Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers

Nikolaos Koutsouleris^{*a*,*CA*}, MD; Eva M. Meisenzahl^{*a*,*}, MD; Stefan Borgwardt^{*b*}, MD; Anita Riecher-Rössler^{*b*}, MD; Thomas Frodl^{*a*,*d*,*e*}, MD; Joseph Kambeitz,^{*a*} MD, PhD; Yanis Köhler^{*a*}; Peter Falkai^{*a*}, MD; Hans-Jürgen Möller,^{*a*} MD; Maximilian Reiser^{*c*}, MD; Christos Davatzikos^{*f*}, Ph.D.

^a Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University, Munich, Germany

^b Department of Psychiatry, University of Basel, Switzerland

^c Department of Radiology, Ludwig-Maximilian-University, Munich, Germany

^d Department of Psychiatry and Psychotherapy, University of Regensburg, Germany

^e Department of Psychiatry, University Dublin, Trinity College Dublin, Ireland

^f Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, United States

* Equally contributed

CA Corresponding Author's address:

Dep. of Psychiatry and Psychotherapy, Luwdig-Maximilian-University, Nussbaumstr. 7, 80336 Munich, GERMANY Phone: ++49 89 4400 55885, Email: Nikolaos.Koutsouleris@med.uni-muenchen.de

Abstract

MRI-based markers of schizophrenia have been repeatedly shown to separate patients from healthy controls at the single-subject level, but it remains unclear whether these markers reliably distinguish schizophrenia from mood disorders across the life span and generalize to new patients as well as to early stages of these illnesses. The current study used structural MRI-based multivariate pattern classification to (1) identify and cross-validate a differential diagnostic signature separating patients with first-episode and recurrent stages of schizophrenia (SZ, N=158) from patients with major depression (MD, N=104), and (2) quantify the impact of major clinical variables, including disease stage, age of disease onset and accelerated brain aging on the signature's classification performance. This diagnostic MRI signature was then evaluated in an independent patient cohort from two different centers to test its generalizability to bipolar disorder (BIP, N=35), firstepisode psychosis (FEP, N=23) and clinically defined at-risk mental states (ARMS, N=89). Neuroanatomical diagnosis was correct in 80% / 72% of MD / SZ patients and involved a pattern of prefronto-temporo-limbic volume reductions and premotor, somatosensory and subcortical increments in SZ vs. MD. Diagnostic performance was not influenced by the presence of depressive symptoms in SZ and psychotic symptoms in MD, but earlier disease onset and accelerated brain aging promoted misclassification in MD due to an increased neuroanatomical SZ likeness of these patients. Furthermore, disease stage significantly moderated neuroanatomical diagnosis as recurrently-ill patients had higher misclassification rates (MD: 23%; SZ: 29%) than first-episode patients (MD: 15%; SZ: 12%). Finally, the trained biomarker classified 74 % of the BIP patients as MD, while 83% / 77% / 61% of the FEP / ultra-high risk / low-risk ARMS individuals were labeled as SZ. Our findings suggest that neuroanatomical information may provide generalisable diagnostic tools distinguishing schizophrenia from mood disorders early in the course of psychosis. Disease course-related variables such as age of disease onset and disease stage as well alterations of structural brain maturation may strongly impact on the neuroanatomical separability of major depression and schizophrenia.

Word count: Abstract (329), Introduction (643), Methods (1677), Results (641), Discussion (1401), Total (4362)

Introduction

Psychiatric diagnoses arise from complex clinical processes and hence are prone to errors (Freedman *et al.*, 2013), depending on the patient's symptoms, the interviewer's experience and the classification systems' normative validity. Biological data so far only served the exclusion of somatic pathologies, leaving the question unanswered whether individualized differential diagnosis could benefit from the analysis of complex neurodiagnostic *patterns* (Fu and Costafreda, 2013; Perkins *et al.*, 2014). Furthermore, pattern analysis could unveil overlaps between and heterogeneity within diagnoses, thus promoting the revision of psychiatric nosology, and ultimately the convergence of neuroscientific and clinical observation (Krystal and State, 2014).

Phenomenological heterogeneity particularly characterizes schizophrenic psychoses and mood disorders (Linscott and Os, 2010; Murray et al., 2005): Affective symptoms are a core feature of prodromal (Addington et al., 2014; Schultze-Lutter et al., 2007) and established schizophrenia (Baynes et al., 2000; Chemerinski et al., 2008; Cotton et al., 2012; Marengo et al., 2000; Romm et al., 2010; Sönmez et al., 2013) while psychotic symptoms frequently coalesce with mania and depression (Goodwin and Jamison, 2007; Ohayon and Schatzberg, 2002). At the brain level, heterogeneity appears as subgrouping and cross-nosological effects, including neuroanatomical correlates of different symptom dimensions (Koutsouleris et al., 2008; Nenadic et al., 2012; Zhang et al., 2014), overlapping and segregating structural abnormalities (Bora et al., 2008, 2010; Du et al., 2012; Hulshoff Pol et al., 2012; Yu et al., 2010) and gradual transitions of brain activation patterns between diagnostic entities (Brandt et al., 2014). Clinically, this heterogeneity may contribute to diagnostic uncertainty along the diversity of possible disease trajectories (Baca-Garcia et al., 2007; Pope et al., 2013; Salvatore et al., 2013). Scientifically, it challenged the detection of diagnostically specific neurobiological markers and hence questioned the validity of the current disease taxonomy (Keshavan and Brady, 2011; Linscott and Os, 2010), suggesting that unipolar depression, bipolar disorder and schizophrenia may represent 'stages' or domains' along a phenotypic and neurobiological disease continuum (Green et al., 2009; Häfner et al., 2005; Lin et al., 2013).

To simultaneously address this debate and close the translational gap between neurobiologi-

cal findings and their clinical application, researchers increasingly employed multivariate pattern analysis (MVPA) to quantify the sensitivity, specificity and generalisability of diagnostic brain signatures (Bray *et al.*, 2009; Fu and Costafreda, 2013) rather than describing them in terms of their constituents' group-level significance (Davatzikos, 2004). Using MVPA, the field recently demonstrated a high separability of different neuropsychiatric conditions *vs. healthy controls*, thus foreshadowing a potential translation of neuroimaging findings into diagnostic tools (Kambeitz *et al.*, 2015; Orrù *et al.*, 2012). However, doubts remain whether MVPA-based biomarkers are really useful in discriminating neuropsychiatric illness from mental well-being, or whether they are rather needed as objective tools for a more reliable *differential diagnosis* (Savitz *et al.*, 2013). Initial findings suggest that neuroimaging may aid in individually separating schizophrenia from bipolar disorder (Schnack *et al.*, 2014) and major depression (Ota *et al.*, 2012; Serpa *et al.*, 2014). However, as these studies focused on pairwise comparisons it remains unclear how the reported neurodiagnostic signatures would perform in patients with 'intermediate' phenotypes and early disease states as well as in populations broadly covering the different age windows of these phenotypes.

An established approach to measure how clinical intersections, disease stages and age windows impact on neurodiagnostic performance is to investigate these variables along a single disease dimension, which is first spanned by 'extreme' or clearly distinct clinical phenotypes and then applied to the 'intermediate' or moderating conditions. This approach has been employed in the dementia field where morphometric patterns distinguishing patients with Alzheimer disease from healthy controls were used to quantify disease progression and severity in patients with Mild Cognitive Impairment (Davatzikos *et al.*, 2009). In the psychosis field, Y. Fan et al. (2008) employed this framework to trace the neuroanatomical schizophrenia (SZ) signature in unaffected firstdegree relatives of SZ patients, indicating that the latter display intermediate neuroanatomical phenotypes between patients and controls. Herein, we took a similar approach to explore the hypothesis that the neuroanatomical signatures of major depression (MD), bipolar disorder (BIP), the atrisk mental states for psychosis (ARMS) and SZ lie along a single direction spanned by MD and SZ

as the two end points of this continuum. Therefore, we first measured the single-subject separability of stable SZ vs. MD in a representative database of 262 patients using MRI-based MVPA and then quantified differential diagnostic scores of independent persons with high-risk or first-episode states of psychosis (N=112) as well as patients with bipolar disorder (N=35). Second, we evaluated whether neurodiagnostic classification was moderated by important variables such as age of disease onset, disease stage and 'accelerated aging' effects (Koutsouleris *et al.*, 2013) as well as cross-sectional psychopathological profiles overlapping between MD and SZ. We expected classification performance to be moderated by gradients of neuroanatomical SZ likeness *increasing* (1) from at-risk states to established schizophrenia, (2) from major depression, over bipolar disorder to schizophrenia, and (3) from later to earlier disease onsets across the life span.

