2015

Confirmatory test of two factors and four subtypes of bipolar disorder based on lifetime psychiatric co-morbidity

A. L. Glowinski
Washington University School of Medicine in St. Louis

et al

Psychological Medicine (2015), **45**, 2181–2196. © Cambridge University Press 2015
doi:10.1017/S0033291715000185

ORIGINAL ARTICLE

Confirmatory test of two factors and four subtypes of bipolar disorder based on lifetime psychiatric co-morbidity

P. O. Monahan[1]*, T. Stump[1], W. H. Coryell[2], J. Harezlak[1], G. A. Marcoulides[3], H. Liu[1], C. M. Steeger[4], P. B. Mitchell[5,6], H. C. Wilcox[7], L. A. Hulvershorn[8], A. L. Glowinski[9], Bipolar Disorder Genome Study (BiGS) Consortium†, Bipolar High Risk Study Group†, P. A. Iyer-Eimerbrink[8], M. McInnis[10] and J. I. Nurnberger Jr[8]

[1] *Department of Biostatistics, Indiana University School of Medicine, Indianapolis, IN, USA*
[2] *Department of Psychiatry, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, IA, USA*
[3] *Research Methods & Statistics Program, Graduate School of Education, University of California–Riverside, Riverside, CA, USA*
[4] *Department of Psychology, College of Arts and Letters, University of Notre Dame, Notre Dame, IN, USA*
[5] *School of Psychiatry, University of New South Wales, Sydney, NSW, Australia*
[6] *Black Dog Institute, Sydney, NSW, Australia*
[7] *Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, MD, USA*
[8] *Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, USA*
[9] *Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA*
[10] *Department of Psychiatry, School of Medicine, University of Michigan, Ann Arbor, MI, USA*

Background. The first aim was to use confirmatory factor analysis (CFA) to test a hypothesis that two factors (internalizing and externalizing) account for lifetime co-morbid DSM-IV diagnoses among adults with bipolar I (BPI) disorder. The second aim was to use confirmatory latent class analysis (CLCA) to test the hypothesis that four clinical subtypes are detectible: pure BPI; BPI plus internalizing disorders only; BPI plus externalizing disorders only; and BPI plus internalizing and externalizing disorders.

Method. A cohort of 699 multiplex BPI families was studied, ascertained and assessed (1998–2003) by the National Institute of Mental Health Genetics Initiative Bipolar Consortium: 1156 with BPI disorder (504 adult probands; 594 first-degree relatives; and 58 more distant relatives) and 563 first-degree relatives without BPI. Best-estimate consensus DSM-IV diagnoses were based on structured interviews, family history and medical records. MPLUS software was used for CFA and CLCA.

Results. The two-factor CFA model fit the data very well, and could not be improved by adding or removing paths. The four-class CLCA model fit better than exploratory LCA models or *post-hoc*-modified CLCA models. The two factors and four classes were associated with distinctive clinical course and severity variables, adjusted for proband gender. Co-morbidity, especially more than one internalizing and/or externalizing disorder, was associated with a more severe and complicated course of illness. The four classes demonstrated significant familial aggregation, adjusted for gender and age of relatives.

Conclusions. The BPI two-factor and four-cluster hypotheses demonstrated substantial confirmatory support. These models may be useful for subtyping BPI disorders, predicting course of illness and refining the phenotype in genetic studies.

Received 16 March 2014; Revised 9 January 2015; Accepted 21 January 2015; First published online 31 March 2015

Key words: Bipolar disorder, co-morbidity, confirmatory factor analysis, confirmatory latent class analysis, subtypes.

* Address for correspondence: P. O. Monahan, Ph.D., Department of Biostatistics, Indiana University, 410 West 10th Street, Suite 3000, Indianapolis, IN 46202-3002, USA.
(Email: pmonahan@iu.edu)

† Authors/group information: The following members of the BiGS Consortium take authorship responsibility: W. Byerley, University of California–San Francisco, San Francisco, CA (W. H. Coryell, M. McInnis and J. I. Nurnberger are also members of this group). The following members of the Bipolar High Risk Study Group take authorship responsibility: W. Reich, Washington University, St Louis, MO; E. Kastelic, Johns Hopkins Medical Institute, Baltimore, MD; and G. Roberts, School of Psychiatry, University of New South Wales, Sydney; and Black Dog Institute, Sydney (W. H. Coryell, M. McInnis, P. B. Mitchell, H. C. Wilcox and J. I. Nurnberger are also members of this group).

Introduction

Bipolar (BP) affective disorder affects 0.5–1.6% of the US adult population and is frequently chronically debilitating (Goodwin & Jamison, 1990; World Health Organization, 2002). BP type I (BPI) illness is highly heritable with up to 80% of risk determined by genetic factors (Gershon *et al.* 1987; Tsuang & Faraone, 1990; Nurnberger & Berrettini, 1998; Potash & DePaulo, 2000; McMahon *et al.* 2001; Smoller & Finn, 2003). The characterization of BPI subtypes of persons may be helpful for early detection and for understanding course of illness, neurobiology and treatment response. Subtypes may be characterized by clinical variables, including patterns of co-morbid psychiatric symptoms and disorders (Nurnberger, 2002; Cassidy *et al.* 2008; Cassano *et al.* 2009). The creation of homogeneous BPI subgroups based on co-morbid conditions has modestly improved the success of genetic mapping (MacKinnon *et al.* 1998; Nurnberger, 2002; Schulze & McMahon, 2003; MacQueen *et al.* 2005; Payne *et al.* 2005; Cheng *et al.* 2006; Saunders *et al.* 2009). For example, three subtypes of BPI, as defined by co-morbidity, have provided some of the strongest evidence of linkage to genomic regions: co-morbid panic disorder, co-morbid psychotic symptoms, and pure BPI with low rates of co-morbidity (MacQueen *et al.* 2005). A subtype of BPI characterized by co-morbid anxiety disorders has been supported by cross-sectional, longitudinal and familial studies (Johnson *et al.* 2000; Schurhoff *et al.* 2000; Birmaher *et al.* 2002; Goodwin & Hamilton, 2002; MacKinnon *et al.* 2002; Wozniak *et al.* 2002). A BPI subtype characterized by high rates of co-morbid drug and alcohol abuse or dependence (Sonne & Brady, 1999; Nurnberger *et al.* 2007) has shown evidence for elevated genetic predisposition to substance use disorders (Winokur *et al.* 1970, 1995, 1996; Helzer & Winokur, 1974; Morrison, 1974, 1975; Gershon *et al.* 1982; Kendler *et al.* 1993; Feinman & Dunner, 1996; Maier & Merikangas, 1996; Duffy *et al.* 1998; DelBello *et al.* 1999; Strakowski & DelBello, 2000; Nurnberger *et al.* 2007). Finally, offspring of BPI patients (compared with controls) have higher rates of affective, anxiety and externalizing disorders such as conduct disorder, oppositional disorder and substance abuse (Gershon *et al.* 1985; Nurnberger *et al.* 1988, 2011; Todd *et al.* 1996; Lapalme *et al.* 1997; Chang *et al.* 2000, 2003; DelBello & Geller, 2001).

Factor analysis and cluster analysis are complementary methods for investigating BPI clinical subtypes. They determine, respectively, whether co-morbid variables (e.g. symptoms or lifetime disorders) can be parsimoniously explained by a fewer number of factors (i.e. clusters of variables) and whether BPI persons can be subdivided into subgroups (i.e. clusters of persons) (see online Supplementary Text, Section B).

Factor analyses have been conducted on psychiatric lifetime disorders among community samples. For example, an internalizing factor and an externalizing factor were found (Kessler *et al.* 2011). However, no published study has used a factor analysis, cluster analysis or a latent cluster analysis (called latent class analysis; LCA) to explore or test hypotheses about the lifetime disorders co-morbid with BPI.

The first purpose of this paper was to disseminate the first published investigation of lifetime disorders (i.e. instead of symptoms) in a confirmatory test of a hypothesized factor analysis model among BPI individuals. Based on literature cited above, we hypothesized a two-factor lifetime co-morbidity model, specified before confirmatory factor analysis (CFA) was performed and then tested with CFA; in which internalizing and externalizing factors are correlated; and for which the internalizing factor explains correlations between anxiety, somatoform and eating disorders; and the externalizing factor explains correlations between alcohol use, drug use, cluster B personality, impulse control, attention-deficit and disruptive disorders.

No published study has tested hypotheses about clinical subtypes, among BPI patients, by using a confirmatory approach to LCA, neither with symptoms nor disorders. Confirmatory LCA (CLCA) provides a powerful method to test and validate hypotheses about subgroups of persons with BPI (Finch & Bronk, 2011).

Thus, the second, and primary, purpose of this paper was to provide a confirmatory test of a hypothesis about subgroups of BPI lifetime co-morbid disorders, specified before CLCA was performed and then tested with CLCA. Based on literature cited above, we hypothesized four co-morbidity subtypes of BPI patients: pure BPI without co-morbidity; co-morbidity with only internalizing disorders; co-morbidity with only externalizing disorders; and co-morbidity with both internalizing and externalizing disorders.

