Individual differences in schizophrenia

Edmund T. Rolls, Wenlian Lu, Lin Wan, Hao Yan, Chuanyue Wang, Fude Yang, Yunlong Tan, Lingjiang Li, Chinese Schizophrenia Collaboration Group, Hao Yu, Peter F. Liddle, Lena Palaniyappan, Dai Zhang, Weihua Yue and Jianfeng Feng

Background
Whether there are distinct subtypes of schizophrenia is an important issue to advance understanding and treatment of schizophrenia.

Aims
To understand and treat individuals with schizophrenia, the aim was to advance understanding of differences between individuals, whether there are discrete subtypes, and how first-episode patients (FEP) may differ from multiple episode patients (MEP).

Method
These issues were analysed in 687 FEP and 1880 MEP with schizophrenia using the Positive and Negative Syndrome Scale for (PANSS) schizophrenia before and after antipsychotic medication for 6 weeks.

Results
The seven Negative Symptoms were correlated with each other and with P2 (conceptual disorganisation), G13 (disturbance of volition), and G7 (motor retardation). The main difference between individuals was in the cluster of seven negative symptoms, which had a continuous unimodal distribution. Medication decreased the PANSS scores for all the symptoms, which were similar in the FEP and MEP groups.

Conclusions
The negative symptoms are a major source of individual differences, and there are potential implications for treatment.

Previous research has indicated that schizophrenia is heterogeneous in clinical symptoms and time course, suggesting the possibility of identifying subtypes differing in cause and/or pathological mechanism.[1] The identification of subtypes could facilitate investigation of cause and mechanism, including the use of genetic and neuroscience approaches.[2–5] Many approaches consider that there are positive symptoms (including hallucinations and delusions), negative symptoms (including low mood, motivation, and emotion), and cognitive symptoms, including poor attention.[6–8] It has been suggested that at least in chronic cases, the positive symptoms split into two subgroups: reality distortion and disorganisation.[7,9] Many studies that have included a wider range of symptoms have found evidence for five clusters of symptoms, namely, reality distortion, disorganisation, negative symptoms, depression and excitation (reviewed by Peralta & Cuesta).[10]

In order to better understand and treat individuals with schizophrenia, the aims of the present investigation were to understand differences between individuals with schizophrenia; whether there are discrete subtypes of schizophrenia; how first-episode patients (FEP) may differ from multiple episode patients (MEP) and how medication influences the symptoms. Two approaches to delineating the heterogeneity of schizophrenia were taken. One approach was an examination of the pattern of correlation between symptoms to identify underlying symptom communities (or dimensions), presumed to reflect specific pathological processes. Another approach was an examination of the correlations between patients, to identify communities of patients (subtypes) each characterised by a symptom profile reflecting a particular combination of symptom dimensions. In the main text, we employ the standard and well-established hierarchical clustering method using k-means.[11] In the supplement, we present qualitatively similar results obtained using a powerful community detection algorithm based on the concept of modularity described by Newman & Girvan,[12] which was also useful in providing an estimate of the optimum number of communities.

Method

Participants and clinical assessment

Two samples comprising a total of 2567 patients with schizophrenia were diagnosed according to DSM-IV criteria.[13] The first sample included 687 FEP and drug-naive patients with schizophrenia. The second study sample included 1880 multi-episode patients (MEP) with schizophrenia with average duration of illness 7.97 years. The patients were recruited from in-patients from four research centres (the Sixth Hospital of Peking University, the Second Xiangya Hospital of Central South University, Beijing Anding Hospital and Beijing HuiLongGuan Hospital) with institutional approval of the investigation and informed consent of the patients.

The symptoms were assessed by trained researchers with the 30-item Positive and Negative Syndrome Scale for schizophrenia (PANSS).[14] For reference, in PANSS, positive symptoms P1–P7 are Items 1–7, the negative symptoms N1–N7 are Items 8–14 and the general symptoms G1–G16 are Items 15–30. In the figures, unless otherwise stated, the symptoms are presented in this order, 1–30, with the symptoms listed in the legend to Fig. 1a. The PANSS was measured first in each of these patients when they were not receiving drug treatment. Then the patients were provided with one out of the seven randomly assigned drugs (risperidone, quetiapine, perphenazine, olanzapine, haloperidol, aripiprazole and ziprasidone) for a 6-week treatment period, and then the PANSS scores were measured again. This design with the same patients assessed

before and after medication allowed measurement of the PANSS when untreated, and also enabled the effects of the medication to be assessed using exactly the same patients. Further details of the patient populations and the medication used are provided in the Data supplement.

Statistical analysis

The data consisted for the MEP group of 1880 patient×30 symptom matrix, and for the FEP group of a 687 patient×30 symptom matrix, for the unmedicated pre-treatment state. Corresponding matrices were available for the post-treatment state, providing the symptoms in the same patients after 6 weeks of treatment. The methods included calculating the symptom correlation matrix and the population correlation matrix. We also examined whether there were discrete populations of patients using the k-means clustering method (performed on the population correlation matrix), and then analysed the symptoms in the sub-populations that were detected. We also investigated the distributions of symptoms across the patient sub-populations.

The data were analysed using the k-means cluster analysis algorithms that are a well-accepted approach standardised for availability within Matlab, using the distance measure based on correlations. This analysis was complemented by a modern community detection method described in the Data supplement, which has the advantage that it can automatically provide an indication of the optimal number of subgroups or communities into which to cluster the data.[15,16] The k-means approach was the method used for analysis described in the main text, because it is a well-known and understood method for clustering data, and because one major difference between the populations that was identified with this approach was in the negative symptoms which have a component that has a unimodal continuous distribution, so we did not wish to rely on a community detection method that might have difficulty if it was forced to determine the number of communities into which to divide the population if it contained at least in part a continuous distribution. However, for comparison key results with the community detection method are provided in the Data supplement.

Results

Multi-episode group pre-treatment (MEP_Pre)

Figure 1a shows the MEP_Pre symptom correlation matrix. The negative symptoms (8–14) form a clear cluster. Some correlation between these negative symptoms and P2 (conceptual disorganisation), G11 (poor attention), G13 (disturbance of volition) and G16 (active social avoidance) are also evident in the symptom correlation matrix shown in Fig. 1a. The community detection method indicated that the symptoms belong to four distinguishable communities (Fig. S1). Symptom community four included the majority of the positive symptoms (P1 P4 P5 P6 P7) and general symptoms (G8 G9 G12 G14). Symptom Community 1 included the majority of the negative symptoms (N1 N2 N3 N4 N6) and general symptoms (G7 G13 G16).

