



Contents lists available at ScienceDirect

## Schizophrenia Research

journal homepage: [www.elsevier.com/locate/schres](http://www.elsevier.com/locate/schres)

## Use of neuroanatomical pattern regression to predict the structural brain dynamics of vulnerability and transition to psychosis

Nikolaos Koutsouleris<sup>a,\*</sup>, Christian Gaser<sup>b</sup>, Ronald Bottlender<sup>a</sup>, Christos Davatzikos<sup>d</sup>,  
Petra Decker<sup>a</sup>, Markus Jäger<sup>a</sup>, Gisela Schmitt<sup>a</sup>, Maximilian Reiser<sup>c</sup>,  
Hans-Jürgen Möller<sup>a</sup>, Eva M. Meisenzahl<sup>a</sup>

<sup>a</sup> Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-University, Munich, Germany

<sup>b</sup> Department of Psychiatry, Friedrich-Schiller-University, Jena, Germany

<sup>c</sup> Department of Radiology, Ludwig-Maximilians-University, Munich, Germany

<sup>d</sup> Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, USA

### ARTICLE INFO

#### Article history:

Received 16 May 2010

Received in revised form 12 August 2010

Accepted 22 August 2010

Available online xxx

#### Keywords:

At-risk mental state

Early psychosis

Deformation-based morphometry

Multivariate analysis

Support-vector regression

### ABSTRACT

**Background:** The at-risk mental state for psychosis (ARMS) has been associated with abnormal structural brain dynamics underlying disease transition or non-transition. To date, it is unknown whether these dynamic brain changes can be predicted at the single-subject level prior to disease transition using MRI-based machine-learning techniques.

**Methods:** First, deformation-based morphometry and partial-least-squares (PLS) was used to investigate patterns of volumetric changes over time in 25 ARMS individuals versus 28 healthy controls (HC) (1) irrespective of the clinical outcome and (2) according to illness transition or non-transition. Then, the baseline MRI data were employed to predict the expression of these volumetric changes at the individual level using support-vector regression (SVR).

**Results:** PLS revealed a pattern of pronounced morphometric changes in ARMS versus HC that affected predominantly the right prefrontal, as well as the perisylvian, parietal and periventricular structures ( $p < 0.011$ ), and that was more pronounced in the converters versus the non-converters ( $p < 0.010$ ). The SVR analysis facilitated a reliable prediction of these longitudinal brain changes in individual out-of-training cases (HC vs ARMS:  $r = 0.83$ ,  $p < 0.001$ ; HC vs converters vs non-converters:  $r = 0.83$ ,  $p < 0.001$ ) by relying on baseline patterns that involved ventricular enlargements, as well as prefrontal, perisylvian, limbic, parietal and subcortical volume reductions.

**Conclusions:** Abnormal brain changes over time may underlie an elevated vulnerability for psychosis and may be most pronounced in subsequent converters to psychosis. Pattern regression techniques may facilitate an accurate prediction of these structural brain dynamics, potentially allowing for an early recognition of individuals at risk of developing psychosis-associated neuroanatomical changes over time.

© 2010 Elsevier B.V. All rights reserved.

### 1. Introduction

The at-risk mental state for psychosis (ARMS) has been associated with subtle brain abnormalities (Borgwardt et al., 2007a,b; Koutsouleris et al., 2009a,b; Lawrie et al., 1999; Meisenzahl et al., 2008b; Pantelis et al., 2003; Seidman et al., 2003) qualitatively similar to those found in established schizophrenia (Gaser et al., 2004; Honea et al., 2005; Koutsouleris et al., 2008; Meisenzahl et al., 2008a). Recent

\* Corresponding author. Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-University, Nussbaumstr. 7, 80336 Munich, Germany. Tel.: +49 89 5160 5772.

E-mail address: [Nikolaos.Koutsouleris@med.uni-muenchen.de](mailto:Nikolaos.Koutsouleris@med.uni-muenchen.de) (N. Koutsouleris).

studies suggested that these alterations are neither confined to single cortical regions nor dispersed across the entire brain, but rather involve interconnected neural systems spanning prefrontal, perisylvian, parietal, limbic and cerebellar regions (Borgwardt et al., 2008; Koutsouleris et al., 2009b; Pantelis et al., 2003; Sun et al., 2009b). Additionally, investigations of longitudinal neuroanatomical changes in clinically defined ARMS individuals (Borgwardt et al., 2008; Pantelis et al., 2003; Sun et al., 2009b; Takahashi et al., 2009) and genetically defined high-risk subjects (Job et al., 2005, 2006; Lawrie et al., 2002) mainly found progressive volume losses within this “prodromal” pattern that primarily affected individuals with a subsequent illness transition. These findings have been interpreted within the concept of a “late neurodevelopmental disturbance”, evolving on the basis of a pre-existing neurobiological predisposition when cortical association areas are placed under increasing functional demand in late adolescence and early adulthood (Pantelis et al., 2005).

It is, however, unclear whether this process is *exclusively* active in those who ultimately develop the disease. Alternatively, the neurobiology of the ARMS may be characterized by a *gradual* progression of structural changes that reaches its full expression across disease transition, but that can also be traced to a lesser extent in vulnerable subjects without subsequent illness (DeLisi, 2008; Koutsouleris et al., 2009b; Takahashi et al., 2009). Thus, in the context of a possible neuroimaging-aided early recognition of psychosis (Bray et al., 2009; Koutsouleris et al., 2009a; Linden and Fallgatter, 2009; Sun et al., 2009c), it has to be evaluated (1) whether *spatiotemporal* neuroanatomical patterns underlying the ARMS can be deconstructed into specific illness-associated trajectories and trajectories conferring an unspecific vulnerability, and (2) whether these trajectories can be predicted at the individual level prior to disease transition.

The existing data suggest that ARMS-associated structural abnormalities are subtle and likely to occur within multicollinear patterns involving networks of interconnected brain regions. Therefore, univariate statistics may be limited in detecting these morphological signatures and in measuring their expression at the single-subject level for the purpose of an individualized early recognition, because these methods degrade these complex patterns into largely overlapping voxel-by-voxel measurements (Davatzikos, 2004). However, these limitations may be overcome by a methodological shift to multivariate analysis methods capable of tracing the high-dimensional structure of ARMS-associated morphological profiles at baseline and over time.

Therefore, we used deformation-based morphometry and partial least squares (PLS) (Gilboa et al., 2005; Kawasaki et al., 2007; McIntosh et al., 1996; Menzies et al., 2007; Nestor et al., 2002; Tura et al., 2008) (1) to investigate system-level covariance patterns of longitudinal structural brain changes in clinically defined ARMS individuals versus healthy controls, and (2) to examine whether these patterns were *exclusively* driven by ARMS individuals subsequently developing a schizophrenia spectrum disorder or, alternatively, whether subjects without disease transition expressed similar brain dynamics, albeit to a lesser degree. We expected these trajectories (1) to involve patterns of volumetric losses covering prefrontal, perisylvian, parietal, limbic and cerebellar structures, and (2) to be differentially expressed according

to illness transition or non-transition. Furthermore, we explored whether the expression of these trajectories could be predicted at the *individual level* using only the MRI data acquired at study inclusion. For this purpose, we employed support-vector regression (SVR) (Schölkopf and Smola, 2002), a multivariate machine-learning technique, due to its good generalization properties. We expected the SVR predictions to rest upon a neuroanatomical baseline pattern covering those brain regions that were subsequently affected by abnormal morphometric change.

## 2. Materials and methods

### 2.1. Participants

A prospective study of 28 healthy volunteers (HC; age at baseline (SD): 25.1 (3.6), 39% females) and 25 ARMS individuals (age at baseline (SD): 23.1 (4.7), 28% females) was conducted using structural MRI scanning and clinical assessments at baseline and after a mean (SD) of 3.7 (1.3) years. All followed subjects were part of a baseline population of 75 HC and 46 ARMS individuals (Koutsouleris et al., 2009a,b) who were recruited using previously described operationalized criteria employed by our and other research groups to study the neurobiology of the ARMS (Frommann et al., 2008; Hurlemann et al., 2008; Koutsouleris et al., 2009a,b; Meisenzahl et al., 2008b; Quednow et al., 2008; Schultze-Lutter et al., 2007). Briefly, different types of prodromal symptoms, including cognitive/perceptive basic symptoms taken from the Bonn Scale for Assessment of Prodromal Symptoms (Kojoh and Hirasawa, 1990; Klosterkötter et al., 2001), as well as the subthreshold psychotic symptoms closely corresponding to the PACE criteria (Yung et al., 2003, 2004), including Attenuated Psychotic Symptoms (APS) as well as Brief Limited Intermittent Psychotic Symptoms (BLIPS) were used to define an ARMS for psychosis (Table 1).

Study inclusion required *either* (1) a positive global functioning & trait marker, *or* (2) at least one positive psychopathological state marker fulfilling specific duration criteria (Table 1). Exclusion criteria (Table 1) were assessed by obtaining the personal and familial history using a semi-structured clinical interview as well as the Structured Clinical Interview for DSM-IV (American Psychiatric Association, 1994). Subjects were excluded from the study if they met the following criteria, including (1) disease transition as defined by Yung et al., (1998), (2) a past or present diagnosis of schizophrenia spectrum and bipolar disorders, as well as delirium, dementia, amnesic or other cognitive disorders, mental retardation and psychiatric disorders due to a somatic factor, following the DSM-IV criteria, (3) alcohol or drug abuse within three months prior to examination, (4) past or present inflammatory, traumatic or epileptic diseases of the central nervous system and (5) any previous treatment with antipsychotics. All participants provided their written informed consent prior to study inclusion. The study was approved by the Local Research Ethics Committee of the Ludwig-Maximilians-University.

Included ARMS individuals were seen weekly in the first month, monthly in the first year, quarterly in the second year and thereafter annually to detect a possible transition to psychosis as defined by Yung et al. (1998). All followed ARMS

**Table 1**

Inclusion/exclusion criteria for the ARMS individuals.

