BJPsych | The British Journal of Psychiatry
1–2. doi: 10.1192/bjp.bp.114.155796

Short report

Cortical folding and the potential for prognostic neuroimaging in schizophrenia

Shuixia Guo,* Sarina Iwabuchi,* Vijender Balain, Jianfeng Feng, Peter Liddle and Lena Palaniyappan

Summary
In 41 patients with schizophrenia, we used neuroanatomical information derived from structural imaging to identify patients with more severe illness, characterised by high symptom burden, low processing speed, high degree of illness persistence and lower social and occupational functional capacity. Cortical folding, but not thickness or volume, showed a high discriminatory ability in correctly identifying patients with more severe illness.

Declaration of interest
L.P.: travel fellowship from Eli Lilly; support in kind from Magstim Co Ltd for a conference presentation. P.L.: honoraria for academic presentations from Janssen-Cilag and Bristol Myers Squibb; advisory panels for Bristol Myers Squibb.

To date there are no objectives tests that aid prognostic prediction in schizophrenia. Historically, clinical outcomes have improved considerably for medical disorders where severity can be quantified reliably (for example malignancies, asthma). Prognostic prediction, in particular the ability to identify those who will do well in the long term, has proved to be a great challenge in schizophrenia.[1] Neuroimaging offers the great promise of providing objective measures of clinical utility in managing psychosis.[2] Recently, the use of multivariate pattern classification in neuroimaging has enabled diagnostic separation at a single patient level.[3] In this study, we investigated whether this approach can reliably discriminate a patient with less severe illness from one with more severe illness. Given the previous observations that cortical thickness,[4] folding patterns[5] and grey matter volume[6] relate to prognosis in schizophrenia, we employed these surface-based morphometric measures to identify illness severity.

Method

A sample of 41 patients with a DSM-IV diagnosis[7] of schizophrenia or schizoaffective disorder was recruited for this study. This sample is described in detail in our previous studies.[8,9] The clinical severity was quantified using a composite index derived from symptom burden, functional ability, cognition and persistence of illness as described in our previous work[8] and in online Supplement DS1. Using this severity index, 20 participants were classified as having a high severity of illness with the remaining 21 having a low severity of illness. The clinical and demographic characteristics of the two groups are presented in online supplement DS1 and Table DS1.

Structural magnetic resonance imaging scans obtained from the participants were processed using Freesurfer (5.1.0) (http://surfer.nmr.mgh.harvard.edu/) as previously described.[10] Reconstructed surfaces were inspected for topological defects and edited in accordance with our previous work[11] by a single rater (L.P.) masked to the severity status at the time of surface editing. Cortical folding was measured using local gyrification index proposed by Schaer *et al*.[12] Cortical thickness was estimated using the standard procedures described by Fischl & Dale.[13] The reconstructed brain surfaces were parcellated using the Destrieux atlas to provide 148 brain regions based on sulcogyral boundaries described by Duvernoy.[14] For each metric, these 148 values were used as features in the classifier.

We used a linear support vector machine (SVM) proposed by Cortes & Vapnik[15] and implemented by the libsvm toolkit (http://www.csie.ntu.edu.tw/~cjlin/libsvm/). SVM is a statistical discrimination procedure that finds a linear separation surface in the high-dimensional multivariate feature space that maximally separates the training data into two classes as specified by the pre-assigned labels (in this case, high and low severity groups). Based on this separation, the class membership of a new participant (test data) can be predicted, and the accuracy of these predictions quantified. Further details are given in online supplement DS1. We computed test performance measures and diagnostic odds ratio using a leave-one-subject-out (LOSO) cross validation. The statistical significance of these measures was determined using permutation testing ($n = 1000$ permutations).

Results

Table 1 displays the accuracy of the classification and the most significant predictors of the best performing classifier. Given that gender and parental socioeconomic status differed between the two groups, we regressed out the variance explained by these two variables, and repeated the SVM analysis. Our results continued to show a superior, statistically significant accuracy for regional gyrification but not for thickness or volume (online Supplement DS1, Fig DS1 and Tables DS3 and DS4).

Discussion

To our knowledge, this is the first study to investigate the prospect of exploiting multivariate neuroanatomical information to predict clinical severity of schizophrenia at the individual level. Using a classifier based on the features of cortical folding, we can identify the degree of illness severity in medicated, community-living patients with clinically stable schizophrenia. This predictive ability appears to be a unique feature of folding patterns, as the classifiers based on thickness and volume do not perform significantly above chance when separating high and low illness severity groups. Furthermore, patients with greater illness severity had reduced cortical folding in most brain regions, suggesting that a distributed defect in cortical morphology influences prognosis. Although the accuracy achieved by the gyrification-based classifier is statistically significant, the performance of this classifier is considerably weaker when compared with the multivariate neuroanatomical classifiers tested in the separation of healthy controls from patients with schizophrenia.[16] There may be several

*These authors contributed equally to the work.