Materials & Methods

Participants

SZ and MD patients were examined at the Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University Munich (LMU) using the Structured Clinical Interview for DSM-IV – Axis I & II Disorders (SCID-I/-II), the review of records and psychotropic medications and a semistandardized assessment of the psychiatric and somatic history. Patients' symptoms were evaluated using standard psychometric scales (**Table 1**). Patients received a consensus diagnosis by two experienced psychiatrists at study inclusion and were excluded in case of an unstable SCID diagnosis over a four-year follow-up period. Further exclusion criteria were (1) a history of (a) schizoaffective and/or bipolar disorder, (b) traumatic brain injury with loss of consciousness, mental retardation, anorexia nervosa, delirium, dementia, amnestic disorders, personality disorders, substance dependence, as defined by DSM-IV, (c) previous electroconvulsive treatments, and (d) somatic conditions affecting the central nervous system, as well as (2) insufficient knowledge of German, IQ < 70, and age <18 or >65. Eleven MD patients fulfilled criteria for psychotic depression (DSM-IV: 296.24/.34). Psychotic psychopathology in the MD group was further quantified by computing a composite *Z* score from the HAMD items 'feelings of guilt', 'hypochondriasis', 'depersonalization and derealisation' and 'paranoid symptoms'. This score was significantly elevated in patients with psychotic MD (MD-P: mean (SD): 1.2 (1.0); MD-NP: -0.2 (0.9); T=4.5, P<0.001). In the SZ group the severity of depressive symptoms was measured by summing the PANSS items 'somatic concern', 'anxiety', 'guilt feelings' and 'depression' and Z transforming this PANSS-D score (El Yazaji *et al.*, 2002; Kontaxakis *et al.*, 2000).

Patients with an illness duration of <1 year, no previous inpatient treatment and <12 months (life-time) psychopharmacological treatment (antipsychotics in SZ, antidepressants in MD) were assigned to first-episode (FE) subgroups, or to recurrently-ill (RE) samples, if they did not fulfil these criteria. These FE criteria were chosen to mitigate potential secondary disease effects (e.g. continuous medication and frequent hospitalization) on brain structure in the respective MD and SZ subgroups. Illness duration was the time between MRI scanning and disease onset defined retrospectively by the onset of symptoms paralleled by a general decline in social and role functioning (Lieberman *et al.*, 2001). Following these definitions, the mean (SD) illness duration in the MD-FE / SZ-FE samples was 0.34 (0.24) / 0.37 (0.68) years, while the respective values for the MD-RE / SZ-RE were 9.19 (8.22) / 7.25 (7.14) years. Diagnosis had no significant main (*F*=1.53, *P*=.217) or interaction effects (F=1.63, P=.203) on illness duration in the FE and RE samples.

The SZ vs. MD classifier was independently validated in 23 patients with first-episode psychosis (Yung *et al.*, 1998) recruited at the Department of Psychiatry, University of Basel and 89 ARMS individuals pooled across the LMU (N=52) and Basel (N=37) early recognition services, which were detailed in previous work (Koutsouleris *et al.*, 2009) (**Supplementary Methods & Table 1**). ARMS individuals were stratified into an (1) early ARMS (ARMS-E, N=21) defined either by predictive basic symptoms OR a Global Functioning-Trait criterion, and (2) late ARMS (ARMS-L, N=68) defined by attenuated or brief limited intermittent psychotic symptoms, which closely corresponded to internationally established high-risk criteria (Klosterkötter *et al.*, 2001; Yung *et al.*, 1998). Psychosis developed in 4.8% / 47.1% of ARMS-E / ARMS-L individuals over a follow-up period of 4.5 years (N=33, 87.9% diagnosed as SZ). At MRI, 61% / 95% of FEP / ARMS individuals were antipsychotic-naïve (FEP: 6 with antipsychotic treatment for <1 month, and 3 for 1-3 months; ARMS: 4 treated

with low-dose atypical antipsychotics for <3 weeks). The diagnosis of FEP patients was evaluated 5 years after baseline and all examined subjects met DSM-IV criteria for schizophrenic psychosis.

Furthermore, classifier validation involved 35 LMU patients with an established SCID diagnosis of bipolar disorder (**Table 1**), who did not meet exclusion criteria 1(b)-(d) and 2. Thirty / Five of these patients fulfilled criteria for bipolar I / II disorder, with bipolar I patients showing depressive (N=11), manic (N=12), mixed episodes (N=3) and euthymic states (N=4). Psychotic episodes were present in 6 bipolar I patients (4 / 2 with manic / depressive states).

Finally, 437 healthy volunteers (HC) previously described in Koutsouleris et al. (2013) and scanned at the same Munich scanner as the patient cohorts were used to correct the patients' MRI data for age and sex effects as detailed below. The study was approved by each center's local ethics committee. Written informed consent was obtained from each participant before inclusion.

MRI data acquisition and preprocessing

Study participants were scanned using two SIEMENS (Erlangen, Germany) MAGNETOM VISION 1.5T scanners located at the University Hospital Basel and the Department of Radiology, Ludwig-Maximilian-University. In Basel, a T1-weighted three-dimensional volumetric spoiled gradient recalled echo sequence generated 176 contiguous slices using the following protocol: time-to-echo (TE), 4 ms; time-to-repetition (TR), 9.7 ms; flip angle, 12; field of view (FOV), 25.6x25.6 cm, matrix, 200x256; voxel dimensions, 1.28 x 1.0 x 1.0 mm. In Munich, a T1-weighted 3D-MPRAGE sequence was employed: TE, 4.9 ms; TR, 11.6 ms; FOV, 230 mm; matrix, 512 x 512; 126 contiguous axial slices; voxel dimensions, 0.45 x 0.45 x 1.5 mm. No calibration of MRI scanners was performed prior to or during the recruitment period.

MRI preprocessing first involved the segmentation of T1-weighted images into gray (GM) and white matter (WM) as well as cerebro-spinal fluid using the VBM8 toolbox (see Koutsouleris et al. (2013) and **Supplementary Methods**) (Gaser, 2009). Then, the high-dimensional DRAMMS (Ou *et al.*, 2011, 2014) algorithm registered each GM map to the single-subject MNI template. Result-ing deformations and warped tissue maps were used to compute GM maps for a Regional Analysis of brain Volumes in Normalized Space (GM-RAVENS) (Davatzikos *et al.*, 2001).

Correction for age and sex effects

To remove age- and sex-related differences between patient groups while retaining diseaseassociated neuroanatomical variation, the following strategy (Dukart *et al.*, 2011) was employed: First, we calculated voxel-level β coefficients for age and sex in our HCs' GM-RAVENS maps using partial correlation analysis. These coefficients described maps of (1) GM volume change from 18 to 65 years-old HCs, and (2) GM volume differences between male and female HCs. Then we residualized the patient data using these coefficients to correct for age- and sex effects not attributable to disease-related factors. This strategy was validated in our **Supplementary Methods**.

Differential diagnostic pattern classification

We implemented a fully automated machine learning pipeline that extracted neuroanatomical features from the GM-RAVENS maps and generated decision rules from these features to individually distinguish MD from SZ patients. To strictly separate the training process from the evaluation of the classifier's generalizability, the pipeline was embedded into a repeated, double cross-validation framework (Filzmoser *et al.*, 2009) (rdCV, **Supplementary Methods**), as detailed previously (Borgwardt *et al.*, 2012; Koutsouleris *et al.*, 2012). More specifically, the following analysis steps were wrapped into a 10x10-fold cross-validation cycle at the outer (CV_2) and the inner (CV_1) levels of rdCV: the training subjects' GM-RAVENS maps were initially corrected for age and sex effects (see above) and then scaled voxel-wise to [0, 1]. To reduce the maps' dimensionality and discard noisy information, Principal Component Analysis (PCA) (Hansen *et al.*, 1999) projected correlated voxel sets to 170 uncorrelated eigenvariates, thus retaining 80% of the variance in each CV_1 training partition. Correction, scaling, and PCA parameters were applied to the CV_1 test data. Then, in each training partition, PCA features entered a recursive feature elimination algorithm (Guyon *et al.*, 2002) that employed a linear SVM (R. Fan *et al.*, 2008) to remove those eigenvariates that impaired separability on the respective CV_1 test data (SVM penalty parameter: *C*=1).

This process was repeated for all CV_1 partitions, thus creating 100 diagnostic models for each CV_2 partition. To obtain CV_2 test predictions, the respective GM-RAVENS data were first processed using the correction, scaling and PCA parameters of each CV_1 training partition, and then

classified using the learned decision rules. Classification produced decision scores measuring the neuroanatomical SZ vs. MD likeness of a given subject. Finally, a CV_2 test case's group membership was predicted by an ensemble classifier that averaged the decision scores of those 1000 CV_1 base learners in the rdCV, in which the subject had not been involved in the training process (**Supplementary Methods**). The BIP, FEP, ARMS-E and ARMS-L samples were processed identically to the CV_2 test subjects. Finally, the classifier's decision function was visualized in **Figure 3** and the underlying patterns of volumetric differences were quantified in **Supplementary Figure 6**.