<p><i>Early At-Risk Mental State (Arms-E)</i> ARMS subjects without APS and/or BLIPS...</p> <p>(1) ... having one or more of the following basic symptoms appeared first at least 12 months prior to study inclusion and several times per week during the last 3 months.</p> <ul style="list-style-type: none"> <li>• Thought interferences</li> <li>• Thought perseveration</li> <li>• Thought pressure</li> <li>• Thought blockages</li> <li>• Disturbances of receptive language, either heard or read</li> <li>• Decreased ability to discriminate between ideas and perception, fantasy and true memories</li> <li>• Unstable ideas of reference (subject-centrism)</li> <li>• Derealisation</li> <li>• Visual perception disturbances</li> <li>• Acoustic perception disturbances</li> </ul> <p><i>and/or</i></p> <p>(2) ... showing a reduction in the Global Assessment of Functioning Score (DSM IV) of at least 30 points (within the past year) combined with at least one of the following trait markers:</p> <ul style="list-style-type: none"> <li>• First-degree relative with a lifetime-diagnosis of schizophrenia or a schizophrenia spectrum disorder</li> <li>• Pre- or perinatal complications</li> </ul>
<p><i>Late At-Risk Mental State (ARMS-L)</i> ARMS subjects with/without basic symptoms, with/without global functioning &amp; trait markers ...</p> <p>(1) ... having at least one of the following Attenuated Positive Symptoms (APS) within the last three months, appearing several times per week for a period of at least one week:</p> <ul style="list-style-type: none"> <li>• Ideas of reference,</li> <li>• Odd beliefs or magical thinking,</li> <li>• Unusual perceptual experiences,</li> <li>• Odd thinking and speech,</li> <li>• Suspiciousness or paranoid ideation</li> </ul> <p><i>and/or</i></p> <p>(2) ... having at least one of the following Brief Limited Intermittent Psychotic Symptoms (BLIPS), defined as the appearance of one of the following psychotic symptoms for less than one week (interval between episodes at least one week), resolving spontaneously:</p> <ul style="list-style-type: none"> <li>• Hallucinations,</li> <li>• Delusions,</li> <li>• Formal thought disorder,</li> <li>• Gross disorganised or catatonic behaviour</li> </ul>
<p><i>Exclusion criteria</i></p> <ul style="list-style-type: none"> <li>• Transition to psychosis as defined by Yung et al., (1998)</li> <li>• Meeting the DSM-IV criteria for a past or present diagnosis of schizophrenia spectrum and bipolar disorders, as well as delirium, dementia, amnesic or other cognitive disorders, mental retardation and psychiatric disorders due to a somatic factor or related to psychotropic substances,</li> <li>• Meeting the DSM-IV criteria for alcohol or drug abuse within three months prior to examination,</li> <li>• Any past or present inflammatory, traumatic or epileptic diseases of the central nervous system</li> </ul>

individuals were offered supportive counselling and clinical management. At follow-up, a complete re-examination was performed and the ARMS individuals were subgrouped either into an ARMS non-transition (ARMS-NT) or a transition (ARMS-T) group, if they met the transition criteria during the follow-up period *and* had a confirmed diagnosis of schizophrenia spectrum disorder according to the ICD-10 research criteria one year after transition. Of the 25 followed ARMS

individuals, 9 developed schizophrenia and 3 schizoaffective psychosis within a range of 40–777 (median: 84) days. The diagnostic outcomes in the ARMS-NT group resulted in no psychiatric diagnosis ( $n = 11$ ) and major depression ( $n = 2$ ). No ARMS-NT individual met the transition criteria during the follow-up period. No neuroleptic agents were prescribed prior to the initial MRI scan. Over the follow-up period, nine ARMS-T but no ARMS-NT individuals received antipsychotics (Table 3). Furthermore, neither lithium nor depot neuroleptics were prescribed over the follow-up period. The mean (SD) duration of illness (disease transition to follow-up scan) was 2.9 (1.7) years in the ARMS-T group.

## 2.2. MRI data acquisition

Baseline and follow-up MR images were obtained on the same 1.5T Magnetom Vision scanner (Siemens, Erlangen, Germany) using a T1-weighted 3D magnetization prepared gradient-echo (MP-RAGE) sequence (TR, 11.6 ms; TE, 4.9 ms; field of view, 230 mm; matrix,  $512 \times 512$ ; 126 contiguous axial slices of 1.5 mm thickness; voxel size,  $0.45 \times 0.45 \times 1.5$  mm). No hardware or software update was performed between baseline and follow-up scanning. All images were checked for image artefacts, gross anatomical abnormalities and signs of clinical pathology by trained neuroradiologists.

## 2.3. Data preprocessing

### 2.3.1. Measuring volumetric brain changes over time

Volumetric brain changes between each individual's pair of serial scans were measured using a deformation-based morphometry pipeline (DBM) implemented in SPM5 (Wellcome Department of Cognitive Neurology, London, UK), running under MATLAB 2008a (The MathWorks, Natick, MA, USA) (suppl. method 1). In summary, the pipeline first realigned the serial scans, removed non-brain tissue and minimized MRI bias field inhomogeneities in each scan. Then, the high-dimensional warping toolbox (HDW) (Chételat et al., 2005; Kipps et al., 2005; Whitford et al., 2006) was used (1) to remove the remaining differential MRI intensity inhomogeneities between the scans, (2) to compute the warping of the follow-up scan to the respective baseline scan and (3) to calculate the Jacobian determinant (abbr. Jacobian) of these between-timepoint deformations as a measure of volumetric change. The obtained deformations were applied to the follow-up volume to minimize its difference to the baseline volume. Then, the soft mean image (Chételat et al., 2005) of baseline and warped follow-up volumes was segmented into GM, WM and CSF partitions using the VBM5 toolbox (<http://dbm.uni-jena.de/software>). The rigid-body realigned GM and WM tissue segments were non-linearly registered to the stereotactic space of the Montreal Neurological Institute (MNI) using the Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (Ashburner, 2007, 2009; Bergouignan et al., 2009; Klein et al., 2009) (DARTEL) toolbox. The obtained DARTEL deformations were used to warp the native-space, between-timepoint Jacobians to MNI space. Finally, the warped Jacobians were split into voxel compression (VC) and expansion (VE) maps (Scahill et al., 2002; Whitwell et al., 2004) and smoothed with a 3 mm FWHM Gaussian kernel.

**Table 2**  
Statistical analysis of sociodemographic and clinical parameters.

Variable	HC	ARMS	T/ $\chi^2$	P	ARMS-NT	ARMS-T	F/T/ $\chi^2$	P
<i>Sociodemographic variables</i>								
N	28	25			13	12		
Age at scan (SD) [years]	25.1 (3.6)	23.2 (4.7)	1.74	n.s.#	24.1 (5.8)	22.2 (3.1)	2.19	n.s.#
Gender (M/F) [%]	60.7 / 39.3	72 / 28	0.75	n.s.†	61.5 / 38.5	83.3 / 16.7	2.04	n.s.†
Handedness (R/L/A) [%]	96.4 / 3.6 / 0	92 / 4.0 / 4.0	1.33	n.s.†	84.6 / 7.7 / 7.7	100 / 0 / 0	3.99	n.s.†
Educational years (SD)	12.5 (1.1)	11.9 (1.3)	1.90	n.s.#	11.9 (1.3)	11.9 (1.2)	1.82	n.s.#
Between-scan interval (SD) [years]	3.7 (1.4)	3.7 (1.1)	0.01	n.s.#	3.9 (0.9)	3.6 (1.3)	0.21	n.s.#
Premorbid IQ (SD)	–	108.5 (14.1)	–	–	113.9 (14.3)	102.4 (14.4)	1.74	n.s.#
Subjects with 1 <sup>st</sup> relatives having schizophrenic psychoses [%]	–	24.0	–	–	15.4	33.3	1.10	n.s.†
Subjects with 1 <sup>st</sup> relatives having affective psychoses [%]	–	16.0	–	–	15.4	16.7	0.01	n.s.†
<i>Global functioning and psychopathology at baseline</i>								
GAF score (SD)	–	57.9 (9.9)	–	–	59.8 (7.8)	52.8 (14.4)	1.24	n.s.#
PANSS total score (SD)	–	50.9 (16.7)	–	–	45.3 (8.9)	61.3 (23.06)	–2.09	n.s.#
PANSS positive score (SD)	–	10.7 (3.2)	–	–	9.6 (2.3)	12.8 (3.7)	–2.31	.035#
PANSS negative score (SD)	–	13.1 (7.7)	–	–	10.3 (4.8)	18.2 (9.7)	–2.29	.038#
PANSS general score (SD)	–	27.2 (7.9)	–	–	25.5 (4.9)	30.3 (11.7)	–0.98	n.s.#

Abbreviations: M male, F female, R right, L left, A ambidextruous, 1<sup>st</sup> degree relatives.

Statistical tests: †Fisher's exact test, # Student t test, # One-way analysis of variance with 3 groups (HC, ARMS-NT, ARMS-T).

Nonsignificant between-group differences regarding age, gender and education were accounted for by adjusting the VC and VE maps for these effects using partial correlations.

**Table 3**  
Clinical and medication data of ARMS-T

Disease transitions [No. (% of all ARMS individuals)]	12 (48)
Time to disease transition (SD) [years]	0.60 (0.71)
Illness duration (SD) [years]	2.91 (1.71)
Diagnoses: subjects [No. (%)]	
Schizophrenia	9 (75)
Schizoaffective psychosis	3 (25)
Medication over the between-scan interval	
No. (%) of treated ARMS-T individuals	9 (75)
ARMS-T individuals [No. (%) ...	
Without neuroleptic agents	3 (25)
With 1 neuroleptic agent	2 (17)
With 2 neuroleptic agents	2 (17)
With 3 neuroleptic agents	2 (17)
With >3 neuroleptic agents	3 (25)
ARMS-T individuals [No. (%) with prescriptions of ...	
Amisulpride	6 (67)
Risperidone	4 (44)
Aripiprazole	4 (44)
Quetiapine	3 (33)
Olanzapine	3 (33)
Ziprasidone	2 (22)
Quetiapine	1 (11)
Medication at the follow-up MRI scan	
No. (%) of treated ARMS-T individuals	8 (67)
ARMS-T individuals [No. (%) ...	
Without neuroleptic agents	4 (33)
With 1 neuroleptic agents	5 (42)
With 2 neuroleptic agents	1 (8)
With 3 neuroleptic agents	2 (17)
With >3 neuroleptic agents	0 (0)
ARMS-T individuals [No. (%) with prescriptions of ...	
Quetiapine	3 (38)
Risperidone	2 (25)
Amisulpride	2 (25)
Aripiprazole	2 (25)
Ziprasidone	1 (13)
Olanzapine	1 (13)
Clozapine	1 (13)

2.3.2. Cross-sectional preprocessing of the baseline MRI data

For the SVR analysis, the baseline MRI data entered a separate processing pipeline in order to exclude any influence of the follow-up on the baseline data. First, the MRI scans were segmented using the VBM5 toolbox. Then, the GM and WM tissue segments were normalized to MNI space using DARTEL. Finally, the obtained deformations fields were written out as cross-sectional Jacobians indicating the volumetric compressions and expansions necessary to register each subject's tissue maps to the DARTEL baseline template (suppl. method 2). We did not smooth the cross-sectional Jacobians due to their intrinsic smoothness (6.3, 7.1, 6.4 mm in the x, y and z axes).