Method

Study design and procedures

The National Institute of Mental Health (NIMH) Genetics Initiative Bipolar Project was active from 1989 to 2007. This article uses data from probands and relatives assessed between 1998 and 2003 using the Diagnostic Interview for Genetic Studies (DIGS) 3.0 interview and resulting Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) diagnoses. At each of 10 sites, multiplex families were ascertained through a proband and a first-degree

relative, both with a DSM-IV diagnosis of BPI or schizo-affective BP type (for the diagnostic process, see online Supplementary Text, Section A). Diagnoses were coded as binary (0 = no, 1 = yes if 'probable' or 'definite').

Statistical methods

MPLUS software (version 5.21) (Muthén & Muthén, 1998–2007) was used to perform CFA and CLCA on 1156 individuals (504 probands and 652 relatives) with BPI disorder using the COMPLEX option to account for the relatedness of individuals within families. Criteria of good overall CFA model fit were the following: comparative fit index (CFI) > 0.95 (Hu & Bentler, 1999), root mean square error of approximation (RMSEA) < 0.06 (Hu & Bentler, 1999), and weighted root mean square residual (WRMR) < 1.00 (Yu, 2002). Loadings above 0.40 indicated adequate fit for individual paths (Nunnally & Bernstein, 1994). We tested (two-sided, 0.05 α) whether fit of the CFA model could be improved by adding or removing paths (see online Supplementary Text, Section C). The robust weighted least squares (WLSMV) estimator and the non-linear probit link were specified for CFA models. For LCA and CLCA models, maximum likelihood with robust standard error estimation was used.

An important deterministic constraint in CLCA models is the hypothesized 'zero' class for which the probabilities of endorsing all co-morbid disorders are fixed to be zero (e.g. 'pure BP' class) (Muthén & Asparouhov, 2006). Our BPI subtyping hypothesis required two other deterministic constraints: an 'externalizing only' class for which the probabilities of endorsing the externalizing disorders were freely estimated and the probabilities of endorsing the internalizing disorders were fixed to zero, and an 'internalizing only' class for which the probabilities of endorsing the internalizing disorders were freely estimated and the probabilities of endorsing the externalizing disorders were fixed to zero. To fix a probability to zero, the threshold of the disorder was constrained to equal 15, which fixes the probability of disorder endorsement to a value extremely close to zero (Finch & Bronk, 2011; Clark *et al.* 2013). We hypothesized a 'both class' of BP persons who have substantial probability of both internalizing and externalizing lifetime disorders for which probabilities were freely estimated for all internalizing and externalizing disorders.

Additionally, we used equality and inequality constraints to test two competing versions of the hypothesized four-class CLCA model, which we specified *a priori* before analyses. In the 'equality' model we used equality constraints to test the hypothesis that the probabilities of the externalizing disorders were not statistically different for the 'both' class and the 'externalizing-only' class, and the probabilities of the internalizing disorders were not statistically different for the 'both' class and the 'internalizing-only' class. In the 'inequality' model, we hypothesized that the probabilities of the externalizing disorders were statistically greater for the 'both' class than the 'externalizing-only' class, and the probabilities of the internalizing disorders were statistically greater for the 'both' class than the 'internalizing class'. Further discussion of the equality and inequality hypotheses is contained in Section B of online Supplementary Text. These two competing four-class CLCA models amount to specifying *a priori* a small specific slice of the universe of possible equality and inequality constraints, deterministic constraints and number of classes.

The fit of CLCA and LCA models was evaluated based on the following criteria. Models with the lowest values for the Bayesian information criterion (BIC), and its sample size-adjusted version (aBIC), were considered the best (Lubke & Muthén, 2005). The BIC and aBIC were used to compare different models with different number of classes and or parameterizations (Finch & Bronk, 2011). The likelihood ratio test (LRT) was used to compare the hypothesized four-class CLCA models with the exploratory four-class model, because the former is nested within the latter (Finch & Bronk, 2011). The LRT is not appropriate for comparing LCA models with differing number of classes (Lubke & Muthén, 2005). Therefore, the Lo–Mendell–Rubin (LMR) (Lo *et al.* 2001) LRT was used to compare nested LCA models that have differing numbers of classes but the same parameterization (e.g. our exploratory one-, two-, three-, four-, five- and six-class LCA models). The LMR *p* value indicates whether a model with one fewer classes can be rejected in favor of the current model. Classes with a sparse number of persons are not practically meaningful (Lubke & Muthén, 2005). Therefore, we rejected models that included classes with a sample size representing less than 5% of the sample. *Post-hoc* modifications to the CLCA models were considered by inspecting graphs of the 'profile' of estimated disorder probabilities for each of the four classes.

Only the disorders that were diagnosed in at least 50 BPI participants were included as individual variables in the CFA and CLCA models to ensure estimation precision. Some clinically similar disorders (e.g. panic disorder with and without agoraphobia) were combined because, if not combined, diagnostic mutual exclusion would have prevented them from loading on the same factor. The combined variables were coded '1' if any of the contributing diagnoses were '1' (probable or definite), and 0 otherwise.

Table 1. *Confirmatory factor analysis of co-morbid lifetime disorders from 1156 BPI probands and BPI relatives*

Variables	Factors	Number[a]	Loading (S.E.)		z	p
	Factor 1: internalizing factor					
1	Panic, agoraphobia, anxiety NOS, GAD, PTSD	315	0.81 (0.08)		9.86	<0.001
2	Any phobia disorder (specific or social)	164	0.50 (0.06)		9.35	<0.001
3	OCD	79	0.55 (0.08)		7.27	<0.001
4	Any eating (anorexia, bulimia, NOS) or somatoform disorder	89	0.55 (0.07)		7.50	<0.001
	Factor 2: externalizing factor					
5	Any alcohol abuse or dependence	444	0.86 (0.07)		13.30	<0.001
6	Any drug abuse or dependence	329	0.79 (0.06)		14.11	<0.001
7	Cluster B personality disorder, impulse control, conduct disorder, ADHD	87	0.50 (0.07)		7.70	<0.001
	Correlation between factor 1 and factor 2		0.34 (0.06)		5.31	<0.001
			Model fit			
			Fit indices		χ^2 Test	
		CFI	RMSEA	WRMR	χ^2 (df)	p
		1.000	0.000	0.556	9.29 (11)	0.60

BPI, Bipolar I; S.E., standard error; z, loading/standard error; NOS, not otherwise specified; GAD, generalized anxiety disorder; PTSD, post-traumatic stress disorder; OCD, obsessive compulsive disorder; ADHD, attention-deficit/hyperactivity disorder; CFI, comparative fit index; RMSEA, root mean square error of approximation; WRMR, weighted root mean square residual; df, degrees of freedom.

[a] Number of 1156 persons with presence (probable or definite) for each variable; each variable is defined as either an individual disorder (e.g. OCD) or presence of any disorder within a combination of disorders (e.g. any phobia disorder).

In addition, clinical judgments regarding *a priori* factor assignment of each variable in the hypothesized CFA model were made based on the relative importance of internalizing and externalizing symptoms for various disorders. The internalizing factor consisted largely of anxiety disorders, and the externalizing factor consisted largely of drug and alcohol use disorders. Due to sparseness of somatoform disorders, we decided *a priori* to include somatoform disorders and eating disorders in one variable, which we hypothesized to be explained by the internalizing factor (see Table 1).

The factor scores and subject clusters for the 1156 BPI individuals were validated by testing their associations with clinical course variables known to be related to prognosis, severity and impairment (see Table 2 footnote for definitions). In each model, a clinical course variable was the dependent variable. The independent variables were the factors or clusters, adjusted for gender of BPI subject. Linear and logistic regression models were estimated using the SAS generalized linear modeling (GENMOD) procedure with the 'generalized estimating equations' (GEE) (Liang & Zeger, 1986) estimation method to account for within-family correlations. For continuous outcomes, rank scores (robust to skewness) of the outcome, the normal error distribution, and the linear link were specified. For binary outcomes, the binomial error distribution and the exchangeable log odds ratio regression structure were specified.

For the familial analysis, only first-degree relatives were included. Four clusters of probands were compared on binary disorders of relatives using GEE with GENMOD, as described above, and on an unordered eight-category dependent variable (eight clusters in relatives) using the SAS SURVEYLOGISTIC procedure, to account for within-family correlations. The familial models were adjusted for gender and age of relatives.

All tests were two-sided with α of 0.05. *Post-hoc* pairwise comparisons were tested only if the omnibus test was significant (i.e. the protected version of Fisher's least significant difference).

