

Confirmatory test of two factors and four subtypes of bipolar disorder based on lifetime
psychiatric comorbidity

Patrick O. Monahan, PhD

Timothy Stump, MA

William H. Coryell, MD

Jaroslav Harezlak, PhD

George A. Marcoulides, PhD

Hai Liu, PhD

Christine M. Steeger, PhD

Philip B. Mitchell, MB BS, MD

Holly C. Wilcox, PhD

Leslie A. Hulvershorn, MD

Anne L. Glowinski, MD, MPE

Bipolar Disorder Genome Study (BiGS) Consortium*

Bipolar High Risk Study Group*

Priya Anapurna Iyer-Eimerbrink, PhD

Melvin McInnis, MD

John I. Nurnberger Jr., MD, PhD

Correspondence: Patrick O. Monahan, Department of Biostatistics, Indiana University, 410 West
10th Street, Suite 3000, Indianapolis, IN 46202-3002. E-mail: pmonahan@iu.edu. Phone: 317-
278-8086. Fax: 317-274-8022.

This is the author's manuscript of the article published in final edited form as:

Monahan, P. O., Stump, T., Coryell, W. H., Harezlak, J., Marcoulides, G. A., Liu, H., ... & Nurnberger, J. I. (2015).
Confirmatory test of two factors and four subtypes of bipolar disorder based on lifetime psychiatric co-morbidity.
Psychological medicine, 45(10), 2181-2196. <http://dx.doi.org/10.1017/S0033291715000185>

Author Affiliations: Department of Biostatistics (Drs Monahan, Harezlak, Liu, and Mr Stump), Department of Psychiatry (Drs Nurnberger, Hulvershorn, Iyer-Eimerbrink), Indiana University School of Medicine, Indianapolis; Department of Psychiatry (Dr Coryell), Roy J. and Lucille A. Carver College of Medicine, The University of Iowa, Iowa City; Research Methods & Statistics Program, Graduate School of Education (Dr Marcoulides), University of California–Riverside, Riverside; Department of Psychology (Ms Steeger), College of Arts and Letters, University of Notre Dame, Notre Dame; School of Psychiatry (Dr Mitchell), University of New South Wales, Sydney, and Black Dog Institute, Sydney; Department of Psychiatry & Behavioral Sciences (Dr Wilcox), Johns Hopkins School of Medicine, Baltimore; Department of Psychiatry (Glowinski), Washington University School of Medicine, St. Louis; Department of Psychiatry (McInnis), School of Medicine, University of Michigan, Ann Arbor.

*Authors/Group Information: The following members of the BiGS Consortium take authorship responsibility: William Byerley, MD, University of California–San Francisco, San Francisco (Drs Coryell, McInnis and Nurnberger are also members of this group). The following members of the Bipolar High Risk Study Group take authorship responsibility: Wendy Reich, PhD, Washington University, St. Louis; Elizabeth Kastelic, MD, Johns Hopkins Medical Institute, Baltimore, Maryland; and Gloria Roberts, School of Psychiatry, University of New South Wales, Sydney, and Black Dog Institute, Sydney (Drs Coryell, McInnis, Mitchell, Wilcox, and Nurnberger are also members of this group).

Acknowledgements: None

Financial support: The analysis and preparation of this manuscript was supported by NIH Grant R01 MH068009. Data and biomaterials were collected in four projects that participated in the National Institute of Mental Health (NIMH) Bipolar Disorder Genetics Initiative. From 1991 to 1998, the principal investigators and coinvestigators were: Indiana University, Indianapolis, IN, U01 MH46282, John Nurnberger, Marvin Miller and Elizabeth Bowman; Washington University, St Louis, MO, U01 MH46280, Theodore Reich, Allison Goate and John Rice; Johns Hopkins University, Baltimore, MD U01 MH46274, J Raymond DePaulo, Jr, Sylvia Simpson and Colin Stine; NIMH Intramural Research Program, Clinical Neurogenetics Branch, Bethesda, MD, Elliot Gershon, Diane Kazuba and Elizabeth Maxwell. Data and biomaterials were collected as part of 10 projects that participated in the National Institute of Mental Health (NIMH) Bipolar Disorder Genetics Initiative. From 1999 to 2007, the principal investigators and coinvestigators were: Indiana University, Indianapolis, IN, R01 MH59545, John Nurnberger, Marvin J Miller, Elizabeth S Bowman, N Leela Rau, P Ryan Moe, Nalini Samavedy, Rif El-Mallakh (at University of Louisville), Husseini Manji (at Wayne State University), Debra A Glitz (at Wayne State University), Eric T Meyer, Carrie Smiley, Tatiana Foroud, Leah Flury, Danielle M Dick, Howard Edenberg; Washington University, St Louis, MO, R01 MH059534, John Rice, Theodore Reich, Allison Goate, Laura Bierut; Johns Hopkins University, Baltimore, MD, R01 MH59533, Melvin McInnis, J Raymond DePaulo, Jr, Dean F MacKinnon, Francis M Mondimore, James B Potash, Peter P Zandi, Dimitrios Avramopoulos and Jennifer Payne; University of Pennsylvania, PA, R01 MH59553, Wade Berrettini; University of California at Irvine, CA, R01 MH60068, William Byerley and Mark Vawter; University of Iowa, IA, R01 MH059548, William Coryell and Raymond Crowe; University of Chicago, IL, R01 MH59535,

Elliot Gershon, Judith Badner, Francis McMahon, Chunyu Liu, Alan Sanders, Maria Caserta, Steven Dinwiddie, Tu Nguyen, Donna Harakal; University of California at San Diego, CA, R01 MH59567, John Kelsoe, Rebecca McKinney; Rush University, IL, R01 MH059556, William Scheftner, Howard M Kravitz, Diana Marta, Annette Vaughn-Brown and Laurie Bederow; NIMH Intramural Research Program, Bethesda, MD, 1Z01MH002810-01, Francis J McMahon, Layla Kassem, Sevilla Detera-Wadleigh, Lisa Austin, Dennis L Murphy.

Declaration of interest: For all authors, there were no financial involvement (including employment, fees, share ownership) or affiliation with any organization whose financial interests may be affected by material in the manuscript, or which might potentially bias it.

Ethical standards: The authors assert that 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.

ABSTRACT

Background: First, to use confirmatory factor analysis (CFA) to test a hypothesis that two factors (internalizing and externalizing) account for lifetime comorbid DSM-IV diagnoses among adults with Bipolar I (BPI) disorder. Second, 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.

Methods: A cohort of 699 multiplex BPI families, ascertained and assessed (1998-2003) by the NIMH 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. Comorbidity, 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.

Key words: bipolar disorder; subtypes; comorbidity; confirmatory factor analysis; confirmatory latent class analysis.

