; permutation test $p = 0.002$).

The likelihood ratio of a positive test result was LR $+ = 3.95$ (95% CI: 1.34 to 5.94, Table 3), meaning that a positive prognostic test in a given ARMS subject would increase the probability of a subsequent transition to psychosis from 45.7% to 77% (posttest probability = posttest odds / posttest odds + 1, posttest odds = pretest odds*LR +).

The misclassified ARMS-NT subjects did not significantly differ from the correctly classified ARMS-NT in any of the socio-demographic or clinical variables (Table 3). On the contrary, the misclassified ARMS-T subjects were significantly different from the correctly classified ARMS-T individuals in terms of gender distribution and use of antipsychotic medication before study entry (Table 3). This may partly explain the lower sensitivity of the SVM-RFE method, since the ARMS-T group consisted of a more inhomogenous group of individuals with regards

to anti-psychotic medication, which in turn might have hindered the identification of a common neuroanatomical signature across subjects in this group. The effect of antipsychotic medication in brain structure is widely acknowledged by the scientific community (Fusar-Poli et al., 2013a, Smieskova et al., 2009, Navari and Dazzan 2009), and might have played a major role in the classification of the ARMS subjects here, despite the fact the exposure was before study entry and relatively brief.

The spatially distributed network that discriminated between the two groups was quite extensive and consisted of GM abnormalities in a spatially distributed network covering all four lobes and the cerebellum. Table 4 presents the most discriminating regions in the classification task, namely the brain regions with the highest (absolute) weight value that contributed relatively higher to the decision function. Specifically, the regions that contributed more in the classification of the ARMS-T subjects included the cerebellum, parts of the superior temporal pole bilaterally, the right anterior cingulate cortex, the right superior

Table 2
Classification performance.

	TP	TN	FP	FN	Sens (%)	Spec (%)	BAC (%)	FPR (%)	PPV (%)	NPV (%)	LR+/LR−
ARMS-T vs ARMS-NT	10	16	3	6	62.5	84.2	74.2	15.7	77	73	3.9/0.45

The diagnostic performance was evaluated by means of sensitivity (Sens), specificity (Spec), balanced accuracy (BAC), false positive rate (FPR) and positive/negative predictive value (PPV/NPV). LR + was also calculated as sensitivity / 1-specificity and LR − = 1-sensitivity / specificity.

Table 3
Misclassification analysis.

	ARMS-T → ARMS-NT	ARMS-T → ARMS-T	p	ARMS-NT → ARMS-T	ARMS-NT → ARMS-NT	p
Socio-demographic variables						
N	6	10		3	16	
Mean age at baseline, y (sd)	29.2 (9)	25.4 (4.5)	ns[a]	24.8 (7.2)	23.8 (6.2)	ns[a]
Sex (male), n (%)	2 (33)	9 (90)	<0.05[b]	1 (33)	9 (56)	ns[b]
Educational level			ns[c]			ns[c]
<9 y, n (%)	1	2		2	6	
9–11 y, n (%)	2	5		1	6	
12–13 y, n (%)	2	3		0	2	
<13 y, n (%)	1	0		0	2	
Cannabis use at baseline			ns[c]			ns[c]
None	4	6		2	9	
Rarely	1	0		0	1	
Several times/month	0	0		1	1	
Several times/week	1	3		0	0	
Daily	0	1		0	5	
Antipsychotics before entry	5	1	<0.05[b]	0	1	ns[b]
Anti-depressants at baseline	3	4	ns[b]	2	3	ns[b]
Clinical variables						
Mean BPRS total score at baseline (sd)	45.7 (11.5)	40.2 (10.2)	ns[c]	38.3 (12.9)	37.5(6.2)	ns[c]
Mean SANS global score at baseline (sd)	9.7(7.7)	9.8 (4.8)	ns[c]	10.3 (5)	6.3 (4.7)	ns[c]
Mean interval between MRI and disease onset, d (sd)	427.5 (483.6)	245.7 (215.3)	ns[a]			

[a] Student's *t*-test.
[b] Fisher's exact test.
[c] Mann-Whitney *U* test.