Ethical standards

All procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Results

Participant characteristics

The 504 probands (from 504 multiplex families) with BPI disorder (96%) or schizo-affective BP type (4%) were adults; age at interview ranged from 18 to 88 years (except one was aged 17 years). The 1157

Table 2. *Factors and clinical course of illness variables among BPI participants (BPI probands and BPI relatives)*[a]

	Factor 1 (internalizing factor)					Factor 2 (externalizing factor)				
	Factor score					Factor score				
	0	1	2, 3, or 4	Omnibus *p*	*Post-hoc* pairwise tests	0	1	2 or 3	Omnibus *p*	*Post-hoc* pairwise tests
Number of BPI participants	711	288	157			609	274	273		
Factor score group number	(1)	(2)	(3)			(1)	(2)	(3)		
Continuous clinical variables										
Median age of onset, years	19	17	15	***	ALL	19	18	16	***	ALL
Median episode frequency per years ill[b]	0.43	0.57	0.87	***	ALL	0.48	0.48	0.58	N.S.	N.A.
Median mania/depression episode ratio[c]	1.00	0.80	1.00	N.S.	N.A.	1.00	1.00	1.00	N.S.	N.A.
Median psychiatric hospitalizations per years ill[d]	0.17	0.17	0.19	N.S.	N.A.	0.16	0.17	0.21	N.S.	N.A.
Median between-episode GAS[e]	70	65	61	***	(1) > (2)(3)	70	70	61	***	(1)(2) > (3)
Binary clinical variables (1 = yes, 0 = no), %										
Female	55	69	81	***	ALL	68	62	50	***	(1)(2) > (3)
Disabled[f]	17	20	25	*	(1)(2) < (3)	17	17	25	*	(1)(2) < (3)
Psychotic symptoms[g]	47	52	62	N.S.	(1) < (3)	45	57	55	**	(1) < (2)(3)
Mood-incongruent psychotic symptoms[h]	11	12	14	N.S.	N.A.	11	13	11	N.S.	N.A.
Rapid cycling[i]	9	10	13	N.S.	N.A.	8	10	13	N.S.	N.A.
Rapid switching[j]	39	53	71	***	ALL	42	46	59	***	(1)(2) < (3)
Mixed states[k]	23	30	36	**	(1) < (2)(3)	23	27	34	**	(1) < (3)

BPI, Bipolar I; ALL, all three pairwise differences were significant; N.S., not significant; N.A., *post-hoc* pairwise tests not applicable because omnibus test not significant; GAS, global assessment score; DIGS, Diagnostic Interview for Genetic Studies.

[a] The sum score, computed separately for internalizing and externalizing variables, was specified in models as a categorical variable with three levels (no disorders, one variable, two or more variables) because validation results were similar for sum scores and coefficient-weighted factor scores. Each row represents a different regression model (linear and logistic, respectively, for continuous and binary clinical course dependent variables), estimated using a generalized linear model (SAS GENMOD) with generalized estimating equations to account for within-family correlations.

[b] Episode frequency = number of clean (i.e. episodes not likely to be caused by a specific organic factor, such as drug abuse, medication, or disease) affective episodes per years of illness.

[c] Mania/depression episode ratio = number of clean manic episodes over the number of clean depressive episodes.

[d] Psychiatric hospitalizations per years ill = number of hospitalizations divided by years of illness.

[e] Between-episode GAS = GAS for the past month if not hospitalized, otherwise GAS equals missing value.

[f] Disabled = occupationally disabled, from the present job question in DIGS 3.0.

[g] Presence of psychotic symptoms = at least one of five psychotic screening items in the DIGS 3.0 psychosis section definitely present and lasted persistently throughout the day for 1 day or intermittently for a period of 3 days.

[h] Mood-incongruent psychotic symptoms were present during either a mania or depression episode assessed in DIGS 3.0.

[i] Rapid cycling = four or more discrete episodes of mania or depression within 12 months demarcated by 8 weeks or more of remission.

[j] Rapid switching = positive on the mania screening item in the DIGS mania section: 'ever switched quickly from high to normal or high to depressed without normal mood between'.

[k] Mixed states = at least three symptoms of the opposite polarity lasting 1 week or more, plus positive on stem question 'During this episode did you have a week or more during which your mood frequently changed between irritability or elation and sadness or depression?' in either the mania or depression sections of DIGS 3.0.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

first-degree relatives of the 504 probands included mostly adults (range 18–93 years) with only eight (1%) adolescents (range 13–17 years). The high percentage of first-degree relatives with BPI or schizo-affective BP type (48.6%) was a function of multiplex family ascertainment.

The main analyses (i.e. testing CFA and CLCA hypotheses and validating against clinical course variables) were based on 1156 persons with BPI disorder or schizo-affective BP type (504 adult probands; 594 first-degree relatives; and 58 more distant relatives) (see online Supplementary Table S1). At time of interview, 44% of those with BPI were married. Their median age of onset was 18.0 years. The majority was female and Caucasian. Prevalence of co-morbid lifetime disorders is shown in online Supplementary Table S1.

Test of the CFA model

All variables demonstrated large standardized factor loadings (0.50 or greater) on their hypothesized factor (Table 1). The CFA model demonstrated excellent fit to the data (CFI = 1.000, RMSEA = 0.000, WRMR = 0.556). The null hypothesis of good fit was not rejected by the χ^2 goodness of fit test [$\chi^2 = 9.29$, degrees of freedom (df) = 11, $p = 0.60$]. The two factors were significantly correlated as hypothesized ($r = 0.34$, $p < 0.001$). A one-factor model showed poor fit (CFI = 0.76, RMSEA = 0.094, WRMR = 1.65, $\chi^2 = 156.63$, df = 14, $p < 0.001$). No statistically significant improvement upon the two-factor model could be found by adding or deleting paths. The eigenvalues further supported a two-factor solution (see online Supplementary Text, Section D).

Associations between factors and clinical course variables

Comparison of factor-scoring methods

Validation with clinical course variables was initially performed for both the CFA-coefficient-weighted factor scores and the commonly used sum scale score (one point for each variable). Results were similar for the two methods (see online Supplementary Table S2). Therefore, validation results are reported next using the sum score because it can be easily computed by hand during clinical encounters and it lent itself to disseminating dose–response relationships according to three clinically sensible categories: no disorders; one variable; and two or more variables (frequencies with three or four variables were sparse).

Associations between factors and clinical course variables

A significant association with internalizing and externalizing sum scores was observed for a majority of the clinical course variables (Table 2). A complete dose–response relationship was observed for age of onset, episode frequency, gender and rapid switching, particularly for the internalizing score (all pairwise differences were significant). In addition, for inter-episode global assessment score (GAS), disability, history of psychotic symptoms and mixed states, one or two of the pairwise differences were significant and the means and percentages of clinical course variables generally showed dose–response trends (Table 2).

There were no significant interactions between the internalizing and externalizing sum scores, except for mixed states ($p = 0.029$). Thus, the relationship with clinical course for one factor generally did not depend on whether participants scored low or high on the other factor.

Test of the CLCA model

The fit of the hypothesized CLCA model was compared with a wide range of exploratory LCA models and *post-hoc* modified CLCA models to strongly test its resilience (Table 3). Both alternatives of the hypothesized CLCA model (models no. 1 and no. 2) fit better (lower BIC and aBIC) than all exploratory (one-class to six-class) models (models no. 3 to no. 8). The five-class and six-class exploratory models were not meaningful due to one or more classes having only 1% membership. The hypothesized CLCA model with full equality constraints (model no. 1) fit the best; its fit was slightly better than the competing CLCA model with full inequality constraints (model no. 2) and also better than all *post-hoc* modified CLCA models (models no. 9 to no. 16).

Further non-essential description of CLCA results in Table 3 is provided in Section D of online Supplementary Text. In summary, none of the *post-hoc* modified four-class CLCA models or any of the one-class to six-class exploratory LCA models fit better than the CLCA model that was hypothesized to have deterministic constraints and full equality constraints (model no. 1).

Associations between CLCA-derived BPI subtypes and clinical course variables

BPI subtypes, based on most likely class membership from the best CLCA model (model no. 1), were compared on clinical course variables. Significant associations were found for seven of the 12 clinical course variables and in the anticipated directions (Table 4). For example, the median age of onset of major affective disorder was highest for those with no co-morbid disorders (19.0 years) and lowest for those with both internalizing and externalizing variables (16.0 years). All pairwise differences were significant for age of

Table 3. *CLCA of seven lifetime co-morbid diagnostic variables (n = 1156 for all models)*