2.4. Multivariate analysis of longitudinal brain changes

We used PLS (Fujiwara et al., 2008; Garthwaite, 1994; Giessing et al., 2007; McIntosh et al., 1996; McIntosh et al., 1999; Menzies et al., 2007; Wold, 1966) because it implements a multivariate, data-driven approach to study morphometric changes across the entire brain with higher sensitivity than the mass-univariate tests operating on a voxel-by-voxel basis. Previous studies of ARMS individuals and patients with established psychosis (Borgwardt et al., 2007a,b, 2008; Honea et al., 2005; Koutsouleris et al., 2009a; Koutsouleris et al., 2009b; Meisenzahl et al., 2008a,b; Pantelis et al., 2003) reported complex patterns of distributed abnormalities suggesting that the underlying pathophysiological process involves networks of inter-related brain regions rather than single circumscribed structures. This disease-related connectivity causes the morphometric effects to be inherently multicollinear, thus violating the assumption of statistical independence among measures underlying univariate statistics (Allen, 1997). In contrast, PLS is specifically designed to deal with multicollinearity both in the brain as well as in the exogenous data, because it breaks down high-dimensional, inter-correlated variables into a PLS model consisting of a small set of uncorrelated latent variables (LV), which maximize the covariance within and between the experimental effects and their predictors (McIntosh et al.,

1996). These latent variables are orthogonal and sorted according to their singular values  $d_{LV}$ , which provide a measure for the strength of covariance between the data blocks. The covariances of each LV pair are described by two singular vectors, termed *X* and *Y saliences*. In the case of longitudinal volumetric imaging data, the *Y saliences* constitute a *singular image* of the longitudinal volumetric effects covarying with the experimental design, while the *X saliences* describe linear combinations of design contrast weights that covary with the singular image and thus represent common or differential brain volume changes across groups. Furthermore, the expression of the singular image in each study participant is characterized by a global *brainscore*, the summed product of the singular image and the respective map of volumetric change.

The significance of the PLS model is *not* evaluated at the voxel-level, as e.g. in the mass-univariate statistics, but instead is assessed at the whole-brain level using a non-parametric permutation test (McIntosh and Lobaugh, 2004) that randomly reassigns the observations to the experimental predictors and recomputes the  $d_{LV}$  of the permuted models. Therefore, no correction for multiple comparisons is needed (McIntosh and Lobaugh, 2004). We performed 5000 permutations to estimate the permutation distribution of  $d_{LV}$  and rejected the null hypothesis that the observed values of  $d_{LV}$  were obtained by chance at  $\alpha=0.05$ . Furthermore, the stability of the covariance pattern elements was determined by estimating the standard errors of the saliences on the LV pairs using 1000 bootstrap samples created by resampling with replacement (Efron and Tibshirani, 1986; McIntosh and Lobaugh, 2004). The minimum amount of different subjects selected by the resampling procedure was set to 50% to exclude low variance samples. Singular image voxels with a ratio of salience to standard error  $>2$ , corresponding to 95% confidence limits, were considered stable as these elements showed little variation of the experimental effects against the varying composition of the bootstrap samples (Sampson et al., 1989). It is of note that bootstrap resampling was employed to determine the *stability*, not the voxel-level significance of the experimental effects. Stable pattern elements were mapped to anatomical regions using Automated Anatomical Labeling (Tzourio-Mazoyer et al., 2002) and their percentage volume change was quantified using the unadjusted VC/VE maps (suppl. Fig. 3).

Our first analysis investigated common and differential patterns of volumetric change in 28 HC vs 25 ARMS subjects. A second analysis in 28 HC vs 13 ARMS-NT vs 12 ARMS-T subjects was conducted to explore whether the detected patterns were driven by disease transition or non-transition. Therefore, two covariance matrices were constructed from a block of dummy regressors indicating the conditions “compression” and “expansion” as well as the group membership, and from a block containing the stacked VC and VE maps of all participants. Stability of between-group effects was calculated using the 95% bootstrapping confidence intervals of the *X saliences* on the compression and expansion conditions. Analyses were performed using PLSgui (<http://www.rotman-baycrest.on.ca/pls/>). Our study participants' individual brainscores on the ARMS or disease-associated patterns of longitudinal volumetric change were defined as the target variable for the subsequent prediction analysis.

## 2.5. Individualized prediction of structural brain changes

Support-vector regression (SVR) was used to predict the expression of ARMS- or transition-associated brain changes at the *individual level* from the respective baseline MRI scans. SVR learning and validation were wrapped into a *repeated double cross-validation* (rdCV) framework (Filzmoser et al., 2009) that allowed (1) constructing ensembles of predictors and optimize their parameters in order to improve prediction stability, and (2) obtaining an unbiased estimate of the generalization capacity using validation data that were strictly separated from the training data at all steps of the analysis.

### 2.5.1. Repeated double cross-validation and ensemble learning

A predictor's generalization ability is typically assessed using *k*-fold cross-validation (CV), where the data is split into *k* folds and each fold iteratively serves once to estimate the predictor's performance obtained from all but this partition. Since in many real-world applications there exists no prior knowledge about the optimal model parameters, CV is used to select those parameters that achieve the best performance on the validation data. However, this strategy violates the principle of separating the training from the testing data and thus, results in model overfitting and an overly optimistic estimation of the predictor's generalization capacity (Varma and Simon, 2006). Therefore, we employed *nested cross-validation* (Stone, 1974) by splitting the data into training and validation sets on an outer (CV2) and an inner CV (CV1) loop based on a stratified 10-fold CV scheme (suppl. method 3). All steps of SVR training, including feature selection and parameter optimization were performed on the CV1 training and validation partitions, while the generalization error was exclusively estimated from the CV2 test samples.

Furthermore, we extended nested CV to *repeated double CV* (Filzmoser et al., 2009) by randomly permuting the study participants 10 times within their groups and repeating the CV cycle for each of these permutations both at the inner and outer CV loops. The rationale for repeated CV on the inner (CV1) loop was to generate robust ensemble SVR models by introducing sample variance into the training process. Through repeated CV on the outer loop (CV2), we obtained a reliable estimate of the predictors' generalization performance. This performance was defined in analogy to Breiman's out-of-bag technique (Breiman and Spector, 1990) as the joint out-of-training prediction (OOT) of the ensemble of predictors, for which a given subject had not participated in the training process (suppl. method 3).

### 2.5.2. Feature extraction and dimensionality reduction

MRI-based pattern regression algorithms are faced with a very large amount of irrelevant or redundant voxels for the prediction problem at hand. The dilemma of picking the most relevant voxels is further exacerbated by the problem of *multivariate prediction* where on the one hand individual highly predictive voxels can be conjointly redundant and on the other hand weakly relevant voxels can yield high predictive power when combined into feature sets (Liu et al., 2005). This “curse of dimensionality” (Bellman, 1961) decreases the signal-to-noise ratio and hence degrades the performance of predictive algorithms to detect relevant

structure in the limited training samples available to neuroimaging studies of the ARMS. Therefore, we implemented a multivariate feature selection strategy to capture the most informative voxel patterns in the training data and to construct optimal features sets for SVR from these patterns. We extended our previous dimensionality reduction method (Koutsouleris et al., 2009a) by including (1) a non-linear multivariate, nearest-neighbor-based filter (RGS algorithm, (Navot, 2006), suppl. method 4) to identify the most relevant voxel sets, (2) principal component analysis (PCA) to further reduce the dimensionality of these voxel sets, and (3) the construction of compact PCA feature sets to Maximize feature Relevance and Minimize Redundancy (MRMR) (Peng et al., 2007). Finally, these compact PCA feature sets, constructed from each CV1 training sample, entered the SVR analysis.

### 2.5.3. $\nu$ -Support-vector regression

SVMs are multivariate machine-learning algorithms (Vapnik, 1999; Burges, 1998) increasingly applied to the MRI-based classification of neuropsychiatric patient populations due to their good generalisation properties (Davatzikos et al., 2005, 2008; Fan et al., 2008a,b; Klöppel et al., 2008; Koutsouleris et al., 2009a; Mourao-Miranda et al., 2005). Beside their application in the context of classification, SVMs have been generalized to the regression estimation of real-valued output functions (Schölkopf and Smola, 2002). The SVR algorithm finds a linear function  $f(x)$  within a linear or non-linear kernel space that provides the optimum fit between the training points and their continuous labels. In analogy to the soft-margin support-vector classification, which allows misclassifications in case of overlapping class distributions, the estimated regression line is surrounded by an  $\varepsilon$ -insensitive tube with the  $\varepsilon$  parameter controlling the precision of regression estimation and thus serving as regularization constraint to avoid model overfitting. However, as the optimal  $\varepsilon$  is typically unknown beforehand, we used the  $\nu$ -SVR algorithm (Schölkopf et al., 2000), which automatically determines the optimal  $\varepsilon$  through the sparsity parameter  $\nu$ . The prediction performance of the trained SVR models was measured both at the CV1 and CV2 levels using the Normalized Root of Mean Squared Deviation (NRMSD, suppl. method 5).

## 3. Results

### 3.1. Sociodemographic and clinical data

Age at baseline, gender, handedness, educational years and between-scan interval did not significantly differ between HC and ARMS, as well as HC, ARMS-NT and ARMS-T

(Table 2). Furthermore, ARMS-T did not differ significantly from the ARMS-NT regarding premorbid verbal IQ (as measured by the Mehrfach-Wortschatz-Test (MWT) (Lehrl, 2005), an established measure in German-speaking populations) and genetic risk as defined by the prevalence of affective and schizophrenic psychoses in the first degree relatives. At study inclusion, there was an overall trend to a more severe psychopathology in the converters vs the non-converters, which was significant for the PANSS positive and negative symptom scale (Table 2).