INTRODUCTION

Bipolar (BP) affective disorder affects 0.5-1.6% of the U.S. adult population and is frequently chronically debilitating (Goodwin and 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; McMahon *et al.*, 2001; Nurnberger and Berrettini, 1998; Potash and DePaulo, 2000; Smoller and Finn, 2003; Tsuang and Faraone, 1990). 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 comorbid psychiatric symptoms and disorders (Cassano *et al.*, 2009; Cassidy *et al.*, 2008; Nurnberger, 2002). The creation of homogeneous BPI subgroups based on comorbid conditions has modestly improved the success of genetic mapping (Cheng *et al.*, 2006; MacKinnon *et al.*, 1998; MacQueen *et al.*, 2005; Nurnberger, 2002; Payne *et al.*, 2005; Saunders *et al.*, 2009; Schulze and McMahon, 2003). For example, three subtypes of BPI, as defined by comorbidity, have provided some of the strongest evidence of linkage to genomic regions: comorbid panic disorder, comorbid psychotic symptoms, and pure BPI with low rates of comorbidity (MacQueen *et al.*, 2005). A subtype of BPI characterized by comorbid anxiety disorders has been supported by cross-sectional, longitudinal and familial studies (Birmaher *et al.*, 2002; Goodwin and Hamilton, 2002; Johnson *et al.*, 2000; MacKinnon *et al.*, 2002; Schurhoff *et al.*, 2000; Wozniak *et al.*, 2002). A BPI subtype characterized by high rates of comorbid drug and alcohol abuse or dependence (Nurnberger *et al.*, 2007; Sonne and Brady, 1999) has shown evidence for elevated genetic predisposition to substance use disorders (DelBello *et al.*, 1999; Duffy *et al.*, 1998; Feinman and Dunner, 1996; Gershon *et al.*, 1982; Helzer and Winokur, 1974; Kendler *et al.*, 1993; Maier and Merikangas, 1996; Morrison, 1974;

Morrison, 1975; Nurnberger *et al.*, 2007; Strakowski and DelBello, 2000; Winokur *et al.*, 1995; Winokur *et al.*, 1996; Winokur *et al.*, 1970). Finally, offspring of BPI patients (compared to controls) have higher rates of affective, anxiety, and externalizing disorders such as conduct disorder, oppositional disorder, and substance abuse (Chang *et al.*, 2003; Chang *et al.*, 2000; DelBello and Geller, 2001; Gershon *et al.*, 1985; Lapalme *et al.*, 1997; Nurnberger *et al.*, 2011; Nurnberger *et al.*, 1988; Todd *et al.*, 1996).

Factor analysis and cluster analysis are complementary methods for investigating BPI clinical subtypes. They determine, respectively, whether comorbid 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 Supplementary Text, Section B).

Factor analyses have been conducted on psychiatric lifetime disorders among *community* samples. For example, an internalizing and 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* comorbid with *BPI*.

The first purpose of this paper is 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 comorbidity 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 and 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 comorbid disorders, specified before CLCA was performed and then tested with CLCA. Based on literature cited above, we hypothesized four comorbidity subtypes of BPI patients: pure BPI without comorbidity, comorbidity with only internalizing disorders, comorbidity with only externalizing disorders, and comorbidity with both internalizing and externalizing disorders.

METHODS

Study Design and Procedures

The NIMH Genetics Initiative Bipolar Project was active from 1989 -2007. This article uses data from probands and relatives assessed between 1998 and 2003 using the DIGS 3.0 interview and resulting 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 Schizoaffective BP Type. (see diagnostic process in Supplementary Text, Section A). Diagnoses were coded as binary (0 = no, 1 = yes if “probable” or “definite”).

Statistical Methods

The MPLUS software (Version 5.21) (Muthen and Muthen, 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) > .95 (Hu and Bentler, 1999), root mean square error of approximation (RMSEA) < .06 (Hu and Bentler, 1999), and weighted root mean square residual (WRMR) < 1.00 (Yu, 2002). Loadings above 0.40 indicated adequate fit for individual paths (Nunnally and Bernstein, 1994). We tested (two-sided, 0.05 alpha) whether fit of the CFA model could be improved by adding or removing paths (see Supplementary Text, Section C). The robust weighted least squares (WLSMV) estimator and the nonlinear probit link were specified for CFA models. For LCA and CLCA models, maximum likelihood with robust standard error (MLR) estimation was used.

An important deterministic constraint in CLCA models is the hypothesized “zero” class for which the probabilities of endorsing all comorbid disorders are fixed to be zero (e.g., “pure BP” class) (Muthen and 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 (Clark *et al.*, 2013; Finch and Bronk, 2011). 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 4-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 Supplementary Text. These two competing 4-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 were evaluated based on the following criteria. Models with the lowest values for the Bayesian Information Criteria (BIC), and its sample size adjusted version (aBIC), were considered the best (Lubke and Muthén, 2005). The BIC and aBIC was used to compare different models with different number of classes and or parameterizations (Finch and Bronk, 2011). The likelihood ratio test (LRT) was used to compare the hypothesized 4-class CLCA models to the exploratory 4-class model, because the former is nested within the latter (Finch and Bronk, 2011). The LRT is not appropriate for comparing LCA models with differing number of classes (Lubke and Muthén, 2005). Therefore, the Lo-Mendell-Rubin (LMR)

(Lo *et al.*, 2001) likelihood ratio test was used to compare nested LCA models that have differing number of classes but the same parameterization (e.g., our exploratory 1-, 2-, 3-, 4-, 5-, and 6-class LCA models). The LMR p-value indicates whether a model with 1 fewer classes can be rejected in favor of the current model. Classes with a sparse number of persons are not practically meaningful (Lubke and 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.

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 and 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 alpha of .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).

RESULTS

Participant characteristics

The 504 probands (from 504 multiplex families) with BPI disorder (96%) or schizoaffective BP type (4%) were adults; age at interview ranged from 18 to 88 (except one was age 17). The 1157 first-degree relatives of the 504 probands included mostly adults (range 18-

93) with only eight (1%) adolescents (range 13-17). The high percentage of first-degree relatives with BPI or schizoaffective 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 schizoaffective BP type (504 adult probands; 594 first-degree relatives; and 58 more distant relatives) (Supplementary Table S1). At time of interview, 44% of those with BPI were married. Their median age of onset was 18.0. A majority was female and Caucasian. Prevalence of comorbid lifetime disorders is shown in Supplementary Table S1.

Test of the confirmatory factor analysis (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, and WRMR = .556). The null hypothesis of good fit was not rejected by the chi-square goodness of fit test (chi-square = 9.29, df = 11, p = 0.60). The two factors were significantly correlated as hypothesized ($r = .34$, $p < .001$). A one-factor model showed poor fit (CFI = .76; RMSEA = .094; WRMR = 1.65; chi-square = 156.63, df = 14, $p < .001$). 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 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 (1 point for each variable). Results were similar for the two methods (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: 0 disorders, 1 variable, and 2 or more variables (frequencies with 3 or 4 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 GAS score, 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 ($p = .029$). 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 Confirmatory Latent Class Analysis (CLCA) model

The fit of the hypothesized CLCA model was compared to 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 #1 and #2) fit better (lower BIC and aBIC) than all exploratory (1-class through 6-class) models (models #3 through #8). The 5-class and 6-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 #1) fit the best; its fit was slightly better than the competing CLCA model with full inequality constraints (model #2) and also better than all post-hoc modified CLCA models (models #9 through #16).

