Cognitive function and lifetime features of depression and bipolar disorder in a large population sample: Cross-sectional study of 143,828 UK Biobank participants

Breda Cullen[a]*; Barbara I. Nicholl[b]; Daniel F. Mackay[c]; Daniel Martin[a]; Zia Ul-Haq[c,d]; Andrew McIntosh[e]; John Gallacher[f]; Ian J. Deary[g]; Jill P. Pell[c]; Jonathan J. Evans[a]; Daniel J. Smith[a].

[a] Mental Health and Wellbeing, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

[b] General Practice and Primary Care, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

[c] Public Health, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

[d] Institute of Public Health and Social Sciences, Khyber Medical University, Peshawar, Pakistan

[e] Division of Psychiatry, University of Edinburgh, Edinburgh, UK

[f] Department of Psychiatry, University of Oxford, Oxford, UK

[g] Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

*Corresponding author:

Dr Breda Cullen

Mental Health and Wellbeing

Room 4, University Section

Ground Floor, Office Block

Queen Elizabeth University Hospital

Glasgow G51 4TF

UK

Breda.Cullen@glasgow.ac.uk

+44(0)141 451 5879

Abstract

Background

This study investigated differences in cognitive performance between middle-aged adults with and without a lifetime history of mood disorder features, adjusting for a range of potential confounders.

Methods

Cross-sectional analysis of baseline data from the UK Biobank cohort. Adults aged 40-69 (n=143,828) were assessed using measures of reasoning, reaction time and memory. Self-reported data on lifetime features of major depression and bipolar disorder were used to construct groups for comparison against controls. Regression models examined the association between mood disorder classification and cognitive performance, adjusting for sociodemographic, lifestyle and clinical confounders.

Results

Inverse associations between lifetime history of bipolar or severe recurrent depression features and cognitive performance were attenuated or reversed after adjusting for confounders, including psychotropic medication use and current depressive symptoms. Participants with a lifetime history of single episode or moderate recurrent depression features outperformed controls to a small (but statistically significant) degree, independent of adjustment for confounders. There was a significant interaction between use of psychotropic medication and lifetime mood disorder features, with reduced cognitive performance observed in participants taking psychotropic medication.

Conclusions

In this general population sample of adults in middle age, lifetime features of recurrent depression or bipolar disorder were only associated with cognitive impairment within unadjusted analyses. These findings

50 underscore the importance of adjusting for potential confounders when investigating mood disorder-related
51 cognitive function.
52
53 **Key words** Cognition; unipolar depression; mania and bipolar disorder; epidemiology; UK Biobank

1. Introduction

Cognitive impairment is a common feature of mood disorders, persisting throughout remission or euthymia, and contributing to functional disability (1-6). Impairment is typically found on tests of attention, working and episodic memory, processing speed and executive function, with standardised effect sizes in the medium range compared to non-mood disordered controls (3-6). Impairments of memory and processing speed are greater during acute illness episodes than in periods of remission or euthymia, whereas deficits in attention and executive function are more likely to persist to a similar degree over time (1, 3, 7). It has been suggested that attention and executive function impairments are endophenotypic features of mood disorder, reflecting dysfunction of prefrontal brain networks (1, 8), but the influence of confounding factors on the relationship between mood disorder features and cognition is not well understood.

A number of demographic and clinical factors have been investigated in association with cognitive impairment in mood disorder. Older age (either at time of assessment or at illness onset) has been linked with poorer cognitive performance (1, 6), and history of more than one illness episode has been associated with greater impairment in some but not all studies (2, 5, 9). The influence of depression severity at time of assessment (measured by symptom rating scales) is unclear, with studies reporting an association with only certain cognitive tasks (1, 5); usually memory, speed and executive function. Studies on the effect of psychotropic medication have produced conflicting results. One study reported an improvement in verbal memory in patients with major depressive disorder (MDD) who responded to a selective serotonin reuptake inhibitor (SSRI) antidepressant (10). Another study of SSRI-responsive MDD patients found normal initial learning performance, but particular difficulties with generalisation of learning from one memory task to another (11). In a study of older adults with MDD, SSRI responders had minor improvements in visuospatial function and psychomotor speed from pre- to post-treatment, whereas SSRI non-responders showed deterioration in psychomotor speed and verbal memory (12). In bipolar disorder, cognitive impairment may be influenced by antipsychotic medications but not mood stabilisers (13). A recent individual participant data

meta-analysis of 2,876 patients with bipolar disorder (5) reported an adverse effect of antipsychotic medication on one measure of memory, but no effects of lithium, antidepressants or anticonvulsants; significantly better performance was also seen on memory tasks in the small number of patients who were drug-free compared with those who were on any psychotropic medication.

