

Polygenic dissection of the bipolar phenotype

M. L. Hamshere, M. C. O'Donovan, I. R. Jones, L. Jones, G. Kirov, E. K. Green, V. Moskvina, D. Grozeva, N. Bass, A. McQuillin, H. Gurling, D. St Clair, A. H. Young, I. N. Ferrier, A. Farmer, P. McGuffin, P. Sklar, S. Purcell, P. A. Holmans, M. J. Owen and N. Craddock

Background

Recent data provide strong support for a substantial common polygenic contribution (i.e. many alleles each of small effect) to genetic susceptibility for schizophrenia and overlapping susceptibility for bipolar disorder.

Aims

To test hypotheses about the relationship between schizophrenia and psychotic types of bipolar disorder.

Method

Using a polygenic score analysis to test whether schizophrenia polygenic risk alleles, en masse, significantly discriminate between individuals with bipolar disorder with and without psychotic features. The primary sample included 1829 participants with bipolar disorder and the replication sample comprised 506 people with bipolar disorder.

Results

The subset of participants with Research Diagnostic Criteria

schizoaffective bipolar disorder ($n=277$) were significantly discriminated from the remaining participants with bipolar disorder ($n=1552$) in both the primary ($P=0.00059$) and the replication data-sets ($P=0.0070$). In contrast, those with psychotic bipolar disorder as a whole were not significantly different from those with non-psychotic bipolar disorder in either data-set.

Conclusions

Genetic susceptibility influences at least two major domains of psychopathological variation in the schizophrenia–bipolar disorder clinical spectrum: one that relates to expression of a 'bipolar disorder-like' phenotype and one that is associated with expression of 'schizophrenia-like' psychotic symptoms.

Declaration of interest

None.

The nosological relationships between schizophrenia, bipolar disorder and mixed forms of illness (particularly schizoaffective disorder) have been the subject of substantial interest and debate since Kraepelin proposed his well-known dichotomy at the end of the 19th century.^{1–11} However, the dichotomy continues to be reflected prominently in recent operational descriptive classifications, including the Research Diagnostic Criteria (RDC),¹² DSM–IV¹³ and ICD–10.¹⁴ A recent large-scale collaborative study by the International Schizophrenia Consortium (ISC) using genome-wide association data (74 000 polymorphisms were used, typed in more than 6900 individuals) provided strong support for a substantial polygenic contribution to schizophrenia that was estimated to explain at least a third of the total variation in liability.¹⁵ The basic principle of that analysis was that a set of many alleles that discriminated case status in one schizophrenia case–control sample also significantly discriminated case status in an independent schizophrenia case–control sample. In the current analysis we use a schizophrenia case–control sample to undertake a polygenic score analysis to explore some basic aspects of the nosological structure of bipolar disorder. The results inform understanding of the nosology of the clinical spectrum of mood and psychotic illness.

Method

Description of the sample

Our data comprised the bipolar disorder sample from the Wellcome Trust Case-Control Consortium (WTCCC).¹⁶ The independent replication data is from the University College London (UCL) bipolar disorder sample collected in the UK.^{17,18} A total of 39 individuals with bipolar disorder were excluded from the WTCCC sample as they, or family members, were also included in the smaller UCL sample (a known overlap of the data

was taken into account in previous analyses¹⁸). Therefore the sample size differs from previous publications of the WTCCC data.^{16,19–21} All samples have been subjected to strict quality assessment, details of which can be found in the original publications.^{15–17}

WTCCC bipolar disorder sample

A detailed description of the WTCCC bipolar disorder sample has been provided elsewhere.¹⁶ All individuals were from the UK and over the age of 16 years. Clinical assessment included semi-structured interview and review of case notes. Ratings of symptom occurrence and course of illness were made including the operational criteria (OPCRIT) item checklist.^{22,23} Diagnoses were based on all available data. The primary diagnostic system used for classifying participants was the Research Diagnostic Criteria (RDC)¹² because it provides greater differentiation between individuals on the basis of the pattern of mood and psychotic symptomatology than do the DSM–IV¹³ or ICD–10.¹⁴ Participants with bipolar disorder had experienced at least one episode of clinically significant elevated mood according to RDC: bipolar I disorder ($n=1283$), schizoaffective disorder, bipolar type ($n=277$), bipolar II disorder ($n=169$) and manic disorder ($n=100$).