As shown in Fig. S2, the community detection algorithm indicated that the sample contained three distinguishable communities of patients. The patient clusters generate by the three-cluster output of the k-mean algorithm had distinct symptom profiles (Fig. 1b). In the MEP_Pre sample, the sub-population labelled PN (positive and high negative symptoms, 694 cases) had high values for the negative symptoms (N1–N7=symptoms 8–14). The sub-population labelled Pn (positive and intermediate negative symptoms, 535 cases) had intermediate values for the negative symptoms. The sub-population

labelled P (positive and low negative symptoms, 651 cases) had low values for the negative symptoms, especially N1 and N2. Little else differed between these clusters, except that the PN cluster had a higher value for P2 (conceptual disorganisation) and relatively small values for P1 and P3–P6; whereas the P cluster had relatively high scores on symptoms P1, P3–P7, G8 and G9.

These analyses provided evidence that the scores of the negative symptoms were an important factor that was different between the three populations detected with k-means, and equally by the community detection algorithm (see Data supplement Fig. S3). Furthermore, Fig. 1a shows that there are moderately high correlations between the seven negative symptoms, N1–N7 (Symptoms 8–14). To investigate how the negative symptoms were distributed across the population of MEP patients, the population × symptom data matrix was sorted by the mean value of symptoms N1–N7. The result shown in Fig. 2a indicates that there is a graded distribution of the negative symptoms across the population of 1880 MEP patients, with patients at one end of the distribution having scores close to 1 (the minimum score possible) and at the other end close to 7 (the maximum score possible). (In a control analysis, it was found that this does not apply to the positive symptoms, which are not continuously graded throughout the range between scores of 1 and 7 as are the negative symptoms shown in Fig. 2a.)

Further evidence was then sought of whether the distribution of the mean score for the seven negative symptoms was continuous and unimodal. Figure 2b and c show that the distribution of the mean of the average of the negative symptom scores in the MEP_Pre group has a unimodal distribution, in one dimension, rather than a bimodal or multimodal distribution. Thus the most important way in which the three sub-populations of MEP–Pre patients differed was in the negative symptom scores, which are continuously distributed from 1 to 7.

Multi-episode group post-treatment (MEP_Post)

The MEP_Post symptom correlation matrix showed, similarly to the MEP_Pre condition illustrated in Fig. 1a, that the negative symptoms (8–14) form a clear cluster (Fig. S5). Three sub-population clusters were again found by the community detection algorithm and the three clusters in the MEP_Post dataset that were detected with the k-means algorithm which had 631 patients in the PN cluster, 643 in the Pn cluster and 606 in the P cluster.

Figure 1c shows the average symptom values in the three patient clusters for the MEP_Post group. The group labelled PN (positive and high negative symptoms) had high values for the negative symptoms (N1–N7=symptoms 8–14). The group labelled Pn (positive and intermediate negative symptoms) had intermediate values for the negative symptoms. The group labelled P (positive and low negative symptoms) had low values for the negative symptoms and relatively low values also for the positive symptoms. It is notable that the scores on all or most of the symptoms have been decreased by the medication (Fig. 4d). Just as in the MEP_Pre condition, a continuously graded distribution of the negative symptoms across the population of 1880 MEP_Post patients was found, with patients at one end of the distribution having scores close to 1 (the minimum score possible) and at the other end close to 7 (the maximum score possible).

First-episode group pre-treatment (FEP_Pre)

Figure 3a shows the FEP_Pre symptom correlation matrix. The negative symptoms (8–14) form a clear cluster.