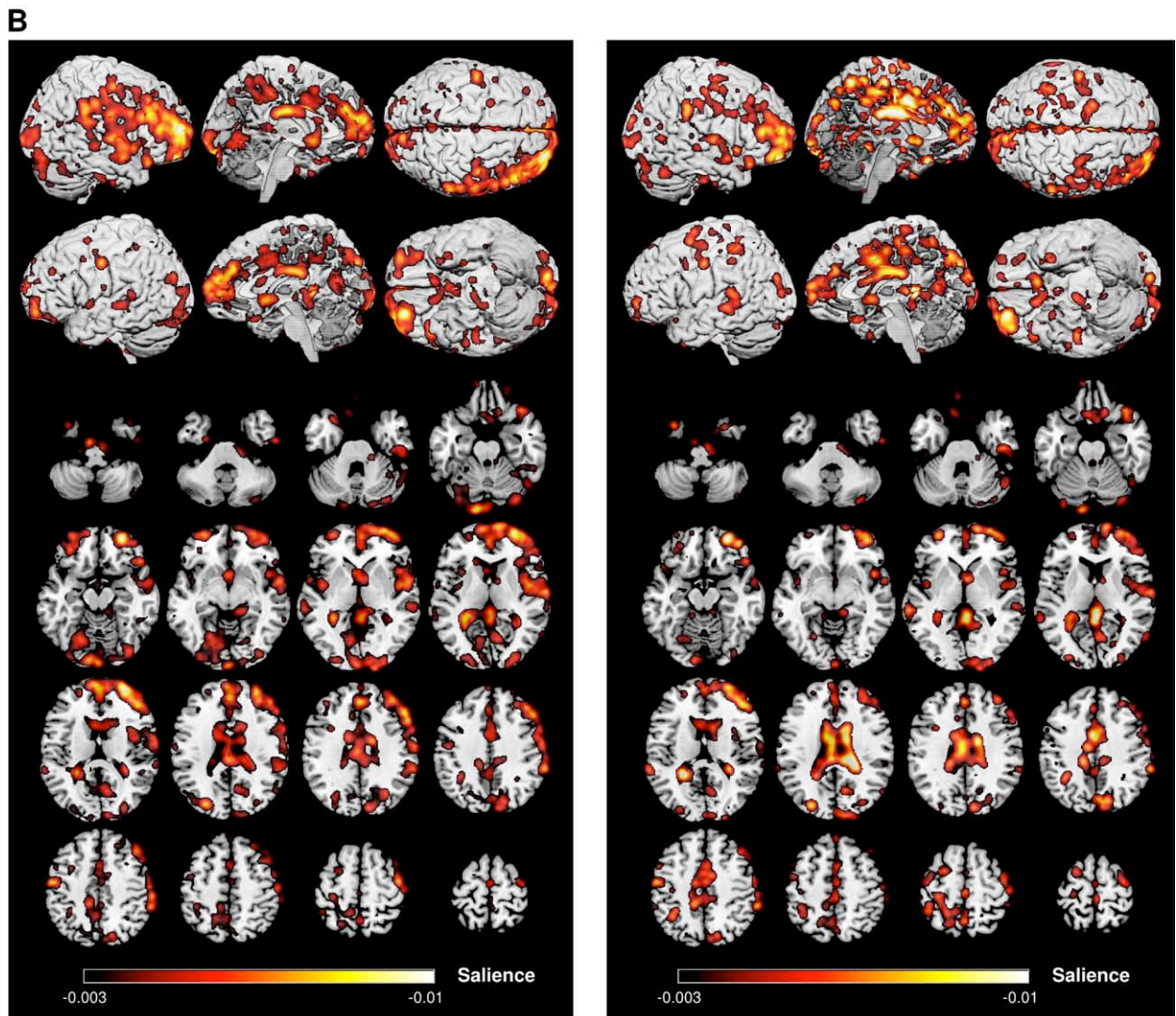
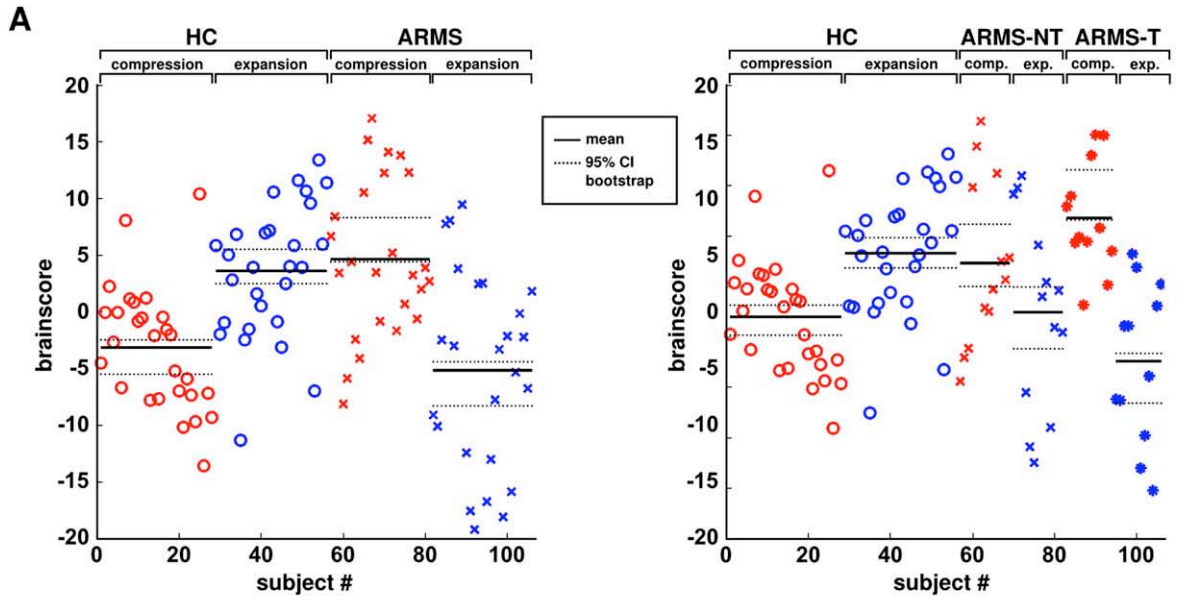
### 3.2. PLS between-group analyses of longitudinal volumetric changes

#### 3.2.1. HC versus ARMS

The first (LV1) of the 4 LV pairs generated by PLS was significant ( $P < 0.011$ ) and explained 98.9% of the covariance between the design matrix and the longitudinal volumetric data. When interpreting LV1, it is important to note that *more negative* values in the VC maps measure increased compression, whereas *more positive* values in the VE maps quantify increased expansion. Thus, *opposite* mean effects between the compression (red) and expansion (blue) conditions in Fig. 1A, left, indicate that the underlying covariance pattern of LV1 maps to brain regions affected by *both* volumetric compression and expansion. In this context, the sign of the X and Y saliences is negligible as it is an arbitrary result of the SVD rotation when the mean centered covariance matrix is decomposed into the LV pairs. Therefore, the *opposite* signs of the brain scores on each volumetric condition indicate a *between-group difference* regarding the expression of longitudinal volumetric change, with the ARMS group showing more pronounced effects than the HC group (suppl. Fig. 3). This difference was stable by the non-overlap of the 95% bootstrapping confidence intervals for each volumetric condition.

The singular image of LV1 (Fig. 1B, left) mapped primarily to anatomical regions where the tissue compression was paralleled by an expansion of the adjacent subarachnoidal or ventricular space. The covariance pattern was distributed across (1) right-hemispheric structures, spanning the frontopolar and dorsolateral areas and covering the entire perisylvian region including the ventrolateral prefrontal and insular cortex, the precentral, supramarginal and superior temporal gyri, (2) the medial prefrontal and lateral orbitofrontal cortex, bilaterally, extending to the cingulate and medial parietal structures, to the caudate nucleus and the posterior part of the corpus callosum, (3) the medial and lateral occipital cortex, and (4) the cerebellum and vermis (Fig. 1B, left and suppl. Fig. 3). Volumetric quantification (suppl. Fig. 3) revealed that the structures most involved were the right

**Fig. 1. A:** Multivariate PLS analysis of volumetric compression and expansion in 28 HC versus 25 ARMS (left side) and 28 HC versus 13 ARMS-NT versus 12 ARMS-T (right side) after removing the effects of age, gender, educational level and between-scan interval from the data using partial correlations. A: Scatter plots show the study participants' brain scores on the compression (red) and expansion (blue) conditions in the first latent variable (LV1) obtained in both PLS analyses. For each condition and group, the mean is represented by a solid line and the 95% bootstrapping confidence interval (CI) by a dotted line. Non-overlapping CIs in the respective condition (red/blue) reveal significant between-group differences regarding volumetric changes over time at  $\alpha = 0.05$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) **B:** In each PLS analysis, a bootstrap ratio (saliency/standard error) threshold of  $\pm 2$ , corresponding to confidence limits of 95%, was defined in order to extract stable voxel elements from the singular image of LV1. These suprathreshold voxels represented reliable between-group differences of volumetric changes over time. For visualization purposes the suprathreshold voxel volumes were smoothed with an 8 mm FWHM Gaussian kernel to reduce the pattern complexity and were overlaid on the single subject template of the Montreal Neurological Institute.



lateral prefrontal, perisylvian, cingulate, occipital and cerebellar cortices, bilaterally.

### 3.2.2. HC versus ARMS-NT versus ARMS-T

The first of six LV pairs was significant ( $P < 0.010$ ), explaining 67.8% covariance of the volumetric data with the experimental predictors. The interpretation and neuroanatomical mapping of LV1 is similar to the first analysis (Fig. 1A, right). The ARMS-T group had the highest brainscores on a pattern of *both* longitudinal compression and expansion, while ARMS-NT loaded on this pattern at an *intermediate* level between HC and ARMS-T. The mean differences between the HC, ARMS-NT and ARMS-T samples were reliable as shown by the non-overlap of the 95% bootstrapping confidence intervals (Fig. 1A, right).

The distribution of the reliable pattern elements was similar to the singular image of the first analysis. The right prefrontal and perisylvian regions were less affected, whereas the middle and posterior part of the cingulate cortex, corpus callosum with the adjacent ventricular space and medial parietal cortex were more involved than in the first PLS analysis (Fig. 1B, right and suppl. Fig. 3).

## 3.3. SVR prediction of longitudinal volumetric changes

### 3.3.1. HC versus ARMS

Using the baseline MRI data, the first SVR analysis (Fig. 2A, left) achieved a significant out-of-training prediction of our study participants' brainscores on the singular image of LV1 ( $r = 0.83, T = 10.8; P < 0.001$ ). As described above, the singular image of LV1 described a differential pattern of pronounced volumetric change in 25 ARMS individuals versus 28 HC. The out-of-training error of the joint SVR ensemble prediction was  $\text{NRMSD}_{\text{OOT}} = 25.5\%$ .

The core neuroanatomical pattern underlying the prediction of ARMS-associated morphometric changes was described in terms of the pattern elements' selection frequency across the CV1 partitions (Fig. 2B, left) and their SV-associated loadings on the principal components obtained during feature extraction (suppl. method 6, suppl. Fig. 4). The predictive baseline pattern involved bilateral ventricular enlargements and distributed GM volume reductions, including (1) the medial temporal lobes (hippocampus, amygdala, and parahippocampus), (2) the orbitofrontal, medial prefrontal and anterior cingulate cortices with extensions to the dorsomedial prefrontal cortices (DMPFC) and the supplementary motor areas (SMA), (3) the peri- and intrasylvian regions (superior temporal, supramarginal, inferior frontal gyrus, Rolandic opercula, and insulae), (4) primarily the right dorsolateral prefrontal (DLPFC), temporal and occipitotem-

poral cortices and (5) the subcortical structures (cerebellum, pons, putamen, and thalamus).