Table 4
List of the most discriminative regions for the classification of ARMS-T vs ARMS-NT.

Lobe	Region/hemisphere	w
Negative weights		
Cerebellum	Cerebellum_Crus2_R	−0.0128
	Cerebellum_3_R	−0.0194
	Cerebellum_4_5_R	−0.0207
	Cerebellum_6_L	−0.0127
	Cerebellum_7b_R	−0.0146
	Cerebellum_10_L	−0.0247
	Vermis_8	−0.0147
Temporal	Temporal_Sup_R	−0.0067
	Temporal_Pole_Sup_L	−0.0084
	Temporal_Mid_L	−0.0075
Frontal	Frontal_Sup_L	−0.0197
	Frontal_Sup_Orb_L	−0.0166
	Frontal_Mid_R	−0.0089
	Frontal_Inf_Tri_R	−0.0106
	Frontal_Sup_Medial_R	−0.0084
	Frontal_Med_Orb_R	−0.0098
	Precentral_R	−0.0176
Parietal	Postcentral_R	−0.0102
	Paracentral_Lobule_R	−0.012
Limbic	Cingulum_Ant_R	−0.017
	Cingulum_Post_L	−0.0081
Basal ganglia	Putamen_R	−0.0172
Perisylvian	Insula_L	−0.0134
	Insula_R	−0.0206
Positive weights		
Temporal	Temporal_Mid_R	0.0069
	Heschl_R	0.0069
Frontal	Frontal_Sup_Orb_R	0.0101
Parietal	Parietal_Inf_R	0.0137
Basal ganglia	Pallidum_L	0.0072

Ant, anterior; Crus, crust; Inf, inferior; L, left hemisphere; Mid, middle; Med, medial; Orb, orbital; Post, posterior; R, right hemisphere; Sup, superior; w, weight vector of corresponding features in the classification process. Note: The SVM weight vector is a linear combination or weighted average of the support vectors and defines the decision boundary. The weight vector is therefore a spatial representation of the decision boundary. Every feature contributes with a certain weight to the decision boundary or classification function. Given a positive and a negative class ($+1 =$ ARMS-T; $−1 =$ ARMS-NT group), a positive weight means the weighted average in that voxel was higher for the ARMS-T group, and a negative weight means the weighted average was higher for ARMS-NT group.

medial frontal and left orbitofrontal cortex and the insula bilaterally, whereas regions with a higher weighted average for the ARMS-NT group consisted of the right inferior parietal lobe, right medial temporal lobe, the right orbitofrontal cortex and the left pallidum.

A discrimination map showing the spatial pattern by which the groups differ is also illustrated in Fig. 2. We emphasize that this spatially distributed pattern should not be interpreted as a statistical map, but rather as a spatial representation of the decision boundary.

4. Discussion

The present findings replicate our previous ones in that MRI-based classification methods were able to predict transition to psychosis in subjects at high clinical risk for developing the disorder using neuroanatomical data at study inclusion. The SVM-RFE classifier achieved 74% accuracy in classifying ARMS-T against ARMS-NT subjects.

The neuroanatomical decision function that discriminated the two groups was associated with GM abnormalities relying on a distributed network of regions covering most cortical and sub-cortical brain structures and the cerebellum. Our present findings agree with findings from a recent voxel-based meta-analysis that reported GM volume reductions in subjects that convert to psychosis in the insular and superior temporal lobe cortices (Fusar-Poli et al., 2011) and also with previous VBM findings on the same dataset (Borgwardt et al., 2007a).

Despite being significant, our classification accuracy here is lower than the accuracy observed in our familial high-risk group (Zarogianni et al., 2016) where the same classification pipeline was used.

Additionally, previous studies using the same ARMS cohort achieved higher classification performances than reported accuracies here (Koutsouleris et al., 2012a; Koutsouleris et al., 2015).

