OPEN

Citation: Transl Psychiatry (2017) 7, e1087; doi:10.1038/tp.2017.59

www.nature.com/tp

ORIGINAL ARTICLE

Two subgroups of antipsychotic-naive, first-episode schizophrenia patients identified with a Gaussian mixture model on cognition and electrophysiology

N Bak[1,2], BH Ebdrup[1,2], B Oranje[1,2,3,4], B Fagerlund[1,2], MH Jensen[1,2], SW Düring[1,2], MØ Nielsen[1,2,4], BY Glenthøj[1,2,4] and LK Hansen[5]

Deficits in information processing and cognition are among the most robust findings in schizophrenia patients. Previous efforts to translate group-level deficits into clinically relevant and individualized information have, however, been non-successful, which is possibly explained by biologically different disease subgroups. We applied machine learning algorithms on measures of electrophysiology and cognition to identify potential subgroups of schizophrenia. Next, we explored subgroup differences regarding treatment response. Sixty-six antipsychotic-naive first-episode schizophrenia patients and sixty-five healthy controls underwent extensive electrophysiological and neurocognitive test batteries. Patients were assessed on the Positive and Negative Syndrome Scale (PANSS) before and after 6 weeks of monotherapy with the relatively selective D_2 receptor antagonist, amisulpride (280.3 ± 159 mg per day). A reduced principal component space based on 19 electrophysiological variables and 26 cognitive variables was used as input for a Gaussian mixture model to identify subgroups of patients. With support vector machines, we explored the relation between PANSS subscores and the identified subgroups. We identified two statistically distinct subgroups of patients. We found no significant baseline psychopathological differences between these subgroups, but the effect of treatment in the groups was predicted with an accuracy of 74.3% ($P = 0.003$). In conclusion, electrophysiology and cognition data may be used to classify subgroups of schizophrenia patients. The two distinct subgroups, which we identified, were psychopathologically inseparable before treatment, yet their response to dopaminergic blockade was predicted with significant accuracy. This proof of principle encourages further endeavors to apply data-driven, multivariate and multimodal models to facilitate progress from symptom-based psychiatry toward individualized treatment regimens.

Translational Psychiatry (2017) **7,** e1087; doi:10.1038/tp.2017.59; published online 11 April 2017

INTRODUCTION

Despite great progress in the characterization of specific functional and structural deficits associated with schizophrenia,[1] the clinical diagnosis according to Diagnostic and Statistical Manual of Mental Disorder and International Classification of Diseases classifications still relies entirely on symptomatology rather than on objective, biological measures. A wide range of group differences ranging from genetic variation over brain structural and functional changes to deficits in early information processing and cognition have been reported in schizophrenia patients when compared with groups of matched healthy controls.[2,3]

Although many of these neurobiological findings have been reproduced at a meta-analytical level,[4] the great heterogeneity between studies suggests that schizophrenia is composed of several biologically different subgroups.[5] The presence of subgroups implies that any given clinical sample of schizophrenia patients comprise multiple subgroups with different patterns of biological disturbances. Such 'mixed' patient samples may hinder identification of effective, individually targeted clinical management. Moreover, the presence of subgroups impairs the development of novel treatment strategies, as potentially important clinical effects may be masked by unknown variance in the clinical sample of schizophrenia patients.[6,7] Finally, observations of deficits in schizophrenia patients are often confounded by effects of chronicity, substance abuse and previous treatment exposure. Subgroups of schizophrenia patients based on biologically valid, objectively measured markers have not yet been identified.

Deficits in early information processing and cognition are among the most robust findings in schizophrenia.[8,9] In early information processing, sensory input is filtered so that only relevant input reaches the conscious, cortical level. Early information processing can be estimated with electrophysiology using several paradigms, such as P50 suppression,[10] the pre-pulse inhibition (PPI) of the startle response[11] and mismatch negativity (MMN).[12] Despite the presence of psychotic symptoms, cognition can be reliably assessed by neurocognitive tests, and pronounced cognitive deficits constitute core deficits in schizophrenia patients.[9]

Handling of complex data sets comprising multiple variables from several modalities requires the application of novel

[1]Centre for Neuropsychiatric Schizophrenia Research, Mental Health Services Glostrup, University of Copenhagen, Copenhagen, Denmark; [2]Centre for Clinical Intervention and Neuropsychiatric Schizophrenia Research, Mental Health Glostrup, University of Copenhagen, Copenhagen, Denmark; [3]Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands; [4]Faculty of Health and Medical Sciences, Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark and [5]Cognitive Systems, DTU Compute, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark. Correspondence: Dr N Bak, Center for Neuropsychiatric Schizophrenia Research or Center for Clinical Intervention and Neuropsychiatric Schizophrenia Research, Mental Health Services Centre Glostrup, University of Copenhagen, Ndr. Ringvej, DK-2600 Copenhagen, Denmark.
E-mail: nikolaj@cnsr.dk
Received 27 October 2016; revised 31 January 2017; accepted 19 February 2017

algorithms. 'Machine learning' enables identification of patterns in complex data, which cannot be modeled by means of more classical statistical methods. Identified patterns can potentially be used to predict future data or outcomes. In essence, the underlying hypothesis for applying machine learning is that a measurable structure in data exists. Machine learning techniques can be divided into supervised and unsupervised learning. In supervised learning, the 'label', which can be a class or an outcome, has to be known for each observation, and the supervised algorithm identifies the underlying structure in the data, which is associated with this label. The identified data structure can then be applied to predict future, independent observations for which the label is unknown. Unsupervised learning, on the other hand, is used to identify new and unknown structures in data. Specifically, unsupervised learning is useful in order to subgroup data if the labels, that is, classes or outcomes are unknown. To ensure generalizable results, data-driven machine learning techniques require valid unbiased methods such as cross-validation.[13]

Previous machine learning studies in schizophrenia patients have reported on subgroup structures (referred to as 'profiles' or 'biotypes') based on electrophysiology (for example, Van Tricht *et al.*,[14] Turetsky *et al.*,[15] Hall *et al.*,[16] John *et al.*[17]), cognition (for example, Geisler *et al.*,[18] Marder *et al.*[19]) or both,[20] but first-episode, antipsychotic-naive patients have not been investigated.