The community detection algorithm detected three clusters or sub-populations of patients, and Fig. 3b shows the average symptom values in the three patient clusters for the FEP_Pre group from
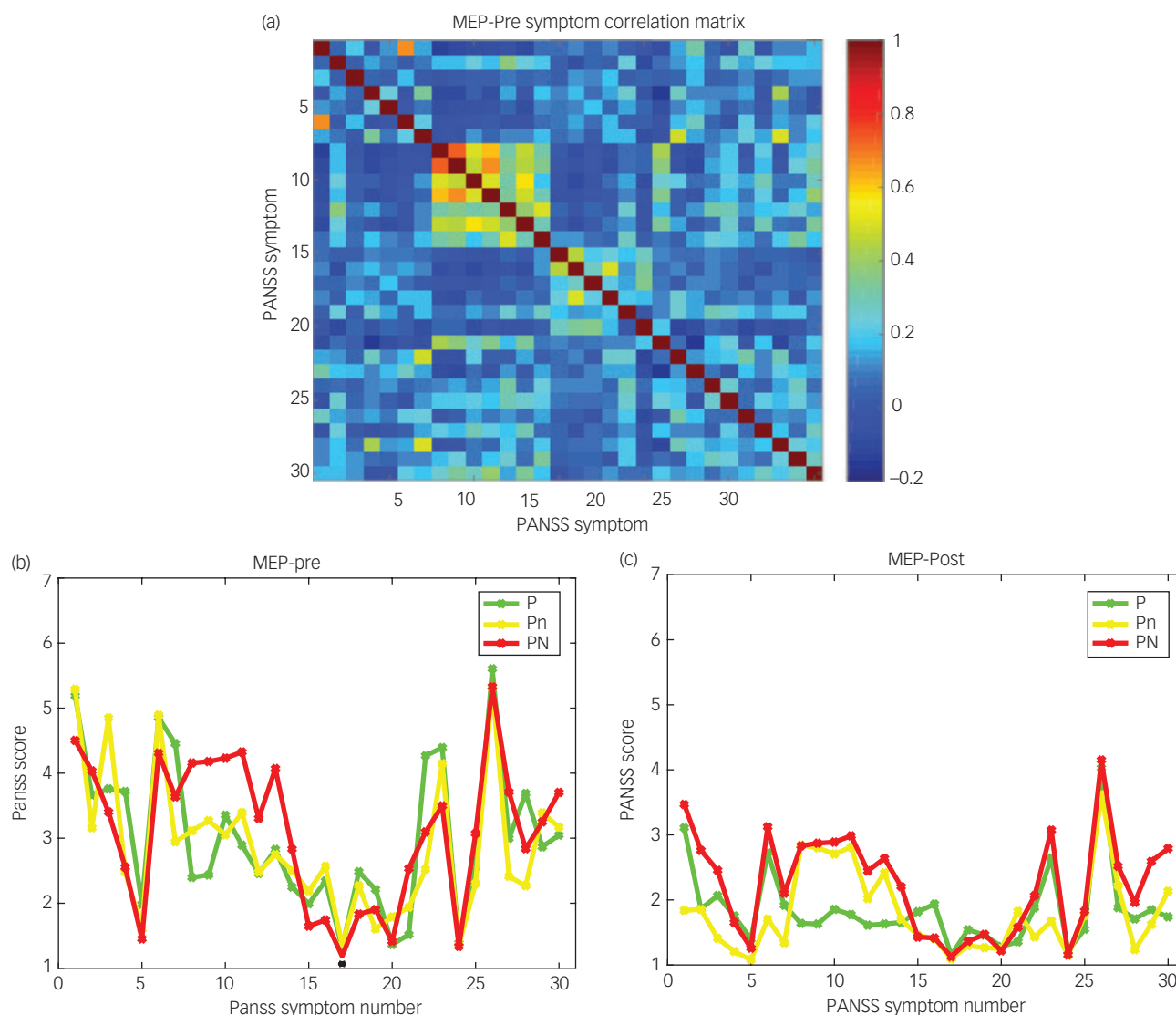
Fig. 1. (a) MEP_Pre symptom correlation matrix. The colour bar indicates the value of the Pearson correlation. In PANSS (14), P1–P7 are Symptoms 1–7, N1–N7 are Symptoms 8–14 and G1–G16 are Symptoms 15–30. The values for each score are 1–7. The symptoms are as follows: Delusions (P1), Conceptual disorganisation (P2), Hallucinations (P3), Hyperactivity (P4), Grandiosity (P5), Suspiciousness/persecution (P6), Hostility (P7), Blunted affect (N1), Emotional withdrawal (N2), Poor rapport (N3), Passive/apathetic social withdrawal (N4), Difficulty in abstract thinking (N5), Lack of spontaneity and flow of conversation (N6), Stereotyped thinking (N7), Somatic concern (G1), Anxiety (G2), Guilt feelings (G3), Tension (G4), Mannerisms and posturing (G5), Depression (G6), Motor retardation (G7), Uncooperativeness (G8), Unusual thought content (G9), Disorientation (G10), Poor attention (G11), Lack of judgment and insight (G12), Disturbance of volition (G13), Poor impulse control (G14), Preoccupation (G15) and Active social avoidance (G16). (b) MEP_Pre average symptom values in the three patient clusters detected by k-means. (c) MEP_Post average symptom values in the three patient clusters detected by k-means.

the k-means analysis. The group labelled PN (positive and high negative symptoms) had high values for the negative symptoms (N1–N7=symptoms 8–14). The group labelled Pn (positive and intermediate negative symptoms) had intermediate values for the negative symptoms. The group labelled P (positive and low negative symptoms) had low values for the negative symptoms. Little else differed between these clusters, except that the PN cluster has a smaller value for P1 (delusions) and P3 (hallucinatory behaviour) (Symptoms 1 and 3) and has higher scores for Symptoms 27–30 (G13–G16). The three clusters contained 228, 233 and 226 patients, respectively. The symptom profiles of the groups detected by the k-means algorithm (Fig. 3b) are very similar to the profiles of the corresponding groups detected by the community detection algorithm (Figs. S9 and S10).

Figure 3a shows that there are moderately high correlations between the seven negative symptoms, N1–N7 (Symptoms 8–14). To investigate how the negative symptoms were distributed across the population of FEP patients, the population × symptom data matrix was sorted by the mean value of Symptoms N1–N7. The result shown in Fig. 4a indicates that there is a graded distribution of the negative symptoms across the population of 687 FEP_Pre patients, with patients at one end of the distribution having scores close to 1 (the minimum score possible) and at the other end close to 7 (the maximum score possible).

Figure 4b and c shows that the distribution of the mean of the average of the negative symptom scores in the FEP_Pre group has a unimodal distribution, in one dimension, rather than a bimodal or multimodal distribution. Thus the most important way in which
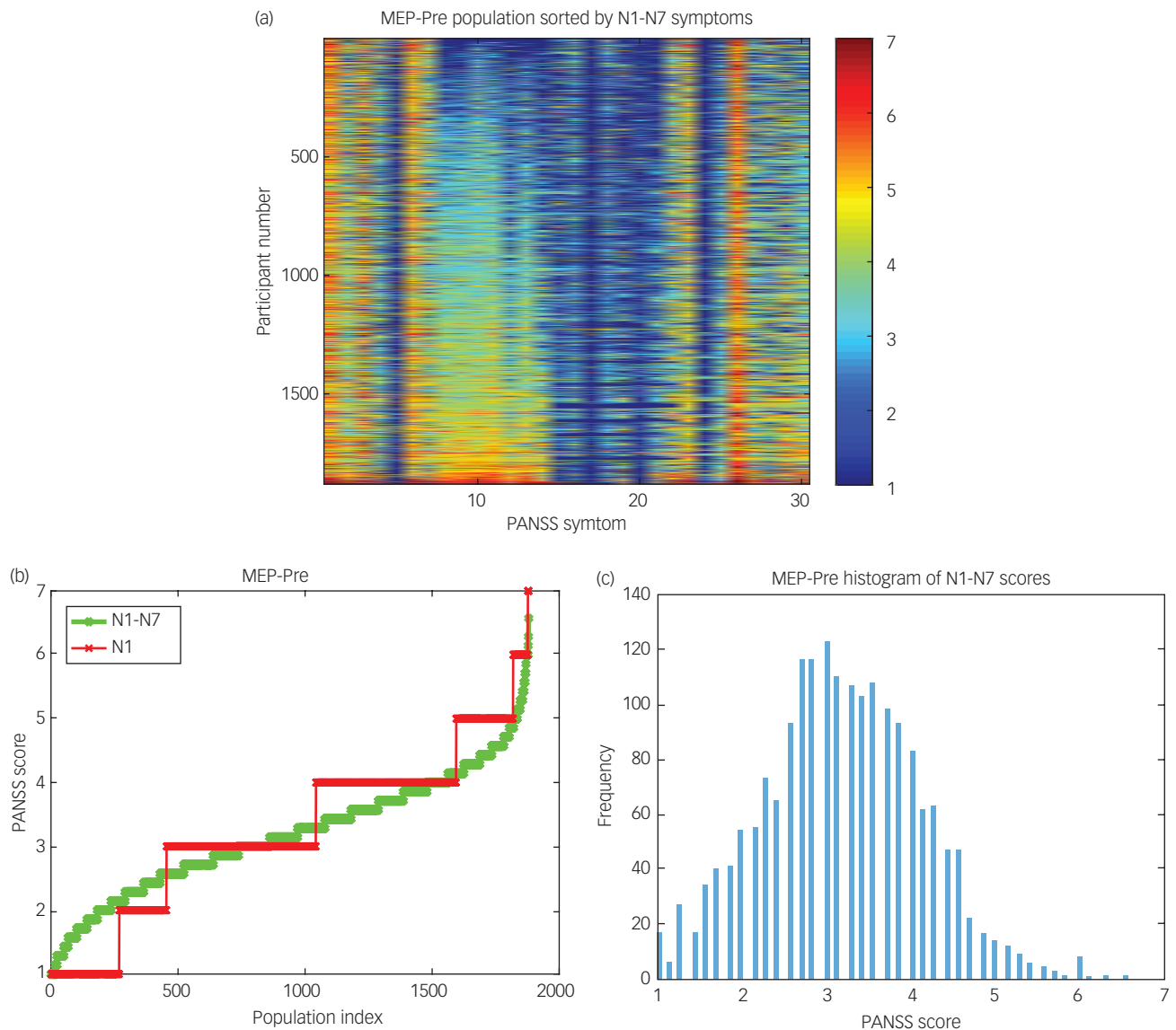
Fig. 2. (a) MEP_Pre group sorted by the average value of the negative symptoms, N1–N7 (Symptoms 8–14). The colour bar on the right shows the PANSS score in the range 1–7. (b) The average score for Symptoms N1–N7 shown sorted by its value in each member of the MEP_Pre population (green). The red plot shows the value for N1. (c) A histogram of the average score for Symptoms N1–N7 in the MEP_Pre group.

