

1 Dissecting clinical heterogeneity of bipolar disorder using multiple polygenic
2 risk scores

3 Brandon J. Coombes, Ph.D.¹; Matej Markota, M.D.²; J. John Mann, M.D.^{3,4}; Colin Colby¹; Eli
4 Stahl, Ph.D.^{5,6}; Ardesheer Talati, Ph.D.^{3,7}; Jyotishman Pathak, Ph.D.⁸; Myrna M. Weissman,
5 Ph.D.^{3,7,9}; Susan L. McElroy, M.D.¹⁰; Mark A. Frye, M.D.²; Joanna M. Biernacka, Ph.D.^{1,2}

6
7 ¹Department of Health Sciences Research, Mayo Clinic, Rochester, MN

8 ²Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN

9 ³Department of Psychiatry, Columbia University Vagelos College of Physicians & Surgeons,
10 New York, NY

11 ⁴Division of Molecular Imaging and Neuropathology, New York State Psychiatric Institute,
12 New York, NY

13 ⁵Pamela Sklar Division of Psychiatric Genomics, Icahn School of Medicine at Mount Sinai,
14 New York, NY

15 ⁶Medical and Population Genomics, Broad Institute, Cambridge, MA

16 ⁷Divisions of Translational Epidemiology, New York State Psychiatric Institute, New York,
17 NY

18 ⁸Department of Healthcare Policy & Research, Weill Medical College, Cornell University,
19 New York, NY

20 ⁹Department of Epidemiology, Mailman School of Public Health, Columbia University, New
21 York, NY

22 ¹⁰Lindner Center of HOPE/University of Cincinnati, Cincinnati, OH

23

24 **Running title:** Dissecting clinical heterogeneity of BD using PRSs

25

26 Corresponding Author: Brandon J. Coombes

27 Address: Mayo Clinic, Harwick Building 7-23. 200 First St. SW, Rochester, MN 55905

28 Phone: (507) 293-0051

29 Fax: (507) 284-9542

30 Email: coombes.brandon@mayo.edu

31

32 **Word Count:** 3286/3500

33 **Abstract (273 words)**

34 Bipolar disorder (BD) has high clinical heterogeneity, frequent psychiatric comorbidities,
35 and elevated suicide risk. To determine genetic differences between common clinical sub-
36 phenotypes of BD, we performed a systematic PRS analysis using multiple polygenic risk
37 scores (PRSs) from a range of psychiatric, personality, and lifestyle traits to dissect
38 differences in BD sub-phenotypes in two BD cohorts: the Mayo Clinic BD Biobank (N = 968)
39 and Genetic Association Information Network (N = 1001). Participants were assessed for
40 history of psychosis, early-onset BD, rapid cycling (defined as four or more episodes in a
41 year), and suicide attempts using questionnaires and the Structured Clinical Interview for
42 DSM-IV. In a combined sample of 1969 bipolar cases (45.5% male), those with psychosis
43 had higher PRS for SCZ (OR = 1.3 per S.D.; $p = 3e-5$) but lower PRSs for anhedonia (OR =
44 0.87; $p = 0.003$) and BMI (OR = 0.87; $p = 0.003$). Rapid cycling cases had higher PRS for
45 ADHD (OR = 1.23; $p = 7e-5$) and MDD (OR = 1.23; $p = 4e-5$) and lower BD PRS (OR = 0.8; $p =$
46 0.004). Cases with a suicide attempt had higher PRS for MDD (OR = 1.26; $p = 1e-6$) and
47 anhedonia (OR = 1.22; $p = 2e-5$) as well as lower PRS for educational attainment (OR =
48 0.87; $p = 0.003$). The observed novel PRS associations with sub-phenotypes align with
49 clinical observations such as rapid cycling BD patients having a greater lifetime prevalence
50 of ADHD. Our findings confirm that genetic heterogeneity underlies the clinical
51 heterogeneity of BD and consideration of genetic contribution to psychopathologic
52 components of psychiatric disorders may improve genetic prediction of complex
53 psychiatric disorders.
54

55 **Introduction**

56 Many psychiatric disorders have moderate to high heritability; however, the genetics of
57 psychiatric disorders are complex and highly polygenic, with each risk variant only
58 conferring a small effect[1]. Psychiatric disorders also have a high level of overlapping
59 clinical heterogeneity, with shared genetic risk explaining some of the clinical overlap, and
60 certain combinations of alleles may contribute to the same psychopathological symptoms
61 in multiple psychiatric disorders. Furthermore, some psychiatric disorders may lie on a
62 continuum rather than being disorders with distinct genetics and biological mechanisms[2,
63 3].

64 To accommodate this genetic complexity, investigations of psychiatric disorders have
65 increasingly relied on polygenic risk scores (PRSs), leveraging knowledge from prior large
66 genome-wide association studies (GWASs) to predict genetic risk of particular disorders in
67 a new sample[4]. When the PRS for one disorder is predictive of a second disorder, this
68 indicates a common polygenic contribution to the two disorders[5].

69 Bipolar disorder (BD) is a complex illness with heterogeneous clinical presentation, and
70 apparent sub-phenotypes often have a different course of illness, prognosis, and treatment
71 response[6–9]. In order to personalize treatment, it is crucial to better understand
72 biological underpinnings of BD clinical sub-phenotypes. One approach is to examine
73 potential relationships of clinical phenotypes to different genetic profiles.