Testing differential diagnostic gradients in the ARMS and patient cohorts

First, the decision scores generated by the differential diagnostic classifier entered ANOVAs that tested the hypotheses of neuroanatomical SZ likeness increasing (1) from the MD, through the BIP, to the SZ group, and (2) from the ARMS-E, through the ARMS-L to the FEP sample. In case of significant omnibus test statistics (P<0.05), post-hoc tests were carried out to evaluate pairwise differences at P<0.05, corrected for multiple comparisons using Tukey's HSD test (**Figure 1**).

Testing clinical and brain structural moderators of neurodiagnostic classification

Second, potential moderating effects of disease stage on classification performance were evaluated at *P*<0.05 by stratifying MD and SZ patients into FE vs. RE subgroups and performing a χ^2 test on the misclassification error in these samples. Then, the impact of age of onset and BrainAGE (Koutsouleris *et al.*, 2013) on decision scores was investigated by median-splitting the SZ and MD groups according to the latter two variables. Main and interactions effects between decision scores and the factors 'Diagnosis', 'Early vs. Late onset', 'Low vs. High BrainAGE' were assessed at *P*<0.05 using the General Linear Model (**Table 4, Figure 2, A**). Further analyses evaluated if classification of early-onset/high-BrainAGE patients vs. late-onset/low-BrainAGE patients equaled diagnostic categorization (**Figure 2, B1 & B2**). Based on these analyses, we assessed the separability within and between onset-defined diagnostic subgroups by performing pairwise SVM analyses as described above (**Table 2**).

Third, we explored potential moderating effects of psychometric psychosis on neurodiagnostic

classification in MD by comparing the decision scores of patients with high (N=17) versus low (N=16) scores on the standardized composite scale of HAMD items 2, 15, 19 and 20. These MD subgroups were identified by thresholding the composite scale at Z > 1 and Z <-1. The same procedure was used to measure the effect of psychometric depression on neurodiagnostic classification in SZ: Identical Z thresholds were applied to the standardized PANSS-D subscale and neuro-anatomical decision scores were compared between the resulting SZ subgroups with high (N=30) and low (N=20) depression scores. Finally, correlations between the decision scores and additional clinical variables of the MD and SZ samples were analyzed in the **Supplementary Table 1**.

Results

Sociodemographic and clinical variables: SZ and MD patients groups did not differ regarding handedness, BMI, schooling years, nicotine or alcohol consumption (**Table 1**). Patient groups differed in the prescribed antipsychotic, antidepressant and mood-stabilizing medications. However, all these variables had no effect on neurodiagnostic decision scores (**Supplementary Table 1**). Group-level differences were observed for age at scan, sex and age of disease onset, but not illness duration. Finally, SZ patients had a higher mean (SD) BrainAGE score of +5.99 (6.00) compared to MD (+4.04 (6.19)).

Neuroanatomical classification and influence of moderating variables: The MRI classifier diagnosed unseen MD and SZ patients with a balanced accuracy (BAC) of 76% (sensitivity / specificity = 79.8% / 72.2%, diagnostic odds ratio = 10.2; **Table 2**). Recurrently-ill patients were more likely misclassified compared to first-episode patients (Error rates MD_{FE} / SZ_{FE} : 15.0% / 11.5%; MD_{RE} / SZ_{RE} : 23.4% / 28.8%; χ^2 =6.6; *P*=.010). The neuroanatomical decision function (**Figure 3**) involved GM reductions in SZ vs. MD covering the perisylvian structures (inferior frontal, insular, supramarginal, angular, superior temporal and temporopolar cortices) with extensions to the orbitofrontal, inferior temporal and medial temporal cortices. Further reductions covered the ventro-medial prefrontal, anterior cingulate, medial parietal, occipital and dorsolateral prefrontal cortices. GM reductions in MD vs. SZ were localized in a spatially distinct pattern including the brainstem

regions, cerebellum, periventricular areas and the somatosensory cortices, extending to the premotor, parietal and supplementary motor areas.

The medians of age of onset / BrainAGE used to stratify patients were 36.3 / +3.57 in the MD and 23.8 / +5.62 in the SZ sample. The GLM evaluating effects of diagnosis, age of onset and BrainAGE factors on diagnostic scores detected significant main effects as well as a significant interaction between the 'diagnosis' and 'early vs. late' factors (**Table 3**). Box plot analyses showed that early disease onset and high BrainAGE increased SZ likeness in both disease groups, with this effect being more pronounced in the MD compared to the SZ (**Figure 2, A**). Using the diagnostic decision scores, late-onset, low-BrainAGE patients were separable from early-onset, high-BrainAGE patients to a similar degree (AUC=.77) as MD from SZ patients (AUC=0.80, **Figure 2, B1 & B2**). Finally, the onset-stratified subgroup classification showed (1) a better separability of early vs. late-onset MD patients (BAC=83.7%) than early vs. late-onset SZ patients (62.3%, **Table 2**), and (2) a particularly low separability of early-onset MD vs. SZ diagnoses, the BAC was lower (72.2%) than in the original whole-group analysis.

The comparison of the neurodiagnostic scores in SZ patients with high vs. low psychometric depression scores (mean (SD): -0.59 (1.30) vs. -0.57 (0.98)) did not yield significant differences (*T*=-0.04; *P*=0.966). Similarly, MD patients with high psychometric psychosis scores did not significantly differ from patients with low scores (mean (SD): 1.64 (1.41) vs. 0.77 (1.17); *T*=1.88; *P*=0.071). Additionally, MD patients with vs. without a DSM-IV diagnosis of psychotic depression did not differ in their neurodiagnostic scores (1.25 (1.63) vs. 0.84 (1.24); *T*=0.98; *P*=0.331) nor in their misclassification rates (20.0% vs. 20.2%; χ^2 =0.00; *P*=1.000).

Presence of a SZ-like neuroanatomical signature in ARMS, FEP and BIP subjects: Decision scores obtained from the independent validation data showed that SZ likeness was most pronounced in the Basel FEP patients (86.9% labeled as SZ, **Figure 1**) followed by the cross-center ARMS-L group (77.9%) and the Munich ARMS-E sample (61.0%). SZ likeness was lower in the Munich BIP group (25.7%) resulting in 74% of these patients being classified as MD. Significant

group differences were detected in all pairwise post-hoc contrasts of the MD vs. BIP vs. SZ comparison (**Figure 1**). In the ARMS-E vs. ARMS-L vs. FEP analysis, we observed a significant increase of SZ likeness in the ARMS-L compared to the ARMS-E group (P=0.042) with the former being on par with the FEP sample (P=1.000).

Discussion

This is to our knowledge the first structural MRI study to report a cross-validated, single-subject separability of 76% in a representative cohort of patients with a stable diagnosis of either schizophrenia or major depression. This finding is in keeping with the balanced accuracy of 78% reported by Ota et al. (2013) who examined 25 age-matched female patients with SZ or MD using fractional anisotropy and GM volumes in predefined regions-of-interest. We observed that neuroanatomical markers successfully generalized to patients with first-episode psychosis who were examined at an independent center using a different MRI protocol and were prospectively diagnosed with schizophrenia. Validation also showed that diagnostic sensitivity extended to the ARMS and grew with symptomatic proximity to overt psychosis. Strikingly, the neurodiagnostic classifier assigned 74% of patients with bipolar disorder to the MD group, suggesting that schizophrenia may be differentiated from mood disorders at the single-subject level. In addition, we did not find evidence that neurodiagnostic classification was significantly influenced by the presence of psychotic symptoms in MD or depressive symptoms in SZ patients, nor by life-style factors or different medications at the time of MRI scanning (Supplementary Table 1). However, we identified a neuroanatomical signature shared by SZ and MD patients with an average disease onset at 26.5 years (Supplementary Table 2) and accelerated brain aging effects (Koutsouleris et al., 2013), which led to a nonseparability of these subgroups.