The primary aim of the study was to identify potential subgroups of schizophrenia on measures of electrophysiology and cognition. For this, we applied an unsupervised machine learning algorithm in an unbiased cross-validation scheme for evaluation. The secondary aim was to investigate whether these subgroups were related to treatment response. For this, we applied a supervised machine learning algorithm including measures of psychopathology before and after treatment with a relatively selective dopamine D_2 antagonist, amisulpride.

MATERIALS AND METHODS

The study was conducted in accordance with the Declaration of Helsinki II, and approved by the Danish National Committee on Biomedical Research Ethics (H-D-2008-088). Clinical trials identifier: NCT01154829. All participants approved participation by signing informed consent.

Participants

We included data from a multimodal first-episode study of antipsychotic-naive schizophrenia patients and healthy controls. The groups were matched on age, gender and socioeconomic status. Patients were excluded if they had a current diagnosis of drug dependency or were treated with antidepressants within the last month or during the study period. In addition, patients were asked to refrain from taking benzodiazepines the evening prior to a test day. At baseline, subjects underwent extensive assessments in multiple modalities, including cognitive and electrophysiological test batteries. Part of the data overlap with previous uni-modal publications on electrophysiology[21–23] as well as publications on functional and structural magnetic resonance imaging,[24–27] oxidative stress[28] and single-photon emission computed tomography data.[29]

After baseline assessments, the patients underwent 6 weeks of antipsychotic monotherapy with the relatively selective D_2 receptor antagonist, amisulpride. Symptom severity in patients was measured with the Positive and Negative Syndrome Scale (PANSS).[30] The subgrouping analyses in the current study included data on electrophysiology (baseline), neurocognition (baseline), PANSS (baseline and follow-up) and amisulpride dose (follow-up).

Participants for analyses and imputation procedure

We included 69 antipsychotic-naive schizophrenia patients and 67 healthy controls. Three patients and two controls did not undergo the electrophysiological and cognitive test batteries. These five subjects were excluded. In addition, 39 subjects had one or more missing variables in electrophysiology or cognition data leaving 92 complete cases (41 patients and 51 controls). On the subset of subjects, who had one or more missing variables, we performed an imputation procedure.[31] After imputation, the data set comprised 97 subjects (44 patients and 53 controls). The subgrouping analyses were performed on these 44 patients. Subsequent supervised machine learning analyses were based on the patients, who had PANSS baseline ($N=43$), follow-up ($N=36$) or both ($N=35$) (due to missing data for one PANSS baseline assessment).

A short description of acquisition and processing of cognitive and electrophysiological data is provided below. Details of the procedures are presented in Supplementary Material.

Cognition

All participants were examined with a comprehensive neurocognitive test battery. The neurocognitive battery took ~2 h to complete and participants were allowed short breaks between tests. All tests were administered by research staff trained in standardized administration and scoring of the battery. Outcome variables from the following neurocognitive tasks were included from: Danish Adult Reading Test,[32] Wechsler Adult Intelligence Scale III,[33] Brief Assessment of Cognition in Schizophrenia[34] and Cambridge Neuropsychological Test Automated Battery.[35] Please, also see Supplementary Material.

Electrophysiology

Participants were examined with the Copenhagen Psychophysiology Test Battery.[21,22,36] The Copenhagen Psychophysiology Test Battery includes PPI, P50 suppression, MMN and selective attention paradigms in a fixed order. Tobacco use was not allowed 1 h before testing in order to avoid acute and/or withdrawal effects of nicotine.[37] Participants were instructed to refrain from intake of caffeinated beverages at the day prior to testing. Testing was performed in a separate room with a sound level < 40 dB situated adjacent to the control room. Participants were seated in a comfortable armchair and were instructed to keep movements to a minimum, keep their eyes fixed on a spot on the wall directly in front of them and stay awake. Auditory stimuli were presented by a computer running Presentation (Neurobehavioral Systems, Albany, NY, USA) software (soundcard: Creative Sound Blaster 5.1, 2008 Creative Technology, Singapore). Stimuli were presented binaurally through stereo insert earphones (Eartone ABR, 1996–2008 Interacoustics A/S, Assens, Denmark; and C and H Distributors, Milwaukee, WI, USA). The soft- and hardware audio settings were calibrated with an artificial ear (Brüel and Kjær, type 2133, Odin Metrology, Thousand Oaks, CA, USA).

Electroencephalography as well as electromyography recordings were performed using BioSemi hardware (Amsterdam, The Netherlands) using a cap with 64 active electrodes. For PPI, the eye-blink component was measured by recording electromyography activity from the right musculus orbicularis oculi with two electrodes. The first of these was aligned with the pupil, the other positioned just laterally. BESA software (version 5.2.4, MEGIS Software, Gräfelfing, Germany) was used for further processing of the data. A background noise consisting of 70 dB white noise was used in all paradigms. Please, also see Supplementary Material.

Statistical analyses

Analyses were performed in MATLAB (The MathWorks, Natick, MA, USA) using the Statistics and Bioinformatics Toolbox Release 2013a.

Identification of subgroups. In order to identify potential subgroups of schizophrenia, we applied unsupervised machine learning analyses on 26 cognitive and 19 electrophysiological variables, denoted 'features' (Figure 1).

A probabilistic principal component analysis based on the 45 standardized (scaled to a mean of zero and unit variance) features from all subjects with complete data sets ($N_{complete}=92$) was performed, identifying the 45-dimensional (D) principal component space. The whole sample was used in this step in order to be able to compare the healthy subjects with patients in a space unbiased toward patients or controls and avoid variance inflation.[38] The 45 components were sorted so that the first component explained most variance in the data. Each of the subsequent components explained most of the remaining variance.