### 3.3.2. HC versus ARMS-NT versus ARMS-T

SVR achieved an out-of-training prediction of the individual brain scores on the singular image obtained in our second PLS analysis with an  $\text{NRMSD}_{\text{OOT}}$  of 25.6% ( $r = 0.83, T = 10.5; P < 0.001$ , Fig. 2A, right). The predictive neuroanatomical pattern (Fig. 2B, right) involved enlargements of the ventricular system as well as predominantly GM volume reductions in (1) the bilateral posterior temporal, occipitotemporal and occipital cortices, (2) the medial parietal and occipital regions, (3) the DLPFC, premotor cortex and SMA, (4) the inferior temporal gyri and medial temporal lobes, (5) subcortical structures (cerebellum, pons, putamen and thalamus).

## 4. Discussion

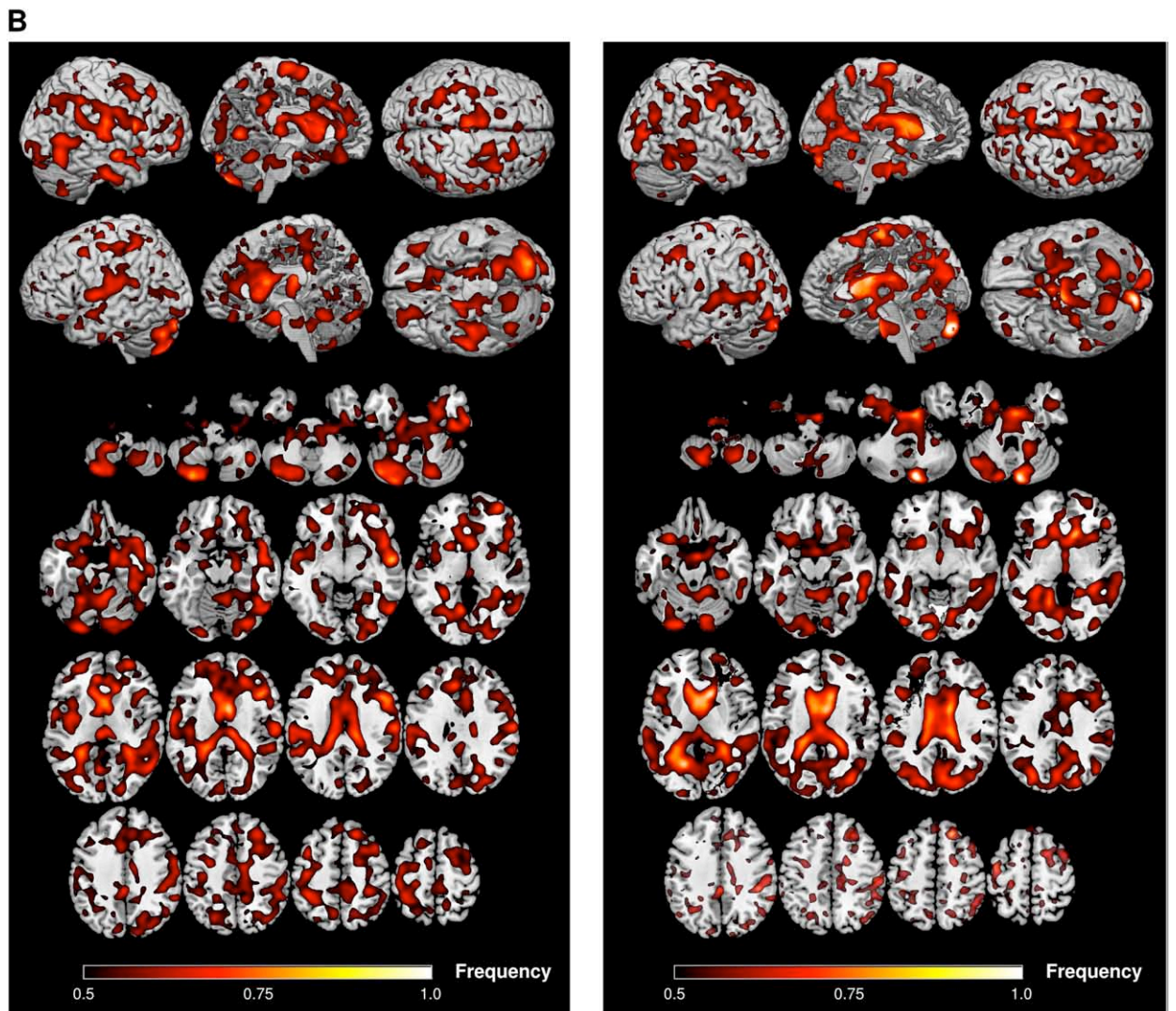
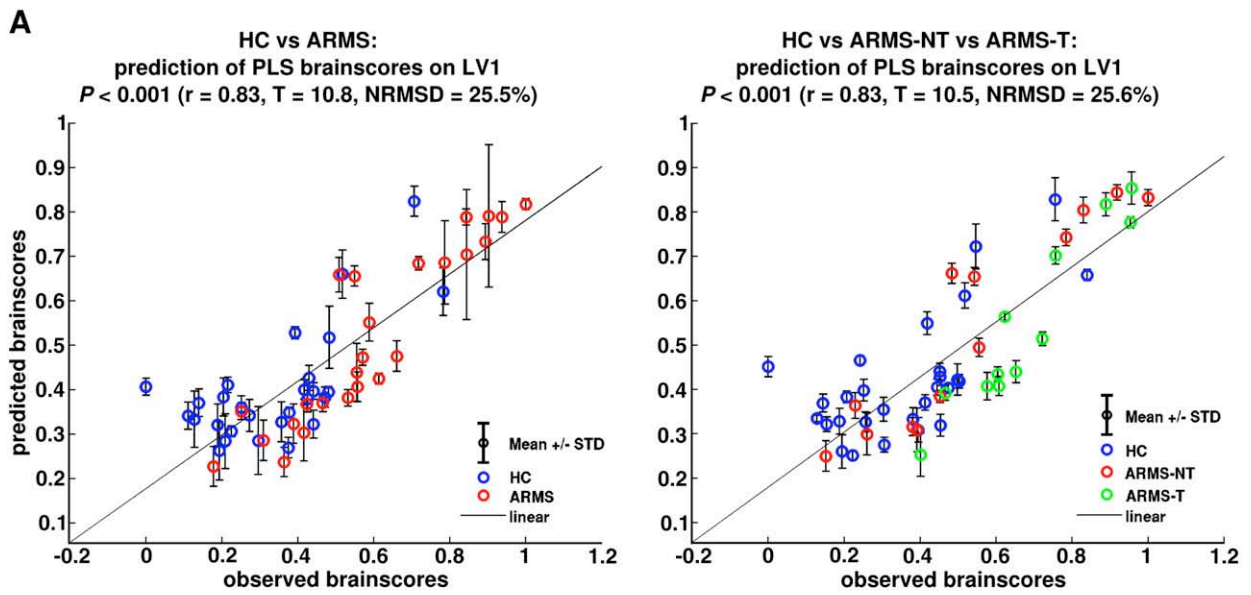
### 4.1. Structural brain dynamics in the ARMS for psychosis

A significant pattern of longitudinal morphometric changes covering predominantly the right prefrontal and perisylvian as well as the midline, ventricular and periventricular structures was observed in ARMS vs HC irrespective of the ARMS individuals' clinical outcome. This finding is consistent with previous clinical studies, reporting patterns of cross-sectional GM volume reductions within the prefrontal, temporal and medial parietal regions of clinical ARMS cohorts (Borgwardt et al., 2007b; Koutsouleris et al., 2009b; Meisenzahl et al., 2008b; Witthaus et al., 2009). Moreover, our results are in keeping with the cross-sectional study of Job et al., (2003) which reported (1) GM reductions in the cingulate cortex of genetically defined high-risk (HR) individuals vs HC and (2) more extended reductions within prefronto-temporo-limbic brain regions of first episode (FE) patients vs HR. Similarly, Witthaus et al., (2009) observed GM volume reductions within the cingulate cortex and the right perisylvian region of clinical ARMS vs HC subjects as well as further and more extended cingulate, perisylvian, orbitofrontal and limbic abnormalities in FE vs ARMS. These cross-sectional data point to an active neurobiological process that (1) affects higher-order cortical association areas in a state of ultra-high risk, and (2) causes further accumulating prefrontal, perisylvian and limbic alterations during the transition from the prodromal phase to the FE of psychosis (Pantelis et al., 2005).

This hypothesis is further supported by our second PLS analysis, which revealed more pronounced morphometric changes over time in ARMS-T vs ARMS-NT subjects within a neuroanatomical pattern that largely overlapped with the volumetric changes found in our first analysis. This finding

**Fig. 2.** A: SVR prediction analysis of morphometric change over time in HC (blue circles) vs ARMS (red circles) individuals (left panel) as well as in 28 HC (blue circles) vs 13 ARMS-NT (red circles) vs 12 ARMS-T (green circles) individuals (right panel). Using the baseline MRI data of each study participant, SVR predicted each individual's brain score on the ARMS-associated longitudinal volumetric changes obtained from the PLS analysis of 28 HC versus 25 ARMS (Fig. 1A and 1B, left panels) and 28 HC versus 13 ARMS-NT versus 12 ARMS-T (Fig. 1A and 1B, right panels). The out-of-training SVR ensemble prediction was plotted as the function of the respective brain scores. Error bars indicate the standard deviation of prediction for each individual subject. B: Voxel selection frequencies within the core neuroanatomical patterns underlying the SVR prediction of morphometric change were shown for the analysis of 28 HC vs 25 ARMS (left) and 28 HC vs 13 ARMS-NT vs 12 ARMS-T (right). The images indicate for every voxel the frequency of being selected as predictive pattern element across all CV1 partitions (see visualization strategy described in suppl. methods 7). For visualization purposes a frequency threshold was set to 0.5 and the resulting suprathreshold voxel volumes were smoothed with an 8 mm FWHM Gaussian kernel to reduce the pattern complexity. These smoothed maps were overlaid on the MNI single subject template.