Our findings partly agree with recent studies that used structural imaging data (Ota *et al.*, 2013) or Near Infrared Spectroscopy (Takizawa *et al.*, 2014) to differentiate between functional psychoses at the single-subject level. However, comparability to these studies is limited because we did not mitigate naturally occurring demographic differences between SZ and MD by studying matched patient samples. Instead, we adjusted our data using a representative database of

healthy controls that fully covered the age range of our patient population (Dukart et al., 2011). Therefore, potential confounds like divergent illness durations and equalized sex distributions were avoided a priori. This approach facilitated the evaluation of the neurodiagnostic pattern, its presence in partly overlapping clinical phenotypes and its clinical moderators across the adult life span: First, we identified a pattern of perisylvian, prefrontal and temporo-limbic GM volume reductions used by the classifier to separate SZ from MD at the single-subject level. Similar patterns were repeatedly described to distinguish patients with schizophrenia from healthy controls (Bora et al., 2011; Honea et al., 2005) and were interpreted in line with a disconnection syndrome (Friston, 1999) underlying the cognitive, perceptual and thought disturbances of psychosis (Modinos et al., 2012; Sans-Sansa et al., 2013). Our finding of a GM volume reduction signature (Supplementary Figure 6) in SZ compared to MD patients, who were on average 11.5 years older, adds to the concept of schizophrenia being a neurodevelopmentally mediated, cognitive illness (Kahn and Keefe, 2013) marked by more unfavorable disease outcomes compared to unipolar depression (Harrow et al., 2000). In contrast, somatosensory, periventricular and subcortical abnormalities distinguished MD from SZ in line with previously reported structural abnormalities in major white matter tracts of depressed patients (Disabato et al., 2014; Guo et al., 2014). Alterations of these regions may subserve core features of depression, such as psychomotor and mood disturbances (Serafini et al., 2010).

Second, the high rate of MD classifications in our BIP sample suggests that bipolar disorder and unipolar depression share a common structural denominator different from schizophrenia in line with Kraepelin's original dichotomic concept of functional psychoses (Kraepelin, 1899). This finding agrees with initial reports of a good single-subject separability of SZ and BIP based on structural (Schnack *et al.*, 2014) or functional MRI (Costafreda *et al.*, 2011). It may also point to an increased sensitivity of MVPA techniques in detecting points of rarity in high-dimensional neuroimaging data compared to univariate methods, which frequently reported considerable overlaps between bipolar disorder and schizophrenia (Arnone *et al.*, 2009). On the other hand, we found a significant difference in the neurodiagnostic scores of the MD and BIP groups (**Figure 1**), which

adds to the growing evidence for a neurobiological signature separating these two largely overlapping conditions (Redlich *et al.*, 2014). Hence, future studies with access to larger sample sizes in the bipolar and schizoaffective categories are needed to directly quantify the neurodiagnostic separability of these 'intermediate' phenotypes (Mathew *et al.*, 2014), thus providing a comprehensive picture of continuities and discontinuities in the functional psychoses spectrum (Laursen *et al.*, 2009).

Third, we found that an *earlier disease occurrence* correlated with lower differential diagnostic accuracy, rendering MD patients with earlier disease onsets inseparable from SZ patients - despite distinct cross-sectional phenotypes. This observation was corroborated by the high separability of age of onset-defined MD samples (83.7%, Table 2) compared to the respective SZ subgroups (62.3%), which suggests that neuroanatomical surrogates of depressive syndromes strongly covary with the disease onset axis. This hypothesis has recently received support from studies showing a pronounced thinning in prefrontal, cingulate, precuneal and inferior temporal cortices of early vs. late-onset MD patients and healthy controls (Truong et al., 2013) (see Supplementary Figure 5). Hence, our results may point to more disrupted neurodevelopmental processes in early depression, potentially manifesting as an accelerated brain aging effect (Koutsouleris et al., 2013). These processes may also be linked to a more severe clinical phenotype of depression, entailing greater illness severity, higher relapse rates, more cognitive disturbances, as well as overall poorer disease outcomes and higher familial co-aggregation with schizophrenia (Dekker et al., 2007; Korten et al., 2012; Maier et al., 1993; Zisook et al., 2004). Hence, the overlaps between schizophrenia and early-onset depression may lead to a diagnostic dilemma, particularly in the early phases of these illnesses when overt psychotic symptoms have not yet evolved or patients are not explored during psychotic phases. Our results indicate that this challenge cannot be resolved by our neuroanatomical classifier, which would frequently diagnose these early-onset depressed patients with schizophrenia. Thus, it remains to be elucidated whether different imaging modalities and combinations thereof may help increasing the diagnostic specificity in the neurobiological classification of early-onset depression.

Fourth, we observed an increasing SZ likeness from ARMS-E to FEP, suggesting that the neurodiagnostic fingerprint of schizophrenia is already detectable in persons with basic symptoms, and further intensifies as attenuated, brief limited intermittent and frank psychotic symptoms emerge. This finding agrees with previous studies reporting longitudinal GM volume changes in the ARMS, indicating a progressive course of neuroanatomical alterations as the at-risk state evolves into overt psychosis (Cannon *et al.*, 2014; Koutsouleris *et al.*, 2010). However, due to the crosssectional design of our study, it remains unclear (1) whether the increase of SZ-likeness along these early states of psychosis also occurs at the level of neuroanatomical disease trajectories, and (2) which protective factors contribute to a non-conversion to psychosis despite the presence of the SZ-specific pattern in a given patient.

Finally, one caveat has to be considered when interpreting our findings: different medication and treatment histories in our SZ vs. MD groups may have influenced the separability of our patients as long-standing antipsychotic treatment has been previously shown to interact with diseaserelated brain changes (Ho et al., 2011). Although life-time medication data was not available for the current dataset, the high diagnostic sensitivity in our minimally treated ARMS and FEP groups argues against major treatment effects on our results. Furthermore, our finding of a significantly higher classification performance in first-episode compared to recurrently-ill patients is at odds with the expectation that relapsing illness stages and - in consequence - accumulating disease-specific treatment effects would increase the neuroanatomical gaps between schizophrenia and unipolar depression. The higher diagnostic error in the recurrently-ill patient sample could be interpreted as a dilution effect, which may arise from increasing neuroanatomical heterogeneity as patients evolve along divergent disease trajectories. Hence, this heterogeneity may result from (1) structural brain variation linked to differential disease courses (Mourao-Miranda et al., 2012) and treatment outcomes (Palaniyappan et al., 2013), (2) distinct neuroanatomical correlates of positive, negative, disorganized and depressive subsyndromes of schizophrenia, as revealed by factor analytic studies (Koutsouleris et al., 2008; Nenadic et al., 2010, 2012; Zhang et al., 2014), and (3) temporal shifts of these profiles over time, with negative and depressive symptoms becoming increasingly

prominent in the course of the disease (Salvatore *et al.*, 2013). Nevertheless, our results did not support a moderating or 'diluting' effect of depressive / psychotic syndromes in SZ / MD on neurodiagnostic classification performance. Thus, the strong impact of longitudinal disease variables such age of disease onset and disease stage may suggest that the identified neuroanatomical biomarker is linked to the temporal and neurodevelopmental characteristics of these clinical phenotypes rather than to their cross-sectional psychopathological features (Gogtay *et al.*, 2011).

In summary, our findings partly confirm and partly question the Kraepelinian dichotomy of functional psychoses into schizophrenia and affective disorders. This is not surprising if one considers the plethora of studies in support of either a phenomenological and neurobiological continuum or a division between these two nosological groups (Kotov *et al.*, 2013). Our results suggest that the diagnostic boundaries drawn by a neuroanatomical disease signature become increasingly porous as patients develop depressive disorders at younger ages, potentially mediated by a cross-nosological disruption of neurodevelopmental processes. This gradient of diagnostic uncertainty does not only challenge clinical and biomarker-based diagnosis, it highlights also the utility of pattern recognition methods to probe the neurological basis of psychiatric illnesses and potentially refine nosological disease constructs.

Conflicts of Interest

None to declare.

Acknowledgments

Role of Funding

This work was supported by The German Association for Psychiatry, Psychotherapy and Psychosomatics with a travel grant (NeuroImaging Prize) to N. Koutsouleris. C. Davatzikos was supported by NIH grant R01-AG14971 for participation in the analyses and writing of the manuscript.

References

Addington J, Shah H, Liu L, Addington D. Reliability and validity of the Calgary Depression Scale for Schizophrenia (CDSS) in youth at clinical high risk for psychosis. [Internet]. Schizophr Res 2014; 153: 64–67. Available from: http://dx.doi.org/10.1016/j.schres.2013.12.014

Arnone D, Cavanagh J, Gerber D, Lawrie SM, Ebmeier KP, McIntosh AM. Magnetic resonance imaging studies in bipolar disorder and schizophrenia: meta-analysis. [Internet]. Br J Psychiatry 2009; 195: 194–201.Available from: http://dx.doi.org/10.1192/bjp.bp.108.059717

Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, Moral AL Fernandez del, Jimenez-Arriero MA, Rivera JL Gonzalez de, et al. Diagnostic stability of psychiatric disorders in clinical practice. [Internet]. Br J Psychiatry 2007; 190: 210–216.Available from: http://dx.doi.org/10.1192/bjp.bp.106.024026

Baynes D, Mulholland C, Cooper SJ, Montgomery RC, MacFlynn G, Lynch G, et al. Depressive symptoms in stable chronic schizophrenia: prevalence and relationship to psychopathology and treatment. Schizophr Res 2000; 45: 47–56.