Further non-essential description of CLCA results in Table 3 is provided in Supplementary Text, Section D. In summary, none of the post-hoc modified 4-class CLCA models or any of the 1-class through 6-class exploratory LCA models fit better than the CLCA model that was hypothesized to have deterministic constraints and full equality constraints (model #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 #1), were compared on clinical course variables. Significant associations were found for 7 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 comorbid disorders (19.0) and lowest for those with both internalizing and externalizing variables (16.0). All pairwise differences were significant for age of 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 score. Persons in class 1 (pure BPI without comorbidity) 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 comorbidity had significantly greater history of rapid switching and mixed states than the cluster without comorbidity (class 1). Rapid switching was particularly elevated for the two clusters with internalizing disorders.

Familial Analysis

There were eight possible comorbidity clusters of first-degree relatives defined by crossing the presence of BPI or other affective disorders with the four comorbidity 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 -value $< .001$). In particular, rates for three of the relative clusters differed between the four proband clusters (4×2 omnibus tests, p -value $< .001$, $.045$, and $.001$). 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 to 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 comorbid factors in the context of BPI. Additional validity of the two-factor comorbidity 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 #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 4-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 comorbidity. “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 bipolar disorder that are typically regarded as evidence of more severe disorder, are differentially clustered in subjects with comorbid disorders.

More work is needed to evaluate whether the presence of comorbid 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 comorbid 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 interepisode functioning, rapid switching, and mixed states. The same pattern is seen in Table 4. 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 comorbid disorders different from subjects with only externalizing comorbid disorders? Firstly, they are more likely to be female (Table 4). Thus, it is important to emphasize that the comparisons of other clinical course variables in Tables 2 and 4, 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 to having at least one internalizing disorder (consider episode frequency, interepisode GAS score, 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 data set 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 EFA and CFA to factor analyze *symptoms* of BP, including the *symptoms* or *signs* of mania (Bräunig *et al.*, 1996; Faraone *et al.*, 2004; Krüger *et al.*, 2010; Perugi *et al.*, 2001; Serretti *et al.*, 1999), other mood or psychotic symptoms (Adida *et al.*, 2008; Adler *et al.*, 2008; Akiskal *et al.*, 2001; Bauer *et al.*, 1991; Berk *et al.*, 2007; Cavanagh *et al.*, 2009; Daneluzzo *et al.*, 2002; Dilsaver *et al.*, 1999; Erkiran *et al.*, 2008; González-Pinto *et al.*, 2003; Gupta *et al.*, 2009; Harvey *et al.*, 2008; Henry *et al.*, 2007; Lindenmayer *et al.*, 2008; Mitchell *et al.*, 2013; Sato *et al.*, 2002; Swann *et al.*, 2001; Swann *et al.*, 2008; Thompson *et al.*, 2010), 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 comorbidity 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 (Cassidy *et al.*, 2001; Dilsaver *et al.*, 1999; Sato *et al.*, 2002; Swann *et al.*, 2001). These studies cluster-analyzed 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 and 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 comorbidity model in BPI individuals. A future *CLCA* on symptoms common to multiple comorbid disorders among BPI patients would add beneficial knowledge to the present *CLCA* findings of distinct lifetime comorbid 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 inpatients.

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 comorbid 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 comorbidity 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 and 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.

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 & 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, PE, Jr., 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, SL, Jr., 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, J, Jr. & 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, JI, Jr., Goldin, LR & Bunney, WE, Jr** (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 & A., G-P** (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. *The 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, Poinot, 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.
- Muthen, B & Asparouhov, T** (2006). Item response mixture modeling: application to tobacco dependence criteria. *Addictive Behaviors* **31**, 1050-1066.
- Muthen, LK & Muthen, BO** (1998-2007). *Mplus User's Guide. Fifth Edition*. Muthen & Muthen: Los Angeles, CA.
- Nunnally, JC & Bernstein, IH** (1994). *Psychometric theory*. McGraw-Hill: New York.
- Nurnberger, JI, Jr.** (2002). Implications of multifactorial inheritance for identification of genetic mechanisms in major psychiatric disorders. *Psychiatric Genetics* **12**, 121-126 [Erratum in: *Psychiatr Genet*. 2003 Mar;13(1):59].
- Nurnberger, JI, Jr. , 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.
- Nurnberger Jr JI & Berrettini, WH** (1998). *Psychiatric Genetics*. Chapman & Hall.
- Nurnberger, JI, Jr., 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, JI, Jr., Kuperman, S, Flury-Wetherill, L, Meyer, ET, Lawson, WB, MacKinnon, DF & Alcoholism, FtNGIBGatCSotGo (2007). Genetics of comorbid mood disorder and alcohol dependence. *Journal of Dual Diagnosis* **3**, 31-46.

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, McInnis, 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, D.** (1996). Familial alcoholism in manic-depressive (bipolar) disease. *American Journal of Medical Genetics* **67**, 197-201.
- Winokur, G, Reich, T, Rimmer, J & Pitts, FN, Jr.** (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. Retrieved from <http://www.statmodel.com/download/Yudissertation.pdf>.

Table 1. Confirmatory Factor Analysis (CFA) of Comorbid Lifetime Disorders from 1156 BPI Probands and BPI Relatives

		<u>Number</u>	<u>Loading</u>	<u>SE</u>	<u>z-value</u>	<u>p-value</u>
<u>Variables</u>	<u>Factor 1 (Internalizing factor)</u>					
1	Panic, agoraphobia, anxiety NOS, GAD, PTSD	315	.81	.08	9.86	< .001
2	Any phobia disorder (specific or social)	164	.50	.06	9.35	< .001
3	Obsessive Compulsive disorder (OCD)	79	.55	.08	7.27	< .001
4	Any eating (anorexia, bulimia, NOS) or somatoform d/o	89	.55	.07	7.50	< .001
<u>Variables</u>	<u>Factor 2 (Externalizing factor)</u>					
5	Any alcohol abuse or dependence	444	.86	.07	13.30	< .001
6	Any drug abuse or dependence	329	.79	.06	14.11	< .001
7	Cluster B PD, impulse control, conduct d/o, ADHD	87	.50	.07	7.70	< .001
	Correlation between Factor 1 and Factor 2		.34	.06	5.31	< .001
Model fit						
			Fit indices		Chi-square test	
		<u>CFI</u>	<u>RMSEA</u>	<u>WRMR</u>	<u>Chi-square (df)</u>	<u>p-value</u>
		1.000	0.000	0.556	9.29 (11)	0.60

Table 1, Continued

Note. z-value = loading / standard error. d/o = disorder. NOS = not otherwise specified. GAD = generalized anxiety disorder.

PTSD = post-traumatic stress disorder. ADHD = attention deficit hyperactivity disorder.

Number = 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).