To identify the statistically distinct PCA subspace of the 45D principal component space, we applied the Akaike information criterion (AIC).[39] AIC is an analytic and asymptotically unbiased estimator of the cross-validation deviance (mean log-likelihood of test data). The asymptotic estimator is appropriate, as we apply this step for the combined cohort of patients and

Feature	Schizophrenia 1	Schizophrenia 2	Controls	PC1	PC2	PC3	PC4
Premorbid IQ (DART)	23.65(10.37)	19.61(7.63)	23.06(6.81)				
Verbal IQ a	106.28(21.17)	98.72(13.93)	113.50(14.11)				
Performance IQ a	103.64(13.97)	97.28(13.44)	108.75(13.32)				
Full scale IQ a	105.72(18.12)	98.11(12.18)	112.50(12.55)				
List learning b	52.46(7.18)	51.06(13.37)	55.23(8.32)				
Digit Sequencing Task b	21.38(4.08)	20.11(3.92)	22.89(3.65)				
Token Motor Task b	68.23(10.81)	72.78(15.19)	76.08(13.30)				
Verbal Fluency ("supermarket") b	23.88(6.93)	26.22(9.08)	31.25(8.00)				
Verbal Fluency ("F") b	13.46(3.67)	12.22(5.41)	15.66(4.22)				
Verbal Fluency ("S") b	14.46(4.59)	12.89(5.86)	16.28(4.76)				
Symbol coding b	59.69(10.13)	54.50(15.56)	65.83(11.66)				
Tower of London b	18.73(1.93)	18.56(2.48)	19.51(2.00)				
Spatial Span c	7.19(1.41)	6.94(1.30)	7.51(1.25)				
Spatial Working Memory (strategy) c	26.69(6.42)	29.56(6.31)	24.36(5.37)				
Spatial Working Memory (total errors) c	11.54(14.92)	15.33(19.44)	7.13(8.43)				
Stockings of Cambridge (problems solved) c	9.65(1.83)	9.06(1.51)	10.21(1.35)				
Stockings of Cambridge (initial thinking time) c	9,470(8,003)	9,872(4,810)	12,472(7,545)				
Intra-Extra Dimensional Set Shift (stages) c	9.00(0.00)	7.89(0.96)	8.87(0.48)				
Intra-Extra Dimensional Set Shift (errors) c	11.42(5.09)	39.94(20.78)	14.60(13.14)				
Intra-Extra Dimensional Set Shift (EDS errors) c	3.38(3.07)	17.89(11.22)	6.25(7.68)				
Reaction Time (simple reaction) c	337.82(56.79)	322.71(35.20)	304.51(33.77)				
Reaction Time (simple movement) c	456.72(127.08)	455.81(113.01)	455.90(141.34)				
Reaction Time (choice reaction) c	410.97(78.84)	381.42(60.91)	356.42(66.47)				
Reaction Time (choice movement) c	402.50(111.79)	409.03(95.58)	413.05(120.22)				
Rapid Visual Processing (A´, 3-5-7) c	0.98(0.02)	0.97(0.02)	0.99(0.01)				
Rapid Visual Processing (A´, 3-5-7, 2-4-6) c	0.96(0.03)	0.95(0.04)	0.97(0.02)				
P50 c-stimulus Amplitude	1.29(0.74)	2.03(1.40)	1.70(0.99)				
P50 t-stimulus Amplitude	0.43(0.49)	0.74(0.88)	0.55(0.55)				
P50 t/c ratio	0.36(0.49)	0.40(0.43)	0.30(0.29)				
P50 c-stimulus Latency	56.72(9.78)	58.44(10.90)	58.88(8.88)				
PPI 85dB 120ms	53.50(40.66)	50.28(29.65)	58.89(37.25)				
PPI 85dB 60ms	57.00(28.00)	50.89(17.38)	58.13(33.01)				
PPI 76dB 120ms	37.77(40.11)	29.56(32.16)	48.87(34.64)				
PPI 76dB 60ms	37.04(39.90)	32.00(27.71)	36.89(48.17)				
PPI Pulse alone	106.19(73.59)	208.06(141.25)	138.28(127.10)				
PPI 85dB 120ms amplitude	40.38(35.81)	93.17(92.69)	56.72(93.12)				
PPI 85dB 60ms Amplitude	39.19(31.92)	92.28(66.52)	55.92(86.18)				
PPI 76dB 120ms Amplitude	58.19(43.75)	135.61(114.29)	74.25(111.11)				
PPI 76dB 60ms Amplitude	59.35(38.96)	153.39(129.16)	83.89(114.38)				
MMN Frequency deviant FCZ	-2.74(1.25)	-2.83(1.42)	-2.49(1.13)				
MMN Duration deviant FCZ	-4.27(2.10)	-3.75(1.40)	-4.01(1.60)				
MMN Frequency and duration deviant FCZ	-3.82(1.72)	-3.86(1.80)	-4.16(1.43)				
MMN Frequency deviant Latency	131.20(38.56)	131.78(43.07)	140.83(41.59)				
MMN Duration deviant Latency	194.33(36.10)	184.44(46.73)	187.55(34.59)				
MMN Frequency and Duration deviant Latency	142.88(37.71)	122.00(32.36)	146.49(40.72)				
Principal Component 1	0.0155(0.1007)	0.1023(0.1127)	-0.0419(0.0788)				
Principal Component 2	-0.0273(0.0613)	0.0271(0.1088)	-0.0008(0.1159)				
Principal Component 3	-0.0496(0.0968)	0.0491(0.0683)	0.0089(0.1087)				
Principal Component 4	0.0418(0.0796)	-0.0697(0.1594)	0.0072(0.0774)				

Figure 1. Cognitive and early information processing data. Mean (s.d.) for the two subgroups and controls. The weights for each feature on the four principal components (PC1-4) are shown as bar charts. Cognitive features from Danish Adult Reading Test (DART); [a]Wechsler Adult Intelligence Scale III; [b]Brief Assessment of Cognition in Schizophrenia; [c]Cambridge Neuropsychological Test Automated Battery. Electrophysiological features from P50 suppression (P50); pre-pulse inhibition (PPI) of the startle effect; mismatch negativity (MMN).

controls.[40] AIC was calculated for dimensions D = 2, 3, 4, …, 45 to determine the number of principal components, which best described the data. AIC identified an optimal number of four components, which were then used in the further analyses (Figure 2).