Individuals were rated for the lifetime occurrence of psychosis. This was done using available data that had been collected at the time of original recruitment into the genetic studies. These included the OPCRIT item checklist^{22,23} and the Bipolar Affective Disorder Dimension Scale (BADDS).²⁴ Lifetime presence of definite psychosis refers to the unambiguous presence of delusions and/or hallucinations on at least one occasion during a person's lifetime and was rated as definitely present ($n=1192$), definitely absent ($n=235$) or unknown ($n=402$). Participants with insufficient available clinical information were scored as missing data. Further information is available in the online supplement.

UCL bipolar disorder sample

These data have been previously analysed together with the US Systematic Enhancement Program for Bipolar Disorders (STEP-BD) samples.^{17,18} A detailed description of this sample can be found in the original publication.¹⁷ Of the 506 participants with bipolar disorder, 74 had experienced symptoms of schizoaffective, bipolar type and 409 had not. The number of individuals experiencing psychotic symptoms was 375, and 117 individuals were non-psychotic. Participants were interviewed by a trained researcher using the Schedules for Affective Disorders and Schizophrenia, Lifetime Version (SADS-L)²⁵ psychiatric interview, the OPCRIT^{22,23} checklist was completed and diagnoses assigned according to RDC.¹²

Genotype data

The set of schizophrenia ‘score’ alleles used to derive the polygenic scores were provided by the ISC and are described in their paper.¹⁵ The WTCCC bipolar disorder data-set comprised 469 557 single nucleotide polymorphisms (SNPs) distributed across the genome. All individual SNP genotypes were obtained through the same analysis pipeline. For the current analysis we selected 377 742 autosomal SNPs that passed stringent quality filters (as described in Moskvin *et al.*²¹ see online supplement for more detail). The UCL bipolar data-set comprised 286 785 SNPs that also passed the quality-control filtering of the WTCCC data (as described above) and the quality control filtering described in Sklar *et al.*¹⁷ (We excluded the G/C and A/T SNPs for which strand alignment was unknown.)

Statistical methods

In general we followed the statistical approach described in the ISC paper.¹⁵ We used the published ISC data as the discovery set to define the score alleles and our WTCCC bipolar disorder sample as the primary target set. We made comparisons of the distributions of the polygenic scores between participants with bipolar disorder with schizoaffective features (the schizoaffective subset) from those without schizoaffective features (non-schizoaffective subset). We then sought to replicate our polygenicity findings in an independent replication target set: the UCL bipolar sample. Having provided this general orientation to the analysis, we will now describe the procedures in more detail.

First we selected the same set of SNPs that were defined in the ISC sample¹⁵ to be in relative linkage equilibrium. (Linkage disequilibrium refers to the correlation that occurs between polymorphisms that lie close together and which, therefore, produces a redundancy of information. Linkage equilibrium refers to the situation where there is no such correlation.) We used the schizophrenia score alleles defined by the ISC discovery sample from comparing participants with schizophrenia with controls.¹⁵ For each SNP we identified the corresponding *P*-value and allelic odds ratio. We also identified which allele was present in the schizophrenia group more frequently than in the controls,

focusing on SNPs significant with $P < 0.5$. We termed these the ‘score’ alleles. A threshold of $P < 0.5$ provided the optimal case–control discrimination of polygenic scores in the original report using this schizophrenia discovery data¹⁵ and also showed good discrimination within our own sample when we tested a range of *P*-value thresholds (see online supplement). We defined our primary target data to be the genotype data for the WTCCC bipolar disorder sample. For each individual in the target data, we obtained the mean per-SNP product of the number of score alleles (as defined in the discovery data) and the log_e odds ratio with the analysis software, PLINK v1.06 (<http://pngu.mgh.harvard.edu/~purcell/plink/index.shtml>) run on Solaris 10.5 × 86_64;²⁶ we called this the polygenic score.¹⁵