Table 2. Factors and Clinical Course of Illness Variables Among BPI Participants (BPI Probands and BPI Relatives)

	Factor 1 (Internalizing factor)					Factor 2 (Externalizing factor)				
	Factor score			Omnibus test	Post-hoc pairwise tests	Factor score			Omnibus test	Post-hoc pairwise tests
	<u>0</u>	<u>1</u>	<u>2, 3, or 4</u>			<u>0</u>	<u>1</u>	<u>2 or 3</u>		
Number of BPI participants:	711	288	157			609	274	273		
Factor score group number:	<u>(1)</u>	<u>(2)</u>	<u>(3)</u>	test	pairwise	<u>(1)</u>	<u>(2)</u>	<u>(3)</u>	test	pairwise
<u>Continuous Clinical Variables</u>	<u>median</u>	<u>median</u>	<u>median</u>	<u>p-value</u>	<u>tests</u>	<u>median</u>	<u>median</u>	<u>median</u>	<u>p-value</u>	<u>tests</u>
Age of onset	19	17	15	***	ALL	19	18	16	***	ALL
Episode frequency per years ill	0.43	0.57	0.87	***	ALL	0.48	0.48	0.58	NS	NA
Mania/depression episode ratio	1.00	0.80	1.00	NS	NA	1.00	1.00	1.00	NS	NA
Psychiatric hospitalizations per years ill	0.17	0.17	0.19	NS	NA	0.16	0.17	0.21	NS	NA
Between-episode GAS score	70	65	61	***	(1)>(2)(3)	70	70	61	***	(1)(2)>(3)
<u>Binary Clinical Variables (1=Yes, 0=No)</u>	<u>n (%)</u>	<u>n (%)</u>	<u>n (%)</u>			<u>n (%)</u>	<u>n (%)</u>	<u>n (%)</u>		
Female	55	69	81	***	ALL	68	62	50	***	(1)(2)>(3)
Disabled	17	20	25	*	(1)(2)<(3)	17	17	25	*	(1)(2)<(3)
Psychotic symptoms	47	52	62	NS	(1)<(3)	45	57	55	**	(1)<(2)(3)
Mood incongruent psychotic symptoms	11	12	14	NS	NA	11	13	11	NS	NA
Rapid cycling	9	10	13	NS	NA	8	10	13	NS	NA
Rapid switching	39	53	71	***	ALL	42	46	59	***	(1)(2)<(3)
Mixed states	23	30	36	**	(1)<(2)(3)	23	27	34	**	(1)<(3)

Table 2, Continued,

Note. 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. Mania/depression episode ratio = number of clean manic episodes over the number of clean depressive episodes. Psychiatric hospitalizations per years ill = number of hospitalizations divided by years of illness. Between-episode global assessment score (GAS) score = GAS score for the past month if not hospitalized, otherwise GAS score equals missing value.

Disabled = occupationally disabled, from the present job question in DIGS 3.0. 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 one day or intermittently for a period of three days. Mood-incongruent psychotic symptoms were present during either a mania or depression episode assessed in DIGS 3.0.

Rapid cycling = four or more discrete episodes of mania or depression within 12 months demarcated by eight weeks or more of remission.

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”. Mixed states = at least three symptoms of the opposite polarity lasting one 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. NS = not significant. ALL = all three pairwise differences were significant. NA = post-hoc pairwise tests not applicable because omnibus test not significant.

* $p < .05$, ** $p < .01$, *** $p < .001$. The sum score, computed separately for internalizing and externalizing variables, was specified in models as a categorical variable with three levels (0 disorders, 1 variable, 2 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 (GEE) to account for within-family correlations.

Table 3. Confirmatory Latent Class Analysis (CLCA) of 7 Lifetime Comorbid Diagnostic Variables (N=1156 for All Models)

Model #	Constraints	No. of classes	Log L	No. free parms	BIC	aBIC	LMR p	Most likely class membership size, n(%)			
								1	2	3	4
<u>Two competing CLCA models</u>											
1	Zero-c, Int-only-c, Ext-only-c, Both-c, Full Eq	4	-3288	10	6647	6616	<.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)
<u>Exploratory LCA models</u>											
3	None	1	-3520	7	7089	7067	N/A	1156 (100)			
4	None	2	-3343	15	6791	6744	<.0001	389 (34)	767 (66)		
5	None	3	-3283	23	6727	6654	<.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)
<u>Post-hoc modifications to CLCA models</u>											
9	Model #1, except Eq(Int only)	4	-3285	13	6661	6619	<.0001	422 (37)	187 (16)	289 (25)	258 (22)
10	Model #1, except Eq(Ext only)	4	-3288	14	6674	6629	<.0001	422 (37)	187 (16)	289 (25)	258 (22)
11	Model #1, except Eq(Alc only)	4	-3285	12	6655	6617	<.0001	422 (37)	187 (16)	289 (25)	258 (22)
12	Model #2, except InEq(Ext only)	4	-3284	17	6687	6633	N/A	422 (37)	187 (16)	289 (25)	258 (22)
13	Model #1 or #2, except Eq(Int only) & InEq(Ext only)	4	-3285	13	6661	6619	N/A	422 (37)	187 (16)	277 (24)	270 (23)
14	Model #1 or #2, except no Eq & no InEq constraints	4	-3284	17	6687	6633	<.0001	422 (37)	187 (16)	289 (25)	258 (22)
15	Model #1, except Int-only+Alc-c	4	-3288	11	6655	6620	<.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)

Table 3, Continued

Note. No. = number. Int = internalizing. Ext = externalizing. Log L = Log Likelihood. No. free parms = Number of free parameters.

Zero-c = zero-class specified by using a deterministic constraint to fix the probabilities of endorsing each of the 7 variables 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.

Int-only-c = internalizing-only class specified using a deterministic constraint to fix the probabilities of endorsing externalizing disorders to be zero.

Both-c = both-class specified by allowing all probabilities for endorsing the 4 internalizing and the 3 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.

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 4 internalizing variables, or only the 3 externalizing variables, or only the 1 alcohol variable, respectively.

InEq(Ext only) = instead of full inequality constraints on all variables, the inequality constraint was placed on only the 3 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. BIC = Bayesian Information Criterion.

aBIC = Sample size adjusted BIC. LMR p = Lo-Mendell Rubin test p-value. N/A = LRM test not available for models with (nonlinear) inequality constraints and not applicable for an exploratory model with 1 class. The most likely class membership size was 547 (47%) for class 5 for model #7, and were 17 (1%) and 101 (9%) for classes 5 and 6, respectively, for model #8.