Model no.	Constraints	No. of classes	Log likelihood	No. of free parameters	BIC	aBIC	LMR *p*	Most likely class membership size, *n* (%)[a] 1	2	3	4
Two competing CLCA models											
1	Zero-c, Int-only-c, Ext-only-c, Both-c, Full Eq	4	−3288	10	6647	6616	<0.0001	422 (37)	187 (16)	289 (25)	258 (22)
2	Zero-c, Int-only-c, Ext-only-c, Both-c, Full InEq	4	−3284	17	6687	6633	N.A.	422 (37)	187 (16)	289 (25)	258 (22)
Exploratory LCA models											
3	None	1	−3520	7	7089	7067	N.A.	1156 (100)			
4	None	2	−3343	15	6791	6744	<0.0001	389 (34)	767 (66)		
5	None	3	−3283	23	6727	6654	<0.0001	101 (9)	725 (63)	330 (28)	
6	None	4	−3266	31	6751	6652	0.19	616 (53)	144 (12)	101 (9)	295 (26)
7	None	5	−3259	39	6793	6669	0.13	7 (1)	124 (11)	376 (32)	102 (9)
8	None	6	−3254	47	6840	6691	0.63	604 (52)	7 (1)	150 (13)	277 (24)
Post-hoc modifications to CLCA models											
9	Model no. 1, except Eq(Int only)	4	−3285	13	6661	6619	<0.0001	422 (37)	187 (16)	289 (25)	258 (22)
10	Model no. 1, except Eq(Ext only)	4	−3288	14	6674	6629	<0.0001	422 (37)	187 (16)	289 (25)	258 (22)
11	Model no. 1, except Eq(Alc only)	4	−3285	12	6655	6617	<0.0001	422 (37)	187 (16)	289 (25)	258 (22)
12	Model no. 2, except InEq(Ext only)	4	−3284	17	6687	6633	N.A.	422 (37)	187 (16)	289 (25)	258 (22)
13	Model no. 1 or no. 2, except Eq(Int only) and InEq(Ext only)	4	−3285	13	6661	6619	N.A.	422 (37)	187 (16)	277 (24)	270 (23)
14	Model no. 1 or no. 2, except no Eq and no InEq constraints	4	−3284	17	6687	6633	<0.0001	422 (37)	187 (16)	289 (25)	258 (22)
15	Model no. 1, except Int-only + Alc-c	4	−3288	11	6655	6620	<0.0001	422 (37)	187 (16)	289 (25)	258 (22)
16	Zero-c, no other constraints	4	−3273	24	6715	6639	0.07	422 (36)	285 (25)	126 (11)	323 (28)

CLCA, Confirmatory latent class analysis; BIC, Bayesian information criterion; aBIC, sample size-adjusted BIC; LMR *p*, Lo–Mendell–Rubin test *p* value; Zero-c, zero-class specified by using a deterministic constraint to fix the probabilities of endorsing each of the seven variables to be zero; Int, internalizing; Ext, externalizing; Int-only-c, internalizing-only class specified using a deterministic constraint to fix the probabilities of endorsing externalizing disorders to be zero; Ext-only-c, externalizing-only class specified using a deterministic constraint to fix the probabilities of endorsing internalizing disorders to be zero; Both-c, both-class specified by allowing all probabilities for endorsing the four internalizing and the three externalizing variables to be freely estimated; Full Eq, full equality constraints in which endorsement probabilities of externalizing variables were constrained to be equal for the externalizing-only class and the both-class, and endorsement probabilities of internalizing variables were constrained to be equal for the internalizing-only class and the both-class; Full InEq, full inequality constraints in which endorsement probabilities of externalizing variables were constrained to be less for the externalizing-only class than the both-class, and endorsement probabilities of internalizing variables were constrained to be less for the internalizing-only class than the both-class; N.A., LMR test not available for models with (non-linear) inequality constraints and not applicable for an exploratory model with one class; Eq(Int only), Eq(Ext only), Eq(Alc only), instead of full equality constraints on all variables, the equality constraint was placed on only the four internalizing variables, or only the three externalizing variables, or only the one alcohol variable, respectively; InEq(Ext only), instead of full inequality constraints on all variables, the inequality constraint was placed on only the three externalizing variables; Int-only + Alc-c, instead of an internalizing-only class, an internalizing-only-plus-alcohol class was specified using a deterministic constraint to fix the probabilities of endorsing externalizing disorders to be zero except for one externalizing variable, alcohol abuse or dependence disorders, for which the probability of endorsement was allowed to be freely estimated along with the internalizing disorders.

[a] The most likely class membership size was 547 (47%) for class 5 for model no. 7, and were 17 (1%) and 101 (9%) for classes 5 and 6, respectively, for model no. 8.

Table 4. *Clusters and course of illness variables among BPI participants (BPI probands and BPI relatives)*[a]

	Clusters based on DSM-IV diagnosed co-morbid disorders					
	None of the co-morbid disorders	Internalizing disorder(s) only	Externalizing disorder(s) only	Internalizing and externalizing disorder(s)	Omnibus *p*	*Post-hoc* pairwise comparisons between clusters
Number of BPI probands and BPI relatives	422	187	289	258		
Cluster group number	(1)	(2)	(3)	(4)		
Continuous clinical variables						
Median age of onset of major affective disorder, years	19.0	17.0	18.0	16.0	***	(1) > (2)(3)(4); (2) > (4); (3) > (4)
Median episode frequency per years ill	0.43	0.62	0.44	0.69	***	(1) < (2)(4); (2) > (3); (3) < (4)
Median manic/depressive episode ratio	1.00	0.75	1.00	1.00	N.S.	N.A.
Median psychiatric hospitalizations per years ill	0.17	0.16	0.19	0.19	N.S.	N.A.
Median between-episode GAS	70	65	70	64	***	(1) > (2)(4); (3) > (4)
Binary clinical variables (1 = yes, 0 = no), %						
Female	61	83	47	66	***	(1) < (2); (1) > (3); (2) > (3)(4); (3) < (4)
Disabled	16	18	18	24	N.S.	N.A.
Psychotic symptoms	41	54	56	56	**	(1) < (3)(4)
Mood-incongruent	11	13	12	12	N.S.	N.A.
Psychotic symptoms						
Rapid cycling	6	10	11	12	N.S.	N.A.
Rapid switching	34	59	46	60	***	(1) < (2)(3)(4); (3) < (4)
Mixed states	19	33	30	32	***	(1) < (2)(3)(4)

BPI, Bipolar I; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th edition; N.S., not significant; N.A., pairwise tests not applicable because omnibus test was not significant; GAS, global assessment score.

[a] Each row represents a different regression model (linear and logistic, respectively, for continuous and binary clinical course dependent variables), estimated using a generalized linear model (SAS GENMOD) with generalized estimating equations to account for within-family correlations.

** $p < 0.01$, *** $p < 0.001$.

onset except for the difference between 'internalizing-only' and 'externalizing-only' classes. The presence of internalizing disorders, either alone or in combination with externalizing disorders, was associated with greater episode frequency and lower GAS. Persons in class 1 (pure BPI without co-morbidity) were less likely to have a history of psychotic symptoms than persons with externalizing disorders only (class 3) or persons with both internalizing and externalizing disorders (class 4). All three groups of BPI individuals with co-morbidity had significantly greater history of rapid switching and mixed states than the cluster without co-morbidity (class 1). Rapid switching was particularly elevated for the two clusters with internalizing disorders.

Familial analysis

There were eight possible co-morbidity clusters of first-degree relatives defined by crossing the presence of BPI or other affective disorders with the four co-morbidity classes that were defined for BPI probands from the best CLCA model. The four clusters of probands were significantly associated with the eight clusters of relatives, indicating significant familial aggregation of the hypothesized clusters after adjusting for gender and age of relatives (Table 5; 4×8 omnibus test, $p < 0.001$). In particular, rates for three of the relative clusters differed between the four proband clusters (4×2 omnibus tests, $p < 0.001$, 0.045 and 0.001). Specifically, relatives were more likely to have pure affective disorder without other internalizing or externalizing disorders (36%) if their proband also had pure BPI (cluster 1). Relatives were also more likely to have affective disorder with both internalizing and externalizing disorders (17%) if the proband also had BPI with both internalizing and externalizing disorders (cluster 4). Finally, relatives were more likely to have affective disorders with other internalizing disorders and no externalizing disorders (18%) if the BPI proband also had internalizing disorders only (cluster 2), compared with relatives of pure BPI probands (cluster 1) or BPI probands with externalizing disorders only (cluster 3).

The remaining familial analyses in Table 5 pertain to individual disorders. To summarize, several disorders and groups of disorders among the relatives differed significantly between the four proband clusters and in the anticipated directions (Table 5).

Discussion

Synthesis of findings

The proposed two-factor CFA model showed an excellent fit to the data, suggesting inter-correlated internalizing and externalizing co-morbid factors in the context of BPI. Additional validity of the two-factor co-morbidity model was supported by significant associations in anticipated directions between higher internalizing and externalizing factor scores and a majority of the clinical course variables that typically indicate worse prognosis, severity and impairment.

For the second set of analyses, the *a priori* hypothesized CLCA model with full equality constraints (model no. 1) fit the data better than a range of 15 other models, including the competing hypothesized CLCA model with full inequality constraints, eight *post-hoc* modifications to both versions of the CLCA model, and six exploratory models. This best-fitting four-class model had a 'zero', 'internalizing-only', 'externalizing-only' and 'both' class. The four subject clusters or BPI subtypes, based on the best-fitting hypothesized CLCA model, demonstrated significant differences on a majority of the clinical course variables, supporting the meaningful interpretation of these subtypes. The findings from familial analysis offered further support for these four clusters of BPI individuals.