74 Historically, the relationship between schizophrenia (SCZ) and BD has shaped classification
75 systems in psychiatry[10]. The corresponding link between phenotype and genetics was
76 recently established with the demonstration that BD patients with a history of psychosis,
77 particularly mood incongruent psychosis and psychosis during mania, have increased

78 genetic risk for SCZ[11–16]. However, it is well recognized that BD genetically overlaps–
79 and has high clinical comorbidity with–other major psychiatric conditions, including major
80 depressive disorder (MDD), attention deficit and hyperactivity disorder (ADHD), anxiety
81 disorders, post-traumatic stress disorder (PTSD), obsessive compulsive disorder (OCD),
82 borderline personality disorder and substance use disorders[17–24]. While significant
83 advances have been made in understanding the genetic relationship between BD psychotic
84 sub-phenotypes and SCZ[11–16], little is known about how genetic risks for other
85 psychiatric disorders as well as important personality and lifestyle traits such as body mass
86 index (BMI), risk-taking, and neuroticism relates to psychosis or other BD clinical sub-
87 phenotypes.

88 The goal of this study was to systematically test if PRSs for major psychiatric conditions
89 and other traits related to BD are predictors of distinct BD sub-phenotypes, in particular
90 with regards to psychosis, age-of-onset, rapid cycling and suicidal behavior. Understanding
91 the shared genetic risk factors between BD clinical sub-phenotypes and other comorbid
92 conditions may contribute to psychiatric clinical classification systems with a more
93 biologically-informed nosological system[25].

94

95 **Methods and Materials**

96 *1. Ethics Statement*

97 Data were collected according to Declaration of Helsinki principles. Mayo Clinic
98 Bipolar Disorder Biobank participants' consent forms and protocols were reviewed
99 and approved by the Mayo Clinic Institutional Review Board (IRB # 08-008794 00).
100 Opt-in written informed consent was obtained.

101

102 2. *Studies*

103 a. *Mayo Clinic Bipolar Disorder Biobank*

104 The Mayo Clinic Bipolar Disorder Biobank collection has been described in
105 previous papers[9, 11, 26]. We restricted our analyses to cases with
106 European ancestry (N = 968), because PRSs derived from GWASs of
107 participants with European ancestry perform much worse in non-European
108 ancestries[27]. Sub-phenotypes were determined using the Structured
109 Clinical Interview for DSM-IV (SCID)[28] as well as a patient questionnaire.
110 Detailed information on the assessment of each sub-phenotype can be found
111 in Supplementary Table 1.

112 b. *Genetic Association Information Network (GAIN)*

113 The Bipolar Disorder Genome Study Consortium conducted a genome-wide
114 association study (GWAS) of BD as part of the Genetic Association
115 Information Network (GAIN)[29]. We obtained the data from dbGaP
116 (phs000017.v3.p1), and restricted our analyses to cases with European
117 ancestry (N = 1001). All cases met criteria for DSM IV-defined bipolar I
118 disorder (BD-I). Subjects recruited at different times were interviewed
119 with the Diagnostic Interview for Genetic Studies 2, 3, or 4 (DIGS 2, 3, 4).
120 Detailed information on the assessment of each sub-phenotype can be found
121 in Supplementary Table 1.

122 3. *Genotyping and Quality Control*

123 a. *Mayo Clinic Bipolar Biobank*

124 Genotyping and genetic data quality control of this sample was previously
125 described as part of a larger case-control study[11]. Briefly, the Illumina
126 HumanOmniExpress platform was used to genotype 1046 BD cases. For
127 quality control purposes, we excluded subjects with <98% call rate and
128 related subjects. SNPs with call rate <98%, MAF <0.01, and those not in
129 Hardy-Weinberg Equilibrium (HWE; $P < 1e-06$) were removed. After these
130 steps 643 011 SNPs and 968 subjects remained.

131 *b. GAIN*

132 Genotyping and quality control procedures for the GAIN-BD data were
133 previously described by Smith *et al.*[30], Briefly, the Affymetrix Genome-
134 Wide Human SNP Array 6.0 platform was used to genotype cases and after
135 excluding SNPs with call rate <98%, MAF <0.01, and those not in HWE, 726
136 315 SNPs and 1001 subjects of European ancestry remained.

137 *c. Imputation*

138 Genotypes in both the GAIN and Mayo samples were imputed to the 1000
139 genomes reference panel, as previously described for the GAIN sample[9].
140 Specifically, SHAPEIT[31] was used for haplotype phasing and imputation
141 was performed using IMPUTE2.2.2[32] with the 1000 genome project
142 reference data (phase 1 data, all populations). Dosage data was converted to
143 best guess genotype for the well-imputed (dosage $R^2 > 0.8$) and common
144 (MAF >0.01) SNPs, resulting in more than 5 million SNPs in both datasets.

145 *4. Polygenic Risk Scores (PRSs)*

146 PRSs were included in the analysis if: 1) there was evidence of significant genetic
147 correlation of the trait with BD and 2) we had at least 80% power to detect PRS
148 association in a general case-only analysis of our data assuming 50% prevalence of
149 the subphenotype. We began by considering PRSs for major psychiatric disorders
150 (BD[33], SCZ[34], MDD[35], ADHD[36], anxiety[37], PTSD[19], OCD[38], anorexia
151 nervosa (AN)[39], alcohol use disorder (AUD)[40], and insomnia[41]) and
152 personality and lifestyle traits related to BD (alcohol consumption[40], educational
153 attainment (EA)[42], risk-taking[43], subjective well-being (SWB)[44],
154 neuroticism[45], anhedonia[46], and body mass index (BMI)[47]). GWAS summary
155 statistics were restricted to well-imputed variants (INFO >0.9) when information on
156 imputation quality was available.

157 Using LD score regression[48], we estimated the genetic correlation of the above
158 traits with BD[33] (Supplementary Table 2). Insomnia and alcohol consumption
159 didn't have significant genetic correlation with BD and were therefore excluded
160 from further analysis.