To replicate the results of polygenic analyses observed in the WTCCC bipolar disorder sample, we also investigated the UCL bipolar disorder sample. As above, we analysed the same SNPs (where available) in relative linkage equilibrium and selected only those that attained a $P < 0.5$ in the ISC schizophrenia case–control sample. We then created polygenic scores in the UCL bipolar sample.

We used logistic regression to compare the polygenic scores in two phenotypically defined subsets of individuals with bipolar disorder and to test specific hypotheses. We expected schizophrenia-defined polygenic scores to be higher in those individuals with schizoaffective or psychotic bipolar disorder when compared with the remaining participants with bipolar disorders, so therefore we use the one-sided alternative hypothesis. One-tailed *P*-values are presented throughout.

Results

Of the 74 062 independent SNPs identified by the ISC, 71 064 were also available in our target data. First we defined the score alleles from the 36 708 (52.0%) SNPs that had $P < 0.5$. Within the WTCCC bipolar disorder sample, polygenic scores in the schizoaffective subset ($n = 277$) were significantly higher than those in the non-schizoaffective subset ($n = 1552$, $P = 0.00059$, Table 1). When we used the (narrower) DSM–IV definition of schizoaffective bipolar disorder we also observed a significant difference in polygenic score between those in the schizoaffective subset ($n = 97$) and those in the non-schizoaffective subset ($n = 1552$, $P = 0.014$). Further, when we considered only those participants with bipolar disorder meeting RDC criteria for schizoaffective bipolar disorder, there was no significant difference in polygenic score between the participants that also met criteria for DSM–IV schizoaffective bipolar disorder compared with those that did not (all of whom met DSM–IV criteria for bipolar I disorder, $P = 0.418$). This shows that the polygenic signal we observe in the RDC schizoaffective subset does not derive solely from those participants that also meet the DSM–IV definition of schizoaffective disorder (see later for discussion).

We next considered a dichotomous comparison of the total bipolar disorder sample based on the presence of lifetime

Table 1 Polygenic score analyses within the bipolar samples, the phenotype of interest is Research Diagnostic Criteria (RDC) schizoaffective bipolar disorder^a

Sample	<i>P</i>
Wellcome Trust Case-Control Consortium bipolar disorder sample	
RDC schizoaffective bipolar disorder subset ($n = 277$) v. non-schizoaffective bipolar disorder subset ($n = 1552$)	0.00059
All RDC schizoaffective bipolar disorder: DSM–IV schizoaffective bipolar disorder subsets ($n = 97$) v. DSM–IV bipolar I disorder subset ($n = 180$)	0.42
Replication University College London bipolar disorder sample	
Schizoaffective bipolar disorder subset ($n = 74$) v. non-schizoaffective bipolar disorder subset ($n = 409$)	0.0070

a. Comparison analyses were performed using logistic regression. All *P*-values are one-sided.

psychotic features. We found no significant difference ($P=0.092$) between polygenic scores when we compared the participants with bipolar disorder with a definite lifetime history of psychotic symptoms ($n=1192$) with those who definitely lacked such a history ($n=235$). The trend was towards higher polygenic scores in those with a definite lifetime history of psychosis.

To seek further support for our findings, we then turned to our replication target sample. We investigated the polygenic scores obtained in the UCL bipolar disorder sample. Of the 74 062 SNPs, 59 987 were available in the UCL data. As before, we define the score alleles from the 30 984 (51.7%) SNPs in the ISC study with $P<0.5$. The schizoaffective subset again differed significantly from the non-schizoaffective subset ($P=0.0070$). The dichotomous comparison psychosis present ($n=375$) *v.* psychosis absent ($n=117$) was non-significant ($P=0.232$).