Bora E, Fornito A, Radua J, Walterfang M, Seal M, Wood SJ, et al. Neuroanatomical abnormalities in schizophrenia: a multimodal voxelwise meta-analysis and meta-regression analysis. [Internet]. Schizophr Res 2011; 127: 46–57. Available from: http://dx.doi.org/10.1016/j.schres.2010.12.020

Bora E, Fornito A, Yücel M, Pantelis C. Voxelwise meta-analysis of gray matter abnormalities in bipolar disorder. [Internet]. Biol Psychiatry 2010; 67: 1097–1105.Available from: http://dx.doi.org/10.1016/j.biopsych.2010.01.020

Bora E, Yucel M, Fornito A, Berk M, Pantelis C. Major psychoses with mixed psychotic and mood symptoms: are mixed psychoses associated with different neurobiological markers? [Internet]. Acta Psychiatr Scand 2008; 118: 172–187.Available from: http://dx.doi.org/10.1111/j.1600-0447.2008.01230.x

Borgwardt S, Koutsouleris N, Aston J, Studerus E, Smieskova R, Rössler A, et al. Distinguishing Prodromal From First-Episode Psychosis Using Neuroanatomical Single-Subject Pattern Recognition. Schizophr Bull 2012; in press

Brandt CL, Eichele T, Melle I, Sundet K, Server A, Agartz I, et al. Working memory networks and activation patterns in schizophrenia and bipolar disorder: comparison with healthy controls. [Internet]. Br J Psychiatry 2014; 204: 290–298.Available from: http://dx.doi.org/10.1192/bjp.bp.113.129254

Bray S, Chang C, Hoeft F. Applications of multivariate pattern classification analyses in developmental neuroimaging of healthy and clinical populations. [Internet]. Front. Hum. Neurosci. 2009; 3: 32.Available from: http://dx.doi.org/10.3389/neuro.09.032.2009

Cannon TD, Chung Y, He G, Sun D, Jacobson A, Erp TGM van, et al. Progressive Reduction in Cortical Thickness as Psychosis Develops: A Multisite Longitudinal Neuroimaging Study of Youth at Elevated Clinical Risk. Biol Psychiatry 2014

Chemerinski E, Bowie C, Anderson H, Harvey PD. Depression in schizophrenia: methodological artifact or distinct feature of the illness? [Internet]. J Neuropsychiatry Clin Neurosci 2008; 20: 431–440.Available from: http://dx.doi.org/10.1176/appi.neuropsych.20.4.431

Costafreda SG, Fu CHY, Picchioni M, Toulopoulou T, McDonald C, Kravariti E, et al. Pattern of neural responses to verbal fluency shows diagnostic specificity for schizophrenia and bipolar disorder. [Internet]. BMC Psychiatry 2011; 11: 18.Available from: http://dx.doi.org/10.1186/1471-244X-11-18

Cotton SM, Lambert M, Schimmelmann BG, Mackinnon A, Gleeson JFM, Berk M, et al. Depressive symptoms in first episode schizophrenia spectrum disorder. [Internet]. Schizophr Res 2012; 134: 20–26.Available from: http://dx.doi.org/10.1016/j.schres.2011.08.018

Davatzikos C, Genc A, Xu D, Resnick SM. Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. [Internet]. Neuroimage 2001; 14: 1361–1369.Available from: http://dx.doi.org/10.1006/nimg.2001.0937

Davatzikos C, Xu F, An Y, Fan Y, Resnick SM. Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index [Internet]. Brain 2009; 132: 2026–2035.Available from: http://brain.oxfordjournals.org/cgi/content/abstract/132/8/2026

Davatzikos C. Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. [Internet]. Neuroimage 2004; 23: 17–20.Available from: http://dx.doi.org/10.1016/j.neuroimage.2004.05.010

Dekker MC, Ferdinand RF, Lang NDJ van, Bongers IL, Ende J van der, Verhulst FC. Developmental trajectories of depressive symptoms from early childhood to late adolescence: gender differences and adult outcome. [Internet]. J Child Psychol Psychiatry 2007; 48: 657–666.Available from: http://dx.doi.org/10.1111/j.1469-7610.2007.01742.x

Disabato BM, Morris C, Hranilovich J, D'Angelo GM, Zhou G, Wu N, et al. Comparison of brain structural variables, neuropsychological factors, and treatment outcome in early-onset versus late-onset late-life depression. [Internet]. Am J Geriatr Psychiatry 2014; 22: 1039–1046.Available from: http://dx.doi.org/10.1016/j.jagp.2013.02.005

Du M-Y, Wu Q-Z, Yue Q, Li J, Liao Y, Kuang W-H, et al. Voxelwise meta-analysis of gray matter reduction in major depressive disorder. [Internet]. Prog Neuropsychopharmacol Biol Psychiatry 2012; 36: 11–16.Available from: http://dx.doi.org/10.1016/j.pnpbp.2011.09.014

Dukart J, Schroeter ML, Mueller K, Alzheimer's Disease Neuroimaging Initiative. Age correction in dementia–matching to a healthy brain. PLoS One 2011; 6: e22193.

El Yazaji M, Battas O, Agoub M, Moussaoui D, Gutknecht C, Dalery J, et al. Validity of the depressive dimension extracted from principal component analysis of the PANSS in drug-free patients with schizophrenia. Schizophr Res 2002; 56: 121–127.

Fan R, Chang K, Hsieh C, Wang X, Lin C. LIBLINEAR: A Library for Large Linear Classification. 2008

Fan Y, Gur RE, Gur RC, Wu X, Shen D, Calkins ME, et al. Unaffected family members and schizophrenia patients share brain structure patterns: a high-dimensional pattern classification study. [Internet]. Biol. Psychiatry 2008; 63: 118–124.Available from: http://dx.doi.org/10.1016/j.biopsych.2007.03.015

Filzmoser P, Liebmann B, Varmuza K. Repeated double cross validation. J. Chemometrics 2009; 23: 160–171.

Freedman R, Lewis DA, Michels R, Pine DS, Schultz SK, Tamminga CA, et al. The initial field trials of DSM-5: new blooms and old thorns. [Internet]. Am J Psychiatry 2013; 170: 1–5.Available from: http://dx.doi.org/10.1176/appi.ajp.2012.12091189

Friston KJ. Schizophrenia and the disconnection hypothesis. Acta Psychiatr Scand Suppl 1999; 395: 68–79.

Fu CHY, Costafreda SG. Neuroimaging-based biomarkers in psychiatry: clinical opportunities of a paradigm shift. Can J Psychiatry 2013; 58: 499–508.

Gaser C. Voxel-Based Morphometry Toolbox, version 8 (VBM8): http://dbm.neuro.uni-jena.de [Internet]. 2009Available from: http://dbm.neuro.uni-jena.de

Gogtay N, Vyas NS, Testa R, Wood SJ, Pantelis C. Age of onset of schizophrenia: perspectives from structural neuroimaging studies. [Internet]. Schizophr Bull 2011; 37: 504–513.Available from: http://dx.doi.org/10.1093/schbul/sbr030

Goodwin F, Jamison K. Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression. New York: Oxford University Press; 2007.

Green EK, Grozeva D, Jones I, Jones L, Kirov G, Caesar S, et al. The bipolar disorder risk allele at CACNA1C also confers risk of recurrent major depression and of schizophrenia. [Internet]. Mol. Psychiatry 2009Available from: http://dx.doi.org/10.1038/mp.2009.49

Grotegerd D, Stuhrmann A, Kugel H, Schmidt S, Redlich R, Zwanzger P, et al. Amygdala excitability to subliminally presented emotional faces distinguishes unipolar and bipolar depression: An fMRI and pattern classification study. [Internet]. Hum Brain Mapp 2013Available from: http://dx.doi.org/10.1002/hbm.22380

Guo W, Liu F, Xun G, Hu M, Guo X, Xiao C, et al. Disrupted white matter integrity in first-episode, drug-naive, late-onset depression. [Internet]. J Affect Disord 2014; 163: 70–75. Available from: http://dx.doi.org/10.1016/j.jad.2014.03.044

Guyon I, Weston J, Barnhill S, Vapnik V. Gene Selection for Cancer Classification using Support Vector Machines [Internet]. Mach. Learn. 2002; 46: 389–422.Available from: http://dx.doi.org/10.1023/A:1012487302797

Häfner H, Maurer K, Trendler G, Heiden W an der, Schmidt M, Könnecke R. Schizophrenia and depression: challenging the paradigm of two separate diseases–a controlled study of schizophrenia, depression and healthy controls. [Internet]. Schizophr. Res. 2005; 77: 11–24.Available from: http://dx.doi.org/10.1016/j.schres.2005.01.004

Hansen LK, Larsen J, Nielsen FA, Strother SC, Rostrup E, Savoy R, et al. Generalizable patterns in neuroimaging: how many principal components? [Internet]. Neuroimage 1999; 9: 534–544.Available from: http://dx.doi.org/10.1006/nimg.1998.0425

Harrow M, Grossman LS, Herbener ES, Davies EW. Ten-year outcome: patients with schizoaffective disorders, schizophrenia, affective disorders and mood-incongruent psychotic symptoms. Br J Psychiatry 2000; 177: 421–426.

