npg

ORIGINAL ARTICLE

Common psychiatric disorders share the same genetic origin: a multivariate sibling study of the Swedish population

E Pettersson, H Larsson and P Lichtenstein

Recent studies have shown that different mental-health problems appear to be partly influenced by the same set of genes, which can be summarized by a general genetic factor. To date, such studies have relied on surveys of community-based samples, which could introduce potential biases. The goal of this study was to examine whether a general genetic factor would still emerge when based on a different ascertainment method with different biases from previous studies. We targeted all adults in Sweden ($n = 3 475 112$) using national registers and identified those who had received one or more psychiatric diagnoses after seeking or being forced into mental health care. In order to examine the genetic versus environmental etiology of the general factor, we examined whether participants' full- or half-siblings had also received diagnoses. We focused on eight major psychiatric disorders based on the *International Classification of Diseases*, including schizophrenia, schizoaffective disorder, bipolar disorder, depression, anxiety, attention-deficit/hyperactivity disorder, alcohol use disorder and drug abuse. In addition, we included convictions of violent crimes. Multivariate analyses demonstrated that a general genetic factor influenced all disorders and convictions of violent crimes, accounting for between 10% (attention-deficit/hyperactivity disorder) and 36% (drug abuse) of the variance of the conditions. Thus, a general genetic factor of psychopathology emerges when based on both surveys as well as national registers, indicating that a set of pleiotropic genes influence a variety of psychiatric disorders.

Molecular Psychiatry advance online publication, 25 August 2015; doi:10.1038/mp.2015.116

INTRODUCTION

Although psychiatric disorders are treated as distinct entities in clinical praxis,[1] comorbidity is the rule rather than exception in psychiatry.[2] For example, in a representative sample of over 9000 individuals in the United States, over half of those who met criteria for one disorder in a given year also met criteria for another disorder.[3] This extensive overlap may indicate that ostensibly discrete disorders partly share similar origins. Two broad factors, internalizing and externalizing, have been found to explain much of the co-occurrence among psychopathologies.[2,4,5] The internalizing dimension puts individuals at risk for depression and anxiety, and the externalizing dimension puts individuals at risk for drug abuse and conduct problems.

More recently, it has been observed that because the internalizing and externalizing dimensions are also positively associated, it may be that an even broader dimension—a general factor of psychopathology—puts individuals at risk for developing all forms of psychopathologies.[6–9] Indeed, models of psychiatric disorders that include a general factor tend to fit data better, compared with models that do not include a general factor.[10] In a community-based study of over 30 000 individuals representative of the United States population, a model based on common mental disorders assessed through diagnostic interviews fit significantly better when a general factor was included.[7] This indicates that risk for mental illness may not only be shared across internalizing and externalizing forms of psychopathology, but also across all disorders to some extent.

Twin studies have demonstrated that the general factor of psychopathology primarily has a genetic origin.[8,9] This suggests that a set of pleiotropic genes influence a variety of disorders and

problems, which has implications for nosology and molecular genetic approaches. For example, aside from identifying genes that are specific to particular disorders, there may also be merit in searching for genetic risk variants that influence all forms of psychiatric diagnoses.

However, to date, all twin studies of the general factor have targeted the general population with a battery of questions about mental illness. Such survey studies could be biased. First, some may decline to participate (selection bias). For example, poorer and more psychiatrically troubled individuals appear less likely to participate in survey research.[11,12] Second, people may not report accurately about their symptoms (presentation bias). For example, retrospective evaluations may underestimate lifetime prevalence rates.[13,14]

The aim of the current study was to explore whether a general genetic factor would emerge when based on individuals in treatment, as identified by registers. Although register data are also vulnerable to biases (for example, not all troubled individuals may seek help and individuals with multiple diagnoses are more likely to come into contact with the mental health system), these differ from those of survey designs. Thus, if different study designs come to similar conclusions, it will strengthen the hypothesis that a general genetic factor is important in psychiatry.