On the subset of subjects ($n = 39$), who had one or more missing variables, we performed an imputation procedure.[31] This method provides an estimate of the imputation error and variance, by simulating missing values in subjects with complete data sets and weighing these by similarity to subjects with missing values. Subjects with missing values and an estimated error below 0.05 (s.d. = 0.1 in all four PCs) in the 4D principal component space were imputed and included in the analyses, while subjects with higher estimated imputation errors were excluded. After the imputation procedure, the analyzed data set comprised 97 subjects (44 patients and 53 controls).

A Gaussian mixture model (GMM) was used to identify group structure (patients only) in the 4D principal component space, using leave-one-out cross-validation[41] based on subjects to estimate the optimal number of groups. The GMM is an unsupervised clustering algorithm that fits data as a number of 'structures' in the Gaussian mixtures. Each structure in the Gaussian mixture represents a subgroup in the data. The structures should not be confused with the principal components that determine the dimensions in which the subgroups are found. The GMM was run on patients alone ($N = 44$), as we aimed to identify subgroups within patients rather than classify patients from controls. To provide an unbiased

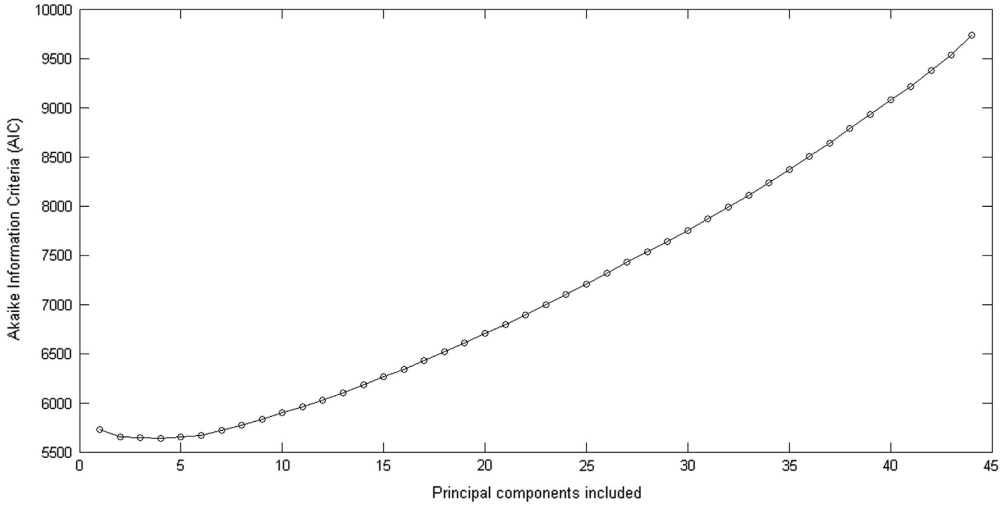

Figure 2. AIC indicating likelihood for number of principal component dimensions. The minimal AIC value is attained at D = 4, models based on larger or smaller dimensions provide poorer fits to the test data. AIC, Akaike information criteria.

estimate of test error, the mean negative log likelihood was estimated for a range of groups, $K = 1,\ldots, 10$ in each cross-validation fold.

To further characterize the identified subgroups, we applied χ^2-test for gender and use of alcohol, tobacco, cannabis or benzodiazepines. Patients fulfilling criteria 'never tried' or 'tried a few times' regarding each item were defined as 'non-users'. Patients fulfilling criteria 'regular use' or 'abuse' were defined as 'users'. We applied independent *t*-tests to test for differences in age, PANSS subscores and amisulpride dose to explore psychopathological subgroup differences. Specifically, we included amisulpride dose at follow-up, positive, negative, and general subscores from baseline, follow-up and change scores (baseline − follow-up; Table 1).

Prediction of treatment response. In order to investigate whether the two identified subgroups were related to treatment response, we used PANSS subscores (positive, negative, general) from baseline, follow-up (6 weeks) and changes in PANSS subscores (baseline − follow-up), which represented the treatment effect. For this, we performed three separate linear support vector machines (SVMs)[42] based on the PANSS subscores (baseline, follow-up or changes; Table 1) with the '*C*' parameter set to 100. The SVM algorithm classifies the data regarding label by finding the multidimensional hyperplane with the largest margin that separates the labels in the input space. The measure of 'accuracy' estimates whether the PANSS scores can predict the identified subgroups. This can be considered an external validation of the subgroups and thereby indicate whether the subgroups have potential clinical relevance. Leave-one-out cross-validation was used to obtain an unbiased estimator of the predictability and the strength of the evidence was evaluated by permutation test. One thousand permutations of group labels were performed, refitting the SVM to randomized labels and testing accuracy in each.

Code availability

Computer code available upon request.

RESULTS

Patients ($N = 44$) and controls ($N = 53$) were well matched on age ($t = 0.222$, $P = 0.825$) and gender ($\chi^2 < 0.001$, $P = 0.983$). At follow-up, the mean amisulpride dose was 280.3 (159.6) mg per day.