A remarkable aspect about the results is the differences between 'pure' BPI and BPI with any co-morbidity. 'Pure' BP runs in families, it has a later age of onset of major affective disorder in our data, generally lower episode frequency and less inter-episode impairment, fewer psychotic symptoms, and less evidence of rapid switching and mixed states. That is, many of the 'complications' of BP disorder that are typically regarded as evidence of more severe disorder are differentially clustered in subjects with co-morbid disorders.

More work is needed to evaluate whether the presence of co-morbid anxiety disorders in children diagnosed with BP may be a marker of very early-onset BP (Wozniak *et al.* 2002). Our findings suggest that the presence of either anxiety disorders (and/or other internalizing disorders) or externalizing disorders (such as substance use) is a marker of earlier onset of major affective disorder in adults diagnosed with BP.

These features may be clearly noted in the two-factor validation findings as well. Here the 'pure' subjects with zero trait presence on the factor score are seen to have later age of onset and fewer psychotic symptoms than those with any externalizing disorder, and lower episode frequency and evidence of mixed states than those with any internalizing disorder.

The other remarkable aspect of this analysis is the 'dose–response' demonstration that the presence of a greater number of co-morbid disorders, either internalizing or externalizing, is associated with a distinct worsening of course of illness. This is particularly true with regard to the incidence of disability, poor

Table 5. *Familial aggregation: clusters and disorders of FDRs by BPI proband clusters*[a]

	Proband clusters based on DSM-IV co-morbid diagnoses					*Post-hoc* pairwise comparisons between clusters					
	None of the co-morbid disorder(s)	Int disorder(s) only	Ext disorder(s) only	Int and Ext disorder(s)	Omnibus p[c]	(1) (2)	(1) (3)	(1) (4)	(2) (3)	(2) (4)	(3) (4)
Cluster group number	(1)	(2)	(3)	(4)							
Number in each BPI proband cluster	177	79	121	127							
Number of their FDRs	406	171	301	279							
Clusters in FDRs					<0.001	N.S.	*	*	N.S.	N.S.	N.S.
No affective disorder + no Int or Ext disorder	23	18	23	15	0.391	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
No affective disorder + Int disorder(s) only	3	2	2	2	0.768	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
No affective disorder + Ext disorder(s) only	7	4	11	8	0.110	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
No affective disorder + Int and Ext disorder(s)	1	2	3	3	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Affective disorder + no Int or Ext disorder	36	27	20	22	<0.001	*	*	*	N.S.	N.S.	N.S.
Affective disorder + Int disorder(s) only	9	18	9	14	0.045	*	N.S.	N.S.	*	N.S.	N.S.
Affective disorder + Ext disorder(s) only	16	17	20	19	0.169	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Affective disorder + Int and Ext disorder(s)	6	13	14	17	0.001	N.S.	*	*	N.S.	N.S.	N.S.
Affective disorders in FDRs											
BPI[b]	49	53	46	48	0.652	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
MDDR	9	11	7	14	0.114	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
SEMD	5	4	5	6	0.753	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Any bipolar disorder NOS[d]	2	5	1	3	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Any depressive disorder NOS[e]	4	4	3	6	0.436	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Any affective disorder	67	74	62	73	0.204	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Internalizing disorders in FDRs											
Anxiety disorders											
Panic disorder without agoraphobia	5	14	13	13	<0.001	*	*	*	N.S.	N.S.	N.S.
Panic disorder with agoraphobia	3	8	5	11	0.002	N.S.	N.S.	*	N.S.	N.S.	*
Agoraphobia without panic attacks	2	1	1	2	0.770	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Anxiety disorder NOS	1	3	1	3	0.340	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
GAD	0	0	0.3	0	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
PTSD	0.3	1	1	1	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Any anxiety disorder except phobia and OCD (v1)[f]	11	25	19	27	<0.001	*	*	*	N.S.	N.S.	N.S.
Social phobia	3	8	4	7	0.143	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Specific phobia	4	12	7	10	0.007	*	N.S.	*	N.S.	N.S.	N.S.
Any phobia (v2)[f]	7	16	10	16	0.003	*	N.S.	*	N.S.	N.S.	*
OCD (v3)[f]	2	10	2	3	0.036	*	N.S.	N.S.	*	*	N.S.

Any anxiety disorders	16	34	26	34	<0.001	*	*	*	N.S.	N.S.	N.S.
Somatization disorder	0.3	0	0.3	2	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Any eating disorder	4	6	3	4	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Any eating or somatoform disorder (v4)[f]	4	6	3	6	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Any Int disorder	18	35	27	36	<0.001	*	*	*	N.S.	N.S.	N.S.
Two, three or four Int variables	4	17	7	13	<0.001	*	N.S.	*	*	N.S.	N.S.
Ext disorders in FDRs											
Alcohol dependence	17	20	27	28	0.002	N.S.	*	*	N.S.	N.S.	N.S.
Alcohol abuse	8	10	14	11	0.077	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Alcohol dependence or abuse (v5)[f]	25	29	40	37	<0.001	N.S.	*	*	*	N.S.	N.S.
Drug dependence	10	15	20	19	0.002	N.S.	*	*	N.S.	N.S.	N.S.
Drug abuse	9	5	12	12	0.038	N.S.	N.S.	N.S.	*	*	N.S.
Drug dependence or abuse (v6)[f]	15	18	27	25	<0.001	N.S.	*	*	*	*	N.S.
Any substance use disorder	29	34	47	45	<0.001	N.S.	*	*	*	*	N.S.
Antisocial personality disorder	1	2	4	3	0.025	N.S.	*	*	N.S.	N.S.	N.S.
Borderline personality disorder	0.3	0	0.3	0.4	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Pathological gambling	1	2	1	3	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Conduct disorder	1	2	4	1	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
ADHD	0.5	0	0	1	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Cluster B, impulse, conduct, ADHD (v7)[f]	4	5	8	7	0.187	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Any Ext disorder	30	36	47	47	<0.001	N.S.	*	*	*	*	N.S.
Two or three Ext variables	12	15	21	20	0.001	N.S.	*	*	*	N.S.	N.S.
Habitual smoking	33	26	35	42	0.034	N.S.	N.S.	*	N.S.	*	N.S.

Data are given as column percentages unless otherwise indicated.

FDR, First-degree relative; MDD, major depressive disorder; BPI, any BPI including BPI with mania and MDD, BPI manic never MDD, BPI mixed never manic, or schizo-affective bipolar; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th edition; Int, internalizing; Ext, externalizing; N.S., not significant; N.A., not applicable; MDDR, MDD recurrent; SEMD, single-episode MDD; NOS, not otherwise specified; GAD, generalized anxiety disorder; PTSD, post-traumatic stress disorder; OCD, obsessive compulsive disorder; ADHD, attention-deficit/hyperactivity disorder; GEE, generalized estimating equations.

[a] Each row represents a different logistic regression model, accounting for within-family correlations by using GEE (SAS GENMOD with GEE) for binary dependent variables and maximum likelihood estimation with variance adjustment based on the cluster structure (SAS SURVEYLOGISTIC) for the unordered eight-category dependent variable (eight clusters in relatives).

[b] The number of FDRs with BPI was mostly predetermined because BPI disorder in at least one FDR of each proband was an ascertainment criterion.

[c] An 'N.A.' in the column for omnibus p indicates that the models could not converge due to low prevalence of these disorders.

[d] Any bipolar disorder NOS, diagnosed with bipolar disorder NOS or bipolar II SEMD or cyclothymia or hypomania.

[e] Any depressive disorder NOS, diagnosed with depressive disorder NOS or dysthymia or adjustment disorder.

[f] The seven variables (v1 to v7) were defined *a priori* before analyses. In the first column, the rows with (v1), (v2), (v3), (v4), (v5), (v6) and (v7) represent the seven variables analysed in the confirmatory factor analysis and confirmatory latent class analysis models.

* $p < 0.05$.

inter-episode functioning, rapid switching and mixed states. The same pattern is seen in Table 3. Those subjects in the group with both externalizing and internalizing disorders are more symptomatic in multiple areas than subjects in the other clusters.

Are subjects with only internalizing co-morbid disorders different from subjects with only externalizing co-morbid disorders? First, they are more likely to be female (Table 3). Thus, it is important to emphasize that the comparisons of other clinical course variables in Tables 2 and 3, many of which were significant, were adjusted for gender.

Additionally, at least one externalizing disorder seems to have less of an impact on course of illness compared with having at least one internalizing disorder (consider episode frequency, inter-episode GAS, rapid switching and mixed states, all of which were different from 'pure' BP in subjects with one internalizing disorder but not in those with one externalizing disorder; Table 2). A single externalizing disorder, on the other hand, is associated with an increased chance of psychotic symptoms, whereas having one or more internalizing disorders is not. The familial analysis is notable in that alcohol and drug use disorders both aggregate in the relatives of subjects with externalizing disorders but not in the relatives of subjects without such disorders.