Discussion

Main findings

Our main interest was to test for phenotypic structure in the bipolar disorder sample using the schizophrenia-derived polygenic score as the tool for exploration. We found that the score discriminated between those with schizoaffective bipolar disorder using the Research Diagnostic Criteria and the remaining participants with bipolar disorder ($P=0.00059$) and that this was replicated using an independent UK bipolar target data-set. In contrast, the schizophrenia-derived polygenic score did not significantly discriminate between those with bipolar disorder with and without psychosis.

Findings from other research

There is a burgeoning and increasingly robust body of evidence from diverse sources that points to a substantial overlap in genetic susceptibility to schizophrenia and bipolar disorder,^{27–29} including large, well-powered studies published recently.^{15,30–33} For example, the largest family study of schizophrenia and bipolar disorder ever undertaken, including over 2 million nuclear families identified from Swedish population and hospital discharge registers showed increased risks of both schizophrenia and bipolar disorder for first-degree relatives of probands with either disorder. Moreover, there was evidence from half-sibs and adopted-away relatives that this is substantially the result of genetic factors.³² Further evidence for overlap of genetic risk comes from the study of offspring in families where one parent is affected by bipolar disorder and the other affected by schizophrenia.³⁴ Large-scale collaborative genome-wide association studies (GWAS) that investigate hundreds of thousands of SNPs in large numbers of cases and controls, have started to deliver genome-wide significant genetic associations for bipolar disorder and schizophrenia and have provided evidence of overlapping genetic susceptibility of the diseases. Studies of approximately 10 000 individuals have shown strong evidence for association with susceptibility to bipolar disorder at variants within two genes involved in ion channel function: *ANKK3* (encoding the protein ankyrin G) and *CACNA1C* (encoding the alpha-1C subunit of the L-type voltage-gated calcium channel). The *CACNA1C* SNP showing maximum association with susceptibility to bipolar disorder showed similar association in UK schizophrenia and unipolar depression samples, indicating that variation at this locus influences susceptibility across the mood–psychosis spectrum.³¹ A similar study in close to 20 000 individuals has shown strong evidence for association with susceptibility to schizophrenia at a variant within *ZNF804A* (encoding a protein of unknown function but which, based on sequence similarity, may act as a

transcription factor).³³ Further, the SNP in *ZNF804A* showing the strongest association with schizophrenia also showed an association with bipolar disorder, demonstrating that variation at this locus also has an effect on illness susceptibility across the traditional diagnostic boundaries.³⁵ Similarly, gene-based analyses have demonstrated overlap in the genes implicated in susceptibility to both disorders.²¹ The data that we report in the current study are consistent with these recent findings. In particular the findings strongly support the existence of many shared genetic susceptibility loci (i.e. a substantial shared polygenic component).¹⁵ This supports the hypothesis that the same set of biological dysfunctions can contribute to susceptibility to a range of clinical phenotypes including prototypical schizophrenia and prototypical bipolar disorder.

Implications

It is of interest that in prior analyses of the WTCCC bipolar disorder data we observed that the RDC schizoaffective bipolar disorder diagnostic subset stood out from the other diagnostic subsets (RDC bipolar I disorder, bipolar II disorder and manic disorder) as having a significantly greater number of strong ($P<10^{-5}$) association signals²⁰ and that variation at genes encoding GABA_A-receptor subunits is associated with risk of RDC schizoaffective bipolar disorder and that this risk is relatively specific to this diagnostic subset.^{19,36} The findings reported here, together with these prior findings, suggests that it may be important, at least from the viewpoint of biological research, to recognise and distinguish cases in which there is a mix of both bipolar and schizophrenia-like symptoms.