Identification of subgroups

The first principal component identified with probabilistic principal component analysis loaded on all cognitive features thus represents the overall cognitive capacity (also referred to as 'Spearman's G').[43] The second principal component loaded primarily on electrophysiological features, specifically the features from PPI of the startle reflex. The third and fourth principal components involved features from both modalities. More specifically, the third component principal component loaded on executive functions and reaction time from the neurocognitive modality and percent inhibition in PPI and latency in MMN. The fourth component loaded on intra-extra dimensional set shifting and all variables from P50 suppression. See Figure 1 for weights of the four principal components.

As indicated by the lowest negative log-likelihood score, the cross-validation of the GMM indicated that a separation into two subgroups of patients provided the most generalizable model, hence the model that best-fit the test data (Figure 3). 'Subgroup 1' consisted of 26 patients and 'Subgroup 2' consisted of 18 patients. Data on the two patient subgroups and the controls are displayed in Figure 1.

Subgroup 1 had, compared to subgroup 2 lower values in PC 1 (representing higher cognitive capacity), lower values in PC 2 (indicating lower PPI and higher PPI amplitudes), and lower values in PC 3 (indicating increased reaction times and increased MMN latency). Conversely, subgroup 1 had higher values in PC 4 (indicating better performance in set-shifting and higher amplitudes in P50 suppression) compared to subgroup 2.

We found no significant differences in between the two patient subgroups with regard to age, gender, PANSS subscores, amisulpride dose, or use of alcohol, tobacco, cannabis or benzodiazepines (Table 1).

Prediction of treatment response

SVM analyses of the predictive value of psychopathology showed that PANSS baseline subscores reached an accuracy of 67.4% (Table 1), which were significantly ($P = 0.017$) higher than 58.1% in 'baseline accuracy'. 'Baseline accuracy' is defined as the accuracy for the simplest classification rule, that is, predicting the majority class for each observation. With PANSS follow-up subscores, SVM reached 50% compared to baseline accuracy of 55.6%. Using changes in PANSS subscores (baseline − follow-up), that is, the treatment effect, the SVM analysis reached an accuracy of 74.3%, significantly ($P = 0.003$) higher than baseline accuracy (54.2%; Figure 4).

DISCUSSION

Our multivariate analyses of multimodal non-biased data in a sample of first-episode antipsychotic-naive schizophrenia patients

Table 1. Demographic and clinical data on controls and the two identified subgroups of patients: PANSS data, baseline, follow-up and difference

Measure	*Group*	N	*Mean (s.d.)*	P	*SVM Acc*	*SVM p*
	Controls	53	30/23			
Gender (m/f)	1	26	17/9	0.168	—	
	2	18	8/10			
	Controls	53	24.79 (5.98)			
Age	1	26	23.46 (4.55)	0.197	—	
	2	18	26.06 (7.39)			
PANSS positive, B	1	25	19.92 (3.93)	0.727		
	2	18	20.39 (4.80)			
PANSS negative, B	1	25	20.32 (6.48)	0.545	0.674	**0.017**
	2	18	21.67 (7.99)			
PANSS general, B	1	25	41.64 (7.84)	0.380		
	2	18	39.22 (10.03)			
PANSS positive, Fu	1	20	13.10 (3.75)	0.091		
	2	16	15.31 (3.84)			
PANSS negative, Fu	1	20	18.15 (6.05)	0.916	0.500	
	2	16	17.94 (5.84)			
PANSS general, Fu	1	20	29.95 (8.13)	0.753		
	2	16	30.81 (8.10)			
PANSS positive, B-Fu	1	19	6.42 (4.90)	0.401		
	2	16	5.13 (3.95)			
PANSS negative, B-Fu	1	19	−0.37 (5.20)	0.056	0.743	**0.003**
	2	16	3.31 (5.79)			
PANSS general, B-Fu	1	19	10.47 (6.82)	0.187		
	2	16	7.13 (7.88)			
Amisulpride doses, Fu	1	17	276.47 (146.97)	0.890	—	
	2	16	284.38 (176.75)			
Alcohol (users)	1	26	21	0.506	—	
	2	18	13			
Tobacco (users)	1	26	17	0.307	—	
	2	18	9			
Cannabis (users)	1	26	9	0.189	—	
	2	18	3			
Benzodiazepines (users)	1	26	0	—	—	
	2	18	0			

Abbreviations: B, baseline; Fu, follow-up; PANSS, Positive and Negative Syndrome Scale; SVM, support vector machine. *P*-value for the *t*-test for the difference between schizophrenia subgroups (χ^2-test for gender and substance use). Support vector machine accuracy (SVM Acc) and *P*-value from the permutation test for significance of the SVM (SVM p). Significant *P*-values ($P < 0.05$) are in bold.

support the notion that the clinical diagnosis of schizophrenia encompasses biologically separable subgroups. Specifically, our data suggest the presence of two distinct subgroups of schizophrenia with regard to early information processing and higher cognitive functions. Univariate analyses showed no significant differences in demographic or clinical data at baseline or at follow-up. However, the SVM analyses showed that subgroup status can significantly be associated with the treatment effect after 6 weeks. Moreover, subgroups status was also associated with baseline PANSS subscores.

The change in PANSS scores reflects the overall clinical treatment response on positive, negative and general symptoms after 6 weeks of amisulpride monotherapy. Specifically, the accuracy of 74.3% was driven primarily by differences in PANSS negative and general scores (Table 1). Interestingly, subgroup 2 tended to improve in negative symptoms, whereas negative symptoms in subgroup 1 did not change after treatment. Amisulpride is the only antipsychotic compound that is approved for treatment of negative symptoms, yet the results from subsequent clinical studies have been inconsistent (for meta-analysis see Leucht[44]). A reduction in negative symptoms of three points on the PANSS scale represents the 'clinical gain'. The modest clinical gain compared with the 'high cost' of the present data acquisition and analyses do not support direct implementation of electrophysiological- and cognitive examinations as part of the routine work-up to predict treatment response to amisulpride. Nevertheless, the previously observed variable clinical trajectories in the response of amisulpride on negative symptoms[45] may be explained by subgroups as identified in the current study.