The results here, using different statistical methods, are largely consistent with our previous results found in a high-risk study using a different dataset entirely; specifically, two groups of childhood disorders (anxiety and externalizing) predicted subsequent major affective illness in adolescents of families with probands with adult BP disorder (Nurnberger *et al.* 2011). In the present study, the two same factors were shown through confirmatory tests of hypotheses to distinguish subtypes of adults with BPI.

Numerous studies have used exploratory factor analysis and CFA to factor analyse symptoms of BP, including the symptoms or signs of mania (Bräunig *et al.* 1996; Serretti *et al.* 1999; Perugi *et al.* 2001; Faraone *et al.* 2004; Krüger *et al.* 2010), other mood or psychotic symptoms (Bauer *et al.* 1991; Dilsaver *et al.* 1999; Akiskal *et al.* 2001; Swann *et al.* 2001, 2008; Daneluzzo *et al.* 2002; Sato *et al.* 2002; González-Pinto *et al.* 2003; Berk *et al.* 2007; Henry *et al.* 2007; Adida *et al.* 2008; Adler *et al.* 2008; Erkiran *et al.* 2008; Harvey *et al.* 2008; Lindenmayer *et al.* 2008; Cavanagh *et al.* 2009; Gupta *et al.* 2009; Thompson *et al.* 2010; Mitchell *et al.* 2013), cognitive and energy features (Cassano *et al.* 2009), temperament (Evans *et al.* 2005), attention-deficit/hyperactivity disorder (ADHD) features (Joo *et al.* 2010) or childhood trauma signs (Garno *et al.* 2005). Resulting factors were shown to be modestly beneficial for genetic mapping in studies of BP, with respect to mood-disturbance factors (Faraone *et al.* 2004; Savitz *et al.* 2008) and an ADHD 'inattention' factor (Joo *et al.* 2010). We anticipate that subtyping using clinical co-morbidity will continue to shed light on the clinical and genetic characteristics of BPI.

To date, there have been only four published studies using cluster analysis or LCA among BPI patients to determine subtypes, all using the exploratory approach (Dilsaver *et al.* 1999; Cassidy *et al.* 2001; Swann *et al.* 2001; Sato *et al.* 2002). These studies cluster-analysed either the factor scores derived from the symptoms and behavioral ratings related to BPI or the symptoms and ratings themselves (Cassidy *et al.* 2001). Their results revealed more similarities than differences (Cassidy & Carroll, 2003; Sato *et al.* 2003). For example, all four studies found a pure (i.e. predominantly euphoric) subtype and a mixed (i.e. depressive or anxious–depressive) subtype. Using a confirmatory approach applied to lifetime disorders, we found supporting evidence for a four-class lifetime co-morbidity model in BPI individuals. A future CLCA on symptoms common to multiple co-morbid disorders among BPI patients would add beneficial knowledge to the present CLCA findings of distinct lifetime co-morbid disorders.

In previous exploratory BPI cluster analyses, differences were found between clusters on acute pharmacological treatment response (Swann *et al.* 2002); masked independent clinical classification of mixed states but not on gender (Dilsaver *et al.* 1999); and gender, suicidality at admission, and social adjustment and residual symptoms at discharge, but not on age of onset (Sato *et al.* 2002). The present study, based on lifetime diagnoses instead of symptoms, established additional differences between classes that demonstrated confirmatory support. It should be noted that our sample size was larger, with more power to detect group differences on clinical course variables, than the previous studies which consisted of 105 (Dilsaver *et al.* 1999), 162 (Swann *et al.* 2001), 327 (Cassidy *et al.* 2001) and 576 (Sato *et al.* 2002) BPI in-patients.

Limitations

It is possible that other variables not measured in the present study, such as traumatic stress, could improve the fit of the CFA and CLCA models. Medications could be a confounding or determining factor for clustering; however, the present dataset was not capable of addressing this issue. Test–retest reliability was not assessed for the co-morbid disorders. This multiplex sample is highly familial and the results may or may not be applicable to sporadic BPI disorders.

Conclusions

There appears to be strong evidence for the two-factor and four-cluster lifetime co-morbidity hypotheses. These hypotheses may be useful for understanding etiology and risk. The factors and clusters could be useful for parsimoniously reducing the number of variables in analyses of family, high-risk and case–control studies of persons living with BPI. These constructs may be useful for subtyping BP disorders and for prognosis by predicting course and severity of illness. Subtypes may also be useful for refining the phenotype in genetic studies. Subtypes may also have value for personalized medicine (Hamburg & Collins, 2010), assuming clinical subtypes could be linked to molecularly distinct subtypes, which may lead to new therapeutic possibilities, either through the development of targeted drugs or the salvaging of abandoned or failed drugs by identifying subgroups of patients likely to benefit from them.

Supplementary material

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0033291715000185

Acknowledgements

The analysis and preparation of this paper were supported by National Institutes of Health grant R01 MH068009. Data and biomaterials were collected in four projects that participated in the NIMH Bipolar Disorder Genetics Initiative. From 1991 to 1998, the principal investigators and co-investigators were: Indiana University, Indianapolis, IN (U01 MH46282) – J. Nurnberger, M. Miller and E. Bowman; Washington University, St Louis, MO (U01 MH46280) – T. Reich, A. Goate and J. Rice; Johns Hopkins University, Baltimore, MD (U01 MH46274) – J. R. DePaulo Jr, S. Simpson and C. Stine; NIMH Intramural Research Program, Clinical Neurogenetics Branch, Bethesda, MD – E. Gershon, D. Kazuba and E. Maxwell. Data and biomaterials were collected as part of 10 projects that participated in the NIMH Bipolar Disorder Genetics Initiative. From 1999 to 2007, the principal investigators and co-investigators were: Indiana University, Indianapolis, IN (R01 MH59545) – J. Nurnberger, M. J. Miller, E. S. Bowman, N. L. Rau, P. R. Moe, N. Samavedy, R. El-Mallakh (at University of Louisville), H. Manji (at Wayne State University), D. A Glitz (at Wayne State University), E. T. Meyer, C. Smiley, T. Foroud, L. Flury, D. M. Dick and H. Edenberg; Washington University, St Louis, MO (R01 MH059534) – J. Rice, T. Reich, A. Goate and L. Bierut; Johns Hopkins University, Baltimore, MD (R01 MH59533) – M. McInnis, J. R. DePaulo Jr, D. F. MacKinnon, F. M. Mondimore, J. B. Potash, P. P. Zandi, D. Avramopoulos and J. Payne; University of Pennsylvania, PA (R01 MH59553) – W. Berrettini; University of California at Irvine, CA (R01 MH60068) – W. Byerley and M. Vawter; University of Iowa, IA (R01 MH059548) – W. Coryell and R. Crowe; University of Chicago, IL (R01 MH59535) – E. Gershon, J. Badner, F. McMahon, C. Liu, A. Sanders, M. Caserta, S. Dinwiddie, T. Nguyen and D. Harakal; University of California at San Diego, CA (R01 MH59567) – J. Kelsoe and R. McKinney; Rush University, IL (R01 MH059556) – W. Scheftner, H. M Kravitz, D. Marta, A. Vaughn-Brown and L. Bederow; NIMH Intramural Research Program, Bethesda, MD (1Z01MH002810-01) – F. J. McMahon, L. Kassem, S. Detera-Wadleigh, L. Austin and D. L. Murphy.

Declaration of Interest

None.

References

Adida M, Clark L, Pomietto P, Kaladjian A, Besnier N, Azorin JM, Jeanningros R, Goodwin GM (2008). Lack of insight may predict impaired decision making in manic patients. *Bipolar Disorders* **10**, 829–837.

Adler M, Liberg B, Andersson S, Isacsson G, Hetta J (2008). Development and validation of the Affective Self Rating Scale for manic, depressive, and mixed affective states. *Nordic Journal of Psychiatry* **62**, 130–135.

Akiskal HS, Hantouche EG, Bourgeois ML, Azorin JM, Sechter D, Allilaire JF, Chatenêt-Duchêne L, Lancrenon S (2001). Toward a refined phenomenology of mania: combining clinician-assessment and self-report in the French EPIMAN study. *Journal of Affective Disorders* **67**, 89–96.

Bauer MS, Crits-Christoph P, Ball WA, Dewees E, McAllister T, Alahi P, Cacciola J, Whybrow PC (1991). Independent assessment of manic and depressive symptoms by self-rating. Scale characteristics and implications for the study of mania. *Archives of General Psychiatry* **48**, 807–812.

Berk M, Malhi GS, Cahill C, Carman AC, Hadzi-Pavlovic D, Hawkins MT, Tohen M, Mitchell PB (2007). The Bipolar Depression Rating Scale (BDRS): its development, validation and utility. *Bipolar Disorders* **9**, 571–579.

Birmaher B, Kennah A, Brent D, Ehmann M, Bridge J, Axelson D (2002). Is bipolar disorder specifically associated with panic disorder in youths? *Journal of Clinical Psychiatry* **63**, 414–419.

Bräunig P, Shugar G, Krüger S (1996). An investigation of the Self-Report Manic Inventory as a diagnostic and severity scale for mania. *Comprehensive Psychiatry* **37**, 52–55.