Other characteristics that are known to differ between mood disorder groups and the general population and which, in themselves, may be associated with variation in cognitive performance, include lifestyle factors (such as smoking and alcohol intake) and socioeconomic status. While many studies of cognitive function in mood disorder have adjusted for age, gender and education level, very few have taken account of other potential lifestyle-related confounders.

Studies in this field typically use small samples, and meta-analyses are limited by heterogeneity in study populations, diversity of cognitive measures used, and differences in levels of statistical adjustment. The UK Biobank project (http://www.ukbiobank.ac.uk/) provides an opportunity to investigate, at scale, the association of cognitive performance with clinically relevant features of depression and bipolar disorder. The aims of this cross-sectional analysis were to investigate differences in cognitive performance between UK Biobank participants with and without a history of depressive and bipolar disorder features, and to examine whether these differences were independent of a broad range of potential confounders.

2. Materials and Methods

2.1. Participants

Adults aged 40 to 69 years who were registered with the National Health Service and living within 25 miles of a study assessment centre were invited by mail to participate in UK Biobank. For the purposes of the

present study, participants were excluded if they self-reported a neurological condition that can affect cognitive performance (see Appendix S1 in the Supplementary Material for a list of excluded conditions).

2.2. Measures and procedure

This study was conducted under generic approval from the NHS National Research Ethics Service (Ref. 11/NW/0382). All participants gave written informed consent. Baseline assessments took place at 22 centres across England, Scotland and Wales between 2006 and 2010. Questionnaires and cognitive assessments were administered in a standardised order via a computerised touchscreen interface, followed by a face-to-face interview with a research nurse to obtain additional data. Assessment took place in a single visit lasting approximately 90 minutes.

2.2.1. Demographic and lifestyle data

The demographic data analysed in the present study included age, gender and neighbourhood-level socio-economic status as measured by the Townsend index of material deprivation (14). Educational qualifications were recorded, and for the present study were dichotomised according to whether or not participants held a university degree. Data on smoking status were used to classify participants into three categories (current, former and never); the latter two were combined into a 'non-smoker' category for some analyses. Current frequency of alcohol consumption was recorded ordinally over six categories from 'never' to 'daily/almost daily'; this was dichotomised to daily/almost daily versus other for some analyses. Current medications were self-reported to the research nurse, and participants were dichotomised according to whether or not they were taking any psychotropic medication (mood stabilisers, antidepressants, antipsychotics, sedatives or hypnotics); Appendix S2 in the Supplementary Material lists these medications.

2.2.2. Cognitive assessment

Five cognitive measures were administered via computerised touchscreen interface. The total time to complete all five cognitive tests was approximately 15 minutes. Full details of each measure are provided in

131 the Supplementary Material (Methods S1). The tests were designed specifically for UK Biobank, in order to
132 allow administration at scale without examiner supervision. Briefly, the assessments included:
133 • Reasoning
134 This test assessed the ability to solve verbal and numeric reasoning problems. The score was the total
135 number of correct answers given within a two minute period.
136 • Reaction time
137 This task recorded response time to visual stimuli (matching pairs). The score for analysis was the mean
138 response time in milliseconds.
139 • Numeric memory
140 Short term memory was assessed using number strings of increasing length. The score for analysis was the
141 maximum string length recalled correctly.
142 • Pairs matching
143 This task assessed visuospatial memory. Six pairs of symbols were presented on-screen in a random pattern.
144 The cards were then turned face down on the screen and participants were asked to touch as many pairs as
145 possible in the fewest tries. The score for analysis was the number of errors made.
146 • Prospective memory
147 The ability to remember and act on an instruction after a filled delay was assessed. For the present analyses,
148 data were dichotomised as either 'correct on first attempt' or not.
149
150 2.2.3. Mood disorder features
151 In the final two years of recruitment, the touchscreen questionnaire included questions about lifetime
152 experience of depressive and manic symptoms, based on symptoms within the Structured Clinical Interview
153 for DSM-IV Axis I Disorders (15), in addition to questions about medical help-seeking for mental health.
154 Responses were used to construct subgroups of participants with clinically significant lifetime features of
155 bipolar disorder and major depression; criteria are listed in Figure 1. Their validity has been described in
156 detail by our group previously (16). Participants were grouped according to whether they met our criteria for

features of each disorder, and were then assigned hierarchically to mutually exclusive categories, such that any individual meeting criteria for more than one subgroup was classified in the following hierarchical order: features of bipolar type I or bipolar type II (combined as bipolar disorder), severe recurrent major depression, moderate recurrent major depression and single episode major depression.

[Figure 1 about here]

The remainder of the sample who did not meet the above criteria formed the control group for statistical analysis. This group included participants with some mild, sub-threshold mood disorder features (shortlived symptoms of mania, or mild features of depression but without having seen a doctor), in addition to those with no clinically significant features of mood disorder.