Research Diagnostic Criteria schizoaffective disorder, bipolar type is a relatively broad definition of ‘middle ground’ cases with features of both bipolar disorder and schizophrenia. In addition to manic episodes, the key requirement is that schizophrenia-like psychotic symptoms should have occurred, but there is no constraint that mood symptoms should be absent at the time. Thus, the emphasis is on the nature of the psychotic symptoms. This is in stark contrast to the DSM–IV approach where the key focus is the temporal relationships between symptoms, the requirement here being psychotic features occur at a time when prominent mood syndrome is absent; the nature of those psychotic symptoms is not constrained. In our subset of 277 participants with RDC schizoaffective bipolar disorder, 180 individuals did not meet the temporal criteria for DSM–IV schizoaffective bipolar disorder. There was no significant difference in polygenic scores between the two schizoaffective subsets (DSM–IV positive and DSM–IV negative, $P=0.418$). Thus, our data suggest that a clinical definition of ‘schizoaffective’ illness that aims to identify individuals with bipolar disorder with underlying similarities to schizophrenia should take account of the type of psychotic symptom (as does, for example, RDC) and not focus solely on the temporal relationship between mood and psychotic symptoms (as does, for example, DSM–IV).

We do not see significant discrimination between those with bipolar disorder in the with- and without-psychosis subsets using the schizophrenia-derived polygenic score. In our data we observed a trend in the direction of larger polygenic scores in those with psychosis so it is possible that with larger samples, significant effects may be observed. However, it is clear that this simple clinical distinction does not readily capture the polygenic similarity with schizophrenia. In contrast, the observation of a significant distinction between RDC schizoaffective and non-schizoaffective bipolar disorder subsets indicates that the nature of psychotic symptoms, rather than simply the presence of psychotic symptoms, is important. This suggests that some

alleles that influence risk of schizophrenia also influence the nature of the psychotic symptoms in bipolar disorder, but not necessarily the occurrence of psychotic symptoms *per se*. However, it should be noted that in comparison with controls, even those in the non-schizoaffective subset carry a significant excess of schizophrenia ‘score’ alleles ($P < 0.001$; data not shown). That observation is not consistent with a simplistic model where schizophrenia risk alleles predispose to a single schizophrenia-like form of bipolar disorder (see online supplement). Instead, our findings point to the existence of genetically influenced phenotypic complexity, with at least two genetically influenced psychopathological domains in those with bipolar disorder: one of which relates to expression of a ‘bipolar disorder’ phenotype (i.e. phenotypic characteristics that will increase the likelihood that an individual will meet criteria for a diagnosis of bipolar disorder) and one that influences the expression of ‘schizophrenia-like psychosis’. We do not suggest that this domain is of fundamental validity (i.e. we do not wish to suggest that there are bipolar genes and schizophrenia-like genes); the important point is that our data point to partly independent domains of psychopathology that happen to be captured to some extent by these broad labels. It is likely that these domains could be usefully further subdivided, and that they may also overlap in genetic susceptibility. Moreover, there will almost certainly be other domains that can be teased apart by approaches such as described here. Further studies, preferably with large samples, will be needed to explore this further.

We have drawn attention to differences in the psychosis-related clinical characteristics of those with high and low scores on a schizophrenia-trained polygenic score, the aim being to inform understanding of nosological structure. However, we should stress that the statistical significance seen in the comparisons are driven by large sample sizes, not large effect sizes and as in the ISC study, the proportion of variance currently explained is negligible ($< 1\%$ of the variance). As such these types of analyses are not currently clinically useful as a ‘test’ for diagnosis or for risk discrimination. As an increasing proportion of the common genetic variation is accurately captured through increased sample sizes and higher density genome coverage, it may be possible to explain $> 30\%$ of the variance. This approach will, therefore, become an increasingly useful research, and even potentially a useful clinical, tool.¹⁵

Strengths and limitations

The limitations of our analysis include those inherent in all genetic studies in psychiatry. Our bipolar disorder sample is large but for the effect sizes observed, it is desirable to have access to substantially larger samples, of the order of 10s of 1000s rather than 1000s. Such samples will be available in the near future within the context of the Psychiatric GWAS Consortium.^{37,38} A further limitation is that measurement of psychopathology is neither straightforward nor without error, and therefore our clinical analyses are limited to relatively broad categorisations.