agrees with previous studies reporting *progressive* volume reductions in the orbitofrontal, medial prefrontal, cingulate, temporal, precuneal and cerebellar regions of clinically defined ARMS-T vs ARMS-NT subjects (Borgwardt et al., 2008; Pantelis et al., 2003; Sun et al., 2009b) and FE patients vs HC (Farrow et al., 2005; Sun et al., 2009a; Whitford et al., 2006). Furthermore, our data is partly consistent with longitudinal MRI studies of high-risk individuals selected for genetic (Job et al., 2005) and clinical reasons (Takahashi et al., 2009) that compared converters, non-converters and healthy controls. These studies showed that abnormal structural brain changes within temporal (Job et al., 2005; Takahashi et al., 2009) and cerebellar regions (Job et al., 2005) can also be traced in a similar, but more subtle form in ARMS subjects with transient psychotic symptoms who ultimately do not become ill. Moreover, Job et al., (2005) also reported GM density losses in their HC group affecting the medial orbitofrontal cortex, a brain region that has also been reported to exhibit abnormal volumetric changes over time in ARMS-T individuals (Borgwardt et al., 2008; Pantelis et al., 2003). In this regard, recent MRI investigations showed that more widespread GM loss also occurs during the normal development of potentially disease-relevant brain regions in healthy populations (Shaw et al., 2008) and that the patterns of GM losses observed in ARMS-NT, ARMS-T and FE cohorts may represent a spatially similar, but exaggerated form of these maturational processes (Sun et al., 2009a,b). Taken together, all these studies, including our own, may point to a gradual deviation from the normal maturational trajectory of higher-order cortical networks but also subcortical structures, which characterizes the ARMS for psychosis and may be most expressed in those individuals who ultimately develop a full-blown psychotic disorder. Different candidate mechanisms may be involved in shaping this abnormal “late-neurodevelopmental” trajectory (Pantelis et al., 2005) including alterations in synaptic pruning, dendritic remodelling, axonal myelination and neuronal cell loss (Davis et al., 2003; Stephan et al., 2006).

#### 4.2. Individualized prediction of brain changes over time

To our knowledge, this is the first study to relate cross-sectional to longitudinal structural brain alterations in the ARMS for psychosis for the purpose of an individualized prediction of vulnerability and disease-related brain trajectories using machine-learning techniques. Based only on the MRI data acquired at study inclusion when all ARMS individuals were unmedicated, our SVR method estimated with significant precision the expression of ARMS- and disease-associated structural brain dynamics at the single-subject level in unseen test individuals.

The visualization of anatomical structures used by SVR to estimate these brain changes revealed morphometric patterns overlapping with previously reported alterations in genetically and clinically defined at-risk populations (Borgwardt et al., 2007a; Job et al., 2003; Koutsouleris et al., 2009b; Lawrie et al., 1999; Meisenzahl et al., 2008b; Pantelis et al., 2003; Sun et al., 2009b; Takahashi et al., 2009) and established psychosis (Meisenzahl et al., 2008a; Koutsouleris et al., 2008; Honea et al., 2005). Moreover, the spatial distribution of predictive brain regions was in line with the

discriminative neuroanatomical patterns detected in our previous work, which used support-vector classification for the purpose of an MRI-based early identification of the ARMS and prediction of disease transition (Koutsouleris et al., 2009a). This overlap suggests that patterns of alterations involving the frontal, temporal and parietal cortices, the basal ganglia and the cerebellum may not only serve as candidate diagnostic markers for a clinically defined ARMS and the transition or non-transition to psychosis, but also as potential predictors of the neurodevelopmental trajectories underlying these clinical phenotypes. However, as the observed PLS brainscores of the ARMS-T and ARMS-NT groups considerably overlapped at the individual level (Fig. 1A, right), our results currently do not allow for an individualized diagnosis of disease transition or non-transition based on the prediction of the underlying brain trajectories. In this regard, the future application of support-vector classification to the longitudinal data of larger samples may better deconstruct the brain development underlying the ARMS into trajectories associated with an elevated vulnerability to psychosis and trajectories linked to an ultimate disease transition. Nevertheless, our results may suggest that it is possible, based on the baseline MRI scan of a person at a clinical risk of psychosis, to predict whether this person will develop a “normal” or rather an ARMS-related trajectory of structural brain changes. We may speculate that this predictive biomarker could facilitate an informed allocation of therapeutic resources to those at particular risk of developing these risk-associated brain changes, and thus could aid in averting or at least delaying the full-blown expression of psychosis. Our findings suggest that this perspective should be further explored within the framework of large-scale prospective studies that allow evaluating the impact of biomarker-based clinical decision-making in high-risk populations.

As hypothesized, several brain regions shown by PLS to be involved in the differential *longitudinal* effects were found to be predictive of the subsequent morphometric changes. These structures consisted primarily of the ventricular system, the basal ganglia and areas distributed across the prefrontal, cingulate, medial parietal and perisylvian cortices. However, the core predictive baseline patterns extended to further brain regions, which were not, or only marginally, involved in the longitudinal PLS patterns, such as the medial temporal lobes, the inferior temporal and occipital gyri as well as the subcortical structures. This partial overlap between the predictive baseline patterns and the patterns of subsequent morphometric change points to complex, possibly non-linear sequences of structural brain changes underlying the ARMS and the transition to psychosis, with the magnitude, direction and spatial distribution of morphometric changes depending on the timepoint within the trajectory as well as on the level of genetically and environmentally mediated vulnerability to psychosis (Douaud et al., 2009). This hypothesis is supported by cross-sectional studies (Borgwardt et al., 2007b; Phillips et al., 2002) reporting non-reductions of medial temporal, subcortical and parietal brain regions in ARMS-T vs ARMS-NT. In this context, further longitudinal imaging investigations of more than two time points are needed to elucidate the complex spatiotemporal brain dynamics underlying the ARMS, the prodromal and early phases of psychosis.

### 4.3. Limitations

Finally, we have to consider several limitations of our current study. First, a replication of our findings is needed and should involve larger sample sizes. Given the known difficulties in recruiting and following ARMS individuals over a sufficiently long period of time, this could be best achieved within a multicenter design. Despite these uncertainties, the application of multivariate analysis techniques in conjunction with non-parametric resampling techniques provides a robust and unbiased means to assess the significance of the observed effects against random noise and to evaluate their stability against a varying sample composition. Furthermore, although the repeated double cross-validation design of the SVR analysis provides an unbiased generalization estimate of the machine learning algorithm's capacity to predict longitudinal morphometric change using only the baseline data, future studies have to examine whether these findings generalize to independent samples recruited across different MRI scanners.

Moreover, we have to consider that the results of our PLS analyses may have been driven by medication effects in the ARMS-T group. In this regard, the influence of antipsychotic medication on longitudinal structural brain changes of ARMS individuals and schizophrenic patients has been controversially discussed (see Borgwardt et al., 2009a, 2009b, for review). Regarding the conflicting literature, we cannot completely exclude an impact of antipsychotic medication on our results. It is, however, unlikely that the covariance patterns of our between-group analyses were solely driven by medication effects, because (1) the non-converters, who did not receive neuroleptic medication during the follow-up period, also showed a significant expression of these patterns, (2) the prediction of morphometric changes over time was based on the baseline MRI data, which was completely uninfluenced by antipsychotic medication effects, and (3) the PLS covariance patterns significantly overlapped with the previously reported cross-sectional baseline differences in unmedicated ARMS subjects (Koutsouleris et al., 2009a,b; Meisenzahl et al., 2008b).

Supplementary data to this article can be found online at doi:10.1016/j.schres.2010.08.032.

### Role of funding source

No funding was provided for the acquisition of MRI data. The German Research Network on Schizophrenia funded by the German Federal Ministry for Education and Research (BMBF) (grant 01 GI 9935) provided financial support for the recruitment and clinical evaluation of the prodromal subjects. Furthermore, the development of methodological procedures used in this study was supported by the BMBF research grants 01EV0709 and 01GW0740 (to Christian Gaser). The funding sources had no involvement in the study design, the collection and analysis of the data or the writing of the manuscript.

### Contributors

Author Nikolaos Koutsouleris participated in the collection of clinical and MRI data, performed the data processing and statistical analysis, participated in the evaluation and discussion of results and wrote the manuscript.

Author Christian Gaser supervised data processing and statistical analysis as well as the writing of the manuscript.

Author Ronald Bottlender was responsible for the recruitment of the ARMS subjects and the acquisition of clinical data, and participated in the evaluation of results and the writing of the manuscript.

Author Christos Davartikos participated in the evaluation of results and the writing of the manuscript.

Author Petra Decker participated in the recruitment of the ARMS subjects and the acquisition of clinical data.

Author Markus Jäger participated in the clinical assessment of the ARMS subjects and the writing of the manuscript.

Author Gisela Schmitt participated in the acquisition of data, the evaluation of results and the writing of the manuscript.

Author Maximilian Reiser participated in MRI data acquisition and the writing of the manuscript.

Author Hans-Jürgen Möller supervised the evaluation and discussion of scientific results and the writing of the manuscript.

Author Eva Meisenzahl designed the study, participated in the collection of clinical and MRI data, and supervised the writing of the manuscript

### Conflict of interest

No conflict of interest exists for any of the authors.

### Acknowledgments

We would like to thank Dr. Reinhold Bader, Linux Cluster Systems for the Munich and Bavarian Universities, for his support in integrating the machine learning algorithms into the batch system of the Linux cluster. Furthermore, we would like to thank Prof. Chih-Jen Lin from the National Taiwan University, Taiwan, for his help in adjusting the LIBSVM software to the needs of neuroimaging analysis.