To our knowledge, this is the first study applying both cognitive and electrophysiological measures in order to subgroup first-episode schizophrenia patients. One previous study identified

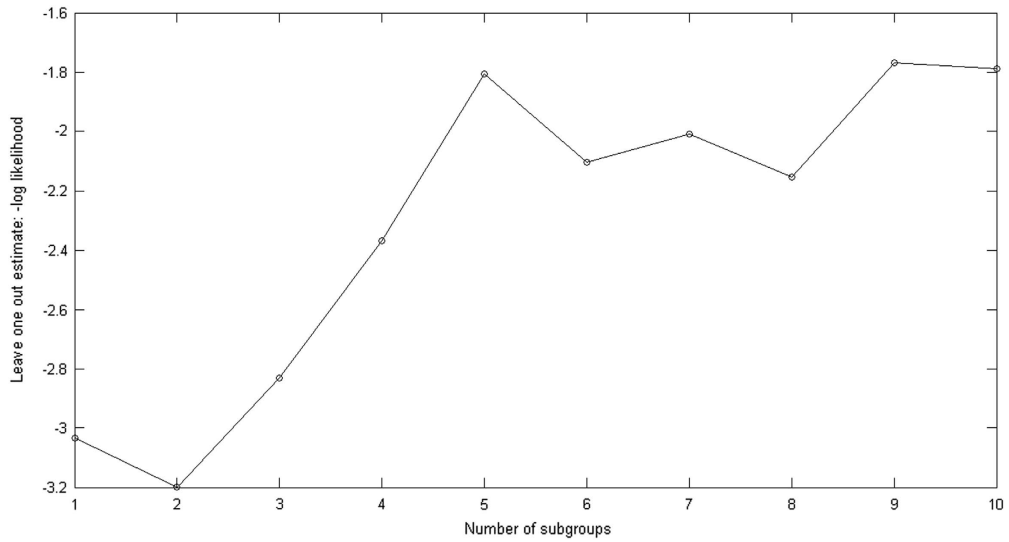

Figure 3. Leave-one-out estimate of test error (negative log likelihood). Calculated for model dimension, that is, number of subgroups = 1,2, …, 10. Lowest value at model dimension = 2, indicating that a model with two subgroups best-fit the test data. Values on *Y* axis are arbitrary.

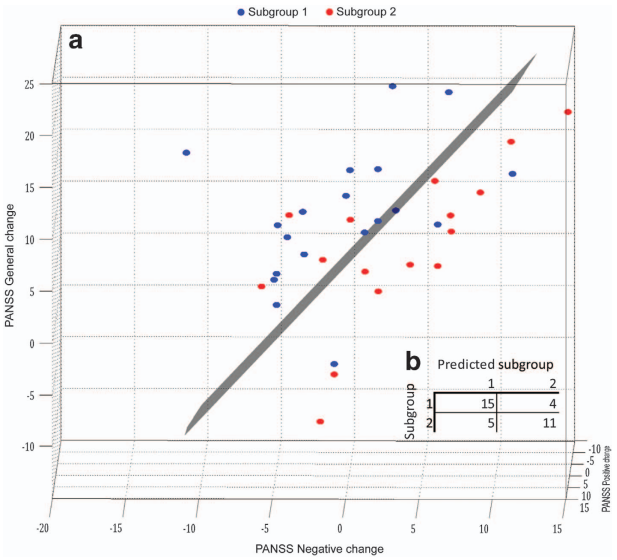

Figure 4. (**a**) The two subgroups in 3D PANSS change subscores with SVM decision surface. The decision surface is determined by PANSS negative and PANSS general subscore changes while PANSS positive change scores seems to have limited to no influence. (**b**) Confusion matrix presenting actual vs predicted group. PANSS, Positive and Negative Syndrome Scale; SVM, support vector machines.

three 'biotypes' across schizophrenia, bipolar disorder and schizoaffective disorder using the same type of measures.[20] Another study identified five subgroups of chronic schizophrenia patients based on aggregate cognitive scores of seven *a priori* defined domains.[46] These studies investigated different diagnostic categories and included chronic, antipsychotic-treated patients and have in essence addressed the validity of the current diagnostic criteria.[47] To our knowledge, the current study is also the first to demonstrate that more homogeneous samples of antipsychotic-naive, first-episode schizophrenia patients display both biologically and clinically relevant subgroup structures.

The two subgroups of schizophrenia identified in this study were identified among 44 first-episode schizophrenia patients. Although, effects of chronicity or previous antipsychotic exposure are ruled out, the findings call for replication in larger, independent samples. Unsupervised algorithms as used in this study usually require more observations than supervised algorithms to reach stable models.[48] In machine learning terms, our number of observations is relatively low, and stresses the importance that the optimal number of latent variables and the optimal number of subgroups is identified properly, as recommended in a recent comprehensive review.[49] The approach in the present study follows these guidelines. Another risk with unsupervised machine learning is that the structures/subgroups identified with might be associated with irrelevant traits, for example, head size, rather than the relevant traits, for example, clinical outcome. Our identified subgroups seem clinically relevant. Although supervised learning can generalize known statistical relations to new data, we have demonstrated the explorative potential of unsupervised learning to identify new generalizable statistical structure.

In the present study, we have only included data from two commonly used modalities in schizophrenia research (cognition and electrophysiology). Clearly, inclusion of more independent modalities may refine the structure and number of subgroups of schizophrenia patients. For example, functional magnetic resonance imaging has been shown to contain information to identify subgroups in schizophrenia.[50] These modalities could include genetic variability, magnetic resonance imaging and *in vivo* receptor imaging with positron emission tomography or single-photon emission computed tomography. Moreover, 'treatment response' in the context of this study is limited to 6 weeks, and the individual long-term course of illness may be subject to other trajectories. Finally, prediction of treatment response to other antipsychotic compounds than amisulpride, cannot be inferred from this study.