MATERIALS AND METHODS

Participants

Using personal identification numbers unique to each individual living in Sweden, we created a population-based cohort by linking together several longitudinal population registers. The National Patient Register captures all

Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. Correspondence: Dr E Pettersson, Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, Stockholm 17177, Sweden.
E-mail: erik.pettersson@ki.se

Table 1. Frequencies of diagnoses

Diagnoses	Full siblings			Maternal half-siblings			Paternal half-siblings		
	No	Yes	Percent (%)	No	Yes	Percent (%)	No	Yes	Percent (%)
Schizophrenia	2 921 658	11 428	0.39	258 145	1285	0.50	281 141	1455	0.51
Schizoaffective	2 929 627	3459	0.12	259 093	337	0.13	282 186	410	0.15
Bipolar	2 916 456	16 630	0.57	257 616	1814	0.70	280 695	1901	0.67
Depression	2 856 451	76 635	2.61	249 161	10 269	3.96	272 147	10 449	3.70
Anxiety	2 827 955	105 131	3.58	243 888	15 542	5.99	267 061	15 535	5.50
Attention-deficit/hyperactivity disorder (ADHD)	2 923 667	9419	0.32	256 937	2493	0.96	280 281	2315	0.82
Drug abuse	2 889 568	43 518	1.48	249 646	9784	3.77	273 239	9357	3.31
Alcohol abuse	2 845 565	87 521	2.98	245 008	14 422	5.56	268 716	13 880	4.91
Convictions of violent crimes	2 835 889	97 197	3.31	237 892	21 538	8.30	261 902	20 694	7.32

public psychiatric inpatient admissions in Sweden since 1973 and outpatient diagnoses since 2001, classified by the attending physician using a non-hierarchical diagnostic structure in accordance with version eight (1969–1986), nine (1987–1996) or ten (1997–present) of the International Classification of Diseases[15] (ICD) system. Diagnoses were recorded as life-time prevalence rates, thus participants could receive a different diagnosis at any point in time.

The Multi-Generation Register enabled identification of all siblings (including twins, however, these were excluded due to power reasons) registered as living in Sweden since 1961 and born in Sweden since 1932. We focused on two siblings per family to simplify the analyses. In order to maximize the follow-up time (and thereby power), we selected the oldest pair. In order to maximize the probability that they experienced a similar shared environment, we only included pairs who were born within 5 years of each other. The final sample consisted of 1 466 543 pairs of full siblings, 129 715 pairs of maternal half-siblings, 141 298 pairs of paternal half-siblings.

Disorders and conditions

ICD-10 groups mental disorders into 11 broad categories.[15] We focused on a subset of these, including drug and alcohol abuse disorders (F10–F19), disorders with psychotic features (F20–F29), mood disorders (F30–F39), anxiety disorders (F40–F48), physiological disturbances (where we only included eating disorders, F50, because the other diagnoses, such as problems with sleep and sex, seemed less related to the major psychiatric disorders), and behavioral and emotional disorders with a childhood onset (where we only focused on hyperkinetic disorders, F90, because the other diagnoses, such as conduct disorder and attachment difficulties, are expected to emerge as different problems in adulthood). We excluded organic disorders (F00–F09), mental retardation (F70–F79) and developmental disorders (F80–f89) because these are generally considered to have different etiologies. We also excluded personality disorders (F60–F69) for power reasons. Because there was no overlap between obsessive compulsive disorder and schizophrenia, or between schizoaffective and eating disorder, in the maternal half-sibling subgroup, we subsumed obsessive compulsive disorder under the anxiety cluster, and opted to retain schizoaffective rather than eating disorder because, to our knowledge, this diagnosis has not been studied as extensively in relation to other disorders. In addition, though not a mental disorder, we also included violent criminal convictions because it may relate to behavioral disorders. The prevalence rates for each subsample are displayed in Table 1, the ICD codes related to each diagnosis and a description of convictions classified as violent are displayed in Supplementary Appendix 1, and the correlations within and between full siblings, maternal half-siblings and paternal half-siblings are presented in Supplementary Appendices 2, 3 and 4, respectively.

Statistical analyses

We used Exploratory Factor Analysis (EFA) to examine the potential presence of a general factor, which describes the overlap among the disorders (operationalized as tetrachoric correlations) with a fewer number of factors. We relied on EFA rather than confirmatory approaches because (a) although we expected internalizing and externalizing dimensions to emerge, we were uncertain how disorders with psychotic symptomatology might pertain to these; (b) we did not expect the data to have simple

structure; and (c) we were unable to rely on the log-likelihood function, which is useful when comparing nested confirmatory models.