the three sub-populations of FEP_Pre patients differed was in the negative symptom scores, which are continuously distributed from 1 to 7.

First-episode group post-treatment (FEP_Post)

The FEP_Post symptom correlation matrix showed, similarly to the FEP_Pre condition illustrated in Fig. 3a, that the negative symptoms (8–14) form a clear cluster (Fig. S12). Three sub-population clusters were again found by the community detection algorithm (Fig. S13), and the three clusters in the FEP_Post dataset that were detected with the k-means algorithm had 288 patients in the PN cluster, 221 in the Pn cluster and 178 in the P cluster.

Figure 3c shows the average symptom values in the three patient clusters for the FEP_Post group. The group labelled PN (positive and high negative symptoms) had high values for the negative symptoms (N1–N7=symptoms 8–14). The group labelled Pn (yellow) had high scores for P1–P7 and intermediate values for the negative symptoms. The group labelled P had low values for the negative

and positive symptoms. It is notable that the scores on all or most of the symptoms have been decreased by the medication (Fig. 4e). Just as in the FEP_Pre condition, a continuously graded distribution of the negative symptoms across the population of 687 FEP_Post patients was found, with patients at one end of the distribution having scores close to 1 (the minimum score possible) and at the other end close to 7 (the maximum score possible).

Effects of treatment on the PANSS scores

The effects of treatment with antipsychotic drugs on the PANSS symptom scores for the MEP and FEP groups are shown in Fig. 4d and e. Interestingly, the treatment reduced almost all the symptoms (with smaller effects for those close to baseline). The effects on the symptoms were highly similar in the MEP and FEP groups. Moreover, the reductions in the scores of the negative symptoms for the different sub-populations PN, Pn and P produced by the medication were similar (making allowance for the fact that the pre-treatment value was different in the three
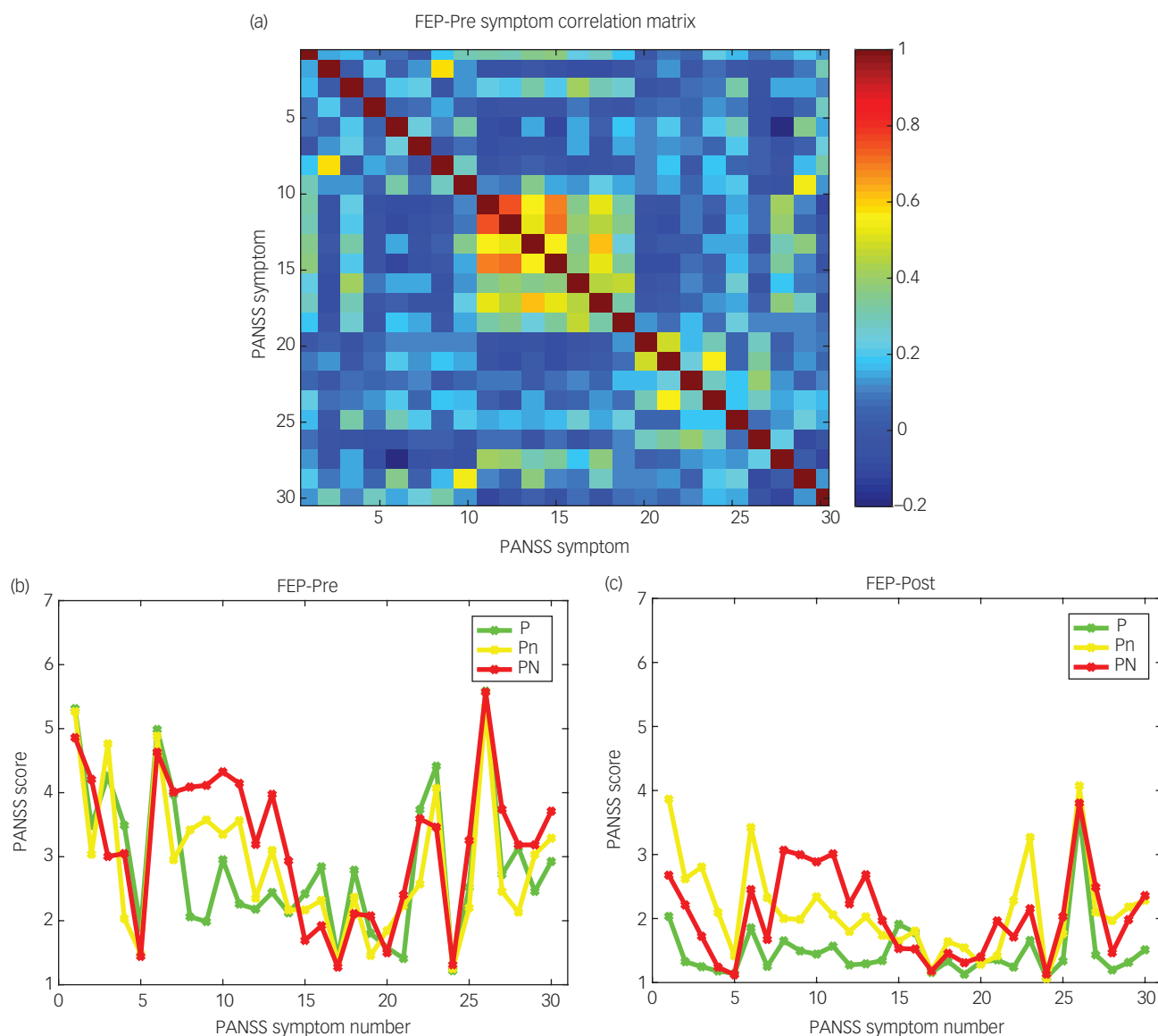
Fig. 3. (a) FEP_Pre symptom correlation matrix. The colour bar indicates the value of the Pearson correlation. In PANSS (14), P1–P7 are Symptoms 1–7, N1–N7 are Symptoms 8–14 and G1–G16 are Symptoms 15–30. The values for each score are 1–7. (b) FEP_Pre average symptom values in the three patient clusters detected by k-means. (c) FEP_Post average symptom values in the three patient clusters detected by k-means.