Table 1 Test performance measures of morphometric multivariate pattern classifiers[a]

Region	Accuracy, % (*P*)	Sensitivity, %	Specificity, %	Likelihood ratio (positive and negative)	Diagnostic odds ratio
Thickness	65.9 (0.07)	70	61.9	1.84, 0.48	3.79
Gyrification	73.2 (0.004)	60	85.7	4.20, 0.47	9.00
Volume	51.2 (0.5)	0	100	–	–

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

reasons for this disparity. The use of median split to divide the sample into high and low severity could have contributed to the lack of strong between-groups discriminative features With larger samples, extreme prognostic groups (lying on either end of the severity continuum) could be used for training the classifier and improve the accuracy. It is worth noting that in clinical practice, it is rarely necessary to apply a test to differentiate a patient with schizophrenia from a healthy control. The classification of a patient with schizophrenia from a healthy control can be done clinically with a high degree of confidence, thus even a high-performance neuroimaging test will have limited clinical utility in this context. On the other hand, at present there are no reliable means of predicting prognostic group membership; even a test that increases the likelihood of identifying prognostic grouping to a moderate extent, could be of significant benefit to patients and clinicians.

We quantified illness severity on the basis of a number of variables; this approach offered a multidomain metric that reflected symptom burden across the three syndromes of schizophrenia, a cognitive function that is most prominently affected in schizophrenia i.e. processing speed, social and functional performance and persistence of illness. Nevertheless, various other metrics relevant for the assessment of severity (such as Clinical Global Impression, quality of life scales, self-rated recovery measures or assessments of daily living) were not collected in this study. Furthermore, from this cross- sectional study it is not possible to extrapolate whether a gyrification-based classifier applied at illness onset could prospectively predict later severity. Nevertheless, when compared with cortical thickness and volume, gyrification has been shown to be relatively stable during adult life.[17] In addition, a large degree of variance in the cortical folding patterns relates to neurodevelopmental integrity during the fetal or early neonatal period.[18] Taken together, these observations suggest that the burden of neurodevelopmental abnormalities in a patient with schizophrenia could be a potential influence on illness severity.

Our results provide preliminary evidence for the utility of cortical folding in single participant-level prognostic imaging in schizophrenia. With larger validation studies that combine high-yield clinical prognostic indicators with gyrification metrics, the predictive value can be further improved, enabling an objective grading of outcome in the management of schizophrenia. This has the promise of assisting targeted service delivery and making personalised recommendations with regard to the required duration of antipsychotic treatment. Most importantly, this approach can be refined to provide accurate information on the chances of a satisfactory clinical recovery and thus potentially empower patients by addressing the uncertainty that surrounds prognosis in psychotic disorders.

Funding

This work was funded by the Medical Research Council (UK) Grant Number: G0601442. L.P. is supported by the Wellcome Trust (Research Training Fellowship WT096002/Z/11). S.G. is supported by the National Natural Science Foundation of China (NSFC) grant (No. 11271121), Program for New Century Excellent Talents in University (NCET-13-0786) and the Natural Science Foundation of Hunan Province (2015JJ1010).

Shuixia Guo, PhD, College of Mathematics and Computer Science, Key Laboratory of High Performance Computing and Stochastic Information Processing (Ministry of Education of China), Hunan Normal University, Changsha, PR China and Department of Computer Science, University of Warwick, Coventry, UK; **Sarina Iwabuchi**, PhD, Division of Psychiatry & Applied Psychology, University of Nottingham and Centre for Translational Neuroimaging, Institute of Mental Health, Nottingham, UK; **Vijender Balain**, MRCPsych, Penticton Regional Hospital, Penticton, British Columbia, Canada; **Jianfeng Feng**, PhD, Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, PR China and Department of Computer Science, University of Warwick, Coventry, UK; **Peter Liddle**, PhD, Division of Psychiatry & Applied Psychology, University of Nottingham and Centre for Translational Neuroimaging, Institute of Mental Health, Nottingham, UK; **Lena Palaniyappan**, PhD, MRCPsych, Division of Psychiatry & Applied Psychology, University of Nottingham, Centre for Translational Neuroimaging, Institute of Mental Health, Nottingham, UK and Penticton Regional Hospital, Penticton, British Columbia, Canada

Correspondence: Lena Palaniyappan, Room-09, C Floor, Institute of Mental Health Building, Triumph Road, Nottingham, NG7 2TU, UK. Email: Lena.Palaniyappan@nottingham.ac.uk

First received 4 Aug 2014, final revision 9 Nov 2014, accepted 25 Jan 2015

References

1 Fenton WS, McGlashan TH. Prognostic scale for chronic schizophrenia. *Schizophr Bull* 1987; **13**: 277–86.

2 Lawrie SM, Olabi B, Hall J, McIntosh AM. Do we have any solid evidence of clinical utility about the pathophysiology of schizophrenia? Do we have any solid evidence of clinical utility about the pathophysiology of schizophrenia? *World Psychiatry* 2011; **10**: 19–31.