By imposing restrictions on how the factors correlate across siblings, one can explore their genetic and environmental etiology. Specifically, by fixing one set of factors to correlate at their expected average genetic overlap across siblings (that is, 0.5 for full siblings and 0.25 for half-siblings), such factors will have 100% heritability, assessing the additive effects of different alleles (that is, genetic factors). To assess the shared environment, that is, non-genetic components making pairs of siblings similar, one fixes another set of factors to correlate at unity across full and maternal half-siblings (because they grow-up in the same household) and at zero for paternal half-siblings (because they grow-up in different households). Lastly, by fixing a set of factors to correlate at zero across all siblings, such factors will measure the non-shared environment, that is, non-genetic components making siblings within a pair dissimilar. This type of EFA is called an Independent Pathway Model (described in general[16] and in detail[17,18]), which we used to explore whether the general factor had a genetic etiology.

In the first step of an EFA, one must determine how many factors to extract. Because we wanted to explore genetic, shared environmental and non-shared environmental factors, we used a Cholesky decomposition to examine how many dimensions existed in each of these three parts separately[19] by relying on four different indices of dimensionality. First, the Eigenvalue-greater-than-one index specifies that a factor should explain more than one variable.[20] Second, parallel analysis simulates random (uncorrelated) data with the same number of observations and variables as the original data, and the number of factors emerging from the random data is used as a lower baseline for how many factors to extract from the original data.[21] Third, the scree plot involves plotting the Eigenvalues against the number of extracted dimensions, and identifying a break, or an 'elbow,' in the plot, after which ensuing factors presumably capitalize on random noise that does not replicate across samples.[22] Fourth, the Minimum Average Partial Index involves iteratively extracting successive Eigenvectors from the observed correlation matrix, and identifying after how many extractions the average of the residual correlations reaches a minimum.[23]

Although estimator and modeling constraints prevented a direct chi-square comparison, a Cholesky decomposition without the shared environment (root-mean-square error of approximation = 0.001, 90% confidence interval: 0.000–0.001; Confirmatory Fit Index = 0.997; Tucker–Lewis Index = 0.996) appeared to fit no worse compared with a Cholesky decomposition that included the shared environment (root-mean-square error of approximation = 0.001, 90% confidence interval: 0.001–0.001; Confirmatory Fit Index = 0.996; Tucker–Lewis Index = 0.995). Therefore, variation attributable to the shared environment was removed from further analyses.

On the basis of the Eigenvalue-greater-than-one rule, indicated by the dashed line in Figure 1, there were two genetic and one non-shared environment factor. The parallel analysis indicated the same number of dimensions (indicated by the same dashed line) because this index approximates the Eigenvalue rule as the sample size grows large. Likewise, the scree plot in Figure 1 also indicated that a break occurred after two genetic and one non-shared environment factor. In contrast to these indices, the Minimum Average Partial Index suggested that there was one genetic and one non-shared environment factor. Because three of the four tests indicated that there were two genetic dimensions, we focused on
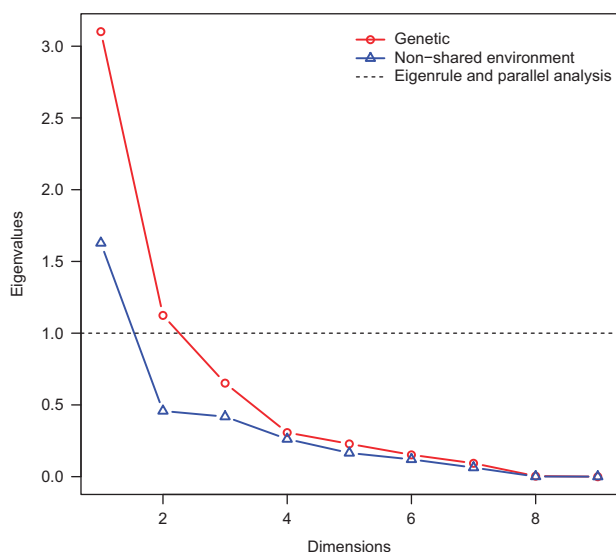
Figure 1. Scree plot of the genetic and non-shared environment variance.