Table 4. Clusters and Course of Illness Variables Among BPI Participants (BPI Probands and BPI Relatives)

	Clusters based on DSM-IV diagnosed co-morbid disorders					
	None of the <u>comorbid disorders</u>	Internalizing <u>d/o(s) only</u>	Externalizing <u>d/o(s) only</u>	Internalizing & <u>Externalizing d/o(s)</u>		
Number of BPI probands & BPI relatives:	422	187	289	258		Post-hoc
Cluster group number:	<u>(1)</u>	<u>(2)</u>	<u>(3)</u>	<u>(4)</u>	Omnibus	pairwise comparisons
<u>Continuous Clinical Variables</u>	<u>median</u>	<u>median</u>	<u>median</u>	<u>median</u>	<u>p-value</u>	<u>between clusters</u>
Age of onset of major affective disorder	19.0	17.0	18.0	16.0	***	(1)>(2)(3)(4); (2)>(4); (3)>(4)
Episode frequency per years ill	0.43	0.62	0.44	0.69	***	(1)<(2)(4); (2)>(3); (3)<(4)
Manic/depressive episode ratio	1.00	0.75	1.00	1.00	NS	NA
Psychiatric hospitalizations per years ill	0.17	0.16	0.19	0.19	NS	NA
Between-episode GAS score	70	65	70	64	***	(1)>(2)(4); (3)>(4)
<u>Binary Clinical Variables (1=Yes, 0=No)</u>	<u>n (%)</u>	<u>n (%)</u>	<u>n (%)</u>	<u>n (%)</u>		
Female	61	83	47	66	***	(1)<(2); (1)>(3); (2)>(3)(4); (3)<(4)
Disabled	16	18	18	24	NS	NA
Psychotic symptoms	41	54	56	56	**	(1)<(3)(4)
Mood incongruent psychotic symptoms	11	13	12	12	NS	NA
Rapid cycling	6	10	11	12	NS	NA
Rapid switching	34	59	46	60	***	(1)<(2)(3)(4); (3)<(4)
Mixed states	19	33	30	32	***	(1)<(2)(3)(4)

Table 4, Continued

Note. NS = not significant. NA = pairwise tests not applicable because omnibus test was not significant. * $p < .05$, ** $p < .01$, *** $p < .001$.

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 (GEE) to account for within-family correlations.

Table 5. Familial Aggregation: Clusters and Disorders of First Degree Relatives (FDRs) by BPI Proband Clusters

	Proband clusters based on DSM-IV co-morbid diagnoses				Omnibus p-value	Post-hoc pairwise comparisons between clusters									
	None of the <u>comorbid d/o(s)</u>	Internalizing <u>d/o(s) only</u>	Externalizing <u>d/o(s) only</u>	Internalizing & <u>Externalizing d/o(s)</u>		(1)	(1)	(1)	(2)	(2)	(3)				
	<u>(1)</u>	<u>(2)</u>	<u>(3)</u>	<u>(4)</u>		<u>(2)</u>	<u>(3)</u>	<u>(4)</u>	<u>(3)</u>	<u>(4)</u>	<u>(4)</u>				
Cluster group number:															
Number in each BPI proband cluster:	177	79	121	127											
Number of their FDRs:	406	171	301	279											
	<u>Column %</u>	<u>Column %</u>	<u>Column %</u>	<u>Column %</u>											
Clusters in FDRs					<.001	NS	*	*	NS	NS	NS				
No Affective d/o + No Int or Ext d/o	23	18	23	15	.391	NA	NA	NA	NA	NA	NA				
No Affective d/o + Int d/o(s) only	3	2	2	2	.768	NA	NA	NA	NA	NA	NA				
No Affective d/o + Ext d/o(s) only	7	4	11	8	.110	NA	NA	NA	NA	NA	NA				
No Affective d/o + Int and Ext d/o(s)	1	2	3	3	NA	NA	NA	NA	NA	NA	NA				
Affective d/o + No Int or Ext d/o	36	27	20	22	<.001	*	*	*	NS	NS	NS				
Affective d/o + Int d/o(s) only	9	18	9	14	.045	*	NS	NS	*	NS	NS				
Affective d/o + Ext d/o(s) only	16	17	20	19	.169	NA	NA	NA	NA	NA	NA				
Affective d/o + Int and Ext d/o(s)	6	13	14	17	.001	NS	*	*	NS	NS	NS				

Table 5, Continued

	<u>Column %</u>	<u>Column %</u>	<u>Column %</u>	<u>Column %</u>							
Affective disorders in FDRs											
Bipolar d/o Type I (BPI)	49	53	46	48	.652	NA	NA	NA	NA	NA	NA
MDD recurrent (MDDR)	9	11	7	14	.114	NA	NA	NA	NA	NA	NA
Single episode MDD (SEMD)	5	4	5	6	.753	NA	NA	NA	NA	NA	NA
Any bipolar d/o NOS	2	5	1	3	NA	NA	NA	NA	NA	NA	NA
Any depressive d/o NOS	4	4	3	6	.436	NA	NA	NA	NA	NA	NA
Any affective d/o	67	74	62	73	.204	NA	NA	NA	NA	NA	NA
Internalizing disorders in FDRs											
Anxiety disorders											
Panic d/o without agoraphobia	5	14	13	13	<.001	*	*	*	NS	NS	NS
Panic d/o with agoraphobia	3	8	5	11	.002	NS	NS	*	NS	NS	*
Agoraphobia without panic attacks	2	1	1	2	.770	NA	NA	NA	NA	NA	NA
Anxiety d/o NOS	1	3	1	3	.340	NA	NA	NA	NA	NA	NA
Generalized anxiety d/o (GAD)	0	0	0.3	0	NA	NA	NA	NA	NA	NA	NA
Post-traumatic stress d/o (PTSD)	0.3	1	1	1	NA	NA	NA	NA	NA	NA	NA
Any anxiety d/o except phobia & OCD (v1)	11	25	19	27	<.001	*	*	*	NS	NS	NS
Social phobia	3	8	4	7	.143	NA	NA	NA	NA	NA	NA
Specific phobia	4	12	7	10	.007	*	NS	*	NS	NS	NS
Any phobia (v2)	7	16	10	16	.003	*	NS	*	NS	NS	*
OCD (v3)	2	10	2	3	.036	*	NS	NS	*	*	NS
Any anxiety disorders	16	34	26	34	<.001	*	*	*	NS	NS	NS
Somatization disorder	0.3	0	0.3	2	NA	NA	NA	NA	NA	NA	NA
Any eating d/o	4	6	3	4	NA	NA	NA	NA	NA	NA	NA
Any eating or somatoform d/o (v4)	4	6	3	6	NA	NA	NA	NA	NA	NA	NA
Any Internalizing disorder	18	35	27	36	<.001	*	*	*	NS	NS	NS
Two, three or four internalizing variables	4	17	7	13	<.001	*	NS	*	*	NS	NS