161 Using the R package *avengeme*[49], we estimated that training sample sizes of
162 20,000 would achieve at least 80% power in our analysis assuming moderate
163 overlap of the trait with the sub-phenotype (genetic covariance = 0.1), high
164 polygenicity (# of independent SNPs = 20,000), and 0.005 α -level to account for
165 multiple testing. The study of OCD included an effective sample size of less than
166 4000 and was thus excluded from further analysis. The final list of PRSs that were
167 tested for association with BD sub-phenotypes is shown in Supplementary Table 2.

168 For traits that satisfied our inclusion criteria, the PRS-CS[50] auto setting was
169 applied to estimate SNP weights using a fully Bayesian shrinkage approach that
170 shrinks SNP effects with a continuous shrinkage prior. This setting allows the
171 algorithm to learn the global shrinkage parameter from the data to create one set of
172 weights per PRS and therefore does not require a validation dataset. This setting
173 also reduces the multiple testing of standard PRS analyses that search over many p-
174 value thresholds[51]. PLINK version 1.9[52] was used to create PRSs using the
175 shrunken SNP weights. The PRSs were then standardized to have a mean of zero
176 and standard deviation (SD) of one.

177 5. Statistical Analyses

178 In each dataset, we performed principal components (PCs) analysis of the
179 genotyped SNPs and kept the first four PCs to be used as within-study nested
180 covariates in subsequent PRS association analyses. In all models, study indicator
181 and an interaction of study and within-study PCs were included as covariates to
182 control for population stratification. All twelve PRSs were individually modeled
183 using a multivariate logistic regression model with each sub-phenotype (psychosis,
184 early-onset BD, rapid cycling, and attempted suicide) as the outcome.

185 We used 10,000 permutations to find the significance threshold to control the false
186 positive rate testing for association with each sub-phenotype with fourteen PRSs (α
187 = 0.005) as well as the family-wise error rate ($\alpha = 0.001$). For each sub-phenotype,
188 we also included all significant PRSs ($p < 0.005$) in a joint model, to estimate the
189 relative contribution of the PRSs after adjusting for other important PRSs. We
190 report the variance explained in the sub-phenotype by each PRS after adjustment

191 for other PRSs using Nagelkerke's pseudo- R^2 statistic. All statistical analyses were
192 performed in R 3.5.2.

193

194 **Results**

195 *Sample Description*

196 Table 1 summarizes the demographic and sub-phenotype information of each study. There
197 was a difference in the sex distribution between the two samples. The GAIN study only
198 included BD type I cases, and the distributions were also significantly different for all sub-
199 phenotypes besides attempted suicide. GAIN BD cases had a higher rate of psychosis and
200 early-onset BD, while Mayo cases had higher rates of rapid cycling, which is more prevalent
201 in women[53].

202 Figure 1 shows a forest plot of the significant PRS associations with each sub-phenotype
203 further broken down by study. Further detailed results for each sub-phenotype can be
204 found in Supplementary Tables 3-6.

205 *Psychosis*

206 Cases with psychosis versus no psychosis had higher PRSs for SCZ (OR = 1.3, 95% CI 1.15 to
207 1.48; p-value = 3.5e-5), but lower PRSs for anhedonia (OR = 0.87, 95% CI 0.79 to 0.95; p-
208 value = 0.003), and BMI (OR = 0.87, 95% CI 0.79 to 0.95; p-value = 0.004). These three
209 PRSs explained 2.6% of the variation in psychosis in the joint model. While anhedonia is a
210 component of MDD and the two PRSs are positively correlated ($r = 0.41$), the PRS for MDD
211 was not associated with psychosis in BD (OR = 0.96, 95% CI 0.87 to 1.06; p-value = 0.45).

212 *Early-onset BD*

213 Higher PRSs for risk-taking (OR = 1.21, 95% CI 1.09 to 1.35; p-value = 0.0005; adj.
214 Nagelkerke's R^2 = 0.8%) and anhedonia (OR = 1.16, 95% CI 1.05 to 1.29; p-value = 0.0047;
215 adj. Nagelkerke's R^2 = 0.8%) were observed in cases with early-onset BD compared to cases
216 that developed BD after age 18.

217 *Rapid Cycling*

218 Cases with rapid cycling versus those without rapid cycling had higher ADHD PRS (OR =
219 1.23, 95% CI 1.11 to 1.36; p-value = $7e-5$; adj. Nagelkerke's R^2 = 0.8%), MDD PRS (OR =
220 1.23, 95% CI 1.11 to 1.36; p-value = $4e-5$; adj. Nagelkerke's R^2 = 0.5%), PTSD PRS (OR =
221 1.28, 95% CI 1.14 to 1.44; p-value = $4e-5$; adj. Nagelkerke's R^2 = 0.7%), and PRS for anxiety
222 (OR = 1.19, 95% CI 1.07 to 1.33; p-value = 0.001; adj. Nagelkerke's R^2 = 0.1%). Cases with
223 rapid cycling also had lower BD PRSs (OR = 0.80, 95% CI 0.68 to 0.93; p-value = 0.004; adj.
224 Nagelkerke's R^2 = 0.9%). The five PRSs explained 3.9% of the variation in rapid cycling
225 when included in one model.

226 *Attempted suicide*

227 The genetic risk for MDD (OR = 1.26, 95% CI 1.15 to 1.39; p-value = $1e-6$; adj. Nagelkerke's
228 R^2 = 0.7%) and anhedonia (OR = 1.22, 95% CI 1.12 to 1.34; p-value = $2e-5$; adj. Nagelkerke's
229 R^2 = 0.3%) was higher in cases with at least one suicide attempt versus those with none.
230 Cases with an attempted suicide also had a lower PRS for educational attainment (OR =
231 0.87, 95% CI 0.79 to 0.96; p-value = 0.0036; adj. Nagelkerke's R^2 = 0.2%). The three PRSs
232 explained a total of 2.3% of the variation when included in one model, but only the MDD
233 PRS remained significant after accounting for the other PRS associations.