### References

- Allen, M.P., 1997. Understanding regression analysis. The problem of multicollinearity. Springer, pp. 176–180. chap.
- American Psychiatric Association, 1994. Diagnostic and statistical manual for mental disorders, fourth ed. . Washington, DC.
- Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. *Neuroimage* 38 (1), 95–113.
- Ashburner, J., 2009. Computational anatomy with the spm software. *Magn. Reson. Imaging* 27 (8), 1163–1174.
- Bellman, R., 1961. Adaptive control processes: a guided tour. Princeton University Press.
- Bergouignan, L., Chupin, M., Czechowska, Y., Kinkingnéhun, S., Lemogne, C., Bastard, G.L., Lepage, M., Garnero, L., Colliot, O., Fossati, P., 2009. Can voxel based morphometry, manual segmentation and automated segmentation equally detect hippocampal volume differences in acute depression? *Neuroimage* 45 (1), 29–37.
- Borgwardt, S.J., Smieskova, R., Fusar-Poli, P., Bendfeldt, K., Riecher-Rössler, A., 2009a. The effects of antipsychotics on brain structure: what have we learnt from structural imaging of schizophrenia? *Psychol. Med.* 39 (11), 1781–1782.
- Borgwardt, S.J., McGuire, P.K., Aston, J., Berger, G., Dazzan, P., Gschwandtner, U., Pflüger, M., D'Souza, M., Radue, E.W., Riecher-Rössler, A., 2007a. Structural brain abnormalities in individuals with an at-risk mental state who later develop psychosis. *Br. J. Psychiatry Suppl.* 51, s69–s75.
- Borgwardt, S.J., McGuire, P.K., Aston, J., Gschwandtner, U., Pflüger, M.O., Stieglitz, R.D., Radue, E.W., Riecher-Rössler, A., 2008. Reductions in frontal, temporal and parietal volume associated with the onset of psychosis. *Schizophr. Res.* 106 (2–3), 108–114.
- Borgwardt, S.J., Riecher-Rössler, A., Dazzan, P., Chitnis, X., Aston, J., Drewe, M., Gschwandtner, U., Haller, S., Pflüger, M., Rechsteiner, E., D'Souza, M., Stieglitz, R.D., Rad, E.W., McGuire, P.K., 2007b. Regional gray matter volume abnormalities in the at risk mental state. *Biol. Psychiatry* 61 (10), 1148–1156.
- Borgwardt, S.J., Riecher-Rössler, A., Smieskova, R., McGuire, P.K., Fusar-Poli, P., 2009b. Superior temporal gray and white matter changes in schizophrenia or antipsychotic related effects? *Schizophr. Res.* 113 (1), 109–110.
- Bray, S., Chang, C., Hoefl, F., 2009. Applications of multivariate pattern classification analyses in developmental neuroimaging of healthy and clinical populations. *Front. Hum. Neurosci.* 3, 32.
- Breiman, L., Spector, P., 1990. Submodel selection and evaluation in regression—the x random case. Tech. rep., Department of Statistics. University of California.
- Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* 2, 121–167.
- Chételat, G., Landeau, B., Eustache, F., Mézenge, F., Viader, F., de la Sayette, V., Desgranges, B., Baron, J.C., 2005. Using voxel-based morphometry to map the structural changes associated with rapid conversion in mci: a longitudinal MRI study. *Neuroimage* 27 (4), 934–946.

- Davatzikos, C., 2004. Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. *Neuroimage* 23 (1), 17–20.
- Davatzikos, C., Shen, D., Gur, R.C., Wu, X., Liu, D., Fan, Y., Huggett, P., Turetsky, B.I., Gur, R.E., 2005. Whole-brain morphometric study of schizophrenia revealing a spatially complex set of focal abnormalities. *Arch. Gen. Psychiatry* 62 (11), 1218–1227.
- Davatzikos, C., Fan, Y., Wu, X., Shen, D., Resnick, S.M., 2008. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. *Neurobiol. Aging* 29 (4), 514–523.
- Davis, K.L., Stewart, D.G., Friedman, J.J., Buchsbaum, M., Harvey, P.D., Hof, P.R., Buxbaum, J., Haroutunian, V., 2003. White matter changes in schizophrenia: evidence for myelin-related dysfunction. *Arch. Gen. Psychiatry* 60 (5), 443–456.
- DeLisi, L.E., 2008. The concept of progressive brain change in schizophrenia: implications for understanding schizophrenia. *Schizophr. Bull.* 34 (2), 312–321.
- Douaud, G., Mackay, C., Andersson, J., James, S., Quested, D., Ray, M.K., Connell, J., Roberts, N., Crow, T.J., Matthews, P.M., Smith, S., James, A., 2009. Schizophrenia delays and alters maturation of the brain in adolescence. *Brain* 132 (Pt 9), 2437–2448.
- Efron, B., Tibshirani, R., 1986. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Stat. Sci.* 1, 54–77.
- Fan, Y., Batmanghelich, N., Clark, C.M., Davatzikos, C., the Alzheimer's Disease Neuroimaging Initiative, 2008a. Spatial patterns of brain atrophy in mci patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. *Neuroimage* 39 (4), 1731–1743.
- Fan, Y., Gur, R.E., Gur, R.C., Wu, X., Shen, D., Calkins, M.E., Davatzikos, C., 2008b. Unaffected family members and schizophrenia patients share brain structure patterns: a high-dimensional pattern classification study. *Biol. Psychiatry* 63 (1), 118–124.
- Farrow, T.F.D., Whitford, T.J., Williams, L.M., Gomes, L., Harris, A.W.F., 2005. Diagnosis-related regional gray matter loss over two years in first episode schizophrenia and bipolar disorder. *Biol. Psychiatry* 58 (9), 713–723.
- Filzmoser, P., Liebmann, B., Varmuza, K., 2009. Repeated double cross validation. *J. Chemom.* 23, 160–171.
- Frommann, I., Brinkmeyer, J., Ruhrmann, S., Hack, E., Brockhaus-Dumke, A., Bechdorf, A., Wölwer, W., Klosterkötter, J., Maier, W., Wagner, M., 2008. Auditory P300 in individuals clinically at risk for psychosis. *Int. J. Psychophysiol.* 70 (3), 192–205.
- Fujiwara, E., Schwartz, M.L., Gao, F., Black, S.E., Levine, B., 2008. Ventral frontal cortex functions and quantified MRI in traumatic brain injury. *Neuropsychologia* 46 (2), 461–474.
- Garthwaite, P., 1994. An interpretation of partial least squares. *Am. Stat. Assoc.* 89, 122–127.
- Gaser, C., Nenadic, I., Buchsbaum, B.R., Hazlett, E.A., Buchsbaum, M.S., 2004. Ventricular enlargement in schizophrenia related to volume reduction of the thalamus, striatum, and superior temporal cortex. *Am. J. Psychiatry* 161 (1), 154–156.
- Giessing, C., Fink, G.R., Rösler, F., Thiel, C.M., 2007. fMRI data predict individual differences of behavioral effects of nicotine: a partial least square analysis. *J. Cogn. Neurosci.* 19 (4), 658–670.
- Gilboa, A., Ramirez, J., Köhler, S., Westmacott, R., Black, S.E., Moscovitch, M., 2005. Retrieval of autobiographical memory in Alzheimer's disease: relation to volumes of medial temporal lobe and other structures. *Hippocampus* 15 (4), 535–550.
- Honea, R., Crow, T.J., Passingham, D., Mackay, C.E., 2005. Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *Am. J. Psychiatry* 162 (12), 2233–2245.
- Hurlmann, R., Jessen, F., Wagner, M., Frommann, I., Ruhrmann, S., Brockhaus, A., Picker, H., Scheef, L., Block, W., Schild, H.H., Möller-Hartmann, W., Krug, B., Falkai, P., Klosterkötter, J., Maier, W., 2008. Interrelated neuropsychological and anatomical evidence of hippocampal pathology in the at-risk mental state. *Psychol. Med.* 38 (6), 843–851.
- Job, D.E., Whalley, H.C., Johnstone, E.C., Lawrie, S.M., 2005. Grey matter changes over time in high risk subjects developing schizophrenia. *Neuroimage* 25 (4), 1023–1030.
- Job, D.E., Whalley, H.C., McConnell, S., Glabus, M., Johnstone, E.C., Lawrie, S.M., 2003. Voxel-based morphometry of grey matter densities in subjects at high risk of schizophrenia. *Schizophr. Res.* 64 (1), 1–13.
- Job, D.E., Whalley, H.C., McIntosh, A.M., Owens, D.G.C., Johnstone, E.C., Lawrie, S.M., 2006. Grey matter changes can improve the prediction of schizophrenia in subjects at high risk. *BMC Med.* 4, 29.
- Kawasaki, Y., Suzuki, M., Kherif, F., Takahashi, T., Zhou, S.Y., Nakamura, K., Matsui, M., Sumiyoshi, T., Seto, H., Kurachi, M., 2007. Multivariate voxel-based morphometry successfully differentiates schizophrenia patients from healthy controls. *Neuroimage* 34 (1), 235–242.
- Kipps, C.M., Duggins, A.J., Mahant, N., Gomes, L., Ashburner, J., McCusker, E.A., 2005. Progression of structural neuropathology in preclinical Huntington's disease: a tensor based morphometry study. *J. Neurol. Neurosurg. Psychiatry* 76 (5), 650–655.
- Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J., Parsey, R.V., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46 (3), 786–802.
- Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox, N.C., Jack, C.R., Ashburner, J., Frackowiak, R.S.J., 2008. Automatic classification of MR scans in Alzheimer's disease. *Brain* 131 (Pt 3), 681–689.
- Klosterkötter, J., Hellmich, M., Steinmeyer, E.M., Schultze-Lutter, F., 2001. Diagnosing schizophrenia in the initial prodromal phase. *Arch. Gen. Psychiatry* 58 (2), 158–164.
- Kojoh, K., Hirasawa, S., 1990. The Bonn scale for the assessment of basic symptoms (BSABS). *Arch. Psychiatr. Diagn. Clin. Eval* 4, 587–597.
- Koutsouleris, N., Meisenzahl, E., Davatzikos, C., Bottlender, R., Frodl, T., Scheuerecker, J., Schmitt, G., Zetzsche, T., Decker, P., Reiser, M., Möller, H.J., Gaser, C., 2009a. Neuroanatomical pattern classification identifies subjects in at-risk mental states of psychosis and predicts disease transition. *Arch. Gen. Psych.* 66 (7), 700–712.
- Koutsouleris, N., Schmitt, G., Gaser, C., Bottlender, R., Scheuerecker, J., McGuire, P., Burgermeister, B., Born, C., Reiser, M., Möller, H.J., Meisenzahl, E., 2009b. Neuroanatomical correlates of different vulnerability states of psychosis in relation to clinical outcome. *Br. J. Psychiatry* 195 (3).
- Koutsouleris, N., Gaser, C., Jäger, M., Bottlender, R., Frodl, T., Holzinger, S., Schmitt, G.J.E., Zetzsche, T., Burgermeister, B., Scheuerecker, J., Born, C., Reiser, M., Möller, H.J., Meisenzahl, E.M., 2008. Structural correlates of psychopathological symptom dimensions in schizophrenia: a voxel-based morphometric study. *Neuroimage* 39 (4), 1600–1612.
- Lawrie, S.M., Whalley, H., Kestelman, J.N., Abukmeil, S.S., Byrne, M., Hodges, A., Rimmington, J.E., Best, J.J., Owens, D.G., Johnstone, E.C., 1999. Magnetic resonance imaging of brain in people at high risk of developing schizophrenia. *Lancet* 353 (9146), 30–33.
- Lawrie, S.M., Whalley, H.C., Abukmeil, S.S., Kestelman, J.N., Miller, P., Best, J.J., Owens, D.G.C., Johnstone, E.C., 2002. Temporal lobe volume changes in people at high risk of schizophrenia with psychotic symptoms. *Br. J. Psychiatry* 181, 138–143.
- Lehrl, S., 2005. Mehrfachwahl-Wortschatz-Intelligenztest MWT-B, fifth ed. Spitta Verlag, Balingen.
- Linden, D.E.J., Fallgatter, A.J., 2009. Neuroimaging in psychiatry: from bench to bedside. *Front. Hum. Neurosci.* 3, 49.
- Liu, H., Dougherty, E.R., Dy, J.G., Torkkola, K., Tuv, E., Peng, H., Ding, C., Long, F., Berens, M., Parsons, L., Zhao, Z., Yu, L., Forman, G., 2005. Evolving feature selection. *IEEE Intell. Syst.* 20 (6), 64–76.
- McIntosh, A.R., Bookstein, F.L., Haxby, J.V., Grady, C.L., 1996. Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage* 3 (3 Pt 1), 143–157.
- McIntosh, A.R., Rajah, M.N., Lobaugh, N.J., 1999. Interactions of prefrontal cortex in relation to awareness in sensory learning. *Science* 284 (5419), 1531–1533.
- McIntosh, A.R., Lobaugh, N.J., 2004. Partial least squares analysis of neuroimaging data: applications and advances. *Neuroimage* 23 (Suppl. 1), S250–S263.
- Meisenzahl, E.M., Koutsouleris, N., Bottlender, R., Scheuerecker, J., Jäger, M., Teipel, S.J., Holzinger, S., Frodl, T., Preuss, U., Schmitt, G., Burgermeister, B., Reiser, M., Born, C., Möller, H.J., 2008a. Structural brain alterations at different stages of schizophrenia: a voxel-based morphometric study. *Schizophr. Res.* 104 (1–3), 44–60.
- Meisenzahl, E.M., Koutsouleris, N., Gaser, C., Bottlender, R., Schmitt, G.J.E., McGuire, P., Decker, P., Burgermeister, B., Born, C., Reiser, M., Möller, H.J., 2008b. Structural brain alterations in subjects at high-risk of psychosis: a voxel-based morphometric study. *Schizophr. Res.* 102 (1–3), 150–162.
- Menzies, L., Achard, S., Chamberlain, S.R., Fineberg, N., Chen, C.H., del Campo, N., Sahakian, B.J., Robbins, T.W., Bullmore, E., 2007. Neurocognitive endophenotypes of obsessive-compulsive disorder. *Brain* 130 (Pt 12), 3223–3236.
- Mourao-Miranda, J., Bokde, A.L.W., Born, C., Hampel, H., Stetter, M., 2005. Classifying brain states and determining the discriminating activation patterns: support vector machine on functional MRI data. *Neuroimage* 28 (4), 980–995.
- Navot, A., 2006. On the role of feature selection in machine-learning. Ph.D. thesis, Hebrew University.
- Nestor, P.G., O'Donnell, B.F., McCarley, R.W., Niznikiewicz, M., Barnard, J., Shen, Z.J., Bookstein, F.L., Shenton, M.E., 2002. A new statistical method for testing