Table 5, Continued

Externalizing disorders in FDRs	<u>Column %</u>	<u>Column %</u>	<u>Column %</u>	<u>Column %</u>								
Alcohol dependence	17	20	27	28	.002	NS	*	*	NS	NS	NS	
Alcohol abuse	8	10	14	11	.077	NA	NA	NA	NA	NA	NA	
Alcohol dependence or abuse (v5)	25	29	40	37	<.001	NS	*	*	*	NS	NS	
Drug dependence	10	15	20	19	.002	NS	*	*	NS	NS	NS	
Drug abuse	9	5	12	12	.038	NS	NS	NS	*	*	NS	
Drug dependence or abuse (v6)	15	18	27	25	<.001	NS	*	*	*	*	NS	
Any substance use disorder	29	34	47	45	<.001	NS	*	*	*	*	NS	
Antisocial personality d/o (ASPD)	1	2	4	3	.025	NS	*	*	NS	NS	NS	
Borderline personality d/o	0.3	0	0.3	0.4	NA	NA	NA	NA	NA	NA	NA	
Pathological gambling	1	2	1	3	NA	NA	NA	NA	NA	NA	NA	
Conduct d/o	1	2	4	1	NA	NA	NA	NA	NA	NA	NA	
Attention deficit hyperactivity d/o (ADHD)	0.5	0	0	1	NA	NA	NA	NA	NA	NA	NA	
Cluster B, impulse, conduct, ADHD (v7)	4	5	8	7	.187	NA	NA	NA	NA	NA	NA	
Any Externalizing disorder	30	36	47	47	<.001	NS	*	*	*	*	NS	
Two or three externalizing variables	12	15	21	20	.001	NS	*	*	*	NS	NS	
Habitual smoking	33	26	35	42	.034	NS	NS	*	NS	*	NS	

Table 5, Continued

Note. d/o = disorder. BPI = any BPI including BPI with mania and MDD, BPI manic never MDD, BPI mixed never manic, or schizoaffective bipolar.

Int = internalizing. Ext = externalizing. MDD = major depressive d/o. The 7 variables (v1 to v7) in bold type face were defined *a priori* before analyses.

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

Any bipolar d/o NOS = diagnosed with bipolar d/o NOS or bipolar II SEMD or cyclothymia or hypomania.

Any depressive d/o NOS = diagnosed with depressive d/o NOS or dysthymia or adjustment d/o.

An "NA" in the omnibus p-value column indicates that the models could not converge due to low prevalence of these disorders.

In the first column, the rows with (v1), (v2), (v3), (v4), (v5), (v6), and (v7) represent the seven variables analyzed in the CFA and CLCA models.

Each row represents a different logistic regression model, accounting for within-family correlations by using generalized estimating equations (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).

Supplementary Table S1. Characteristics and Comorbid Disorders of 1156 BPI Participants

Participant Characteristics (n = 1156),	<u>N</u>	<u>Min</u>	<u>Max</u>	<u>Mdn</u>	<u>Mean</u>	<u>SD</u>
Age at interview (years)	1121	14	89	42.0	42.4	12.5
[range]	[18 - 89 (99.5%); 14-17 (n=6; 0.5%)]					
Age of onset of major affective disorder	1141	5	75	18.0	19.6	8.3
Education (years)	1122	7	27	14.0	14.5	2.8
Episode frequency	1022	0	51	0.5	1.2	2.9
Manic-depressive episode ratio	802	0	85	1.0	2.0	4.7
Hospitalizations per year	831	0	9	0.2	0.3	0.6
Between-episode GAS score	931	10	100	68.0	67.3	15.5
	<u>Number (column %)</u>					
Familial relationship (n = 1156)						
Proband	504 (43.6)					
First-degree relative	594 (51.4)					
Second-degree relative	16 (1.4)					
Relative, but not first or second degree	42 (3.6)					
Female (n = 1155)	716 (62.0)					
Disabled (n = 1124)	208 (18.5)					
Ever psychosis symptoms (n = 1113)	560 (50.3)					
Mood incongruent psychotic symptoms (n = 1111)	131 (11.8)					
Rapid cycling (n = 1086)	103 (9.5)					
Rapid switching (n = 1087)	510 (46.9)					
Mixed states (n = 1111)	185 (16.7)					

Supplementary Table S1, continued

Self-reported race/ethnicity (n = 1117)	<u>Number (column %)</u>
White (non-Hispanic)	1033 (92.5)
Black (non-Hispanic)	32 (2.9)
Hispanic	19 (1.7)
American Indian/Alaskan Native	12 (1.1)
Asian/Pacific Islander	2 (0.2)
Other/Unknown	19 (1.7)
Marital Status (n = 1123)	
Married	491 (43.7)
Never married	328 (29.2)
Divorced	238 (21.2)
Widowed	26 (2.3)
Separated	40 (3.6)
DSM-IV disorders (n = 1156),	
Major affective disorders	
Bipolar I with mania and major depressive disorder (MDD)	1031 (89.2)
Bipolar I manic, never MDD	45 (3.9)
Bipolar I mixed, never manic	27 (2.3)
Schizoaffective bipolar type	53 (4.6)
Anxiety disorders	
Panic disorder without agoraphobia	163 (14.1)
Panic disorder with agoraphobia	115 (10.0)
Agoraphobia without panic attacks	27 (2.3)
Anxiety disorder NOS	25 (2.2)

Supplementary Table S1, continued

	<u>Number (column %)</u>
Anxiety disorders, continued	
Generalized anxiety disorder (GAD)	4 (0.4)
Post traumatic stress disorder (PTSD)	15 (1.3)
Any panic or other anxiety disorders except OCD or phobias	418 (36.2)
Obsessive compulsive disorder (OCD)	79 (6.8)
Social phobia disorder	82 (7.1)
Specific phobia disorder	106 (9.2)
Eating disorders	
Anorexia	30 (2.6)
Bulimia	51 (4.4)
Eating disorder NOS	15 (1.3)
Any eating disorder	83 (7.2)
Somatoform disorders	
Conversion disorder	1 (0.1)
Somatization disorder	6 (0.5)
Substance use disorders	
Alcohol dependence or abuse	444 (38.4)
Drug dependence or abuse	329 (28.5)
Any substance use disorder	529 (45.8)
Cluster B Personality disorders	
Anti-social personality disorder (ASPD)	43 (3.7)
Borderline disorder	4 (0.4)

Supplementary Table S1, continued

	<u>Number (column %)</u>
Impulse control disorders	
Pathological gambling	21 (1.8)
Kleptomania	1 (0.1)
Impulse control disorder NOS	2 (0.2)
Attention-deficit and disruptive behavior disorders	
Conduct disorder	23 (2.0)
Attention-deficit/hyperactivity disorder (ADHD)	7 (0.6)

Note. BP = Bipolar. Mdn = Median. NOS = not otherwise specified.

Race/ethnicity was self-reported using options defined by the investigators.

Supplementary Table S2. Validation of Coefficient-Weighted Versus Sum Factor Scores

	<u>Factor 1 (Internalizing factor)</u>		<u>Factor 2 (Externalizing factor)</u>	
	<u>Model</u>	<u>Sum</u>	<u>Model</u>	<u>Sum</u>
Number of BPI participants:	1156	1156	1156	1156
<u>Continuous Clinical Variables</u>	<u>r-sp</u>	<u>r-sp</u>	<u>r-sp</u>	<u>r-sp</u>
Age of onset	-.28***	-.24***	-.27***	-.23***
Episode frequency per years ill	.16***	.16***	.09**	.05
Mania/depression episode ratio	-.05	-.05	-.02	.02
Psychiatric hospitalizations per years ill	.03	.00	.07*	.08*
Between-episode GAS score	-.19***	-.17***	-.18***	-.15***
<u>Binary Clinical Variables (1=Yes, 0=No)</u>				
Female	.11**	.19***	-.07*	-.14***
Disabled	.09**	.07*	.08**	.07*
Psychotic symptoms	.13**	.10*	.13***	.10**
Mood incongruent psychotic symptoms	.02	.02	.02	.01
Rapid cycling	.08*	.05	.09**	.08*
Rapid switching	.25***	.22***	.19***	.13***
Mixed states	.14***	.11***	.13**	.10*

Note. r-sp = Spearman rank correlation coefficient. "Model" columns = MPLUS factor score iteratively estimated from CFA model using WLSMV estimation. "Sum" columns = sum score based on sum of 0/1 variable scores.