Ho B-C, Andreasen NC, Ziebell S, Pierson R, Magnotta V. Long-term antipsychotic treatment and brain volumes: a longitudinal study of first-episode schizophrenia. [Internet]. Arch Gen Psychiatry 2011; 68: 128–137. Available from: http://dx.doi.org/10.1001/archgenpsychiatry.2010.199

Honea R, Crow TJ, Passingham D, Mackay CE. Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. [Internet]. Am. J. Psychiatry 2005; 162: 2233–2245.Available from: http://dx.doi.org/10.1176/appi.ajp.162.12.2233

Hulshoff Pol HE, Baal GCM van, Schnack HG, Brans RGH, Schot AC van der, Brouwer RM, et al. Overlapping and segregating structural brain abnormalities in twins with schizophrenia or bipolar disorder. [Internet]. Arch Gen Psychiatry 2012; 69: 349–359. Available from: http://dx.doi.org/10.1001/archgenpsychiatry.2011.1615

Kahn RS, Keefe RSE. Schizophrenia is a cognitive illness: time for a change in focus. [Internet]. JAMA Psychiatry 2013; 70: 1107–1112.Available from: http://dx.doi.org/10.1001/jamapsychiatry.2013.155

Kambeitz J, Kambeitz-Ilankovic L, Leucht S, Wood S, Davatzikos C, Malchow B, et al. Detecting Neuroimaging Biomarkers for Schizophrenia: A Meta-Analysis of Multivariate Pattern Recognition Studies. [Internet]. Neuropsychopharmacology 2015Available from: http://dx.doi.org/10.1038/npp.2015.22

Keshavan MS, Brady R. Biomarkers in schizophrenia: we need to rebuild the Titanic. World Psychiatry 2011; 10: 35–36.

Klosterkötter J, Hellmich M, Steinmeyer EM, Schultze-Lutter F. Diagnosing schizophrenia in the initial prodromal phase. Arch. Gen. Psychiatry 2001; 58: 158–164.

Kontaxakis VP, Havaki-Kontaxaki BJ, Stamouli SS, Margariti MM, Collias CT, Christodoulou GN. Comparison of four scales measuring depression in schizophrenic inpatients. Eur Psychiatry 2000; 15: 274–277.

Korten NCM, Comijs HC, Lamers F, Penninx BWJH. Early and late onset depression in young and middle aged adults: Differential symptomatology, characteristics and risk factors? [Internet]. Journal of Affective Disorders 2012; 138: 259–267.Available from: http://www.sciencedirect.com/science/article/pii/S0165032712000894

Kotov R, Leong SH, Mojtabai R, Erlanger AC, Fochtmann LJ, Constantino E, et al. Boundaries of schizoaffective disorder: revisiting Kraepelin. JAMA Psychiatry 2013; 70: 1276–1286.

Koutsouleris N, Borgwardt S, Studerus E, Smieskova R, Meisenzahl E, Riecher A. MRI-based prediction of psychosis across populations: findings from the Munich and Basel early recognition services. 2012; in prep.

Koutsouleris N, Davatzikos C, Borgwardt S, Gaser C, Bottlender R, Frodl T, et al. Accelerated Brain Aging in Schizophrenia and Beyond: A Neuroanatomical Marker of Psychiatric Disorders. [Internet]. Schizophr Bull 2013Available from: http://dx.doi.org/10.1093/schbul/sbt142

Koutsouleris N, Gaser C, Bottlender R, Davatzikos C, Decker P, Jäger M, et al. Use of neuroanatomical pattern regression to predict the structural brain dynamics of vulnerability and transition to psychosis. [Internet]. Schizophr Res 2010; 123: 175–187.Available from: http://dx.doi.org/10.1016/j.schres.2010.08.032

Koutsouleris N, Gaser C, Jäger MM, Bottlender R, Frodl T, Holzinger S, et al. Structural correlates of psychopathological symptom dimensions in schizophrenia: A voxel-based morphometric study. [Internet]. Neuroimage 2008; 39: 1600–1612.Available from: http://dx.doi.org/10.1016/j.neuroimage.2007.10.029

Koutsouleris N, Meisenzahl E, Davatzikos C, Bottlender R, Frodl T, Scheuerecker J, et al. Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. [Internet]. Arch Gen Psychiatry 2009; 66: 700–712.Available from: http://dx.doi.org/10.1001/archgenpsychiatry.2009.62

Kraepelin E. Psychiatrie. 6th ed. Leipzig, Germany: Barth; 1899.

Krystal JH, State MW. Psychiatric Disorders: Diagnosis to Therapy. [Internet]. Cell 2014; 157: 201–214. Available from: http://dx.doi.org/10.1016/j.cell.2014.02.042

Laursen TM, Agerbo E, Pedersen CB. Bipolar disorder, schizoaffective disorder, and schizophrenia overlap: a new comorbidity index. [Internet]. J Clin Psychiatry 2009; 70: 1432–1438.Available from: http://dx.doi.org/10.4088/JCP.08m04807

Lieberman J, Chakos M, Wu H, Alvir J, Hoffman E, Robinson D, et al. Longitudinal study of brain morphology in first episode schizophrenia. Biol. Psychiatry 2001; 49: 487–499.

Lin A, Reniers RLEP, Wood SJ. Clinical staging in severe mental disorder: evidence from neurocognition and neuroimaging. [Internet]. Br J Psychiatry Suppl 2013; 54: s11–s17.Available from: http://dx.doi.org/10.1192/bjp.bp.112.119156

Linscott RJ, Os J van. Systematic reviews of categorical versus continuum models in psychosis: evidence for discontinuous subpopulations underlying a psychometric continuum. Implications for DSM-V, DSM-VI, and DSM-VII. [Internet]. Annu Rev Clin Psychol 2010; 6: 391–419.Available from: http://dx.doi.org/10.1146/annurev.clinpsy.032408.153506

Maier W, Lichtermann D, Minges J, Hallmayer J, Heun R, Benkert O, et al. Continuity and discontinuity of affective disorders and schizophrenia. Results of a controlled family study. Arch Gen Psychiatry 1993; 50: 871–883.

Marengo J, Harrow M, Herbener ES, Sands J. A prospective longitudinal 10-year study of schizophrenia's three major factors and depression. Psychiatry Res. 2000; 97: 61–77.

Mathew I, Gardin TM, Tandon N, Eack S, Francis AN, Seidman LJ, et al. Medial temporal lobe structures and hippocampal subfields in psychotic disorders: findings from the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) study. [Internet]. JAMA Psychiatry 2014; 71: 769–777.Available from: http://dx.doi.org/10.1001/jamapsychiatry.2014.453

Modinos G, Costafreda SG, Tol M-J van, McGuire PK, Aleman A, Allen P. Neuroanatomy of auditory verbal hallucinations in schizophrenia: A quantitative meta-analysis of voxel-based morphometry studies. [Internet]. Cortex 2012Available from: http://dx.doi.org/10.1016/j.cortex.2012.01.009

Mourão-Miranda J, Almeida JRC, Hassel S, Oliveira L de, Versace A, Marquand AF, et al. Pattern recognition analyses of brain activation elicited by happy and neutral faces in unipolar and bipolar depression. [Internet]. Bipolar Disord 2012; 14: 451–460.Available from: http://dx.doi.org/10.1111/j.1399-5618.2012.01019.x

Mourao-Miranda J, Reinders AATS, Rocha-Rego V, Lappin J, Rondina J, Morgan C, et al. Individualized prediction of illness course at the first psychotic episode: a support vector machine MRI study. [Internet]. Psychol Med 2012; 42: 1037–1047.Available from: http://dx.doi.org/10.1017/S0033291711002005

Murray V, McKee I, Miller P, Young D, Muir W, Pelosi A, et al. Dimensions and classes of psychosis in a population cohort: a four-class, four-dimension model of schizophrenia and affective psychoses. Psychol Med 2005; 35: 499–510.