In the studies conducted by Koutsouleris et al. (2012b and 2015), their MRI-based classifier achieved 84.2% and 75% accuracy respectively, in correctly classifying ARMS-T against ARMS-NT individuals drawn from the FePsy study. Differences in the observed classification accuracies may be partly explained by differences in the preprocessing of the MRI scans (where RAVENS maps and the VBM 8 toolbox were used) and partly by the chosen implementation of the SVM classifier, which relied upon the construction of SVM ensembles that incorporated feature selection, model training and predictive learning wrapped together in a repeated nested cross-validation framework. Ensemble learning approaches are usually selected on the basis that they can achieve higher predictive performance than single classifiers, by combining multiple weak learning models that decide upon the classification of a new instance through majority voting (Polikar, 2006).

Compared to the diagnostic performance of our classifier in the familial high-risk cohort of the Edinburgh High Risk Study, the classification performance in the ARMS groups of the FePsy study was notably lower, contrary to what would be expected since the ARMS groups represent help-seeking individuals, most of whom already manifest putative transient and/or sub-threshold psychotic symptoms. Interestingly, 5 out of the 6 ARMS-T subjects that were misclassified received antipsychotic medication some time before study inclusion (Table 3) while the other misclassified ARMS-T subject was prescribed tranquilizers (Lorazepam). Many studies have reported the effect of antipsychotic medication on grey matter volume in the direction of significant regional reductions (Fusar-Poli et al., 2013b; Smieskova et al., 2009), thus possibly suggesting a neuroanatomical heterogeneity expressed with divergent pathophysiological trajectories between subjects receiving and subjects not receiving any anti-psychotic treatment.

Despite the lower diagnostic performance, our MRI-based classifier managed to increase the diagnostic certainty from 45.7% to 77% in case of a positive test result, suggesting that an MRI-based pattern classification system could, with refinement, become a useful part of a multi-step diagnostic procedure that would reliably quantify the risk for conversion to psychosis (Schmidt et al., 2017; Studerus et al., 2017) and inform appropriate care and treatment strategies.

Certain limitations of this study have to be considered. Again the sample size in this investigation is small. The rate of transition to psychosis amounted to nearly 46%, which is generally in keeping with other clinically at-risk cohorts (Koutsouleris et al., 2009; Yung et al., 2003; Klosterkötter et al., 2001). However, it is not clear how the classifier would perform if presented with an ARMS cohort with significantly lower conversion rates. Finally, the administration of antipsychotic and antidepressant medication might have confounded our results, despite the fact that any drug treatment was administered before study inclusion.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.schres.2017.08.061.

References

Andreasen, N.C., 1989. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. Br. J. Psychiatry Suppl. 7, 49–58.

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry–the methods. NeuroImage 11 (6 Pt 1), 805–821.

Borgwardt, S.J., McGuire, P.K., Aston, J., et al., 2007a. Structural brain abnormalities in individuals with an at-risk mental state who later develop psychosis. Br. J. Psychiatry Suppl. 51, s69–75.

Borgwardt, S.J., Riecher-Rössler, A., Dazzan, P., et al., 2007b. Regional gray matter volume abnormalities in the at risk mental state. Biol. Psychiatry 61 (10), 1148–1156.

Borgwardt, S.J., McGuire, P.K., Aston, J., et al., 2008. Reductions in frontal, temporal and parietal volume associated with the onset of psychosis. Schizophr. Res. 106 (2–3), 108–114.

Borgwardt, S., Koutsouleris, N., Aston, J., et al., 2013. Distinguishing prodromal from first-episode psychosis using neuroanatomical single-subject pattern recognition. Schizophr. Bull. 39 (5), 1105–1114.

Dazzan, P., Soulsby, B., Mechelli, A., et al., 2012. Volumetric abnormalities predating the onset of schizophrenia and affective psychoses: an MRI study in subjects at ultra high risk of psychosis. Schizophr. Bull. 38 (5), 1083–1091.

Fusar-Poli, P., Perez, J., Broome, M., et al., 2007. Neurofunctional correlates of vulnerability to psychosis: a systematic review and meta-analysis. Neurosci. Biobehav. Rev. 31 (4), 465–484.

Fusar-Poli, P., Borgwardt, S., Crescini, A., et al., 2011. Neuroanatomy of vulnerability to psychosis: a voxel-based meta-analysis. Neurosci. Biobehav. Rev. 35 (5), 1175–1185.