The clinical course of illness variables are explained in the footnote of Table 3.

*p < .05, **p < .01, ***p < .001.

Supplementary Text. Additional Background, Methods, and Results

A. Details regarding the diagnostic process

Further details about the study design and procedures

For the NIMH Genetics Initiative Bipolar Project, all diagnostic data and DNA has been deposited with the NIMH repository and is publicly available (zork.wustl.edu). In 2007 the collaborating groups formed the Bipolar Genome Study (BiGS), which is currently involved in sequencing studies based partially on the sample described here.

DIGS Structured Interview

The DIGS interview (Nurnberger *et al.*, 1994) was developed as part of the NIMH Genetics Initiative. In 1998 the DIGS was revised (DIGS 3.0) to collect additional data on multiple episodes of mania, depression, self-injurious behavior, and drug use including tobacco use. Unipolar and BP affective disorders were fully assessed as were phenomenology such as age of onset, number of episodes, irritability, mixed states, and suicidal ideation or behavior. A complete lifetime psychiatric history was obtained from each participant (probands and relatives). This history covered psychotic disorders, anxiety disorders, substance dependence (alcohol, nicotine, and illicit drugs), and an assessment of childhood and adult attention deficit symptoms. Habitual smoking was defined as smoking one pack per day or more for at least six months, either currently or in the past (Culverhouse *et al.*, 2005).

Family History Assessment and Medical Records

A structured family history assessment, the FIGS (available at zork.wustl.edu), was administered to all probands and relatives to obtain information on affective disorders in first-degree relatives. Some participants were described by multiple FIGS assessments and all assessments were included in the best estimate diagnostic package. Medical records were requested for all participants who had mental health treatment.

Diagnostic Procedures

First, the interviewer assigned a clinical diagnosis after reviewing and editing the DIGS and writing a narrative summary. Second, an editor (normally the study coordinator) examined the DIGS for missing data and inconsistencies. The best-estimate package included the DIGS narrative summary, FIGS family history information, and medical records. Next, a senior clinician (most were MD psychiatrists) reviewed all available information in the package and assigned best estimate diagnoses. A second clinician reviewed the package independently. The diagnoses of the two clinicians were compared and discrepancies were resolved by discussion. Non-resolved cases were referred to a tiebreaker clinician who determined the final diagnoses. Test-retest reliability coefficients for the major affective diagnoses were greater than 0.70 (Nurnberger *et al.*, 1994). Age of onset was defined as age at first diagnosis for each of the comorbid disorders and age at first diagnosis with any major affective disorder (except hypomania) for the major affective disorders.

Interviewer Training and Quality Assurance

Interviewers were trained to administer the DIGS. Study participants were recontacted as needed to complete questions. Senior editors at each site reviewed all interviews for

completeness and accuracy. They also regularly met with principal investigators to resolve difficult questions regarding interviews and to ensure consistency of data across all sites.

B. A brief introduction to CFA and CLCA in the context of the present paper

Both categorical (i.e., different aggregations of symptoms are considered distinct dichotomous disorders) and dimensional (i.e., symptoms common to different disorders are considered dimensions that span categorical diagnoses) models of BPI provide insights into pathophysiology (Akiskal, 1983; Ketter *et al.*, 2004). An analysis of symptoms common to different comorbid disorders in BPI persons is consistent with the dimensional viewpoint and acknowledges the complexity and heterogeneity of disorders (Kendell and Jablensky, 2003). The lifetime comorbid disorder strategy offers clinical utility and accounts for the longitudinal course of illness rather than specific episodes and thus provides a way to subdivide BP by trait as opposed to state characteristics. This paper uses the lifetime disorder approach for the factor and latent class analyses because it has received much less attention than the episodic symptom approach in prior factor and cluster analysis studies of comorbidity in BPI.

Exploratory factor analysis (EFA) is used to explore the sample data to determine the minimum number of factors needed to parsimoniously explain the correlations among observed comorbid disorders and to determine which variables cluster together into (i.e., have regression coefficients that “load high on”) each factor. Confirmatory factor analysis (CFA) is used to test *a priori* hypotheses regarding the number of factors and the “assignment” of each variable to a factor. Variables are typically assigned to each factor by allowing, separately for each variable, the standardized coefficient or “loading” for the paths from the factors to be estimated for only one factor and to be fixed to zero for the other factors.

Observed cluster analysis or latent class analysis (LCA) can be used to identify subtypes of persons based on defined characteristics. LCA is a latent-variable approach to cluster analysis in which the measurement error involved with identifying latent (unobserved) clusters (i.e., classes) is accounted for in the analysis. The classes are based on the profile of the probabilities of endorsing each salient characteristic (in this case the comorbid disorders). In the LCA model, the prevalence or proportion of each latent class is estimated. Each person's probability of membership in each class is estimated and used to classify them into their most likely group membership.

CLCA is performed by specifying and testing specific hypotheses, expressed as a set of parameter constraints, regarding the number of latent classes and the relationship between the observed characteristics and the latent factors (Finch and Bronk, 2011). We used all three types of model constraints: deterministic, equality, and inequality constraints (Finch and Bronk, 2011). In CLCA, fit statistics are used to compare one or more hypothesized models with exploratory LCA models using a clinically relevant range of number of classes, and with other potential post-hoc modifications of the hypothesized model (Finch and Bronk, 2011).

We are not aware of published work that has tested, within BPI patients, the equality and inequality hypotheses, which we specified and tested in the two versions of the CLCA model. However, there is some minor evidence for the plausibility of those hypotheses. For example, in the Yale Family Study, relatives of probands diagnosed with both alcohol dependence and anxiety disorders had a trend for slightly higher alcohol dependence (20%) compared to relatives of probands with only alcohol dependence (15%) but had no elevated rate for anxiety disorders compared to relatives of probands with only anxiety disorders (19% vs 21%, respectively) (Merikanagas *et al.*, 1996; Merikangas *et al.*, May 1996).

However, we did not hypothesize what we felt to be unreasonable inequality hypotheses. For example, we did *not* hypothesize that the probabilities of externalizing and internalizing disorders in relatives would be *less* in the "both" class versus the externalizing or internalizing only classes. By choosing constraints for the hypothesized CLCA model, such as deterministic, and equality or inequality, we were able to hypothesize and test two alternatives (a particular equality constraint vs a particular inequality constraint) of a very specific model (4-class model with deterministic constraints) within a much larger universe of possible models and then compare the hypothesized model to many other possible exploratory models and other models with post-hoc modifications of the hypothesized model.