sub-populations). The treatment consisted of one out of seven randomly assigned drugs (risperidone, quetiapine, perphenazine, olanzapine, haloperidol, aripiprazole and ziprasidone) for a 6-week treatment period. Further details on the effects of the treatments are available from the authors on request and will be provided elsewhere.[16]

Comparison of MEP and FEP groups

It is evident from Fig. 4d and e that in the pre-treatment condition, the FEP and MEP groups had almost identical scores, and this is confirmed by the explicit comparison shown in Fig. S19. It is evident from Fig. 4d and e that in the post-treatment condition, the FEP and MEP groups had almost identical scores (though lower than in the pre-treatment groups), and this is confirmed by the explicit comparison shown in Fig. S20.

Community detection results

The community detection algorithm fast_mo[15] and its modified form fast_mo_sgn[16] applied to the population correlation matrix produced very similar results as shown in the Data supplement (Figs. S1–S14) to those described in the main text using k-means, with the main difference that the numbers of patients in the different clusters or communities differed somewhat.

Factor analysis and multidimensional scaling on the symptom correlation matrices

To provide further insight into how the different PANSS symptoms are related to each other and how they separate from each other, factor analyses and multidimensional scaling were performed on the symptom correlation matrices, as described in the Data
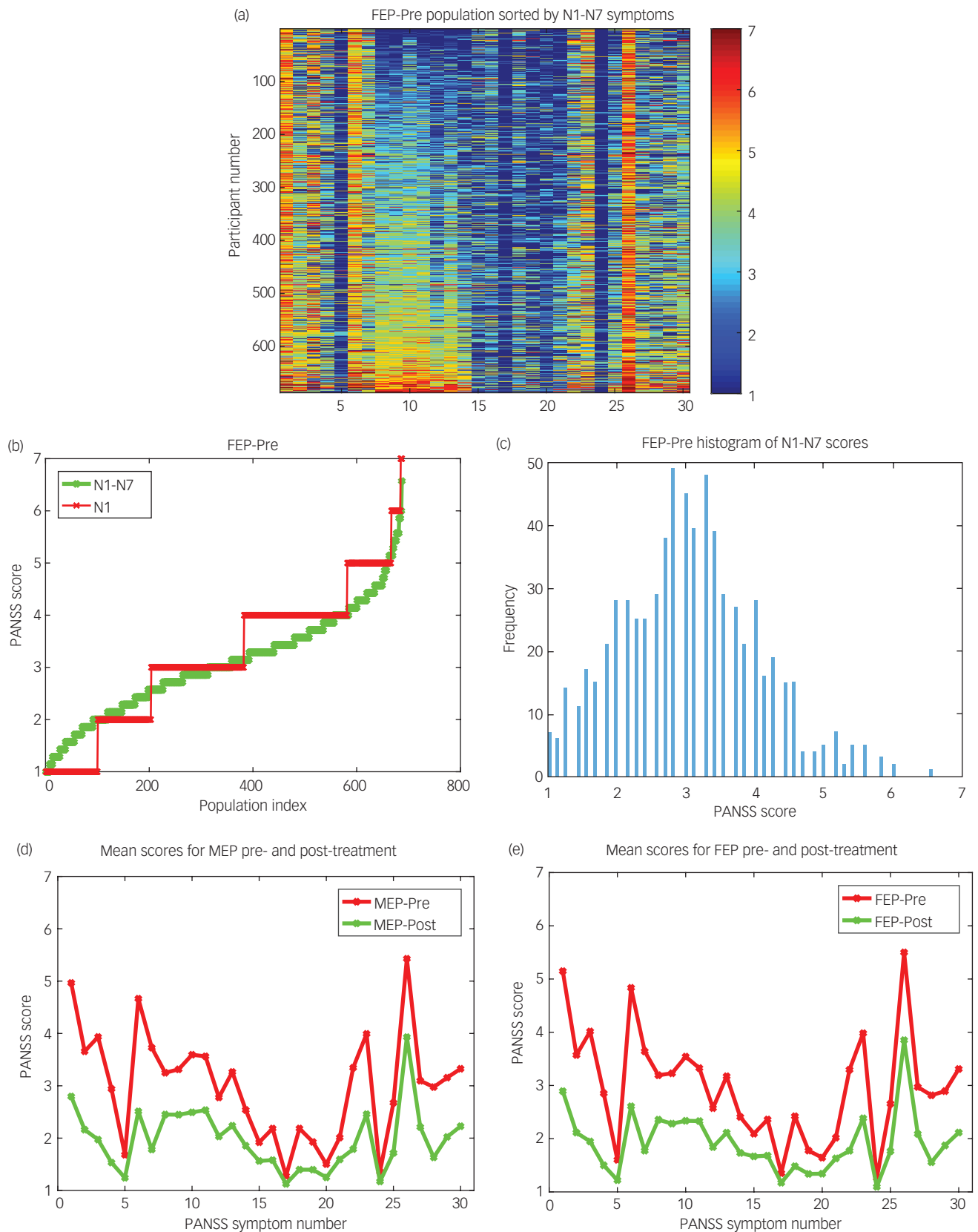
Fig. 4. (a) FEP_Pre group sorted by the average value of the negative symptoms, N1–N7 (Symptoms 8–14). The colour bar on the right shows the PANSS score in the range 1–7. (b) The average score for Symptoms N1–N7 shown sorted by its value in each member of the FEP_Pre population (green). The red plot shows the value for N1. (c) A histogram of the average score for Symptoms N1–N7 in the FEP_Pre group. (d and e) PANSS symptom scores before and after treatment for (d) the MEP and (e) the FEP groups.