234

235 **Discussion**

236 To our knowledge, this is the first PRS dissection of clinical sub-phenotypes in BD that is
237 comprehensive with respect to the range of psychiatric, personality, and lifestyle
238 phenotypes for which genetic liabilities were estimated and used to predict the BD sub-
239 phenotypes. Previous analyses of BD sub-phenotypes of psychosis, early onset BD and
240 suicide focused on genetic liability to the major psychiatric diagnoses of BD, SCZ, and
241 MDD[6, 11, 12, 16, 54]. Here, we took an expanded agnostic approach to PRS analysis by
242 using many different PRSs beyond just these three to more systematically test for PRS
243 association with clinically important sub-phenotypes of BD, including rapid cycling.
244 Importantly, the contribution of each PRS to a sub-phenotype was assessed after adjusting
245 for the other PRSs' contributions, thereby assessing how predictive a genetic risk is above
246 and beyond other genetic risks. The forest plots in Figure 1 show that our results were
247 highly comparable in the two cohorts, lending greater confidence to the conclusions.
248 Overall, we find that the different BD clinical sub-phenotypes have different profiles of PRS
249 associations with major psychiatric conditions.

250 ***BD with psychosis***

251 Previously, using the Mayo Clinic sample, we showed that BD patients with a history of
252 psychosis during mania had higher genetic risk for SCZ [11]. Here, this finding is replicated
253 in the GAIN cohort. This finding was also reported by Ruderfer *et al.*[16], in a larger study
254 that included both the Mayo and GAIN BD cohorts. However, in addition to this
255 relationship, in the present study we also found that BD cases that have not experienced
256 psychotic symptoms had higher genetic scores for anhedonia and BMI. Association of
257 higher genetic risk for anhedonia with a subtype of BD without psychotic features implies
258 that a patient with more genetic predisposition for anhedonia during major depressive

259 episodes is less likely to include episodes with psychotic features. In fact, rates of psychotic
260 features are higher in BD compared with MDD and familial studies show a greater
261 heritability of psychotic features in BD relative to in other mood disorders[55].
262 Interestingly, we did not observe a significant association with MDD PRS despite a strong
263 genetic correlation between anhedonia and MDD. This may underscore the importance of
264 relying on core symptoms in these analyses, instead of using more complex and syndromal
265 entities like MDD. The relationship between BMI and psychosis is complex and influenced
266 by heritable, environmental, and iatrogenic factors. Over the course of illness, most
267 patients with BD and psychosis gain weight, which contributes to morbidity and
268 mortality[56, 57]. Our finding that BD patients with psychosis have lower genetic
269 predisposition to elevated BMI than BD patients without psychosis suggest that weight gain
270 in those with psychosis likely occurs primarily as a side effect of medications. However, the
271 complex relationship between BD and greater body weight needs to be further explored in
272 the context of sub-phenotypes and use of atypical antipsychotics or lithium.

273 ***Early-onset BD***

274 We found evidence that higher genetic liability for risk-taking behavior was associated with
275 early-onset BD, but no evidence that genetic risk for SCZ or BD were associated with age of
276 onset of illness. A previous study of polygenic associations with age-of-onset of BD also
277 showed no association of SCZ or BD genetic risk with both a dichotomous sub-phenotype,
278 as defined in our study, or continuous age-of-onset[58]. Risk-taking is a hallmark feature of
279 normative adolescence, but is also commonly seen in mania. There are several possible
280 explanations for the risk-taking PRS and early-onset BD association found in this study.
281 Perhaps the simplest explanation is that youth with particularly high propensity for risk-

282 taking behaviors come to clinical attention earlier and subsequently have BD identified at
283 an earlier age. However, there are several potential limitations that may have affected
284 these findings. Due to the way the data were collected, for this study, age of onset was
285 dichotomized based on a cutoff age of 18, which may have reduced power. Also, our early-
286 onset BD definition did not differentiate between age of first manic and first depressive
287 episodes. Furthermore, given the high genetic and clinical overlap between BD and other
288 conditions investigated here (e.g. ADHD), a study of age-of-onset of any psychiatric
289 disorder/symptom rather than just BD could be informative. It is of note that earlier onset
290 MDD is associated with more pronounced aggressive/impulsive traits[59]. Nevertheless,
291 the observed association of risk-taking PRS with early vs. late onset BD is intriguing and
292 warrants further investigation.

293 ***BD with rapid cycling***

294 Previous clinical studies have shown a strikingly higher clinical comorbidity rate of ADHD
295 in BD patients with rapid cycling compared to non-rapid cycling BD patients[60]. While the
296 general genetic association between ADHD and BD has been described before[17], our
297 results are the first study to show possible genetic underpinnings for this specific rapid
298 cycling BD and ADHD association. We also found a strong association of MDD genetic risk
299 with rapid cycling. This implies that genetic variation related to ADHD and MDD may also
300 be related to episode frequency in BD, and that comorbid ADHD and more depressive
301 episodes would be clinically associated with the rapid cycling form of BD, though the
302 predominant directionality of mood episodes was not discernible from the available data.
303 Rapid cycling BD has been reported to have more episodes of major depression and a
304 higher rate of parental MDD compared with non-rapid cycling BD[61], which is consistent

305 with our PRS association findings. Finally, rapid cycling cases had lower BD PRS as
306 reported in a previous investigation[15]. This could simply reflect that prevalence of rapid
307 cycling in cases ascertained for the sample used in the GWAS of BD by the Psychiatric
308 Genomics Consortium (PGC) was lower than in the two samples included here, but still
309 demonstrates a systematic difference in genetics of rapid cycling and non-rapid cycling BD.