Overall, this proof of principle study supports the presence of biological, clinically relevant subgroups of schizophrenia and implies that stratification of patients is required to recognize specific treatment needs in individual subgroups. The current results encourage further endeavors to apply data-driven, multivariate and multimodal models to facilitate progress from

symptom-based psychiatry toward individualized treatment regimens.

CONFLICT OF INTEREST

BHE is supported by the University of Copenhagen/Mental Health Services, Capital Region of Denmark, and has received lecture fees from and/or is part of Advisory Boards of Bristol-Myers Squibb, Eli Lilly and Company, Janssen-Cilag, Otsuka Pharma Scandinavia and Takeda Pharmaceutical Company. BYG is the leader of a Lundbeck Foundation Center of Excellence for Clinical Intervention and Neuropsychiatric Schizophrenia Research (CINS), which is partially financed by an independent grant from the Lundbeck Foundation based on international review and partially financed by the Mental Health Services in the Capital Region of Denmark, the University of Copenhagen and other foundations. All grants are the property of the Mental Health Services in the Capital Region of Denmark and administrated by them. The remaining authors declare no conflict of interest.

ACKNOWLEDGMENTS

This work was supported by the Mental Health Services, The Capital Region, Denmark; and The Lundbeck Foundation (grant number R155-2013-16337). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1 Tandon R, Keshavan MS, Nasrallah HA. Schizophrenia, "Just the Facts": what we know in 2008 part 1: overview. *Schizophr Res* 2008; **100**: 4–19.
2 Stober G, Ben-Shachar D, Cardon M, Falkai P, Fonteh AN, Gawlik M *et al.* Schizophrenia: from the brain to peripheral markers. A consensus paper of the WFSBP task force on biological markers. *World J Biol Psychiatry* 2009; **10**: 127–155.
3 Andreasen NC. A unitary model of schizophrenia: Bleuler's "fragmented phrene" as schizencephaly. *Arch Gen Psychiatry* 1999; **56**: 781–787.
4 Haijma SV, Van Haren N, Cahn W, Koolschijn PCMP, Hulshoff Pol HE, Kahn RS. Brain volumes in schizophrenia: a meta-analysis in over 18 000 subjects. *Schizophr Bull* 2013; **39**: 1129–1138.
5 Howes OD, Kapur S. A neurobiological hypothesis for the classification of schizophrenia: type A (hyperdopaminergic) and type B (normodopaminergic). *Br J Psychiatry* 2014; **205**: 1–3.
6 Lewis DA, Gonzalez-Burgos G. Pathophysiologically based treatment interventions in schizophrenia. *Nat Med* 2006; **12**: 1016–1022.
7 Insel TR, Scolnick EM. Cure therapeutics and strategic prevention: raising the bar for mental health research. *Mol Psychiatry* 2006; **11**: 11–17.
8 Thibaut F, Boutros N, Jarema M, Oranje B, Hasan A, Daskalakis ZJ *et al.* Consensus paper of the WFSBP Task Force on Biological Markers: criteria for biomarkers and endophenotypes of schizophrenia part I: neurophysiology. *World J Biol Psychiatry* 2015; **16**: 280–290.
9 Schaefer J, Giangrande E, Weinberger DR, Dickinson D. The global cognitive impairment in schizophrenia: consistent over decades and around the world. *Schizophr Res* 2013; **150**: 42–50.
10 Adler LE, Pachtman E, Franks RD, Pecevich M, Waldo MC, Freedman R. Neurophysiological evidence for a defect in neuronal mechanisms involved in sensory gating in schizophrenia. *Biol Psychiatry* 1982; **17**: 639–654.
11 Braff DL. Sensorimotor gating and schizophrenia. *Arch Gen Psychiatry* 1990; **47**: 181.
12 Shelley AM, Ward PB, Catts SV, Michie PT, Andrews S, McConaghy N. Mismatch negativity: an index of a preattentive processing deficit in schizophrenia. *Biol Psychiatry* 1991; **30**: 1059–1062.
13 Hansen LK, Larsen J. Unsupervised Learning and Generalization. *Proceedings of International Conference on Neural Networks (ICNN'96)* [Internet]. IEEE, 1996, pp 25–30.
14 Van Tricht MJ, Bour LJ, Koelman JHTM, Derks EM, Braff DL, de Wilde OM *et al.* Qualitative and quantitative aspects of information processing in first psychosis: latent class analyses in patients, at-risk subjects, and controls. *Psychophysiology* 2015; **52**: 585–593.
15 Turetsky BI, Bilker WB, Siegel SJ, Kohler CG, Gur RE. Profile of auditory information-processing deficits in schizophrenia. *Psychiatry Res* 2009; **165**: 27–37.
16 Hall M-H, Smoller JW, Cook NR, Schulze K, Hyoun Lee P, Taylor G *et al.* Patterns of deficits in brain function in bipolar disorder and schizophrenia: a cluster analytic study. *Psychiatry Res* 2012; **200**: 272–280.
17 John ER, Prichep LS, Alper KR, Mas FG, Cancro R, Easton P *et al.* Quantitative electrophysiological characteristics and subtyping of schizophrenia. *Biol Psychiatry* 1994; **36**: 801–826.
18 Geisler D, Walton E, Naylor M, Roessner V, Lim KO, Charles Schulz S *et al.* Brain structure and function correlates of cognitive subtypes in schizophrenia. *Psychiatry Res* 2015; **234**: 74–83.
19 Marder SR, Asarnow RF, Van Putten T. Information processing and neuroleptic response in acute and stabilized schizophrenic patients. *Psychiatry Res.* 1984; **13**: 41–49.
20 Clementz BA, Sweeney JA, Hamm JP, Ivleva EI, Ethridge LE, Pearlson GD *et al.* Identification of distinct psychosis biotypes using brain-based biomarkers. *Am J Psychiatry* 2016; **173**: 373–384.
