Online supplement to Guo et al. Cortical folding and the potential for prognostic neuroimaging in schizophrenia. Br J Psychiatry doi: 10.1192/bjp.bp.114.155796

Online Supplement DS1

Clinical sample

The data reported here comes from a previously reported[9] sample of 41 patients satisfying DSM-IV criteria for schizophrenia/schizoaffective disorder, recruited from community-based mental health teams in Nottinghamshire and Leicestershire, UK. All participants were diagnosed as per the clinical consensus derived in accordance with the procedure described byLeckman et al.,[18] using all available information including a review of case files and a standardized clinical interview (Symptoms and Signs in Psychotic Illness-SSPI[19]). All patients were in a stable phase of illness (defined as no more than 10 points change in Global Assessment of Functioning in the preceding 6 weeks before the scan) with no change in prescribed psychotropic medications in the 6 weeks prior to the study. The median Defined Daily Dose (DDD)[20] was calculated for all prescribed psychotropic medications. Participants with age <18 or >50, with neurological disorders, current substance dependence, or intelligence quotient < 70 using Quick Test[21] were excluded.

Severity Index

At an individual level, the burden of schizophrenia cannot be adequately quantified using a single metric of clinical severity. Nevertheless, several indicators such as socio-occupational functioning, cognitive performance, the everyday experience of psychotic symptoms and the persistence of these symptoms across the course of illness, can provide a composite measure of illness severity, especially when assessed during a period of relative clinical stability. We quantified current occupational and social dysfunction using the Social and Occupational Functioning Assessment Scale (SOFAS)[22] and assessed speed of cognitive processing, a consistent and prominent cognitive deficit in schizophrenia using the Digit Symbol Substitution Test [DSST].[23] DSST was administered using a written and an oral format with a mean score computed from the two measures. In addition to current SSPI scores (on the day of MRI scan) to measure the symptoms of reality distortion, disorganisation and psychomotor poverty, we also collected retrospective information regarding the longitudinal severity (persistence) of psychotic symptoms by applying the SSPI scale over the entire recorded period of illness using clinical case notes to derive a single numerical score representing total persistence of psychotic symptoms across the life-course. High interrater reliability was achieved for the persistence measure among three psychiatrists involved in this study (Intra-class correlation coefficient=0.87[0.73-0.94]; n=25 subjects).

We then undertook a principal component analysis to extract the first unrotated factor explaining the largest proportion of variance from the measures of illness severity (3 SSPI syndrome scores, total persistence score, SOFAS score, mean DSST score). Positive loading of illness severity factor was seen in patients with persistent illness, poor functional ability, poor processing speed and higher symptom burden of disorganisation, psychomotor poverty and reality distortion. Negative loading indicated less persistent illness, with better functional ability, higher processing speed and lower symptom burden across the three syndromes. Based on the factor scores we divided the patient sample into those showing greater illness severity (positive loading on the severity factor; n=20) and less illness severity (negative loading on the severity factor; n=21). Demographic features of these two groups are presented in Table DS1.

Table DS1 Demographic features of the subgroups based on the severity of illness

	Low severity (n=21)	**High severity (n=20)**	**T/X²**
Gender (male/female)	13/8	18/2	x^2=2.99, p=0.08
Handedness (right/left)	19/2	18/2	x^2=0.00, p=1.0
Age in years (SD)	31.4(9.1)	35.9 (9.1)	T=-1.6 p=0.12
Parental NS-SEC (SD)	1.9(1.3)	3.1(1.5)	T=-2.8, p=0.01
Global mean gyrification	2.97(0.17)	2.93(0.15)	T=0.79, p=0.43
DDD (SD)	1.2(1.03)	1.4(1.2)	T=-0.58, p=0.56

DDD: Define Daily Dose of antipsychotics. NS-SEC: Parental Socio-Economic Status (National Statistics Scale) SD: Standard Deviation