- hypotheses of neuropsychological/MRI relationships in schizophrenia: partial least squares analysis. *Schizophr. Res.* 53 (1–2), 57–66.
- Pantelis, C., Velakoulis, D., McGorry, P.D., Wood, S.J., Suckling, J., Phillips, L.J., Yung, A.R., Bullmore, E.T., Brewer, W., Soulsby, B., Desmond, P., McGuire, P.K., 2003. Neuroanatomical abnormalities before and after onset of psychosis: a cross-sectional and longitudinal MRI comparison. *Lancet* 361 (9354), 281–288.
- Pantelis, C., Yücel, M., Wood, S.J., Velakoulis, D., Sun, D., Berger, G., Stuart, G.W., Yung, A., Phillips, L., McGorry, P.D., 2005. Structural brain imaging evidence for multiple pathological processes at different stages of brain development in schizophrenia. *Schizophr. Bull.* 31 (3), 672–696.
- Peng, Y., Li, W., Liu, Y., 2007. A hybrid approach for biomarker discovery from microarray gene expression data for cancer classification. *Cancer Inform.* 2, 301–311.
- Phillips, L.J., Velakoulis, D., Pantelis, C., Wood, S., Yuen, H.P., Yung, A.R., Desmond, P., Brewer, W., McGorry, P.D., 2002. Non-reduction in hippocampal volume is associated with higher risk of psychosis. *Schizophr. Res.* 58 (2–3), 145–158.
- Quednow, B.B., Frommann, I., Berning, J., Kühn, K.U., Maier, W., Wagner, M., 2008. Impaired sensorimotor gating of the acoustic startle response in the prodrome of schizophrenia. *Biol. Psychiatry* 64 (9), 766–773.
- Sampson, P.D., Streissguth, A.P., Barr, H.M., Bookstein, F.L., 1989. Neurobehavioral effects of prenatal alcohol: Part II. partial least squares analysis. *Neurotoxicol. Teratol.* 11 (5), 477–491.
- Scahill, R.L., Schott, J.M., Stevens, J.M., Rossor, M.N., Fox, N.C., 2002. Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI. *Proc. Natl Acad. Sci. USA* 99 (7), 4703–4707.
- Schölkopf, B., Smola, A., 2002. Learning with kernels. Support vector machines, regularization, optimization and beyond. MIT.
- Schölkopf, B., Smola, A., Williamson, R.C., Bartlett, P.L., 2000. New support vector algorithms. *Neural Comput.* 12, 1207–1245.
- Schultze-Lutter, F., Ruhrmann, S., Pickler, H., von Reventlow, H.G., Daumann, B., Brockhaus-Dumke, A., Klosterkötter, J., Pukrop, R., 2007. Relationship between subjective and objective cognitive function in the early and late prodrome. *Br. J. Psychiatry Suppl.* 51, s43–s51.
- Seidman, L.J., Pantelis, C., Keshavan, M.S., Faraone, S.V., Goldstein, J.M., Horton, N.J., Makris, N., Falkai, P., Caviness, V.S., Tsuang, M.T., 2003. A review and new report of medial temporal lobe dysfunction as a vulnerability indicator for schizophrenia: a magnetic resonance imaging morphometric family study of the parahippocampal gyrus. *Schizophr. Bull.* 29 (4), 803–830.
- Shaw, P., Kabani, N.J., Lerch, J.P., Eckstrand, K., Lenroot, R., Gogtay, N., Greenstein, D., Clasen, L., Evans, A., Rapoport, J.L., Giedd, J.N., Wise, S.P., 2008. Neurodevelopmental trajectories of the human cerebral cortex. *J. Neurosci.* 28 (14), 3586–3594.
- Stephan, K.E., Baldeweg, T., Friston, K.J., 2006. Synaptic plasticity and dysconnection in schizophrenia. *Biol. Psychiatry* 59 (10), 929–939.
- Stone, M., 1974. Cross-validator choice and assessment of statistical predictions. *J. Royal Stat. Soc. B* 36 (1), 111–147.
- Sun, D., Stuart, G.W., Jenkinson, M., Wood, S.J., McGorry, P.D., Velakoulis, D., van Erp, T.G.M., Thompson, P.M., Toga, A.W., Smith, D.J., Cannon, T.D., Pantelis, C., 2009a. Brain surface contraction mapped in first-episode schizophrenia: a longitudinal magnetic resonance imaging study. *Mol. Psychiatry* 14 (10), 976–986.
- Sun, D., Phillips, L., Velakoulis, D., Yung, A., McGorry, P.D., Wood, S.J., van Erp, T.G.M., Thompson, P.M., Toga, A.W., Cannon, T.D., Pantelis, C., 2009b. Progressive brain structural changes mapped as psychosis develops in 'at risk' individuals. *Schizophr. Res.* 108 (1–3), 85–92.
- Sun, D., van Erp, T.G.M., Thompson, P.M., Bearden, C.E., Daley, M., Kushan, L., Hardt, M.E., Nuechterlein, K.H., Toga, A.W., Cannon, T.D., 2009c. Elucidating a magnetic resonance imaging-based neuroanatomic biomarker for psychosis: classification analysis using probabilistic brain atlas and machine learning algorithms. *Biol. Psychiatry*.
- Takahashi, T., Wood, S.J., Yung, A.R., Soulsby, B., McGorry, P.D., Suzuki, M., Kawasaki, Y., Phillips, L.J., Velakoulis, D., Pantelis, C., 2009. Progressive gray matter reduction of the superior temporal gyrus during transition to psychosis. *Arch. Gen. Psychiatry* 66 (4), 366–376.
- Tura, E., Turner, J.A., Fallon, J.H., Kennedy, J.L., Potkin, S.G., 2008. Multivariate analyses suggest genetic impacts on neurocircuitry in schizophrenia. *NeuroReport* 19 (6), 603–607.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 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.
- Vapnik, V.N., 1999. An overview of statistical learning theory. *IEEE Trans. Neural Netw.* 10 (5), 988–999.
- Varma, S., Simon, R., 2006. Bias in error estimation when using cross-validation for model selection. *BMC Bioinform.* 7, 91.
- Whitford, T.J., Grieve, S.M., Farrow, T.F.D., Gomes, L., Brennan, J., Harris, A.W.F., Gordon, E., Williams, L.M., 2006. Progressive grey matter atrophy over the first 2–3 years of illness in first-episode schizophrenia: a tensor-based morphometry study. *Neuroimage* 32 (2), 511–519.
- Whitwell, J.L., Anderson, V.M., Scahill, R.L., Rossor, M.N., Fox, N.C., 2004. Longitudinal patterns of regional change on volumetric MRI in frontotemporal lobar degeneration. *Dement. Geriatr. Cogn. Disord.* 17 (4), 307–310.
- Witthaus, H., Kaufmann, C., Bohner, G., Ozgürdal, S., Gudowski, Y., Gallinat, J., Ruhrmann, S., Brüne, M., Heinz, A., Klingebiel, R., Juckel, G., 2009. Gray matter abnormalities in subjects at ultra-high risk for schizophrenia and first-episode schizophrenic patients compared to healthy controls. *Psychiatry Res.* 173 (3), 163–169.
- Wold, H., 1966. Estimation of principal components and related models by iterative least squares. In: Krishnaiah, P.R. (Ed.), *Multivariate analysis*. Academic Press, New York, pp. 391–420.
- Yung, A.R., Phillips, L.J., McGorry, P.D., McFarlane, C.A., Francey, S., Harrigan, S., Patton, G.C., Jackson, H.J., 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., Francey, S.M., McFarlane, C.A., Hallgren, M., McGorry, P.D., 2003. Psychosis prediction: 12-month follow up of a high-risk ("prodromal") group. *Schizophr. Res.* 60 (1), 21–32.
- Yung, A.R., Phillips, L.J., Yuen, H.P., McGorry, P.D., 2004. Risk factors for psychosis in an ultra high-risk group: psychopathology and clinical features. *Schizophr. Res.* 67 (2–3), 131–142.