Table 2. Factor structure of common mental disorders

Diagnosis	Genetic factors			Non-shared environment factor
	General factor	Psychotic problems	Non-psychotic problems	Mood problems
Schizophrenia	**0.58**	**0.56**	0.09	0.02
Schizoaffective	**0.55**	**0.67**	-0.05	0.29
Bipolar	**0.41**	**0.40**	0.06	**0.72**
MDD	**0.35**	0.14	0.25	**0.86**
Anxiety	**0.38**	0.04	**0.39**	**0.46**
ADHD	**0.31**	− 0.11	**0.46**	**0.38**
Drug	**0.60**	0.02	**0.65**	0.15
Alcohol	**0.46**	0.01	**0.51**	0.18
Crime	**0.43**	0.01	**0.47**	− 0.02

Abbreviations: ADHD, attention-deficit/hyperactivity disorder; MDD, major depressive disorder. Note: Loadings ⩾0.30 are bolded.

such a model. Because all four tests indicated that there was only one non-shared environment factor, we did not extract more (or less) than that.

The second step of an EFA concerns how to rotate the factors. We opted to rotate the variance shared between the two genetic factors (which correlated at $r = 0.45$ based on the Geomin rotation) toward a single factor.[24] This way, the solution consisted of a common (general) genetic factor, and two independent specific genetic factors, referred to as subfactors henceforth. The multivariate analyses were carried out in Mplus.[25] Rotations were done in R[26] with the GPArotations package,[27] and the parallel analysis and Minimum Average Partial Index were conducted using the psych package.[28] Because half-siblings tend to display a higher rate of disorders compared with full siblings, thresholds were allowed to vary across the three sibling structures (full, maternal half-siblings and paternal half-siblings). The effect of gender was regressed out of all analyses.

Code availability
The computer code is available in Supplementary Appendix 5.

RESULTS
The EFA solution, displayed in Table 2, fit the data very well (root mean square error of approximation = 0.001, 90% confidence
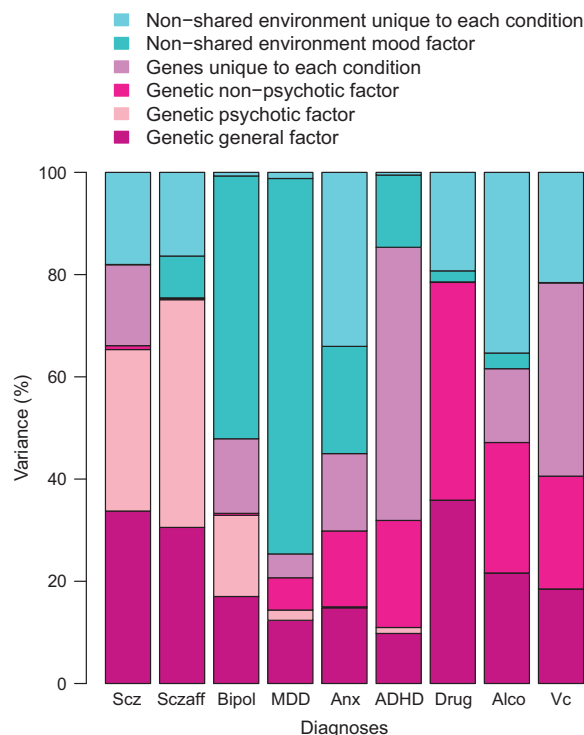
Figure 2. Variance attributable to genetic and non-shared environment sources for each diagnosis. ADHD, attention-deficit/hyperactivity disorder; Alco, alcohol abuse; Anx, anxiety; Bipol, bipolar; Drug, drug abuse; MDD, major depressive disorder; Scz, schizophrenia; Sczaff, schizoaffective disorder; Vc, conviction of violent crimes.

interval: 0.001–0.001; Confirmatory Fit Index = .99; $\chi^2 = 996.65$, degrees of freedom = 550, $P > 0.001$). As hypothesized, all disorders and violent criminal convictions loaded in the same direction on the general genetic factor (range 0.31–0.60), indicating that all conditions partly shared the same genetic origin. Aside from the general genetic factor, schizoaffective disorder (loading = 0.67), schizophrenia (loading = 0.56) and bipolar disorder (loading = 0.40) loaded together, indicating that psychotic problems shared a genetic pathway independent of the general factor. The second genetic subfactor included loadings on drug abuse (loading = 0.65), alcohol abuse (loading = 0.51), violent criminal convictions (loading = 0.47), ADHD (loading = 0.46) and anxiety (loading = 0.39). Although the strongest loadings were on typical externalizing problems, we interpreted this factor as non-psychotic problems because anxiety also loaded on it. The non-shared environment factor, which we interpreted as mood problems, included loadings on major depression (loading = 0.86), bipolar (loading = 0.72) and anxiety (loading = 0.46).