310 ***BD with a history of a suicide attempt***

311 Our finding of increased MDD genetic load in BD patients with a history of suicide attempts
312 is consistent with a recent study that included both of the Mayo and GAIN data, which
313 showed that genetic risk factors for MDD increase the risk for suicide trans-
314 diagnostically[54]. BD with a history of suicide attempt having a higher MDD genetic
315 liability is consistent with the clinical observation that suicide attempts are most common
316 during major depressive episodes or mixed states and rare during manic episodes or while
317 euthymic[62, 63]. Interestingly, even after adjusting for MDD PRS, we also found that
318 genetic liability for anhedonia is marginally associated with suicide, suggesting that
319 anhedonia may be a particularly relevant factor contributing to suicidality, compared to
320 other components that comprise the MDD syndrome. This is consistent with findings from
321 non-genetic studies, which found that association of anhedonia with suicidality is
322 independent of the association with depression and psychotic features[64].

323 ***Methodological limitations***

324 PRSs used in this study are based on data from previously published large scale
325 investigations and are limited by the diagnostic accuracy, recruitment criteria, and
326 methodology of previous studies. The most recent PGC study of BD[33] included the cases
327 and controls from the GAIN and Mayo samples, and most PGC studies of other disorders

328 used controls from the GAIN study. Sample overlap of testing datasets with training data
329 can create substantial biases in PRS analyses. However, here we studied genetic
330 differences within cases and thus, the sample overlap is not expected to bias our results,
331 because there is no correlation between case-control status used to build the training
332 models and the within-case sub-phenotypes. Another limitation in this study is the lack of
333 data on the number, duration and severity of major depressive and manic episodes to more
334 precisely map the clinical picture onto the PRS profile. Finally, it is important to note that
335 no PRS explained a large amount of variation in our analysis. Thus, while the associations
336 identified in this study provide evidence of genetic differences that may underlie clinical
337 subtypes of BD, these PRSs cannot yet be used for purposes of personalized psychiatry.

338

339 **Conclusion**

340 Our findings contribute to the understanding of the underlying genetic causes of clinical
341 heterogeneity of BD and of comorbidity between BD and other major psychiatric
342 conditions. We find evidence that psychopathologic components of BD, including psychotic
343 symptoms, rapid cycling and suicidal behavior are linked to the PRSs for related disorders
344 including schizophrenia, ADHD and MDD, respectively. Finally, larger studies are needed to
345 more precisely map genetic risk factors to clinical sub-phenotypes. Harmonization of sub-
346 phenotypes across studies is a well-recognized challenge. Nevertheless, such efforts are
347 critical in helping to classify psychiatric disorders more accurately and identify risk of
348 suicide, psychosis, and other adverse outcomes in patients.

349 Supplementary information is available at MP's website.

350

351 **Acknowledgments**

352 This work was supported by the Marriott Foundation and the Thomas and Elizabeth
353 Grainger Fund in Bipolar Functional Genomics and Drug Development awarded to the
354 Mayo Clinic. Establishment of the Bipolar Disorder Biobank was supported by a generous
355 gift from the Marriot Family and the Mayo Clinic Center for Individualized Medicine. The
356 contributions from authors JJM, AT, and MMW was supported by NIMH R0MH121921
357 (Wickramaratne and Mann, M.P.I.s).

358

359 **Conflicts of Interest**

360 Dr Frye has received grant support from Assurex Health, Myriad, Pfizer, National Institute
361 of Mental Health (RO1 MH079261), National Institute of Alcohol Abuse and Alcoholism
362 (P20AA017830), Mayo Foundation; has been a consultant to Janssen Global Services, LLC,
363 Mitsubishi Tanabe Pharma Corporation, Myriad, Sunovion, and Teva Pharmaceuticals; has
364 received CME/Travel Support/presentation from CME Outfitters Inc. and Sunovian; Mayo
365 Clinic has a financial interest in AssureRx and OneOme. Dr. McElroy is a consultant to or
366 member of the scientific advisory boards of Bracket, MedAvante, Naurex, Shire, and
367 Sunovion. She is a principal or co-investigator on studies sponsored by the Agency for
368 Healthcare Research & Quality (AHRQ), AstraZeneca, Cephalon, Forest, Marriott
369 Foundation, National Institute of Mental Health, Orexigen Therapeutics, Inc., Shire, and
370 Takeda Pharmaceutical Company Ltd. She is also an inventor on United States Patent No.
371 6,323,236 B2, Use of Sulfamate Derivatives for Treating Impulse Control Disorders, and
372 along with the patent's assignee, University of Cincinnati, Cincinnati, Ohio, has received
373 payments from Johnson & Johnson, which has exclusive rights under the patent. In the last

374 three years, Dr Weissman has received research funds from NIMH, Templeton Foundation,
375 Brain and Behavior and the Sackler Foundation and has received royalties for publications
376 of books on interpersonal psychotherapy from Perseus Press, Oxford University Press, on
377 other topics from the American Psychiatric Association Press and royalties on the social
378 adjustment scale from Multihealth Systems. None of these represent a conflict of interest.