Nenadic I, Gaser C, Sauer H. Heterogeneity of brain structural variation and the structural imaging endophenotypes in schizophrenia. [Internet]. Neuropsychobiology 2012; 66: 44–49.Available from: http://dx.doi.org/10.1159/000338547

Nenadic I, Sauer H, Gaser C. Distinct pattern of brain structural deficits in subsyndromes of schizophrenia delineated by psychopathology. [Internet]. Neuroimage 2010; 49: 1153–1160.Available from: http://dx.doi.org/10.1016/j.neuroimage.2009.10.014

Ohayon MM, Schatzberg AF. Prevalence of depressive episodes with psychotic features in the general population. Am J Psychiatry 2002; 159: 1855–1861.

Orrù G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. [Internet]. Neurosci Biobehav Rev 2012; 36: 1140–1152.Available from: http://dx.doi.org/10.1016/j.neubiorev.2012.01.004

Ota M, Ishikawa M, Sato N, Hori H, Sasayama D, Hattori K, et al. Discrimination between schizophrenia and major depressive disorder by magnetic resonance imaging of the female brain. [Internet]. J Psychiatr Res 2013; 47: 1383–1388.Available from: http://dx.doi.org/10.1016/j.jpsychires.2013.06.010

Ou Y, Akbari H, Bilello M, Da X, Davatzikos C. Comparative Evaluation of Registration Algorithms in Different Brain Databases with Varying Difficulty: Results and Insights. [Internet]. IEEE Trans Med Imaging 2014Available from: http://dx.doi.org/10.1109/TMI.2014.2330355

Ou Y, Sotiras A, Paragios N, Davatzikos C. DRAMMS: Deformable registration via attribute matching and mutualsaliency weighting. [Internet]. Med Image Anal 2011; 15: 622–639.Available from: http://dx.doi.org/10.1016/j.media.2010.07.002

Palaniyappan L, Marques TR, Taylor H, Handley R, Mondelli V, Bonaccorso S, et al. Cortical folding defects as markers of poor treatment response in first-episode psychosis. [Internet]. JAMA Psychiatry 2013; 70: 1031–1040.Available from: http://dx.doi.org/10.1001/jamapsychiatry.2013.203

Perkins DO, Jeffries CD, Addington J, Bearden CE, Cadenhead KS, Cannon TD, et al. Towards a Psychosis Risk Blood Diagnostic for Persons Experiencing High-Risk Symptoms: Preliminary Results From the NAPLS Project. [Internet]. Schizophr Bull 2014Available from: http://dx.doi.org/10.1093/schbul/sbu099

Pope MA, Joober R, Malla AK. Diagnostic stability of first-episode psychotic disorders and persistence of comorbid psychiatric disorders over 1 year. Can J Psychiatry 2013; 58: 588–594.

Redlich R, Almeida JJR, Grotegerd D, Opel N, Kugel H, Heindel W, et al. Brain morphometric biomarkers distinguishing unipolar and bipolar depression. A voxel-based morphometry-pattern classification approach. [Internet]. JAMA Psychiatry 2014; 71: 1222–1230.Available from: http://dx.doi.org/10.1001/jamapsychiatry.2014.1100

Romm KL, Rossberg JI, Berg AO, Barrett EA, Faerden A, Agartz I, et al. Depression and depressive symptoms in first episode psychosis. [Internet]. J Nerv Ment Dis 2010; 198: 67–71.Available from: http://dx.doi.org/10.1097/NMD.0b013e3181c81fc0

Salvatore P, Baldessarini RJ, Khalsa H-MK, Amore M, Di Vittorio C, Ferraro G, et al. Predicting diagnostic change among patients diagnosed with first-episode DSM-IV-TR major depressive disorder with psychotic features. [Internet]. J Clin Psychiatry 2013; 74: 723–31; quiz 731.Available from: http://dx.doi.org/10.4088/JCP.12m08328

Sans-Sansa B, McKenna PJ, Canales-Rodríguez EJ, Ortiz-Gil J, López-Araquistain L, Sarró S, et al. Association of formal thought disorder in schizophrenia with structural brain abnormalities in language-related cortical regions. [Internet]. Schizophr Res 2013; 146: 308–313.Available from: http://dx.doi.org/10.1016/j.schres.2013.02.032

Savitz JB, Rauch SL, Drevets WC. Clinical application of brain imaging for the diagnosis of mood disorders: the current state of play. [Internet]. Mol Psychiatry 2013; 18: 528–539.Available from: http://dx.doi.org/10.1038/mp.2013.25

Schnack HG, Nieuwenhuis M, Haren NEM van, Abramovic L, Scheewe TW, Brouwer RM, et al. Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar

disorder and healthy subjects. [Internet]. Neuroimage 2014; 84: 299–306.Available from: http://dx.doi.org/10.1016/j.neuroimage.2013.08.053

Schultze-Lutter F, Ruhrmann S, Picker H, Reventlow HG von, Brockhaus-Dumke A, Klosterkötter J. Basic symptoms in early psychotic and depressive disorders. [Internet]. Br. J. Psychiatry Suppl. 2007; 51: s31–s37.Available from: http://dx.doi.org/10.1192/bjp.191.51.s31

Serafini G, Pompili M, Innamorati M, Fusar-Poli P, Akiskal HS, Rihmer Z, et al. Affective temperamental profiles are associated with white matter hyperintensity and suicidal risk in patients with mood disorders. [Internet]. J Affect Disord 2010Available from: http://dx.doi.org/10.1016/j.jad.2010.07.020

Serpa M, Ou Y, Schaufelberger M, Doshi J, Ferreira L, Machado-Vieira R, et al. Neuroanatomical classification in a population-based sample of psychotic major depression and bipolar I disorder with 1 year of diagnostic stability. Biomed Res Int 2014: in press.

Sönmez N, Romm KL, Andreassen OA, Melle I, Røssberg JI. Depressive symptoms in first episode psychosis: a oneyear follow-up study. [Internet]. BMC Psychiatry 2013; 13: 106.Available from: http://dx.doi.org/10.1186/1471-244X-13-106

Takizawa R, Fukuda M, Kawasaki S, Kasai K, Mimura M, Pu S, et al. Neuroimaging-aided differential diagnosis of the depressive state. [Internet]. Neuroimage 2014; 85 Pt 1: 498–507.Available from: http://dx.doi.org/10.1016/j.neuroimage.2013.05.126

Truong W, Minuzzi L, Soares CN, Frey BN, Evans AC, MacQueen GM, et al. Changes in cortical thickness across the lifespan in major depressive disorder. [Internet]. Psychiatry Res 2013; 214: 204–211.Available from: http://dx.doi.org/10.1016/j.pscychresns.2013.09.003

Yu K, Cheung C, Leung M, Li Q, Chua S, McAlonan G. Are Bipolar Disorder and Schizophrenia Neuroanatomically Distinct? An Anatomical Likelihood Meta-analysis. [Internet]. Front Hum Neurosci 2010; 4: 189.Available from: http://dx.doi.org/10.3389/fnhum.2010.00189

Yung AR, Phillips LJ, McGorry PD, McFarlane CA, Francey S, Harrigan S, et al. Prediction of psychosis. A step towards indicated prevention of schizophrenia. Br. J. Psychiatry Suppl. 1998; 172: 14–20.

Zhang T, Koutsouleris N, Meisenzahl E, Davatzikos C. Heterogeneity of Structural Brain Changes in Subtypes of Schizophrenia Revealed Using Magnetic Resonance Imaging Pattern Analysis. Schizophr Bull 2014; in press

Zisook S, Rush AJ, Albala A, Alpert J, Balasubramani GK, Fava M, et al. Factors that differentiate early vs. later onset of major depression disorder. [Internet]. Psychiatry Res 2004; 129: 127–140.Available from: http://dx.doi.org/10.1016/j.psychres.2004.07.004

Figures Legends

Figure 1: Box plot comparison and ANOVAs of SVM decision values including the MD & SZ training database (light grey) and independent validation data consisting of BIP, ARMS and FEP samples (dark grey). Box plots describe decision value distributions in terms of 5%, 25%, 50%, 75% and 95%-CIs. Frequency of SZ-positive diagnosis is measured as percentage of subjects per group labeled as SZ by the classifier (top of the box plot chart). *P* values of post-hoc comparisons in both ANOVAs are provided below and were corrected for multiple comparisons using Tukey's HSD method (SPSS version 20, IBM Inc.).