Support Vector Machine Analysis

For pattern classification analysis, we used Support Vector Machine (SVM), a supervised learning algorithm that addresses the problem of discriminating two groups on the basis of a large number of features. SVM toolkit libsvm written by Lin Chih-Jen from Taiwan university (http://www.csie.ntu.edu.tw/~cjlin/libsvm/) was used with a radial basis function as kernel function (t=2) and parameter C = 10 to trade-off learning and extend ability while other parameters are kept as default values, in line with our previous work.[24] To measure the test performance and to validate the classifier, a leave-one-subject-out cross validation approach was employed, where the classifier is trained on all subjects except one, which is used as test data. Balanced accuracy, specificity, sensitivity, and predictive values for each classifier were obtained and statistical significance of these measures was determined by way of permutation testing using leave-one-subject-out method (1000 permutations with random assignment of high/low illness severity labels to the training data). On the basis of this permutation analysis, mean discrimination accuracy, sensitivity and specificity was obtained for the entire sample. Details of the test performance measures are provided in Table DS2.

Table DS2: Test performance measures

Accuracy = Number of subjects correctly classified to either groups / total number of subjects in the sample
Sensitivity = Number of subjects correctly identified to have high illness severity on the basis of the classifier / total number of subjects with high illness severity
Specificity = Number of subjects correctly identified to have low illness severity on the basis of the classifier / total number of subjects with low illness severity
Likelihood ratio of positive test (LR+) = sensitivity / (1-specificity)
Likelihood ratio of negative test (LR-) = (1-sensitivity) / specificity
Diagnostic odds ratio = LR+ / LR-

Effect of removing the variance related to gender and parental socioeconomic status

A baseline comparison revealed that the two illness severity groups differed at a trend level on gender distribution, and significantly on the parental socioeconomic status. To study the effect of these two variables on the overall prognostic accuracy achieved using regional morphometric measures, we estimated the residual of variances in the morphometric measures that were not explained by the linear effect of gender and parental NSSEC scores, and used these residuals to repeat the SVM analysis. The results of this analysis are shown in Table DS3. Statistically significant discrimination accuracy persisted for regional gyrification but not for regional thickness or volume.

Table DS3: Discrimination accuracy after regressing the linear effect of parental NSSEC and gender.

	Accuracy (p values)	**Sensitivity**	**Specificity**	**Likelihood Ratio (Positive & Negative)**	**Diagnostic Odds Ratio**
Regional Thickness	31.71%(0.87)	35%	28.57%	0.49,2.28	0.22
Regional Gyrification	68.29%(0.031)	60%	76.19%	2.52,0.53	4.80
Regional Volume	51.22%(0.5)	0%	100%	-	-

Effect of removing the variance related to antipsychotic dose

We presented the result of comparing the defined daily dose of antipsychotics between the high and low severity groups. There was no significant difference between the two groups ($t = –0.58$, $P = 0.56$), suggesting that the classification accuracy is unlikely to be driven by differential antipsychotic usage. Furthermore, we observe that the gyrification-based classification continued to perform superiorly when compared to thickness and volume metrics when regressing the linear effect of antipsychotic dose as shown below in Table DS4.

Table DS4: Discrimination accuracy after regressing the linear effect of antipsychotic dose equivalents.

	Accuracy (p values)	**Sensitivity**	**Specificity**	**Likelihood Ratio (Positive & Negative)**	**Diagnostic Odds Ratio**
Regional Thickness	46.34%(0.58)	50%	42.86%	0.875,1.17	0.75
Regional Gyrification	60.98%(0.07)	55%	66.67%	1.65,0.675	2.44
Regional Volume	51.22%(0.5)	0%	100%	-	-

In Table 1 of the manuscript, we presented some test performance measures of morphometric multivariate pattern classifiers. Further information is shown below in Table DS5.

Table DS5: Most significant predictors among regional gyrification values in test performance measures of morphometric multivariate pattern classifiers[a]

	Direction of difference	Low severity group, mean (s.d.)	High severity group, mean (s.d.)	*T*	*P*
Right lateral occipitotemporal sulcus	Low>High	2.69 (0.10)	2.59 (0.10)	2.8051	0.0078
Left inferior orbitofrontal gyrus	Low>High	3.83 (0.52)	3.48 (0.31)	2.4911	0.0171
Left middle temporal gyrus	Low>High	3.29 (0.23)	3.12 (0.19)	2.4481	0.0190
Right inferior temporal gyrus	Low>High	2.64 (0.10)	2.56 (0.10)	2.3143	0.0260
Right inferior occipital gyrus and sulcus	Low>High	2.67 (0.13)	2.58 (0.12)	2.2993	0.0269
Left inferior temporal sulcus	Low>High	2.87 (0.18)	2.75 (0.14)	2.2329	0.0314
Left superior occipital gyrus	Low>High	2.71 (0.14)	2.61 (0.19)	2.0458	0.0476
Right posterior midcingulate sulcus and gyrus	High>Low	2.08 (0.11)	2.15 (0.12)	–2.0377	0.0484

a. *P*-values are not corrected for multiple comparisons.