supplement. These confirm that the major source of variation between individuals with schizophrenia is in the negative symptoms.

Discussion

The results of the analysis on the datasets with 687 FEP with schizophrenia before treatment (FEP_Pre) and after treatment (FEP_Post), and 1880 MEP before treatment (MEP_Pre) and after treatment (MEP_Post), provide evidence that a major difference between patients is the extent of the negative symptoms (Figs 1–4). There is little evidence for very discrete populations, with instead a gradation in the amount of the negative symptoms found across the patient populations (Figs 1–4). Indeed, the distributions for the mean of the negative symptoms N1–N7 are unimodal (Figs 2 and 4), so there is little support for discrete subtypes of schizophrenia, on the basis of PANSS scores. Indeed, when the k-means and the community detection algorithms did separate the patient populations into three groups, these groups were distinguished mainly by the amount of the negative symptoms, as illustrated in Figs 1, 3, S3, S4, S10 and S11. Figures 4d, e and S19 show that the PANSS symptoms are almost identical before treatment in the FEP and MEP groups. Figures 4d, e and S20 show that the PANSS symptoms are almost identical after treatment in the FEP and MEP groups. The results shown in Fig. 4d and e show that medication decreases the scores on essentially all PANSS symptom scores.

Our finding of a continuous unimodal distribution of negative symptoms is consistent with many cross-sectional studies that have described the heterogeneity of schizophrenia in terms of multiple dimensions (reviewed by Peralta & Cuesta[10]). Our findings indicate that the negative dimension plays a particularly important role in differentiating patients, but do not provide evidence for discrete subtypes of illness. However, our cross-sectional analyses do not address the question of whether or not taking account of persistence of symptoms or treatment response might help identify discrete subtypes. Some previous studies that have taken account of the tendency of symptoms to persist have identified a subtype of illness designated deficit schizophrenia, characterised by persistent primary negative symptoms.[17,18] Other studies, reviewed by Gillespie et al,[19] have concluded that treatment-resistant schizophrenia might be categorically distinct from treatment-responsive schizophrenia.

Evidence from many treatment trials, reviewed by Barry et al,[20] indicates that negative symptoms tend to be poorly responsive to treatment. However, we found in our very large patient sample that positive and negative symptoms reduced to a similar degree during treatment. An issue that must be considered is the characteristics of the patient sample, including selection criteria. It is illustrative to compare our findings with Tollefson et al[21] in a trial of olanzapine compared with haloperidol in a large sample of approximately 2000 patients. (Although the focus of Tollefson et al study was on the differences between responses to the two different antipsychotics, those differences were small and for the purpose of this discussion the differences between the antipsychotics is of limited relevance.) Tollefson et al reported a reduction of six points in PANSS negative symptom scores after 6 weeks of treatment with olanzapine and five points after 6-weeks treatment with haloperidol. We found a reduction in PANSS negative symptom scores of 8.2 points in MEP and 8.3 points in FEP cases. It is possible that the fact that we excluded cases with treatment resistant illness, whereas Tollefson et al did not, contributed to the difference in the magnitude of the reduction produced by antipsychotics. In addition, it is possible that social, cultural or ethnic influences contributed to the greater reduction in negative symptoms in our sample. We discuss this in greater detail in the Data supplement in the 'Social, cultural and ethnic influences on response to treatment' section.

It should be noted that we did not attempt to distinguish primary from secondary negative symptoms. Our finding of a significant correlation between the reduction in positive symptoms and the reduction in negative symptoms raises the possibility that a proportion of the negative symptoms reported in our study might be secondary negative symptoms which are expected to decrease during treatment. However, the results of an analysis of covariance presented in the Data supplement show that the reduction in the negative symptoms produced by the medication was in part independent of the reduction in the positive symptoms.

Furthermore, it is noteworthy that we found MEP and FEP cases exhibited similar decreases in symptom severity during 6-weeks treatment. In contrast, other moderately large studies have reported that FEP cases show a greater reduction in symptoms during treatment than MEP cases.[22] In a longitudinal study extending over the first 5 years of illness, Lieberman et al[23] found that treatment response diminished with increasing number of relapses. It is possible that the exclusion of participants with treatment-resistant illness from our MEP sample contributed to our finding of a relatively large reduction of negative symptoms in MEP cases.

Further analyses showed the following

The negative symptoms (N1–N7) are moderately well correlated with each other (Figs 1a and 3a). As demonstrated in the Data supplement, the results of applying the community detection algorithm fast_mo_sgn to the symptoms showed that in the MEP and FEP groups, both Pre- and Post-treatment, four symptom communities were detected (with one minor exception). One community generally included most of the positive symptoms and another community included most of the negative symptoms (Figs S1, S5, S8 and S12). The analysis of symptom communities identified a community of symptoms dominated by anxiety and depression (G6) in both FEP (Fig. S8) and MEP (Fig. S1) groups, both before and after treatment. This is consistent with many studies that have reported a depression syndrome in schizophrenia (see the review by Peralta & Cuesta[10]). Furthermore, in the MEP_Pre cases, there was a community of symptoms reflecting disorganisation of thought and behaviour, together with cognitive symptoms such as poor attention, resembling the disorganisation syndrome identified in symptomatic cases of chronic schizophrenia by Bilder et al[9] and Liddle.[7] Further analysis of just the negative symptoms found that they could be separated into three communities (N1, N2 and N4; N3 and N6; and N5 and N7), and that these symptom communities were all similarly reduced by medication (Data supplement).