Cassano GB, Mula M, Rucci P, Miniati M, Frank E, Kupfer DJ, Oppo A, Calugi S, Maggi L, Gibbons R, Fagiolini A (2009). The structure of lifetime manic–hypomanic spectrum. *Journal of Affective Disorders* **112**, 59–70.

Cassidy F, Carroll BJ (2003). Symptom factors and clinical subtypes in mania. *American Journal of Psychiatry* **160**, 392.

Cassidy F, Pieper CF, Carroll BJ (2001). Subtypes of mania determined by grade of membership analysis. *Neuropsychopharmacology* **25**, 373–383.

Cassidy F, Yatham LN, Berk M, Grof P (2008). Pure and mixed manic subtypes: a review of diagnostic classification and validation. *Bipolar Disorders* **10**, 131–143.

Cavanagh J, Schwannauer M, Power M, Goodwin GM (2009). A novel scale for measuring mixed states in bipolar disorder. *Clinical Psychology and Psychotherapy* **16**, 497–509.

Chang K, Steiner H, Ketter T (2003). Studies of offspring of parents with bipolar disorder. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics* **123**, 26–35.

Chang KD, Steiner H, Ketter TA (2000). Psychiatric phenomenology of child and adolescent bipolar offspring. *Journal of the American Academy of Child and Adolescent Psychiatry* **39**, 453–460.

Cheng R, Juo SH, Loth JE, Nee J, Iossifov I, Blumenthal R, Sharpe L, Kanyas K, Lerer B, Lilliston B, Smith M, Trautman K, Gilliam TC, Endicott J, Baron M (2006). Genome-wide linkage scan in a large bipolar disorder sample from the National Institute of Mental Health Genetics Initiative suggests putative loci for bipolar disorder, psychosis, suicide, and panic disorder. *Molecular Psychiatry* **11**, 252–260.

Clark SL, Muthén B, Kaprio J, D'Onofrio BM, Viken R, Rose RJ (2013). Models and strategies for factor mixture analysis: an example concerning the structure underlying psychological disorders. *Structural Equation Modeling* **20**, 681–703.

Daneluzzo E, Arduini L, Rinaldi O, Di Domenico M, Petruzzi C, Kalyvoka A, Rossi A (2002). PANSS factors and scores in schizophrenic and bipolar disorders during an index acute episode: a further analysis of the cognitive component. *Schizophrenia Research* **56**, 129–136.

DelBello MP, Geller B (2001). Review of studies of child and adolescent offspring of bipolar parents. *Bipolar Disorders* **3**, 325–334.

DelBello MP, Strakowski SM, Sax KW, McElroy SL, Keck Jr PE, West SA, Kmetz GF (1999). Familial rates of affective and substance use disorders in patients with first-episode mania. *Journal of Affective Disorders* **56**, 55–60.

Dilsaver SC, Chen YR, Shoaib AM, Swann AC (1999). Phenomenology of mania: evidence for distinct depressed, dysphoric, and euphoric presentations. *American Journal of Psychiatry* **156**, 426–430.

Duffy A, Grof P, Grof E, Zvolsky P, Alda M (1998). Evidence supporting the independent inheritance of primary affective disorders and primary alcoholism in the families of bipolar patients. *Journal of Affective Disorders* **50**, 91–96.

Erkiran M, Sönmez G, Evren C, Aytaçlar S, Oral T (2008). The factor analytic symptom structure of manic episode and its relationship with affective temperaments [article in Turkish]. *Turkish Journal of Psychiatry* **19**, 157–166.

Evans L, Akiskal HS, Keck PE, McElroy Jr SL, Sadovnick AD, Remick RA, Kelsoe JR (2005). Familiality of temperament in bipolar disorder: support for a genetic spectrum. *Journal of Affective Disorders* **85**, 153–168.

Faraone SV, Su J, Tsuang MT (2004). A genome-wide scan of symptom dimensions in bipolar disorder pedigrees of adult probands. *Journal of Affective Disorders* **82** (Suppl. 1), S71–S78.

Feinman JA, Dunner DL (1996). The effect of alcohol and substance abuse on the course of bipolar affective disorder. *Journal of Affective Disorders* **37**, 43–49.

Finch WH, Bronk KC (2011). Conducting confirmatory latent class analysis using MPlus. *Structural Equation Modeling* **18**, 132–151.

Garno JL, Goldberg JF, Ramirez PM, Ritzler BA (2005). Bipolar disorder with comorbid cluster B personality disorder features: impact on suicidality. *Journal of Clinical Psychiatry* **66**, 339–345.

Gershon ES, Berrettini W, Nurnberger Jr J, Goldin LR (1987). Genetics of affective illness. In *Psychopharmacology: The Third Generation of Progress* (ed. H. Y. Meltzer), pp. 481–491. Raven Press: New York.

Gershon ES, Hamovit J, Guroff JJ, Dibble E, Leckman JF, Sceery W, Targum SD, Nurnberger Jr JI, Goldin LR, Bunney Jr WE (1982). A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. *Archives of General Psychiatry* **39**, 1157–1167.

Gershon ES, McKnew D, Cytryn L, Hamovit J, Schreiber J, Hibbs E, Pellegrini D (1985). Diagnoses in school-age children of bipolar affective disorder patients and normal controls. *Journal of Affective Disorders* **8**, 283–291.

González-Pinto A, Ballesteros J, Aldama A, Pérez de Heredia JL, Gutierrez M, Mosquera F, González-Pinto A (2003). Principal components of mania. *Journal of Affective Disorders* **76**, 95–102.

Goodwin FK, Jamison KR (1990). *Manic–Depressive Illness*. Oxford University Press: New York.

Goodwin RD, Hamilton SP (2002). The early-onset fearful panic attack as a predictor of severe psychopathology. *Psychiatry Research* **109**, 71–79.

Gupta SC, Sinha VK, Praharaj SK, Gandotra S (2009). Factor structure of manic symptoms. *Australian and New Zealand Journal of Psychiatry* **43**, 1141–1146.

Hamburg MA, Collins FS (2010). The path to personalized medicine. *New England Journal of Medicine* **363**, 301–304.

Harvey PD, Endicott JM, Loebel AD (2008). The factor structure of clinical symptoms in mixed and manic episodes prior to and after antipsychotic treatment. *Bipolar Disorders* **10**, 900–906.

Helzer JE, Winokur G (1974). A family interview study of male manic depressives. *Archives of General Psychiatry* **31**, 73–77.

Henry C, M'baïlara K, Poinsot R, Falissard B (2007). Construction and validation of a dimensional scale for mood disorders: Multidimensional Assessment of Thymic States (MAThyS) [article in French]. *Encephale* **33**, 768–774.

Hu L-T, Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria *versus* new alternatives. *Structural Equation Modeling* **6**, 1–55.

Johnson JG, Cohen P, Brook JS (2000). Associations between bipolar disorder and other psychiatric disorders during adolescence and early adulthood: a community-based longitudinal investigation. *American Journal of Psychiatry* **157**, 1679–1681.

Joo EJ, Greenwood TA, Schork N, McKinney RA, Sadovnick AD, Remick RA, Keck PE, McElroy SL, Kelsoe JR (2010). Suggestive evidence for linkage of ADHD features in bipolar disorder to chromosome 10p14. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics* **153B**, 260–268.

Kendler KS, Heath AC, Neale MC, Kessler RC, Eaves LJ (1993). Alcoholism and major depression in women: a twin study of the causes of comorbidity. *Archives of General Psychiatry* **50**, 690–698.

Kessler RC, Ormel J, Petukhova M, McLaughlin KA, Green JG, Russo LJ, Stein DJ, Zaslavsky AM, Aguilar-Gaxiola S, Alonso J, Andrade L, Benjet C, de Girolamo G, de Graaf R, Demyttenaere K, Fayyad J, Haro JM, Hu CY, Karam A, Lee S, Lepine JP, Matchsinger H, Mihaescu-Pintia C, Posada-Villa J, Sagar R, Ustün TB (2011). Development of lifetime comorbidity in the World Health Organization World Mental Health Surveys. *Archives of General Psychiatry* **68**, 90–100.

Krüger S, Quilty L, Bagby M, Lippold T, Bermpohl F, Bräunig P (2010). The Observer-Rated Scale for Mania (ORSM): development, psychometric properties and utility. *Journal of Affective Disorders* **122**, 179–183.

Lapalme M, Hodgins S, LaRoche C (1997). Children of parents with bipolar disorder: a metaanalysis of risk for mental disorders. *Canadian Journal of Psychiatry* **42**, 623–631.

Liang KY, Zeger SL (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Lindenmayer JP, Bossie CA, Kujawa M, Zhu Y, Canuso CM (2008). Dimensions of psychosis in patients with bipolar mania as measured by the positive and negative syndrome scale. *Psychopathology* **41**, 264–270.

Lo Y, Mendell N, Rubin DB (2001). Testing the number of components in a normal mixture. *Biometrika* **88**, 767–778.