379 **References**

- 380 1. Smoller JW, Andreassen OA, Edenberg HJ, Faraone S V, Glatt SJ, Kendler KS.
381 Psychiatric genetics and the structure of psychopathology. *Mol Psychiatry*.
382 2019;24:409–420.
- 383 2. Owen MJ, O'Donovan MC. Schizophrenia and the neurodevelopmental
384 continuum:evidence from genomics. *World Psychiatry*. 2017;16:227–235.
- 385 3. Tesli M, Espeseth T, Bettella F, Mattingsdal M, Aas M, Melle I, et al. Polygenic risk
386 score and the psychosis continuum model. *Acta Psychiatr Scand*. 2014;130:311–317.
- 387 4. Martin AR, Daly MJ, Robinson EB, Hyman SE, Neale BM. Predicting Polygenic Risk of
388 Psychiatric Disorders. *Biol Psychiatry*. 2019;86:97–109.
- 389 5. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic
390 risk scores. *Nat Rev Genet*. 2018;19:581–590.
- 391 6. Charney AW, Ruderfer DM, Stahl EA, Moran JL, Chambert K, Belliveau RA, et al.
392 Evidence for genetic heterogeneity between clinical subtypes of bipolar disorder.
393 *Transl Psychiatry*. 2017;7.
- 394 7. Alda M. The phenotypic spectra of bipolar disorder. *Eur Neuropsychopharmacol*.
395 2004;14 Suppl 2:S94-9.

- 396 8. Alda M, Hajek T, Calkin C, O'Donovan C. Treatment of bipolar disorder: new
397 perspectives. *Ann Med*. 2009;41:186–196.
- 398 9. Winham SJ, Cuellar-Barboza AB, McElroy SL, Oliveros A, Crow S, Colby CL, et al.
399 Bipolar disorder with comorbid binge eating history: A genome-wide association
400 study implicates APOB. *J Affect Disord*. 2014;165:151–158.
- 401 10. Kraepelin E. Manic-depressive insanity and paranoia: classics in psychiatry. Salem
402 (NH): Ayer Company Publishers; 1921.
- 403 11. Markota M, Coombes BJ, Larrabee BR, McElroy SL, Bond DJ, Veldic M, et al.
404 Association of schizophrenia polygenic risk score with manic and depressive
405 psychosis in bipolar disorder. *Transl Psychiatry*. 2018;8:188.
- 406 12. Ruderfer DM, Fanous AH, Ripke S, McQuillin A, Amdur RL, Gejman P V., et al.
407 Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and
408 schizophrenia. *Mol Psychiatry*. 2014;19:1017–1024.
- 409 13. Hamshere ML, O'Donovan MC, Jones IR, Jones L, Kirov G, Green EK, et al. Polygenic
410 dissection of the bipolar phenotype. *Br J Psychiatry*. 2011;198:284–288.
- 411 14. Allardyce J, Leonenko G, Hamshere M, Pardiñas AF, Forty L, Knott S, et al. Association
412 Between Schizophrenia-Related Polygenic Liability and the Occurrence and Level of
413 Mood-Incongruent Psychotic Symptoms in Bipolar Disorder. *JAMA Psychiatry*.
414 2018;75:28–35.
- 415 15. Ruderfer DM, Ripke S, McQuillin A, Boocock J, Stahl EA, Pavlides JMW, et al. Genomic
416 Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell*.
417 2018;173:1705-1715.e16.
- 418 16. Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics

- 419 Consortium. Electronic address: douglas.ruderfer@vanderbilt.edu, Bipolar Disorder
420 and Schizophrenia Working Group of the Psychiatric Genomics Consortium. Genomic
421 Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell*.
422 2018;173:1705-1715.e16.
- 423 17. van Hulzen KJE, Scholz CJ, Franke B, Ripke S, Klein M, McQuillin A, et al. Genetic
424 Overlap Between Attention-Deficit/Hyperactivity Disorder and Bipolar Disorder:
425 Evidence From Genome-wide Association Study Meta-analysis. *Biol Psychiatry*.
426 2017;82:634–641.
- 427 18. Wozniak J, Biederman J, Monuteaux MC, Richards J, Faraone S V. Parsing the
428 comorbidity between bipolar disorder and anxiety disorders: a familial risk analysis.
429 *J Child Adolesc Psychopharmacol*. 2002;12:101–111.
- 430 19. Duncan LE, Ratanatharathorn A, Aiello AE, Almli LM, Amstadter AB, Ashley-Koch AE,
431 et al. Largest GWAS of PTSD (N=20 070) yields genetic overlap with schizophrenia
432 and sex differences in heritability. *Mol Psychiatry*. 2018;23:666–673.
- 433 20. Yilmaz Z, Halvorsen M, Bryois J, Yu D, Thornton LM, Zerwas S, et al. Examination of
434 the shared genetic basis of anorexia nervosa and obsessive-compulsive disorder. *Mol*
435 *Psychiatry*. 2018. 7 August 2018. <https://doi.org/10.1038/s41380-018-0115-4>.
- 436 21. Witt SH, Streit F, Jungkunz M, Frank J, Awasthi S, Reinbold CS, et al. Genome-wide
437 association study of borderline personality disorder reveals genetic overlap with
438 bipolar disorder, major depression and schizophrenia. *Transl Psychiatry*.
439 2017;7:e1155.
- 440 22. Schulze TG, Akula N, Breuer R, Steele J, Nalls MA, Singleton AB, et al. Molecular
441 genetic overlap in bipolar disorder, schizophrenia, and major depressive disorder.