Fusar-Poli, P., Bonoldi, I., Yung, A.R., et al., 2012a. Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. Arch. Gen. Psychiatry 69 (3), 220–229.

Fusar-Poli, P., Deste, G., Smieskova, R., et al., 2012b. Cognitive functioning in prodromal psychosis: a meta-analysis. Arch. Gen. Psychiatry 69 (6), 562–571.

Fusar-Poli, P., Byrne, M., Badger, S., et al., 2013a. Outreach and support in south London (OASIS), 2001–2011: ten years of early diagnosis and treatment for young individuals at high clinical risk for psychosis. Eur. Psychiatry 28 (5), 315–326.

Fusar-Poli, P., Smieskova, R., Kempton, M.J., et al., 2013b. Progressive brain changes in schizophrenia related to antipsychotic treatment? A meta-analysis of longitudinal MRI studies. Neurosci. Biobehav. Rev. 37 (8), 1680–1691.

Fusar-Poli, P., Cappucciati, M., Rutigliano, et al., 2015. At risk or not at risk? A meta-analysis of the prognostic accuracy of psychometric interviews for psychosis prediction. World Psychiatry 14 (3), 322–332.

Fusar-Poli, P., Cappucciati, M., Borgwardt, S., et al., 2016. Heterogeneity of psychosis risk within individuals at clinical high risk: a meta-analytical stratification. JAMA Psychiat. 73, 113–120.

Guyon, I., Weston, J., Barnhill, S., et al., 2002. Gene selection for cancer classification using support vector machines. Mach. Learn. 46 (1–3), 389–422.

Klosterkötter, J., Hellmich, M., Steinmeyer, E.M., et al., 2001. Diagnosing schizophrenia in the initial prodromal phase. Arch. Gen. Psychiatry 58 (2), 158–164.

Koutsouleris, N., Meisenzahl, E.M., Davatzikos, C., et al., 2009. Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. Arch. Gen. Psychiatry 66 (7), 700–712.

Koutsouleris, N., Borgwardt, S., Meisenzahl, E.M., et al., 2012a. Disease prediction in the at-risk mental state for psychosis using neuroanatomical biomarkers: results from the FePsy study. Schizophr. Bull. 38 (6), 1234–1246.

Koutsouleris, N., Davatzikos, C., Bottlender, R., et al., 2012b. Early recognition and disease prediction in the at-risk mental states for psychosis using neurocognitive pattern classification. Schizophr. Bull. 38 (6), 1200–1215.

Koutsouleris, N., Riecher-Rössler, A., Meisenzahl, E.M., et al., 2015. Detecting the psychosis prodrome across high-risk populations using neuroanatomical biomarkers. Schizophr. Bull. 41 (2), 471–482.

Lencz, T., Smith, C.W., McLaughlin, D., et al., 2006. Generalized and specific neurocognitive deficits in prodromal schizophrenia. Biol. Psychiatry 59 (9), 863–871.

Marshall, M., Lewis, S., Lockwood, A., et al., 2005. Association between duration of untreated psychosis and outcome in cohorts of first-episode patients: a systematic review. Arch. Gen. Psychiatry 62 (9), 975–983.

McGlashan, T.H., Zipursky, R.B., Perkins, D., et al., 2003. The PRIME North America randomized double-blind clinical trial of olanzapine versus placebo in patients at risk of being prodromally symptomatic for psychosis. I. Study rationale and design. Schizophr. Res. 61 (1), 7–18.

Mechelli, A., Riecher-Rössler, A., Meisenzahl, E.M., et al., 2011. Neuroanatomical abnormalities that predate the onset of psychosis: a multicenter study. Arch. Gen. Psychiatry 68 (5), 489–495.

Navari, S., Dazzan, P., 2009. Do antipsychotic drugs affect brain structure? A systematic and critical review of MRI findings. Psychol. Med. 39 (11):1763–1777. http://dx.doi.org/10.1017/S0033291709005315.

Nelson, Yuen, K., Yung, A.R., 2011. Ultra high risk (UHR) for psychosis criteria: are there different levels of risk for transition to psychosis? Schizophr. Res. 125 (1).