The results we present here are robust to differences in the precise methodology used to derive and apply the polygenic score (see online supplement), which gives confidence that our findings reflect basic properties of the data. We note that population stratification, a potential confounder in case–control studies, is not a likely explanation for our findings for several reasons. First, analyses that take into account principal components of our genotype data obtained from the analysis software, EIGENSTRAT (implemented as part of the EIGENSOFT version 2.0 (<http://genepath.med.harvard.edu/~reich/Software.htm>) run on Redhat (RHEL 5) $\times 86_64$),^{39,40} continue to show significant

differences (see online supplement). Second, extensive analyses within the original ISC excluded population stratification as an explanation for the broad effects observed between cases and controls.¹⁵ Third, it is implausible that exactly the same stratification differences would occur between the case and comparison data-sets in both discovery and target samples. Further, we verified that there was no polygenic signal when we trained on non-psychiatric disease WTCCC data-sets that can be expected to be phenotypically unrelated (Crohn’s disease) to mood–psychotic illness. Thus, at least within the limits of our sample sizes and methodology, the significant effect observed in our data seems to be specific to our psychiatric data-sets.

In summary, we have used an analytic approach that considers the aggregate genetic association evidence across a very large set of common polymorphisms spread across the genome in order to gain insights into the nosological relationships within the clinical mood–psychosis spectrum. We found that genetic susceptibility influences at least two major domains of psychopathological variation in the schizophrenia–bipolar disorder clinical spectrum: one that relates to expression of a ‘bipolar disorder-like’ phenotype and one that is associated with expression of ‘schizophrenia-like’ psychotic symptoms. This analysis supports the move in classificatory thinking away from the traditional discrete dichotomous categories and towards approaches that better accommodate and recognise the common co-occurrence of both domains of variation. Using dimensions and recognising ‘middle ground’ categories, such as schizoaffective disorder, are both ways to achieve this.

M. L. Hamshere, PhD, **M. C. O’Donovan**, PhD, FRCPsych, **I. R. Jones**, PhD, MRCPsych, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, UK; **L. Jones**, PhD, Department of Psychiatry, University of Birmingham, National Centre for Mental Health, Birmingham, UK; **G. Kirov**, PhD, MRCPsych, **E. K. Green**, PhD, **V. Moskvina**, PhD, **D. Grozeva**, MSc, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, UK; **N. Bass**, MD, **A. McQuillin**, PhD, **H. Gurling**, MD, FRCPsych, Department of Mental Health Sciences, University College London, UK; **D. St Clair**, MD, PhD, University of Aberdeen, Institute of Medical Sciences, Aberdeen, UK; **A. H. Young**, PhD, FRCPsych, School of Neurology, Neurobiology and Psychiatry, Royal Victoria Infirmary, Newcastle upon Tyne, UK and UBC Institute of Mental Health, University Boulevard Vancouver, Canada; **I. N. Ferrier**, MD, FRCPsych, School of Neurology, Neurobiology and Psychiatry, Royal Victoria Infirmary, Newcastle upon Tyne, UK; **A. Farmer**, PhD, FRCPsych, **P. McGuffin**, PhD, FRCPsych, Institute of Psychiatry, King’s College London, UK; **P. Sklar**, MD, PhD, **S. Purcell**, PhD, Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetic Research, and Department of Psychiatry, Massachusetts General Hospital, Boston, and Stanley Center for Psychiatric Genetics, Broad Institute, Cambridge, Massachusetts, USA; **P. A. Holmans**, PhD, **M. J. Owen**, PhD, FRCPsych, **N. Craddock**, PhD, FRCPsych, MRCPsych, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, UK