3 Borgwardt S, Fusar-Poli P. Third-generation neuroimaging in early schizophrenia: translating research evidence into clinical utility. *Br J Psychiatry* 2012; **200**: 270–2.

4 Szeszko PR, Narr KL, Phillips OR, McCormack J, Sevy S, Gunduz-Bruce H, et al. Magnetic resonance imaging predictors of treatment response in first-episode schizophrenia. *Schizophr Bull* 2012; **38**: 569–78.

5 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. *JAMA Psychiatry* 2013; **70**: 1031–40.

6 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. *Psychol Med* 2012; **42**: 1037–47.

7 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorder (4th edn) (DSM-IV)*. APA, 1994.

8 Palaniyappan L, Park B, Balain V, Dangi R, Liddle P. Abnormalities in structural covariance of cortical gyrification in schizophrenia. *Brain Struct Funct* 2014; April 26 (Epub ahead of print).

9 Palaniyappan L, Liddle PF. Diagnostic discontinuity in psychosis: a combined study of cortical gyrification and functional connectivity. *Schizophr Bull* 2014; **40**: 675–84.

10 Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis: i. segmentation and surface reconstruction. *Neuroimage* 1999; **9**: 179–94.

11 Palaniyappan L, Liddle PF. Aberrant cortical gyrification in schizophrenia: a surface-based morphometry study. *J Psychiatry Neurosci* 2012; **37**: 110–9.

12 Schaer M, Cuadra MB, Tamarit L, Lazeyras F, Eliez S, Thiran J-P. A surface-based approach to quantify local cortical gyrification. *IEEE Trans Med Imaging* 2008; **27**: 161–70.

13 Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci USA* 2000; **97**: 11050–5.

14 Destrieux C, Fischl B, Dale A, Halgren E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* 2010; **53**: 1–15.

15 Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995; **20**: 273–97.

16 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. *Neurosci Biobehav Rev* 2012; **36**: 1140–52.

17 Raznahan A, Shaw P, Lalonde F, Stockman M, Wallace GL, Greenstein D, et al. How does your cortex grow? *J Neurosci* 2011; **31**: 7174–7.

18 Dubois J, Benders M, Borradori-Tolsa C, Cachia A, Lazeyras F, Ha-Vinh Leuchter R, et al. Primary cortical folding in the human newborn: an early marker of later functional development. *Brain* 2008; **131**: 2028–41.

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.

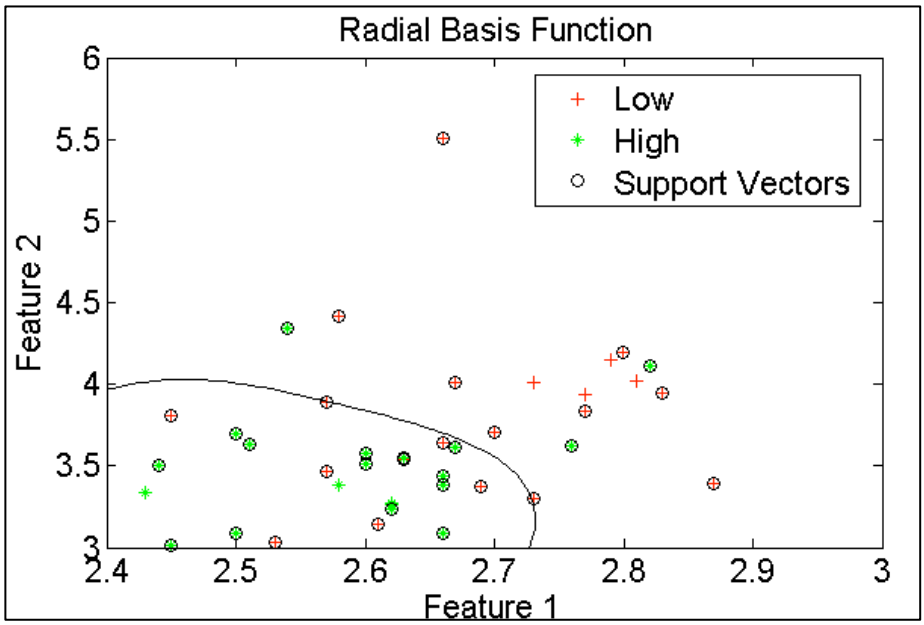

Figure DS1: A scatter plot of the 2 most discriminating features that separate patients with poor vs. good outcome in schizophrenia. Green crosses represent patients with high severity of illness (poor outcome) red crosses indicate those who have low severity (good outcome). The radial basis function 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.

Cortical folding and the potential for prognostic neuroimaging in schizophrenia

Shuixia Guo, Sarina Iwabuchi, Vijender Balain, Jianfeng Feng, Peter Liddle and Lena Palaniyappan

BJP published online July 23, 2015 Access the most recent version at DOI: 10.1192/bjp.bp.114.155796