Orrù, G., Pettersson-Yeo, W., Marquand, A.F., et al., 2012. Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. Neurosci. Biobehav. Rev. 36 (4), 1140–1152.

Pantelis, C., Velakoulis, D., McGorry, P.D., et al., 2003. Neuroanatomical abnormalities before and after onset of psychosis: a cross-sectional and longitudinal MRI comparison. Lancet 361 (9354), 281–288.

Perkins, D.O., Gu, H., Boteva, K., et al., 2005. Relationship between duration of untreated psychosis and outcome in first-episode schizophrenia: a critical review and meta-analysis. Am. J. Psychiatry 162 (10), 1785–1804.

Phillips, K.A., Van Bebber, S., Issa, A.M., 2006. Diagnostics and biomarker development: priming the pipeline. Nat. Rev. Drug Discov. 5 (6), 463–469.

Polikar, R., 2006. Ensemble based systems in decision making. IEEE Circ. Syst. Mag. 6 (3): 21–45. http://dx.doi.org/10.1109/MCAS.2006.1688199.

Riecher-Rössler, A., Gschwandtner, U., Borgwardt, S., et al., 2006. Early detection and treatment of schizophrenia: how early? Acta Psychiatr. Scand. Suppl. 429, 73–80.

Riecher-Rössler, A., Gschwandtner, U., Aston, J., et al., 2007. The Basel early-detection-of-psychosis (FEPSY)-study–design and preliminary results. Acta Psychiatr. Scand. 115 (2), 114–125.

Riecher-Rössler, A., Aston, J., Ventura, J., et al., 2008. The Basel Screening Instrument for Psychosis (BSIP): development, structure, reliability and validity. Fortschr. Neurol. Psychiatr. 76 (4), 207–216.

Schmidt, A., Cappucciati, M., Radua, J., et al., 2017. Improving prognostic accuracy in subjects at clinical high risk for psychosis: systematic review of predictive models and meta-analytical sequential testing simulation. Schizophr. Bull. 43 (2), 375–388.

Smieskova, R., Fusar-Poli, P., Allen Pet al., 2009. The effects of antipsychotics on the brain: what have we learnt from structural imaging of schizophrenia?–a systematic review. Curr. Pharm. Des. 15 (22), 2535–2549.

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

Studerus, E., Ramyead, A., Riecher-Rössler, A., 2017. Prediction of transition to psychosis in patients with a clinical high risk for psychosis: a systematic review of methodology and reporting. Psychol. Med. 47 (7), 1163–1178.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., et al., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15 (1), 273–289.

Ventura, J., Lukoff, D., Nuechterlein, K.H., et al., 1993. Training and quality assurance with the brief psychiatric rating scale: the Drift Busters; Appendix 1 The Brief Psychiatric Rating Scale (expanded version). Int. J. Methods Psychiatr. Res. 3, 221–224.

Yung, A.R., McGorry, P.D., McFarlane, C.A., et al., 1996. Monitoring and care of young people at incipient risk of psychosis. Schizophr. Bull. 22 (2), 283–303.

Yung, A.R., Phillips, L.J., McGorry, P.D., et al., 1998. Prediction of psychosis. A step towards indicated prevention of schizophrenia. Br. J. Psychiatry Suppl. 172 (33), 14–20.

Yung, A.R., Phillips, L.J., Yuen, H.P., et al., 2003. Psychosis prediction: 12-month follow up of a high-risk ("prodromal") group. Schizophr. Res. 60 (1), 21–32.

Yung, A.R., Nelson, B., Stanford, C., et al., 2008. Validation of "prodromal" criteria to detect individuals at ultra high risk of psychosis: 2 year follow-up. Schizophr. Res. 105 (1–3), 10–17.

Zarogianni, E., Moorhead, T.W.J., Lawrie, S.M., 2013. Towards the identification of imaging biomarkers in schizophrenia, using multivariate pattern classification at a single-subject level. Neuroimage Clin. 3, 279–289.

Zarogianni, E., Storkey, A.J., Johnstone, E.C., et al., 2016. Improved individualized prediction of schizophrenia in subjects at familial high risk, based on neuroanatomical data, schizotypal and neurocognitive features. Schizophr. Res.