The results shown in Figs S19 and S20 show that the PANSS symptoms are very similar between FEP and MEP groups without medication; and that the medication produces changes in both groups which make the FEP_Post and MEP_Post groups very similar after medication. These findings have major implications for 'staging' hypotheses of schizophrenia. Any 'staging' hypothesis would need to carefully factor out the effects of treatment for its results to be interpretable. It may be that very long-term patients with schizophrenia do show symptom differences.[5] But if so, that would be related to the long-term nature of the disease not clearly evident in the MEP and FEP groups described here.

K-means cluster analysis can be used to divide the patients into 2–4 sub-populations (with results for three sub-populations presented), which differ from each other mainly in the extent of the negative symptoms. Community detection analysis (using fast_mo_sgn)[16] (modified from[15]) provides a very similar set of three sub-populations of both the MEP and FEP group, though the numbers of patients in each population are not identical. To the extent that the distribution of the negative symptoms is in part continuous, the cluster/community detection may not imply that the sub-populations are very

discrete. Instead, a continuous distribution of the negative symptom scores from 1 to 7 across the population, as illustrated in Figs 1–4, needs to be considered. One way to assess patients clinically might be on the mean of the seven negative symptom PANSS scores, and clinicians may wish to consider different treatments with the graded distribution across the population in mind, with for example some findings becoming available from Lu *et al* (in preparation).[16]

The finding that the main difference in the PANSS scores of different patients is in the extent of the negative symptoms is of considerable interest. If a k-means analysis is performed using only the values of the negative symptoms, then three population communities are found which differ in the mean value of the negative symptoms (Figs S4 and S11). This adds to the evidence that the values of the negative symptom scores are an important factor in producing different sub-populations or communities, and moreover that what distinguishes these communities includes a difference in the mean negative symptom scores. Furthermore, the evidence in the Data supplement indicates that the negative symptoms are independent of the positive and general symptoms (i.e. the correlations between these symptoms pre-treatment is close to zero), so that there must be at least two partly independent underlying brain processes that are involved.

Subtypes of schizophrenia are no longer included in DSM-5, and this large study with unmedicated patients is of interest in that context, for no highly discrete subtypes of patients were detected in this study, but instead there was a continuous distribution of the mean value of the negative symptoms in both the FEP and MEP populations. The subtypes of symptoms illustrated in Figs S1, S5, S8 and S12 do show how a modern community detection algorithm clusters together different symptoms, with one symptom community including, for example, most of the positive symptoms, and a second community including most of the negative symptoms.

The present investigation adds to previous analyses of subtypes of schizophrenia (e.g.[11,24–26]) by using a very much larger sample of patients (2567) than in any previous investigation; by investigating the symptoms in the same patients in both the unmedicated and medicated states; by comparing the symptoms in FEP and MEP and by including a more powerful machine learning, community detection, approach. This investigation differs from the analysis of a partly overlapping FEP patient population by Lu *et al* (in preparation)[16] in that we used here the same participants unmedicated and medicated to understand the effects of the medication, which may produce clearer results than comparing two different groups when unmedicated or not; in the use of more classical clustering approaches to define how the symptoms cluster and how the members of the population cluster; and in the emphasis on the continuous unimodal distribution of the negative symptoms which differs between individuals, instead of focusing on communities that are thought to be distinct because they have been categorised separately by a community detection algorithm which looks for distinct communities.[15]

Apart from the large sample size in this study, a strength is that the patients came from several centres, and that similar results were obtained from the different centers.[16] Another strength is assessment of the PANSS when unmedicated and after 6 weeks of medication in the same patients. Several limitations may be mentioned. PANSS data were used to identify the communities in the present investigation and the communities may become further refined when further measures of differences between patients are included. Some PANSS items, such as difficulty in abstract thinking (N5) and poor attention (G11), include what may be different aspects of impoverished and disorganised thinking within single items, and separate assessment might lead to a richer categorisation.[7,27] Further we suggest that the classification of possible subtypes of schizophrenia could be enhanced by also including information from functional neuroimaging and from genetics.

In conclusion, in a very large sample of cases of schizophrenia, we have demonstrated that variation in the negative symptoms is an important aspect of the variation between patients with schizophrenia. The severity of the negative symptoms has a continuous unimodal distribution supporting a dimensional description of the heterogeneity of the illness, at least with respect to the negative symptoms. Furthermore, in this sample from which treatment resistant cases were excluded, the severity of the negative symptoms decreased substantially during 6 weeks of antipsychotic treatment in both FEP and MEP cases.

Edmund T. Rolls, MA, DPhil, DSc, Department of Computer Science, University of Warwick, Coventry, UK; Oxford Centre for Computational Neuroscience, Oxford, UK; **Wenlian Lu**, PhD, Centre for Computational Systems Biology, Fudan University, Shanghai, PR China; **Lin Wan**, PhD, National Center for Mathematics and Interdisciplinary Sciences, The Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, PR China; **Hao Yan**, MD, Institute of Mental Health, the Sixth Hospital, Peking University, Beijing, PR China; Key Laboratory of Mental Health, Ministry of Health & National Clinical Research Center for Mental Disorders, Peking University, Beijing, PR China; **Chuanyue Wang**, MD, Beijing Anding Hospital, Capital Medical University, Beijing, PR China; **Fude Yang**, M.M., Beijing HuiLongGuan Hospital, Peking University, Beijing, PR China; **Yunlong Tan**, Beijing HuiLongGuan Hospital, Peking University, Beijing, PR China; **Lingjiang Li**, MD, Institute of Mental Health, The Second Xiangya Hospital of Central South University, Changsha, PR China; **Chinese Schizophrenia Collaboration Group** (see Data supplement); **Hao Yu**, PhD, Institute of Mental Health, the Sixth Hospital, Peking University, Beijing, PR China; Key Laboratory of Mental Health, Ministry of Health & National Clinical Research Center for Mental Disorders, Peking University, Beijing, PR China; **Peter F. Liddle**, MBBS, MRCPsych, PhD, Centre for Translational Neuroimaging, Institute of Mental Health, Division of Psychiatry & Applied Psychology, University of Nottingham, Nottingham, UK; Sir Peter Mansfield MR Centre, University of Nottingham, Nottingham, UK; **Lena Palaniyappan**, MBBS, PhD, Department of Psychiatry, University of Western Ontario, London, Ontario, Canada; Department of Medical Biophysics, University of Western Ontario, London, Ontario, Canada; Robarts & Lawson Health Research Institutes, London, Ontario, Canada; **Dai Zhang**, MD, Institute of Mental Health, the Sixth Hospital, Peking University, Beijing, PR China; Key Laboratory of Mental Health, Ministry of Health & National Clinical Research Center for Mental Disorders, Peking University, Beijing, PR China; Peking-Tsinghua Joint Center for Life Sciences/PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China; **Weihua Yue**, MD, Institute of Mental Health, the Sixth Hospital, Peking University, Beijing, PR China; Key Laboratory of Mental Health, Ministry of Health & National Clinical Research Center for Mental Disorders, Peking University, Beijing, PR China; **Jianfeng Feng**, PhD, Department of Computer Science, University of Warwick, Coventry, UK