- 442 World J Biol Psychiatry. 2014;15:200–208.
- 443 23. Chen Q, Hartman CA, Haavik J, Harro J, Klungsøyr K, Hegvik TA, et al. Common
444 psychiatric and metabolic comorbidity of adult attention-deficit/hyperactivity
445 disorder: A population-based cross-sectional study. PLoS One. 2018;13.
- 446 24. McKowen JW, Frye MA, Altshuler LL, Gitlin MJ. Patterns of alcohol consumption in
447 bipolar patients comorbid for alcohol abuse or dependence. Bipolar Disord.
448 2005;7:377–381.
- 449 25. Salagre E, Dodd S, Aedo A, Rosa A, Amoretti S, Pinzon J, et al. Toward Precision
450 Psychiatry in Bipolar Disorder: Staging 2.0. Front Psychiatry. 2018;9.
- 451 26. Frye MA, McElroy SL, Fuentes M, Sutor B, Schak KM, Galardy CW, et al. Development
452 of a bipolar disorder biobank: differential phenotyping for subsequent biomarker
453 analyses. Int J Bipolar Disord. 2015;3:14.
- 454 27. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current
455 polygenic risk scores may exacerbate health disparities. Nat Genet. 2019;51:584–
456 591.
- 457 28. Association AP, others. Diagnostic and statistical manual of mental disorders (DSM-
458 5®). American Psychiatric Pub; 2013.
- 459 29. GAIN Collaborative Research Group, Manolio TA, Rodriguez LL, Brooks L, Abecasis G,
460 Collaborative Association Study of Psoriasis, et al. New models of collaboration in
461 genome-wide association studies: the Genetic Association Information Network. Nat
462 Genet. 2007;39:1045–1051.
- 463 30. Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, et al. Genome-
464 wide association study of bipolar disorder in European American and African

- 465 American individuals. *Mol Psychiatry*. 2009;14:755–763.
- 466 31. Delaneau O, Marchini J, Zagury J-F. A linear complexity phasing method for
467 thousands of genomes. *Nat Methods*. 2011;9:179–181.
- 468 32. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate
469 genotype imputation in genome-wide association studies through pre-phasing. *Nat*
470 *Genet*. 2012;44:955.
- 471 33. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-
472 wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet*.
473 2019;51:793–803.
- 474 34. Consortium SWG of the PG, Ripke S, Neale BM, Corvin A, Walters JTR, Farh K-H, et al.
475 Biological insights from 108 schizophrenia-associated genetic loci. *Nature*.
476 2014;511:421–427.
- 477 35. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide
478 meta-analysis of depression identifies 102 independent variants and highlights the
479 importance of the prefrontal brain regions. *Nat Neurosci*. 2019;22:343–352.
- 480 36. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. Discovery of
481 the first genome-wide significant risk loci for attention deficit/hyperactivity
482 disorder. *Nat Genet*. 2019;51:63–75.
- 483 37. Purves KL, Coleman JRI, Meier SM, Rayner C, Davis KAS, Cheesman R, et al. A major
484 role for common genetic variation in anxiety disorders. *Mol Psychiatry*. 2019. 20
485 November 2019. <https://doi.org/10.1038/s41380-019-0559-1>.
- 486 38. International Obsessive Compulsive Disorder Foundation Genetics Collaborative
487 (IOCDF-GC) and OCD Collaborative Genetics Association Studies (OC GAS). Revealing

- 488 the complex genetic architecture of obsessive-compulsive disorder using meta-
489 analysis. *Mol Psychiatry*. 2018;23:1181–1188.
- 490 39. Watson HJ, Yilmaz Z, Thornton LM, Hübel C, Coleman JRI, Gaspar HA, et al. Genome-
491 wide association study identifies eight risk loci and implicates metabo-psychiatric
492 origins for anorexia nervosa. *Nat Genet*. 2019;51:1207–1214.
- 493 40. Kranzler HR, Zhou H, Kember RL, Vickers Smith R, Justice AC, Damrauer S, et al.
494 Genome-wide association study of alcohol consumption and use disorder in 274,424
495 individuals from multiple populations. *Nat Commun*. 2019;10:1–11.
- 496 41. Lane JM, Jones SE, Dashti HS, Wood AR, Aragam KG, van Hees VT, et al. Biological and
497 clinical insights from genetics of insomnia symptoms. *Nat Genet*. 2019.
- 498 42. Lee JJ, Wedow R, Okbay A, Kong E, Maghziyan O, Zacher M, et al. Gene discovery and
499 polygenic prediction from a genome-wide association study of educational
500 attainment in 1.1 million individuals. *Nat Genet*. 2018;50:1112–1121.
- 501 43. Strawbridge RJ, Ward J, Cullen B, Tunbridge EM, Hartz S, Bierut L, et al. Genome-wide
502 analysis of self-reported risk-taking behaviour and cross-disorder genetic
503 correlations in the UK Biobank cohort. *Transl Psychiatry*. 2018;8.
- 504 44. Okbay A, Baselmans BML, De Neve JE, Turley P, Nivard MG, Fontana MA, et al. Genetic
505 variants associated with subjective well-being, depressive symptoms, and
506 neuroticism identified through genome-wide analyses. *Nat Genet*. 2016;48:624–633.
- 507 45. Nagel M, Jansen PR, Stringer S, Watanabe K, De Leeuw CA, Bryois J, et al. Meta-
508 analysis of genome-wide association studies for neuroticism in 449,484 individuals
509 identifies novel genetic loci and pathways. *Nat Genet*. 2018;50:920–927.
- 510 46. Ward J, Lyall LM, Bethlehem RAI, Ferguson A, Strawbridge RJ, Lyall DM, et al. Novel