Figure 2 displays the variance accounted for in each disorder attributable to different sources. As can be seen, between 10 (for ADHD) and 36 percent (for drug abuse) of the observed phenotypic variance could be attributed to genetic effects shared across all disorders. This indicates that if a sibling displayed some kind of psychopathology, the co-sibling was at increased risk for not only the same condition, but also for all other forms of psychopathology. There were also genetic effects unique to each condition, with ADHD and violent criminal convictions exhibiting the greatest influence by genes not in common with the other disorders.

The non-shared environment factor indicated that there was overlap among mood disorders within sibling pairs after controlling for genes. In addition, there were substantial non-shared

environment components unique to each disorder, particularly for alcohol abuse disorder and anxiety.

Sensitivity analyses

We also explored models with one (Supplementary Appendix 6) and three (Supplementary Appendix 7) genetic factors. The general genetic factor emerged in both of these models (loadings ranged from 0.31 to 0.88), indicating the general genetic factor is not merely an artifact based on the number of extracted factors. In Supplementary Appendix 8, we explored the assumptions about the shared environment, especially regarding how it contributed to similarity among paternal half-siblings. In short, regardless of how we parameterized the shared environment, it played a vanishingly small role for the covariation among the disorders, in accordance with our main analyses.

Because the rotation toward a general factor is indeterminate when based on two factors, we examined whether this influenced the general genetic factor in Supplementary Appendix 9. Results demonstrated that a general genetic factor tended to emerge regardless of how this indeterminacy was resolved. Furthermore, the general genetic factor appeared highly similar when based on an indeterminate as well as a determinate solution, indicating that the general genetic factor was not an artifact related to rotational constraints.

DISCUSSION

Results demonstrated that common psychiatric disorders and a history of violent criminal convictions partly shared the same genetic origin, dovetailing with past twin,[8,9] family[29] and genomic studies.[30,31] Furthermore, independent of a shared genetic origin, there were specific genetic pathways influencing disorders with psychotic and non-psychotic features, respectively. In contrast to the genetic general factor, the non-shared environment factor did not influence all disorders but primarily only those related to mood problems.

General factor interpretation

Although the general factor appears to be an important phenomenon in both psychology and psychiatry and predicts a host of maladaptive outcomes,[6,32,33] it is difficult to interpret because, by definition, it consists of the variance in common among all conditions. Perhaps as a consequence, researchers have suggested quite different interpretations of the general factor, including that it represents an evolved tendency favoring more cooperative and stable personalities;[34] psychotic thinking;[6] and the personality trait neuroticism,[10] which taps the tendency to experience negative and high arousal emotions.[35] We proffer an additional speculation, namely, that the general factor measures overall distress or impairment, akin to the Global Assessment Functioning Index of the Diagnostic Statistical Manual-IV.[36] Although similar to neuroticism, general distress represents a broader construct in that it encapsulates unpleasant feelings that are both high and low in arousal.[35,37]

Limitations and future directions

First, we were unable to explore if the relative contribution of genetic and environmental influences on the general factor differed by sex because, after splitting the sample into men and women, some of the bivariate tables contained zero observations (such that the tetrachoric correlations could not be estimated properly). The registers used in this study continuously add new data; future studies should investigate this possibility.

Second, as noted above, the rotation method we used is underdetermined but this did not appear to influence the results (Supplementary Appendix 9). On a related note, because we could not use an estimator that maximized the log-likelihood function, we could not compare the dimensionality of the correlation matrices using metrics such as Akaike's Information Criterion.[38] However, given that the general genetic factor emerged regardless of whether we extracted one, two or three genetic factors, the results did not appear to hinge on a particular dimensionality index.