21 Düring S, Glenthøj BY, Andersen GS, Oranje B. Effects of dopamine D2/D3 blockade on human sensory and sensorimotor gating in initially antipsychotic-naive, first-episode schizophrenia patients. *Neuropsychopharmacology* 2014; **39**: 3000–3008.
22 Düring S, Glenthøj BY, Oranje B. Effects of blocking D2/D3 receptors on mismatch negativity and P3a amplitude of initially antipsychotic naïve, first episode schizophrenia patients. *Int J Neuropsychopharmacol* 2015; **19**: pyv109.
23 Bertelsen B, Oranje B, Melchior L, Fagerlund B, Werge TM, Mikkelsen JD *et al.* Association study of CHRNA7 promoter variants with sensory and sensorimotor gating in schizophrenia patients and healthy controls: a danish case-control study. *Neuromolecular Med* 2015; **17**: 423–430.
24 Nielsen MØ, Rostrup E, Wulff S, Bak N, Lublin H, Kapur S *et al.* Alterations of the brain reward system in antipsychotic nave schizophrenia patients. *Biol Psychiatry* 2012; **71**: 898–905.
25 Nielsen MO, Rostrup E, Wulff S, Bak N, Broberg BV, Lublin H *et al.* Improvement of brain reward abnormalities by antipsychotic monotherapy in schizophrenia. *Arch Gen Psychiatry* 2012; **69**: 1195–1204.
26 Nielsen MØ, Rostrup E, Wulff S, Glenthøj B, Ebdrup BH. Striatal reward activity and antipsychotic-associated weight change in patients with schizophrenia undergoing initial treatment. *JAMA Psychiatry* 2016; **73**: 121–128.
27 Ebdrup B, Raghava J, Nielsen M, Rostrup E, Glenthøj B. Frontal fasciculi and psychotic symptoms in antipsychotic-naive patients with schizophrenia before and after 6 weeks of selective dopamine D2/3 receptor blockade. *J Psychiatry Neurosci* 2016; **41**: 133–141.
28 Nordholm D, Poulsen HE, Hjorthøj C, Randers L, Nielsen MØ, Wulff S *et al.* Systemic oxidative DNA and RNA damage are not increased during early phases of psychosis: a case control study. *Psychiatry Res* 2016; **241**: 201–206.
29 Wulff S, Pinborg LH, Svarer C, Jensen LT, Nielsen MØ, Allerup P *et al.* Striatal D(2/3) binding potential values in drug-naïve first-episode schizophrenia patients correlate with treatment outcome. *Schizophr Bull* 2015; **41**: 1143–1152.
30 Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987; **13**: 261–276.
31 Bak N, Hansen LK. Data driven estimation of imputation error—a strategy for imputation with a reject option. *PLoS ONE* 2016; **11**: e0164464.
32 Nelson HE, O'Connell A. Dementia: the estimation of premorbid intelligence levels using the new adult reading test. *Cortex* 1978; **14**: 234–244.
33 Wechsler D. *WAIS-III Administration and Scoring Manual*. The Psychological Corporation: San Antonio, TX, 1997.
34 Keefe RS, Goldberg TE, Harvey PD, Gold JM, Poe MP, Coughenour L. The Brief Assessment Of Cognition In Schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophr Res* 2004; **68**: 283–297.
35 Levaux M-N, Potvin S, Sepehry AA, Sablier J, Mendrek A, Stip E. Computerized assessment of cognition in schizophrenia: promises and pitfalls of CANTAB. *Eur Psychiatry* 2007; **22**: 104–115.
36 Van Lutterveld R, Oranje B, Kemner C, Abramovic L, Willems AE, Boks MPM *et al.* Increased psychophysiological parameters of attention in non-psychotic individuals with auditory verbal hallucinations. *Schizophr Res* 2010; **121**: 153–159.
37 Braff DL, Geyer MA, Swerdlow NR. Human studies of prepulse inhibition of startle: normal subjects, patient groups, and pharmacological studies. *Psychopharmacology* 2001; **156**: 234–258.
38 Abrahamsen TJ, Hansen LK. A cure for variance inflation in high dimensional kernel principal component analysis. *J Mach Learn Res* 2011; **12**: 2027–2044.
39 Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr* 1974; **19**: 716–723.
40 Akaike H. Factor analysis and AIC. *Psychometrika* 1987; **52**: 317–332.
41 Toussaint G. Bibliography on estimation of misclassification. *IEEE Trans Inf Theory* 1974; **20**: 472–479.
42 Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995; **20**: 273–297.
43 Spearman C. *The Abilities of Man: Their Nature and Measurement*. Macmillan: London, UK, 1927.
44 Leucht S. Amisulpride, an unusual "atypical" antipsychotic: a meta-analysis of randomized controlled trials. *Am J Psychiatry* 2002; **159**: 180–190.
45 Levine SZ, Leucht S. Treatment response heterogeneity in the predominant negative symptoms of schizophrenia: analysis of amisulpride vs placebo in three clinical trials. *Schizophr Res* 2014; **156**: 107–114.
46 Dawes SE, Jeste DV, Palmer BW. Cognitive profiles in persons with chronic schizophrenia. *J Clin Exp Neuropsychol* 2011; **33**: 929–936.

47 Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry* 2012; **17**: 1174–1179.
48 Ljung L. *System Identification: Theory for the User*. PRT Prentice Hall: Englewood Cliffs, NJ, 1987.
49 Marquand AF, Wolfers T, Mennes M, Buitelaar J, Beckmann CF. Beyond lumping and splitting: a review of computational approaches for stratifying psychiatric disorders. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2016; **1**: 433–447.
50 Brodersen KH, Deserno L, Schlagenhauf F, Lin Z, Penny WD, Buhmann JM *et al*. Dissecting psychiatric spectrum disorders by generative embedding. *NeuroImage Clin* 2014; **4**: 98–111.

Supplementary Information accompanies the paper on the Translational Psychiatry website (http://www.nature.com/tp)