- 511 genome-wide associations for anhedonia, genetic correlation with psychiatric
512 disorders, and polygenic association with brain structure. *Transl Psychiatry*.
513 2019;9:1–9.
- 514 47. UK Biobank — Neale lab. <http://www.nealelab.is/uk-biobank>. Accessed 24 February
515 2020.
- 516 48. Bulik-Sullivan B, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD score
517 regression distinguishes confounding from polygenicity in genome-wide association
518 studies. *Nat Genet*. 2015;47:291–295.
- 519 49. Dudbridge F, Cole CB, Palla L. avengeme: Analysis 1.0., of polygenic scoring methods.
520 R package version. 2019.
- 521 50. Ge T, Chen CY, Ni Y, Feng YCA, Smoller JW. Polygenic prediction via Bayesian
522 regression and continuous shrinkage priors. *Nat Commun*. 2019;10.
- 523 51. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data.
524 *Gigascience*. 2019;8.
- 525 52. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation
526 PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7.
- 527 53. Kupka RW, Luckenbaugh DA, Post RM, Leverich GS, Nolen WA. Rapid and non-rapid
528 cycling bipolar disorder: A meta-analysis of clinical studies. *J Clin Psychiatry*.
529 2003;64:1483–1494.
- 530 54. Mullins N, Bigdeli TB, Børglum AD, Coleman JRI, Demontis D, Mehta D, et al. GWAS of
531 Suicide Attempt in Psychiatric Disorders and Association With Major Depression
532 Polygenic Risk Scores. *Am J Psychiatry*. 2019;176:651–660.
- 533 55. Black DW, Nasrallah A. Hallucinations and delusions in 1,715 patients with unipolar

- 534 and bipolar affective disorders. *Psychopathology*. 1989;22:28–34.
- 535 56. Strassnig M, Kotov R, Cornaccio D, Fochtmann L, Harvey PD, Bromet EJ. Twenty-year
536 progression of body mass index in a county-wide cohort of people with
537 schizophrenia and bipolar disorder identified at their first episode of psychosis.
538 *Bipolar Disord*. 2017;19:336–343.
- 539 57. Caravaggio F, Brucato G, Kegeles LS, Lehembre-Shiah E, Arndt LY, Colibazzi T, et al.
540 Exploring the relationship between body mass index and positive symptom severity
541 in persons at clinical high risk for psychosis. *J Nerv Ment Dis*. 2017;205:893–895.
- 542 58. Kalman JL, Papiol S, Forstner AJ, Heilbronner U, Degenhardt F, Strohmaier J, et al.
543 Investigating polygenic burden in age at disease onset in bipolar disorder: Findings
544 from an international multicentric study. *Bipolar Disord*. 2018. 2018.
- 545 59. Brent DA, Oquendo M, Birmaher B, Greenhill L, Kolko D, Stanley B, et al. Familial
546 transmission of mood disorders: Convergence and divergence with transmission of
547 suicidal behavior. *J Am Acad Child Adolesc Psychiatry*. 2004;43:1259–1266.
- 548 60. Aedo A, Murru A, Sanchez R, Grande I, Vieta E, Undurraga J. Clinical characterization
549 of rapid cycling bipolar disorder: Association with attention deficit hyperactivity
550 disorder. *J Affect Disord*. 2018;240:187–192.
- 551 61. Kupka RW, Luckenbaugh DA, Post RM, Suppes T, Altshuler LL, Keck PE, et al.
552 Comparison of rapid-cycling and non-rapid-cycling bipolar disorder based on
553 prospective mood ratings in 539 outpatients. *Am J Psychiatry*. 2005;162:1273–1280.
- 554 62. Oquendo MA, Waternaux C, Brodsky B, Parsons B, Haas GL, Malone KM, et al. Suicidal
555 behavior in bipolar mood disorder: clinical characteristics of attempters and
556 nonattempters. *J Affect Disord*. 2000;59:107–117.

- 557 63. Valtonen HM, Suominen K, Mantere O, Leppämäki S, Arvilommi P, Isometsä ET.
558 Prospective study of risk factors for attempted suicide among patients with bipolar
559 disorder. *Bipolar Disord.* 2006;8:576–585.
- 560 64. Ducasse D, Loas G, Dassa D, Gramaglia C, Zeppegno P, Guillaume S, et al. Anhedonia is
561 associated with suicidal ideation independently of depression: A meta-analysis.
562 *Depress Anxiety.* 2018;35:382–392.
- 563
- 564

565 **Figure Legends**

566
567 **Figure 1.** Forest plot of significant PRS associations with each sub-phenotype stratified by
568 study (black = Combined; green = GAIN; maroon = Mayo). Each bar represents a 95%
569 confidence interval of the increased log(odds) in the sub-phenotype associated with one SD
570 increase in the PRS. P-values for each PRS included in the model by itself (P.m) or with
571 other significant PRSs (P.j) and adjusted Nagelkerke's R² (R2) are listed in the margins for
572 each PRS.

573 **Table 1.** Table of sub-phenotypes and sex for each study. P-value is for a chi-square test of
 574 differences between GAIN and Mayo BD-I subject sub-phenotypes.
 575

Variable	Value	All N = 1969	GAIN N = 1001	Mayo N = 968	p-value
Sex	Male	895 (45.5%)	500 (50.0%)	395 (40.8%)	0.003
	Female	1074 (54.5%)	501 (50.0%)	573 (59.2%)	
BD Type	BD-II	263 (13.4%)	-	263 (27.2%)	-
	BD-I	1706 (86.6%)	1001	705 (72.8%)	
Psychosis	No Psychosis	880 (47.3%)	344 (34.9%)	536 (61.3%)	3e-7
	Psychosis	980 (52.7%)	642 (65.1%)	338 (38.7%)	
	Missing	109	15	94	
Age-of-onset	> 18 yrs	1292 (70.0%)	569 (60.3%)	723 (80.0%)	1e-13
	< 19 yrs	555 (30.1%)	374 (39.7%)	181 (20.0%)	
	Missing	122	58	64	
Rapid Cycling	No	903 (50.1%)	508 (60.3%)	395 (41.1%)	3e-11
	Yes	901 (49.9%)	334 (39.7%)	567 (58.9%)	
	Missing	165	159	6	
Suicide attempts	None	1115 (57.2%)	554 (56.5%)	561 (58.0%)	0.619
	1+	833 (42.8%)	427 (43.5%)	406 (42.0%)	
	Missing	21	20	1	

576

■ Combined ■ GAIN ■ Mayo