C. Tests of modification indices for CFA results

A two-sided 0.05 alpha level was used to test whether a path in the CFA model should be added (using MPLUS modification indices) or removed (using a manual routine). An alpha of 0.05 for tests of modification indices represents a robust test of the hypothesized model because “ $p < 0.05$ ” is a rigorous decision rule for rejecting the null hypothesis of good fit for individual paths in a hypothesized CFA model and subsequently modifying the confirmatory model based on sample data. Practitioners of CFA often recommend using 0.01 alpha for determining whether to modify (add or delete) individual paths to avoid over-fitting by capitalizing on chance error in the sample data.

D. Additional results

Model stability

In all CLCA and LCA models, the maximum (i.e., best) log likelihood value was replicated for multiple sets of starting values, indicating good model stability. We attempted to estimate factor mixture models which combines CFA and CLCA models into a single model (Muthén, 2008); however, convergence was not achieved which is not uncommon for binary variables.

Additional results to accompany the CFA findings

Because the fit of the two-factor CFA model was extremely good, this suggests that a model with three or more factors would add very little additional benefit. To determine whether the sample data also suggests only two factors, the eigenvalues were inspected. The scree plot (not shown) of the eigenvalues indicated that there were two dominant factors and not more than two, because the decreases between eigenvalues became much smaller starting with the third eigenvalue: 2.59, 1.55, 0.73, 0.71, 0.62, 0.50, and 0.31.

Additional description of results of CLCA and LCA from Table 4

The 3-factor exploratory LCA model fit significantly better than the 2-factor exploratory LCA model (LRM, $p < .0001$). The 4-factor exploratory LCA model was not a significant improvement over the 3-factor exploratory LCA model (LMR, $p = 0.19$), and which of those models (3-factor or 4-factor exploratory) was preferred depended on whether the BIC or aBIC was used. However, according to the BIC and aBIC, both of the competing and *a priori*

hypothesized 4-factor CLCA models fit better than any exploratory LCA model, including the 2-factor, 3-factor and 4-factor exploratory LCA models.

Post-hoc inspection of parameter estimates and graphical output were used to determine whether the fit of the hypothesized CLCA models could be improved. In models #9 through #14, the same deterministic constraints were used as in models #1 and #2 but the equality or inequality constraints were relaxed instead of using full equality constraints. Equality constraints were attempted for only the internalizing disorders (model #9) or only the externalizing disorders (model #10) or only alcohol (model #11). Inequality constraints were attempted for only the externalizing disorders (model #12). Equality constraints were specified for only the internalizing disorders and inequality constraints were specified for only the externalizing disorders (model #13). In model #14, no equality or inequality constraints were applied.

In models #15 and #16, the deterministic constraints were modified. Instead of an “internalizing only” class, model #15 allowed an “internalizing-only-plus-alcohol” class in which the externalizing variables were fixed to zero except for one externalizing variable, alcohol, for which a non-zero probability of endorsement was allowed to be estimated along with the internalizing disorders. In model #16, the zero-class or “pure BP” was the only deterministic constraint; in other words, all disorders were allowed to be estimated with non-zero probability for all of the three other classes.

The estimable LMR tests showed that the 4-factor CLCA models (#1, #2), and their post-hoc modifications (#9, #10, #11, #14, #15), were significantly better ($p < .0001$) fits to the data than a similarly constrained 3-factor model. The use of only a “zero” class for deterministic constraints was not statistically better than a 3-class model with a similar constraint (model #16,

LMR, $p = .07$). However, none of the post-hoc modified models fit better than the hypothesized CLCA model with deterministic constraints and full equality constraints (model #1).

Sensitivity analysis

Habitual smoking is different from other substance use disorders in its social and legal aspects. Therefore, although habitual smoking was reported as one of the individual outcome variables among first-degree relatives in the familial analysis, habitual smoking was excluded from the hypothesized two-factor model and the four-cluster model and from the CFA and CLCA models used to test these hypotheses. However, as a sensitivity analysis, re-estimation of CFA and CLCA models were conducted after including habitual smoking as a fourth indicator of the externalizing factor. Habitual smoking was defined as smoking one pack per day or more for at least six months, either currently or in the past (Culverhouse *et al.*, 2005). Results showed approximately similar (i.e., good) fit as the CFA and CLCA models that we reported which excluded smoking. However, the factor and class associations with the clinical course variables were not as strong when habitual smoking was included.

In another sensitivity analysis, the CFA and CLCA analyses were re-run after excluding the sparsest disorders: generalized anxiety disorder (GAD), post-traumatic stress disorder (PTSD), somatoform disorders, impulse control disorders, conduct disorder and attention deficit hyperactivity disorder (ADHD). Results were very similar to those reported in this paper.

References for Supplementary Text

- Akiskal, HS** (1983). Diagnosis and classification of affective disorders: new insights from clinical and laboratory approaches. *Psychiatric Developments* **1**, 123-160.
- Culverhouse, R, Bucholz, KK, Crowe, RR, Hesselbrock, V, Nurnberger, JI, Jr., Porjesz, B, Schuckit, MA, Reich, T & Bierut, LJ** (2005). Long-term stability of alcohol and other substance dependence diagnoses and habitual smoking: an evaluation after 5 years. *Archives of General Psychiatry* **62**, 753-760.
- Finch, WH & Bronk, KC** (2011). Conducting confirmatory latent class analysis using MPlus. *Structural Equation Modeling* **18**, 132-151.
- Kendell, R & Jablensky, A** (2003). Distinguishing between the validity and utility of psychiatric diagnoses. *American Journal of Psychiatry* **160**, 4-12.
- Ketter, TA, Wang, PW, Becker, OV, Nowakowska, C & Yang, Y** (2004). Psychotic bipolar disorders: dimensionally similar to or categorically different from schizophrenia? *Journal of Psychiatric Research* **38**, 47-61.
- Merikanagas, KR, Stevens, D & Fenton, B** (1996). Comorbidity of alcoholism and anxiety disorders: The role of family studies. *Alcohol Health & Research World* **20**, 100-106.
- Merikangas, KR, Stevens, D, Fenton, B, O'Malley, S, Woods, S, Stolar, M & Risch, N** (May 1996). Comorbidity and cotransmission of anxiety disorders and alcoholism: Results of the Yale Family Study. In *Proceedings of the American Psychiatric Association*.
- Muthén, B** (2008). Latent variable hybrids: Overview of old and new models. In *Advances in latent variable mixture models* (ed. G. R. Hancock and K. M. Samuelsen), pp. 1-24. Information Age Publishing, Inc.: Charlotte, NC.

Nurnberger, JI, Jr. , Blehar, MC, Kaufmann, CA, York-Cooler, C, Simpson, SG, J., H-F, Severe, JB, Malaspina, D & Reich, T (1994). Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. *Archives of General Psychiatry* **51**, 849-859; discussion 863-864.