Lubke GH, Muthén B (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods* **10**, 21–39.

MacKinnon DF, Xu J, McMahon FJ, Simpson SG, Stine OC, McInnis MG, DePaulo JR (1998). Bipolar disorder and panic disorder in families: an analysis of chromosome 18 data. *American Journal of Psychiatry* **155**, 829–831.

MacKinnon DF, Zandi PP, Cooper J, Potash JB, Simpson SG, Gershon E, Nurnberger J, Reich T, DePaulo JR (2002). Comorbid bipolar disorder and panic disorder in families with a high prevalence of bipolar disorder. *American Journal of Psychiatry* **159**, 30–35.

MacQueen GM, Hajek T, Alda M (2005). The phenotypes of bipolar disorder: relevance for genetic investigations. *Molecular Psychiatry* **10**, 811–826.

Maier W, Merikangas K (1996). Co-occurrence and cotransmission of affective disorders and alcoholism in families. *British Journal of Psychiatry* **30** (Suppl.), 93–100.

McMahon FJ, Simpson SG, McInnis MG, Badner JA, MacKinnon DF, DePaulo JR (2001). Linkage of bipolar disorder to chromosome 18q and the validity of bipolar II disorder. *Archives of General Psychiatry* **58**, 1025–1031.

Mitchell PB, Hadzi-Pavlovic D, Evoniuk G, Calabrese JR, Bowden CL (2013). A factor analytic study in bipolar depression, and response to lamotrigine. *CNS Spectrums* **18**, 214–224.

Morrison JR (1974). Bipolar affective disorder and alcoholism. *American Journal of Psychiatry* **131**, 1130–1133.

Morrison JR (1975). The family histories of manic–depressive patients with and without alcoholism. *Journal of Nervous and Mental Disease* **160**, 227–229.

Muthén B, Asparouhov T (2006). Item response mixture modeling: application to tobacco dependence criteria. *Addictive Behaviors* **31**, 1050–1066.

Muthén LK, Muthén BO (1998–2007). *Mplus User's Guide*, 5th edn. Muthén & Muthén: Los Angeles, CA.

Nunnally JC, Bernstein IH (1994). *Psychometric Theory*. McGraw-Hill: New York.

Nurnberger Jr JI (2002). Implications of multifactorial inheritance for identification of genetic mechanisms in major psychiatric disorders [Erratum in: *Psychiatric Genetics* 2003, **13**, 59]. *Psychiatric Genetics* **12**, 121–126.

Nurnberger Jr JI, Berrettini WH (1998). *Psychiatric Genetics*. Chapman & Hall: London.

Nurnberger Jr JI, Hamovit J, Hibbs E, Pellegrini D, Guroff J, Maxwell E, Smith A, Gershon ES (1988). A high-risk study of primary affective disorder: selection of subjects, initial assessment and 1- to 2-year follow-up. In *Relatives at Risk for Mental Disorder* (ed. D. L. Dunner, E. S. Gershon and J. E. Barrett), pp. 161–177. Raven Press: New York.

Nurnberger Jr JI, Kuperman S, Flury-Wetherill L, Meyer ET, Lawson WB, MacKinnon DF (2007). Genetics of comorbid mood disorder and alcohol dependence. *Journal of Dual Diagnosis* **3**, 31–46.

Nurnberger Jr JI, McInnis M, Reich W, Kastelic E, Wilcox HC, Glowinski A, Mitchell P, Fisher C, Erpe M, Gershon ES, Berrettini W, Laite G, Schweitzer R, Rhoadarmer K, Coleman VV, Cai X, Azzouz F, Liu H, Kamali M, Brucksch C, Monahan PO (2011). A high-risk study of bipolar disorder. Childhood clinical phenotypes as precursors of major mood disorders. *Archives of General Psychiatry* **68**, 1012–1020.

Payne JL, Potash JB, DePaulo Jr JR (2005). Recent findings on the genetic basis of bipolar disorder. *Psychiatric Clinics of North America* **28**, 481–498.

Perugi G, Maremmani I, Toni C, Madaro D, Mata B, Akiskal HS (2001). The contrasting influence of depressive and hyperthymic temperaments on psychometrically derived manic subtypes. *Psychiatry Research* **101**, 249–258.

Potash JB, DePaulo Jr JR (2000). Searching high and low: a review of the genetics of bipolar disorder. *Bipolar Disorders* **2**, 8–26.

Sato T, Bottlender R, Kleindienst N, Möller HJ (2002). Syndromes and phenomenological subtypes underlying acute mania: a factor analytic study of 576 manic patients. *American Journal of Psychiatry* **159**, 968–974.

Sato T, Bottlender R, Kleindienst N, Schröter A, Möller HJ (2003). Dr. Sato and colleagues reply. *American Journal of Psychiatry* **160**, 392–393.

Saunders EF, Zhang P, Copeland JN, Mclnnis MG, Zöllner S (2009). Suggestive linkage at 9p22 in bipolar disorder weighted by alcohol abuse. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics* **150B**, 1133–1138.

Savitz J, van der Merwe L, Ramesar R (2008). Personality endophenotypes for bipolar affective disorder: a family-based genetic association analysis. *Genes, Brain, and Behavior* **7**, 869–876.

Schulze TG, McMahon FJ (2003). Genetic linkage and association studies in bipolar affective disorder: a time for optimism. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics* **123C**, 36–47.

Schurhoff F, Bellivier F, Jouvent R, Mouren-Simeoni MC, Bouvard M, Allilaire JF, Leboyer M (2000). Early and late onset bipolar disorders: two different forms of manic–depressive illness? *Journal of Affective Disorders* **58**, 215–221.

Serretti A, Rietschel M, Lattuada E, Krauss H, Held T, Nöthen MM, Smeraldi E (1999). Factor analysis of mania. *Archives of General Psychiatry* **56**, 671–672.

Smoller JW, Finn CT (2003). Family, twin, and adoption studies of bipolar disorder. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics* **123C**, 48–58.

Sonne SC, Brady KT (1999). Substance abuse and bipolar comorbidity. In *The Psychiatric Clinics of North America* (ed. H. S. Akiskal and C. A. Bewick), pp. 609–627. W.B. Saunders Company: Philadelphia.

Strakowski SM, DelBello MP (2000). The co-occurrence of bipolar and substance use disorders. *Clinical Psychology Review* **20**, 191–206.

Swann AC, Bowden CL, Calabrese JR, Dilsaver SC, Morris DD (2002). Pattern of response to divalproex, lithium, or placebo in four naturalistic subtypes of mania. *Neuropsychopharmacology* **26**, 530–536.

Swann AC, Janicak PL, Calabrese JR, Bowden CL, Dilsaver SC, Morris DD, Petty F, Davis JM (2001). Structure of mania: depressive, irritable, and psychotic clusters with different retrospectively-assessed course patterns of illness in randomized clinical trial participants. *Journal of Affective Disorders* **67**, 123–132.

Swann AC, Steinberg JL, Lijffijt M, Moeller FG (2008). Impulsivity: differential relationship to depression and mania in bipolar disorder. *Journal of Affective Disorders* **106**, 241–248.

Thompson PM, Gonzalez JM, Singh V, Schoolfield JD, Katz MM, Bowden CL (2010). Principal domains of behavioral psychopathology identified by the Bipolar Inventory of Signs and Symptoms Scale (BISS). *Psychiatry Research* **175**, 221–226.

Todd RD, Reich W, Petti TA, Joshi P, DePaulo JR, Nurnberger J, Reich T (1996). Psychiatric diagnoses in the child and adolescent members of extended families identified through adult bipolar affective disorder probands. *Journal of the American Academy of Child and Adolescent Psychiatry* **35**, 664–671.

Tsuang MT, Faraone SV (1990). *The Genetics of Mood Disorders*. Johns Hopkins University Press: Baltimore.

Winokur G, Coryell W, Akiskal HS, Maser JD, Keller MB, Endicott J, Mueller T (1995). Alcoholism in manic–depressive (bipolar) illness: familial illness, course of illness, and the primary–secondary distinction. *American Journal of Psychiatry* **152**, 365–372.

Winokur G, Coryell W, Endicott J, Keller M, Akiskal HS, Solomon D (1996). Familial alcoholism in manic–depressive (bipolar) disease. *American Journal of Medical Genetics* **67**, 197–201.

Winokur G, Reich T, Rimmer J, Pitts Jr FN (1970). Alcoholism. III. Diagnosis and familial psychiatric illness in 259 alcoholic probands. *Archives of General Psychiatry* **23**, 104–111.

World Health Organization (2002). *The World Health Report 2002*. World Health Organization: Geneva.

Wozniak J, Biederman J, Monuteaux MC, Richards J, Faraone SV (2002). Parsing the comorbidity between bipolar disorder and anxiety disorders: a familial risk analysis. *Journal of Child and Adolescence Psychopharmacology* **12**, 101–111.

Yu C-Y (2002). Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. Doctoral Dissertation, University of California, Los Angeles (http://www.statmodel.com/download/Yudissertation.pdf).