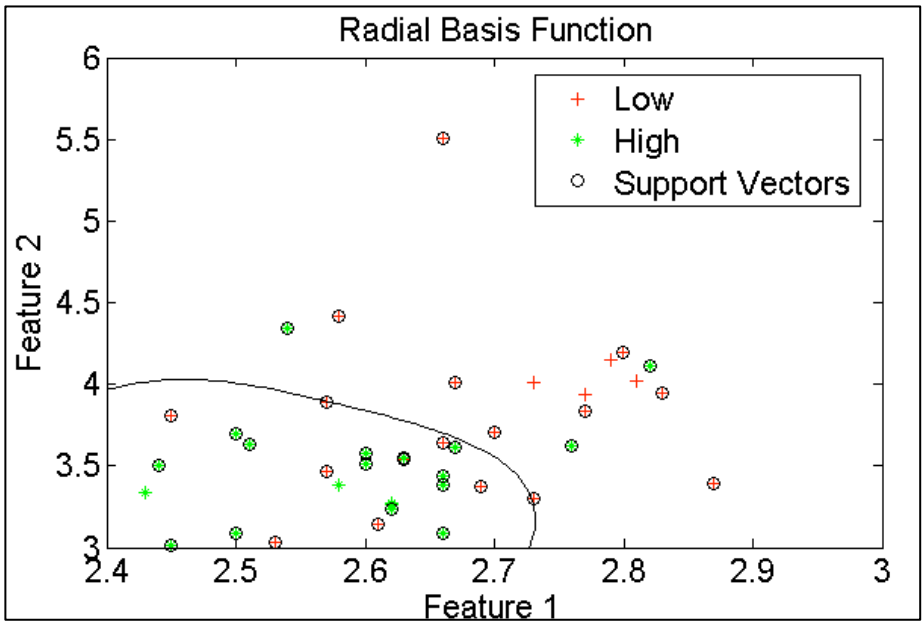

Figure DS1: A scatter plot of the 2 most discriminating features that separate patients with poor vs. good outcome in schizophrenia. Green crosses represent patients with high severity of illness (poor outcome) red crosses indicate those who have low severity (good outcome). The radial basis function for separating the two classes is shown by a continuous line (plane), along with circled members whose group separation specifies the plane. The upper portion mostly includes good outcome subjects, except for 3 misclassifications, indicating high sensitivity (ability to correctly predict outcome among those who have good outcome).

Additional references

18 Leckman JF, Sholomskas D, Thompson D, Belanger A, Weissman MM. Best Estimate of Lifetime Psychiatric Diagnosis: A Methodological Study. *Arch Gen Psychiatry* 1982; **39**: 879–83.

19 Liddle PF, Ngan ETC, Duffield G, Kho K, Warren AJ. Signs and Symptoms of Psychotic Illness (SSPI): a rating scale. *The British Journal of Psychiatry* 2002; **180**: 45–50.

20 WHO Collaborating Centre for Drug Statistics and Methodology. Guidelines for ATC Classification and DDD Assignment. 2003.

21 Ammons RB, Ammons CH. The Quick Test (QT): Provisional manual. *Psychological Reports* 1962; **11**: 111–61.

22 Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: a review of measures of social functioning. *Am J Psychiatry* 1992; **149**: 1148–56.
23 Dickinson D, Ramsey ME, Gold JM. Overlooking the Obvious: A Meta-analytic Comparison of Digit Symbol Coding Tasks and Other Cognitive Measures in Schizophrenia. *Arch Gen Psychiatry* 2007; **64**: 532–42.
24 Guo S, Palaniyappan L, Yang B, Liu Z, Xue Z, Feng J. Anatomical distance affects functional connectivity in patients with schizophrenia and their siblings. *Schizophr Bull* 2014; **40**: 449–59.