Correspondence: Nick Craddock, MRC Centre for Neuropsychiatric Genetics and Genomics, School of Medicine, Cardiff University, Heath Park, Cardiff, CF14 4XN, UK. Email: craddockn@cardiff.ac.uk

First received 3 Sep 2010, final revision 21 Dec 2010, accepted 17 Jan 2011

Funding

Funding for recruitment and phenotype assessment has been provided by the Wellcome Trust and the Medical Research Council. This study makes use of data generated by the Wellcome Trust Case-Control Consortium. Funding for the project was provided by the Wellcome Trust under award 076113.

Acknowledgements

We are indebted to all individuals who have participated in, or helped with, our research, particularly those involved in the Bipolar Disorder Research Network (bdm.org). We thank MDF The Bipolar Organisation for the help of its staff and members. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk.

References

- 1 Marneros A, Akiskal HS. *The Overlap of Affective and Schizophrenic Spectra*. Cambridge University Press, 2006.

- 2 Crow TJ. The continuum of psychosis and its genetic origins. The sixty-fifth Maudsley lecture. *Br J Psychiatry* 1990; **156**: 788–97.
- 3 Kendell RE. Diagnosis and classification of functional psychoses. *Br Med Bull* 1987; **43**: 499–513.
- 4 Kendell RE, Brockington IF. The identification of disease entities and the relationship between schizophrenic and affective psychoses. *Br J Psychiatry* 1980; **137**: 324–31.
- 5 Brockington IF, Kendell RE, Wainwright S, Hillier VF, Walker J. The distinction between the affective psychoses and schizophrenia. *Br J Psychiatry* 1979; **135**: 243–8.
- 6 Taylor MA. Are schizophrenia and affective disorder related? A selective literature review. *Am J Psychiatry* 1992; **149**: 22–32.
- 7 Van Os J, Gilvarry C, Bale R, Van Horn E, Tattan T, White I, et al. A comparison of the utility of dimensional and categorical representations of psychosis. *Psychol Med* 1999; **29**: 595–606.
- 8 Kendler KS, Karkowski LM, Walsh D. The structure of psychosis: latent class analysis of probands from the Roscommon Family Study. *Arch Gen Psychiatry* 1998; **55**: 492–9.
- 9 Murray RM, Sham P, Van Os J, Zanelli J, Cannon M, McDonald C. A developmental model for similarities and dissimilarities between schizophrenia and bipolar disorder. *Schizophr Res* 2004; **71**: 405–16.
- 10 Angst J, Scharfetter C, Stassen HH. Classification of schizo-affective patients by multidimensional scaling and cluster analysis. *Psychiatr Clin (Basel)* 1983; **16**: 254–64.
- 11 Brockington IF, Meltzer HY. The nosology of schizoaffective psychosis. *Psychiatr Dev* 1983; **1**: 317–38.
- 12 Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability. *Arch Gen Psychiatry* 1978; **35**: 773–82.
- 13 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (4th edn) (DSM-IV)*. APA, 1994.
- 14 World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. WHO, 1992.
- 15 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–52.
- 16 Wellcome Trust Case-Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007; **447**: 661–78.
- 17 Sklar P, Smoller JW, Fan J, Ferreira MA, Perlis RH, Chambert K, et al. Whole-genome association study of bipolar disorder. *Mol Psychiatry* 2008; **13**: 558–69.
- 18 Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, et al. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet* 2008; **40**: 1056–8.
- 19 Craddock N, Jones L, Jones IR, Kirov G, Green EK, Grozeva D, et al. Strong genetic evidence for a selective influence of GABA(A) receptors on a component of the bipolar disorder phenotype. *Mol Psychiatry* 2008; **15**: 146–53.