Correspondence: Jianfeng Feng, Centre for Computational Systems Biology, School of Mathematical Sciences, Fudan University, Shanghai, China; Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK. E-mail: jianfeng64@gmail.com; Professor Weihua Yue, The Sixth Hospital, Peking University, Beijing, 100191, China. Email: dryue@bjmu.edu.cn; Edmund Rolls, University of Warwick, Coventry, UK. Email: Edmund.Rolls@warwick.ac.uk.

First received 4 May 2017, final revision 31 Jul 2017, accepted 23 Sep 2017

References

1 Tsuang MT, Lyons MJ, Faraone SV. Heterogeneity of schizophrenia. Conceptual models and analytic strategies. *Br J Psychiatry* 1990; **156**: 17–26.

2 Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–7.

3 Rolls ET. *Cerebral Cortex: Principles of Operation*. Oxford University Press, 2016.

4 Rolls ET, Loh M, Deco G, Winterer G. Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. *Nat Rev Neurosci* 2008; **9**: 696–709.

5 Li T, Wang Q, Zhang J, Rolls ET, Yang W, Palaniyappan L, et al. Brain-wide analysis of functional connectivity in first-episode and chronic stages of schizophrenia. *Schizophr Bull* 2017; **43**: 436–48.

6 Liddle PF. Syndromes of schizophrenia on factor analysis (letter). *Br J Psychiatry* 1992; **161**: 861.

7 Liddle PF. The symptoms of chronic schizophrenia. A re-examination of the positive–negative dichotomy. *Br J Psychiatry* 1987; **151**: 145–51.

8 Owen MJ, Sawa A, Mortensen PB. Schizophrenia. *Lancet* 2016; **388**: 86–97.

9 Bilder RM, Mukherjee S, Rieder RO, Pandurangi AK. Symptomatic and neuropsychological components of defect states. *Schizophr Bull* 1985; **11**: 409–19.

10 Peralta V, Cuesta MJ. A dimensional and categorical architecture for the classification of psychotic disorders. *World Psychiatry* 2007; **6**: 100–1.

11 Strauss JS, Bartko JJ, Carpenter WT Jr. The use of clustering techniques for the classification of psychiatric patients. *Br J Psychiatry* 1973; **122**: 531–40.

12 Newman ME, Girvan M. Finding and evaluating community structure in networks. *Physical Rev E* 2004; **69**: 026113.

13 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM-IV). APA, 1984.

14 Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987; **13**: 261–76.

15 Le Martelot E, Hankin C. Fast multi-scale detection of relevant communities in large-scale networks. *Comput J* 2013; **56**: 1136–50.

16 Lu W, Wan L, Rolls ET, Liddle PF, Ma L, Yan H, et al. Novel subtyping of schizophrenia predicts response to antipsychotics. 2018. [paper in preparation]

17 Kirkpatrick B, Galderisi S. Deficit schizophrenia: an update. *World Psychiatry* 2008; **7**: 143–7.

18 Ahmed AO, Strauss GP, Buchanan RW, Kirkpatrick B, Carpenter WT. Are negative symptoms dimensional or categorical? Detection and validation of deficit schizophrenia with taxometric and latent variable mixture models. *Schizophr Bull* 2015; **41**: 879–91.

19 Gillespie AL, Samanaite R, Mill J, Egerton A, MacCabe JH. Is treatment-resistant schizophrenia categorically distinct from treatment-responsive schizophrenia? A systematic review. *BMC Psychiatry* 2017; **17**: 12.

20 Barry SJ, Gaughan TM, Hunter R. Schizophrenia. *BMJ Clin Evid* Jun 2012 (http://clinicalevidence.bmj.com/x/systematic-review/1007/overview.html). Accessed 26 April 2017.

21 Tollefson GD, Beasley CM Jr, Tran PV, Street JS, Krueger JA, Tamura RN, et al. Olanzapine versus haloperidol in the treatment of schizophrenia and schizoaffective and schizophreniform disorders: results of an international collaborative trial. *Am J Psychiatry* 1997; **154**: 457–65.

22 Jager M, Riedel M, Messer T, Laux G, Pfeiffer H, Naber D, et al. Psychopathological characteristics and treatment response of first episode compared with multiple episode schizophrenic disorders. *Eur Arch Psychiatry Clin Neurosci* 2007; **257**: 47–53.

23 Lieberman JA, Koreen AR, Chakos M, Sheitman B, Woerner M, Alvir JM, et al. Factors influencing treatment response and outcome of first-episode schizophrenia: implications for understanding the pathophysiology of schizophrenia. *J Clin Psychiatry* 1996; **57** (Suppl 9): 5–9.

24 Carpenter WT Jr, Stephens JH. An attempted integration of information relevant to schizophrenic subtypes. *Schizophr Bull* 1979; **5**: 490–506.

25 Carpenter WT Jr, Bartko JJ, Carpenter CL, Strauss JS. Another view of schizophrenia subtypes. A report from the international pilot study of schizophrenia. *Arch Gen Psychiatry* 1976; **33**: 508–16.

26 Strauss JS, Carpenter WT Jr, Bartko JJ. The diagnosis and understanding of schizophrenia. Part III. Speculations on the processes that underlie schizophrenic symptoms and signs. *Schizophr Bull* 1974; **11**: 61–9.

27 Baxter RD, Liddle PF. Neuropsychological deficits associated with schizophrenic syndromes. *Schizophr Res* 1998; **30**: 239–49.