Figure 2: Box plot and ROC analyses comparing the separability of diagnosis-based vs. age of onset & BrainAGE-based patient groups. A1: Effects of age of onset (left) and BrainAGE (right) on diagnostic separability in MD vs SZ patients. **B1:** Diagnostic separability of diagnostic groups (MD vs SZ, left) vs. separability in cross-nosological patient groups (right) defined by early vs. late disease onset (EO vs LO) and high vs low BrainAGE (Br+ vs. Br-). **B2:** ROC analyses of MRI-based decision scores in the classification of diagnosis (left) and EO/Br+ vs. LO/Br- groups.

Figure 3: Voxel probability map (VPM) of reliable contributions to the MD vs. SZ decision boundary. Voxels with a probability of >50% were overlaid on the single subject MNI template using the MRIcron software package (<u>http://www.mccauslandcenter.sc.edu/mricro/mricron/</u>). Methodological descriptions on how the VPMs were computed can be found in the Supplementary Methods section.

Table 1: Sociodemographic and clinical characteristics of study groups. Descriptive analyses between MD and SZ patient groups were performed by means of chi²-tests for categorical data (†) and *t* tests for continuous data (‡) *t* tests. **Abbreviations:** *BMI* Body-Mass-Index, *BPRS* Brief Psychiatric Rating Scale, *BrainAGE* Brain Age Gap Estimation score, *cig.* cigarettes, *HDRS* Hamilton Depression Rating Scale, *P* significance, *PANSS* Positive and Negative Symptom Scale, *yrs.* years, *SD* standard deviation, *YMRS* Young Mania Rating Scale.

	Training and cross-validation database				Independent validation database				
Sociodemographic & Clinical Variables	MD	SZ	т	Ρ	BIP	FEP	ARMS-E	ARMS-L	
Ν	104	158			35	23	21	68	
N Basel [%] / Munich [%] Mean age at baseline [yrs.]	0 / 100	0 / 100			0 / 100	100 / 0	0 / 100	54 / 46	
(SD)	42.3 (12.0)	30.8 (10.0)	8.1‡	<.001	39 (9.6)	26.8 (6.5)	25.6 (5.6)	24.6 (5.9)	
Sex (male) [%]	50	74	15.8†	<.001	51	74	48	68	
Handedness (right) [%]	95	91	1.5†	ns	85	78	81	90	
BMI [kg/m²] (SD)	24.7 (4.6)	24.3 (4.4)	0.7‡	ns	25.6 (3.6)		21.1 (2.4)	23.0 (3.3) ^m	
Schooling [yrs.] (SD)	10.6 (2.0)	10.6 (2.1)	0.2‡	ns	11.6 (1.6)	10.0 (1.6)	11.5 (2.9)	11.1 (1.5)	
Nicotine [cig./day]	9.7 (13.4)	13.2 (13.7)	-1.96‡	ns	11.4 (13.1)		7.0 (9.9)	7.0 (9.8) ^m	
Alcohol [g/day] Mean age of disease onset	11.2 (21.1)	11.2 (25.5)	-0.01‡	ns	5.6 (20.6)		2.9 (5.6)	7.7 (15.3) ^m	
[yrs.] (SD; Median)	36.5 (12.0)	25.5 (8.0)	8.1‡	<.001	26.1 (9.1)				
Mean illness duration [yrs.] (SD) Current treatment with typical	6.0 (7.8)	4.5 (7.0)	1.4‡	ns	13.9 (9.2)				
antipsychotics [%] Current treatment with atypical	10.0	30.7	17.7†	<.001	0.0	0.0			
antipsychotics [%]	9.0	67.3	86.7†	<.001	40.0	39.1			
lents [mg/d]	43.1 (162.0)	346.3 (373.4)	-8.6‡	<.001	189.2 (322.9)	244.0 (163)			
pressants [%]	73.1	7.9	156.7†	<.001	16.7	21.7	23.8	12.9	
stabilizers [%]	13.0	3.3	8.1†	<.01	66.7	0.0	0.0	0.0	
[%]	7.0	0.0	10.†	<.01	16.7	0.0	0.0	0.0	
Mean BPRS (SD)						52.7 (13.6)		41.9 (10.6) ^b	
Mean PANSS total (SD)		52.6 (29.2)				/	56.8 (14.0)	62.0 (22.2)	
Mean PANSS positive (SD)		11.9 (8.0)					9.86 (2.6)	13.7 (4.5)	
Mean PANSS negative (SD)		15.2 (9.7)					14.9 (6.7)	15.7 (8.5)	
pathology (SD)		25.6 (16.1)					32.0 (7.9)	32.6 (11.1)	
Mean SANS (SD)		45.0 (26.8)				10.0 (5.3)	,	9.5 (5.4) ^b	
Mean HDRS (SD)	21.3 (9.5)	/			9.7 (9.8)	,			
Mean YMRS (SD)	- ()				11.0 (12.0)				
BrainAGE [yrs.] (SD)	4.0 (6.2)	6.0 (6.0)	-2.55‡	<.05	3.8 (6.5)	5.1 (8.5)	-1.5 (7.7)	2.7 (6.8)	

Table 2: Diagnostic performance: The performance of the MRI diagnostic system was evaluated by means of sensitivity (Sens), specificity (Spec), balanced accuracy (BAC), false positive rate (FPR), positive / negative predictive value (PPV / NPV) and Diagnostic Odds Ratio (DOR). These measures were calculated from the confusion matrix containing the number of true positives (TP), false negatives (FN), true negatives (TN) and false positives (FP). **Abbreviations:** *MD-E* early-onset MD, *MD-L* late-onset MD, *SZ-E* early-onset SZ.

Dataset	ΤР	TN	FP	FN	Sens [%]	Spec [%]	BAC [%]	FPR [%]	PPV [%]	NPV [%]	DOR
Cross-validation MD [+1] vs. SZ [-1] (Munich)	83	114	44	21	79.8	72.2	76.0	27.8	65.4	84.4	10.2
Age-of-onset stratified multi- group classifier											
MD-E vs. SZ-E	38	51	24	14	73.1	68.0	70.5	30.2	61.3	78.5	5.8
MD-E vs. MD-L	42	45	7	10	80.8	86.5	83.7	13.5	85.7	81.8	27.0
MD-E vs. SZ-L	31	42	34	21	59.6	55.3	57.4	44.7	47.7	66.7	1.82
SZ-E vs. MD-L	67	52	0	8	89.3	100.0	94.7	0.0	100.0	86.7	
SZ-E vs. SZ-L	47	47	29	28	62.7	61.8	62.3	38.2	61.8	62.7	2.72
MD-L vs. SZ-L	46	64	12	6	88.5	84.2	86.3	15.8	79.3	91.4	40.9
MD vs. SZ (collapsed)	79	104	47	25	76.0	68.9	72.4	31.1	62.7	80.6	6.99

Table 3: Moderators of MRI-based differential diagnosis. Main and two-way interaction effects of diagnosis, early vs. late disease onset and low vs. high BrainAGE on diagnostic scores were analyzed using univariate linear modelling in SPSS (version 20, IBM Inc.). **Abbreviations**: *E* estimate (mean difference or marginal mean), SE standard error, *95%-CI Low / Up* 95%-confidence interval with lower and upper bounds, *F* F-statistic.

Dataset	Е	SE	95%-Cl Low / Up	F	Ρ
Main effects					
MD vs. SZ [mean difference]	1.49	0.13	1.24 / 1.75	134.0	<.001
MD [marginal mean]	0.89	0.10	0.70 / 1.09		
SZ [marginal mean]	-0.60	0.08	-0.76 / -0.44		
Early vs. Late-Onset	-1.03	0.13	-1.29 / -0.78	62.6	<.001
Early-Onset	-0.37	0.09	-0.55 / -0.19		
Late-Onset	0.66	0.09	0.48 / 0.85		
Low vs. High BrainAGE	0.41	0.13	0.15 / 0.67	9.8	.002
Low BrainAGE	0.35	0.09	0.17 / 0.53		
High BrainAGE	-0.06	0.09	-0.24 / 0.13		
Two-way interaction effects					
MD vs. SZ×Early vs. Late-Onset				8.3	0.004
MD×Early-Onset	0.19	0.14	-0.09 / 0.47		
MD×Late-Onset	1.59	0.14	1.32 / 1.88		
SZ×Early-Onset	-0.93	0.12	-1.16 / -0.70		
SZ×Late-Onset	-0.27	0.12	-0.50 / -0.04		
MD vs. SZ×Low vs. High BrainAGE				3.7	0.056
MD×Low BrainAGE	1.22	0.14	0.95 / 1.50		
MD×High BrainAGE	0.56	0.12	0.29 / 0.84		
SZ×Low BrainAGE	-0.52	0.12	-0.74 / -0.29		
SZ×High BrainAGE	-0.68	0.12	-0.91 / -0.44		