- 20 Hamshere ML, Green EK, Jones IR, Jones L, Moskvina V, Kirov G, et al. Genetic utility of broadly defined bipolar schizoaffective disorder as a diagnostic concept. *Br J Psychiatry* 2009; **195**: 23–9.
- 21 Moskvina V, Craddock N, Holmans P, Nikolov I, Pahwa JS, Green E, et al. Gene-wide analyses of genome-wide association data sets: evidence for multiple common risk alleles for schizophrenia and bipolar disorder and for overlap in genetic risk. *Mol Psychiatry* 2009; **14**: 252–60.
- 22 McGuffin P, Farmer A, Harvey I. A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system. *Arch Gen Psychiatry* 1991; **48**: 764–70.
- 23 Craddock M, Asherson P, Owen MJ, Williams J, McGuffin P, Farmer AE. Concurrent validity of the OPCRIT diagnostic system. Comparison of OPCRIT diagnoses with consensus best-estimate lifetime diagnoses. *Br J Psychiatry* 1996; **169**: 58–63.
- 24 Craddock N, Jones I, Kirov G, Jones L. The Bipolar Affective Disorder Dimension Scale (BADDs) – a dimensional scale for rating lifetime psychopathology in bipolar spectrum disorders. *BMC Psychiatry* 2004; **4**: 19.
- 25 Endicott J, Spitzer RL. A diagnostic interview: the schedule for affective disorders and schizophrenia. *Arch Gen Psychiatry* 1978; **35**: 837–44.
- 26 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–75.
- 27 Berrettini W. Bipolar disorder and schizophrenia: not so distant relatives? *World Psychiatry* 2003; **2**: 68–72.
- 28 Cardno AG, Rijdsdijk FV, Sham PC, Murray RM, McGuffin P. A twin study of genetic relationships between psychotic symptoms. *Am J Psychiatry* 2002; **159**: 539–45.
- 29 Craddock N, Owen MJ. Rethinking psychosis: the disadvantages of a dichotomous classification now outweigh the advantages. *World Psychiatry* 2007; **6**: 84–91.
- 30 Craddock N, O'Donovan MC, Owen MJ. Psychosis genetics: modeling the relationship between schizophrenia, bipolar disorder, and mixed (or 'schizoaffective') psychoses. *Schizophr Bull* 2009; **35**: 482–90.
- 31 Green EK, Grozeva D, Jones I, Jones L, Kirov G, Caesar S, et al. The bipolar disorder risk allele at CACNA1C also confers risk of recurrent major depression and of schizophrenia. *Mol Psychiatry* 2010; **15**: 1016–22.
- 32 Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* 2009; **373**: 234–9.
- 33 O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V, et al. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* 2008; **40**: 1053–5.
- 34 Gottesman II, Laursen TM, Bertelsen A, Mortensen PB. Severe mental disorders in offspring with 2 psychiatrically ill parents. *Arch Gen Psychiatry* 2010; **67**: 252–7.
- 35 Williams HJ, Norton N, Dwyer S, Moskvina V, Nikolov I, Carroll L, et al. Fine mapping of ZNF804A and genome-wide significant evidence for its involvement in schizophrenia and bipolar disorder. *Mol Psychiatry* 2010; Apr 6 (Epub ahead of print).
- 36 Breuer R, Hamshere ML, Strohmaier J, Mattheisen M, Degenhardt F, Meier S, et al. Independent evidence for the selective influence of GABA(A) receptors on one component of the bipolar disorder phenotype. *Mol Psychiatry* 2010; Jun 15 (Epub ahead of print).
- 37 Cichon S, Craddock N, Daly M, Faraone SV, Gejman PV, Kelsoe J, et al. Genomewide association studies: history, rationale, and prospects for psychiatric disorders. *Am J Psychiatry* 2009; **166**: 540–56.
- 38 Craddock N, Kendler K, Neale M, Nurnberger J, Purcell S, Rietschel M, et al. Dissecting the phenotype in genome-wide association studies of psychiatric illness. *Br J Psychiatry* 2009; **195**: 97–9.
- 39 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**: 904–9.
- 40 Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet* 2006; **2**: e190.