Third, assessment of disorders from registers is relatively crude, and do not permit the richness of structured interviews or surveys. Furthermore, register data will not include troubled individuals who chose not to seek help, leading to an underestimation of prevalence rates. Also, individuals with multiple diagnoses are more likely to get in contact with the mental health system, which can lead to an overestimation of associations among disorders. On the other hand, register data could also underestimate comorbidity in comparison with structured interviews. For example, clinicians may rely too heavily on exclusionary criteria in order to identify a diagnosis, favor a particular diagnosis and thereby miss other potential diagnoses, and assign diagnoses partly based on past conditions. Nevertheless, we note that our results of a general genetic factor of psychopathology are consistent with previous survey studies[8,9] having completely different types of assumptions, suggesting that results are not entirely due to biases.

Fourth, in the main analyses we used the assumption that maternal half-siblings shared all of the shared environment and that paternal half-siblings shared none. We acknowledge that this assumption is a simplification, but sensitivity analyses using different assumptions[39] showed that the shared environment had a vanishingly small influence on the observed overlap among the disorders (Supplementary Appendix 8). Thus, for our research question the exact parameterization of the shared environment does not seem to be essential.

Fifth, it is a considerable limitation that we did not have access to the more fine-grained distinctions among different anxiety disorders introduced in ICD-10 over the entire assessment period. If so, we may have identified an internalizing disorders factor. On a related note, even though personality disorders are related to common psychiatric diagnoses,[40] we were unable to include them due to power reasons. Taken together, it would be important to replicate these findings in an independent sample with more fine-grained diagnoses.

CONCLUSION

Relying on a population-based sibling study, we showed that a variety of common psychiatric conditions partly shared the same genetic origin. This is the first study of its kind to demonstrate such broad genetic overlap among disorders in individuals who sought or were forced to seek mental health care, dovetailing with survey studies of the general population. Given its ubiquitous influence and predictive power, this broad genetic risk factor warrants further investigation and consideration.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

1 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th (edn). American Psychiatric Publishing: Arlington, VA, 2013.

2 Kessler RC, Ormel J, Petukhova M, McLaughlin KA, Green JG, Russo LJ *et al*. Development of lifetime comorbidity in the World Health Organization world mental health surveys. *Arch Gen Psychiatry* 2011; **68**: 90–100.

3 Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005; **62**: 617–627.

4 Krueger RF, Caspi A, Moffitt TE, Silva PA. The structure and stability of common mental disorders (DSM-III-R): a longitudinal-epidemiological study. *J Abnorm Psychol* 1998; **107**: 216–227.

5 Kendler KS, Aggen SH, Knudsen GP, Roysamb E, Neale MC, Reichborn-Kjennerud T. The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV axis I and all axis II disorders. *Am J Psychiatry* 2011; **168**: 29–39.

6 Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S *et al*. The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clin Psychol Sci* 2013; **2**: 119–137.

7 Lahey BB, Applegate B, Hakes JK, Zald DH, Hariri AR, Rathouz PJ. Is there a general factor of prevalent psychopathology during adulthood? *J Abnorm Psychol* 2012; **121**: 971–977.

8 Lahey BB, Van Hulle CA, Singh AL, Waldman ID, Rathouz PJ. Higher-order genetic and environmental structure of prevalent forms of child and adolescent psychopathology. *Arch Gen Psychiatry* 2011; **68**: 181–189.

9 Pettersson E, Anckarsater H, Gillberg C, Lichtenstein P. Different neurodevelopmental symptoms have a common genetic etiology. *J Child Psychol Psychiatry* 2013; **54**: 1356–1365.

10 Tackett JL, Lahey BB, van Hulle C, Waldman I, Krueger RF, Rathouz PJ. Common genetic influences on negative emotionality and a general psychopathology factor in childhood and adolescence. *J Abnorm Psychol* 2013;.**122**: 1142–1153.

11 Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol* 2007; **17**: 643–653.

12 Kessler RC, Berglund P, Chiu WT, Demler O, Heeringa S, Hiripi E *et al*. The US National Comorbidity Survey Replication (NCS-R) design and field procedures. *Int J Methods Psychiatr Res* 2004; **13**: 69–92.

13 Takayanagi Y, Spira AP, Roth KB, Gallo JJ, Eaton WW, Mojtabai R. Accuracy of reports of lifetime mental and physical disorders results from the baltimore epidemiological catchment area study. *JAMA Psychiatry* 2014; **71**: 273–280.

14 Moffitt TE, Caspi A, Taylor A, Kokaua J, Milne BJ, Polanczyk G *et al*. How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychol Med* 2010; **40**: 899–909.

15 World Health Organization. *ICD-10: The ICD-10 Classification of Mental and Behavioural Disorders*. Geneva, Switzerland, 1992.

16 Kendler KS, Heath AC, Martin NG, Eaves LJ. Symptoms of anxiety and symptoms of depression - same genes, different environments. *Arch Gen Psychiatry* 1987; **44**: 451–457.

17 McArdle JJ, Goldsmith HH. Alternative common factor models for multivariate biometric analyses. *Behav Genet* 1990; **20**: 569–608.

18 Neale MC, Cardon LR. *Methodology for Genetic Studies of Twins and Families*. Kluwer Academic Publishers: Dordrecht, The Netherlands, 1992.

19 Franic S, Dolan CV, Borsboom D, Hudziak JJ, van Beijsterveldt CEM, Boomsma DI. Can genetics help psychometrics? Improving dimensionality assessment through genetic factor modeling. *Psychol Methods* 2013; **18**: 406–433.

20 Kaiser HF. The application of electronic-computers to factor-analysis. *Educ Psychol Meas* 1960; **20**: 141–151.

21 Horn JL. A rationale and test for the number of factors in factor-analysis. *Psychometrika* 1965; **30**: 179–185.

22 Cattell RB. Scree test for number of factors. *Multivariate Behav Res* 1966; **1**: 245–276.

23 Velicer WF. Determining number of components from matrix of partial correlations. *Psychometrika* 1976; **41**: 321–327.

24 Schmid J, Leiman JM. The development of hierarchical factor solutions. *Psychometrika* 1957; **22**: 53–61.

25 Muthén LK, Muthén BO. Mplus User's Guide, 7th (edn). Muthén & Muthén: Los Angeles, CA, USA, 1998-2012.

26 Team RC. R: *A Language and Environment for Statistical Computing*, 3.0.1 (edn). R Foundation for Statistical Computing: Vienna, Austria, 2013.

27 Bernaards CA, Jennrich RI. Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educ Psychol Meas* 2005; **65**: 676–696.

28 Revelle W. psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, 2015. http://CRAN.R-project.org/package=psych Version = 1.5.4.

29 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**: 234–239.

30 Smoller JW, Craddock N, Kendler K, Lee PH, Neale BM, Nurnberger JI *et al*. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013; **381**: 1371–1379.

31 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**: 748–752.

32 Hopwood CJ, Malone JC, Ansell EB, Sanislow CA, Grilo CM, McGlashan TH *et al*. Personality assessment in DSM-5: empirical support for rating severity, style, and traits. *J Pers Disord* 2011; **25**: 305–320.

33 Lahey BB, Rathouz PJ, Keenan K, Stepp SD, Loeber R, Hipwell AE. Criterion validity of the general factor of psychopathology in a prospective study of girls. *J Child Psychol Psychiatry* 2015; **56**: 415–422.

34 Rushton JP, Bons TA, Hur YM. The genetics and evolution of the general factor of personality. *J Res Pers* 2008; **42**: 1173–1185.

35 Tellegen A, Watson D, Clark LA. On the dimensional and hierarchical structure of affect. *Psychol Sci* 1999; **10**: 297–303.

36 Association AP. *Diagnostic and Statistical Manual of Mental Disorders*, 4th (edn), text rev. Washington, DC: Author 2000.

37 Tellegen A, Ben-Porath YS, McNulty JL, Arbisi PA, Graham JR, Kaemmer B. *The MMPI-2 Restructured Clinical (RC) Scales: Development, Validation, and Interpretation*. University of Minnesota Press: Minneapolis, MN, USA, 2003.

38 Akaike H. New look at statistical-model identification. *IEEE Trans Automat Contr* 1974; **Ac19**: 716–723.

39 Kendler KS, Maes HH, Sundquist K, Ohlsson H, Sundquist J. Genetic and family and community environmental effects on drug abuse in adolescence: a Swedish national twin and sibling study. *Am J Psychiatry* 2014; **171**: 209–217.

40 Kendler KS, Aggen SH, Knudsen GP, Roysamb E, Neale MC, Reichborn-Kjennerud T. The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV axis I and All Axis II disorders. *Am J Psychiat* 2011; **168**: 29–39.

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