Current depressive symptoms were assessed by four questions about depressive symptom experience in the past two weeks: frequency of depressed mood/hopelessness, unenthusiasm/uninterest, tenseness/restlessness, and tiredness/lethargy. Participants self-rated each symptom on a four-point scale from 'not at all' (scoring 0) to 'nearly every day' (scoring 3), summated to produce an overall score ranging from 0 to 12, with higher scores indicating more depressive symptoms during the preceding two weeks. Neuroticism was assessed using 12 yes/no questions from the Eysenck Personality Questionnaire Revised Neuroticism scale (17), producing a total score ranging from 0 to 12, with higher scores indicating greater neuroticism.

2.3. Statistical analysis

Associations between lifetime features of mood disorder and cognitive performance were tested using a series of regression models, with successive adjustment for confounders (see Methods S2 in the Supplementary Material for details). Regression results are reported as unstandardised coefficients,

incidence rate ratios (IRR) or odds ratios (OR), with 95% confidence intervals (CI). Statistical significance was defined conservatively as $p<0.01$. All analyses were performed using Stata version 12.1 (18).

3. Results

3.1. Composition and characteristics of the study participants

There was a response rate of approximately 6% to the study invitations (19). Of 172,745 participants who were assessed with regard to lifetime depressive and manic symptoms, 149,843 (86.7%) provided sufficient data to allow classification into subgroups according to presence or absence of clinically significant mood disorder features. Participants self-reporting a history of neurological disorder that can impair cognition (n=6,015; 4.0%) were excluded, leaving a study population of 143,828.

Of these, 1,521 (1.1%) met criteria for features of bipolar disorder; 8,354 (5.8%) for severe recurrent major depression; 14,386 (10.0%) for moderate recurrent major depression and 7,607 (5.3%) for a single episode of major depression. The remaining 111,960 (77.8%) formed the control group. Of these, 86,190 had no clinically significant mood disorder features (narrow control group), 23,384 had depressive symptoms that did not fulfil the criteria and 2,386 had manic symptoms that did not fulfil the criteria. Figure S1 in the Supplementary Material shows a flowchart of the sample composition.

Table 1 shows the demographic, lifestyle and psychological characteristics of the sample. Characteristics were very similar between the wide and narrow control groups (see Table 2), and so the results from the wide control group (including those with subthreshold mood disorder features) will be the focus here. The youngest group was the bipolar disorder group, who were approximately 2.7 years younger than the control group. Compared with the control group, women were over-represented in the three major depression groups but not in the bipolar group. A disproportionate number within the bipolar and severe recurrent

depression groups were at the more deprived end of the socioeconomic distribution, based on Townsend score quintiles, but the proportion holding a degree qualification was slightly higher for all mood disorder groups than for the control group. The prevalence of smoking was highest in the bipolar group and lowest in the control group, with evidence of a gradient across the groups according to severity of mood disorder. Conversely, abstention from alcohol was highest in the bipolar and severe recurrent depression groups.

Table 1. Sample characteristics, by lifetime mood disorder features

	Bipolar disorder *n* = 1521	**Recurrent major depression (severe) *n* = 8354**	**Recurrent major depression (moderate) *n* = 14 386**	**Single episode major depression *n* = 7607**	**Control *n* = 111 960**	**Test statistic**	**df**	***p* value**
Age[a] mean (SD), years	54.28 (8.08)	55.56 (8.06)	55.35 (7.93)	56.24 (8.02)	56.96 (8.20)	204.79[b]	4, 143 823	<.001
Gender[a] female *n* (%)	745 (49.0)	4846 (58.0)	9940 (69.1)	4860 (63.9)	56 517 (50.5)	2200[c]	4	<.001
Townsend score quintile								
Missing data *n*	1	17	29	19	169			
1 (least deprived) *n* (%)	162 (10.7)	1098 (13.2)	2273 (15.8)	1283 (16.9)	19 071 (17.1)			
2	196 (12.9)	1330 (16.0)	2769 (19.3)	1553 (20.5)	22 721 (20.3)			
3	258 (17.0)	1561 (18.7)	2952 (20.6)	1615 (21.3)	23 238 (20.8)	781.78[c]	16	<.001
4	373 (24.5)	1959 (23.5)	3268 (22.8)	1767 (23.3)	25 016 (22.4)			
5 (most deprived)	531 (34.9)	2389 (28.7)	3095 (21.6)	1370 (18.1)	21 745 (19.5)			
Educated to degree level								
Missing data *n*	11	61	85	38	1107			
Yes *n* (%)	546 (36.2)	3044 (36.7)	5029 (35.2)	2792 (36.9)	38 356 (34.6)	30.90[c]	4	<.001
Smoking status								
Missing data *n*	2	5	2	5	43			
Current *n* (%)	322 (21.2)	1392 (16.7)	1655 (11.5)	787 (10.4)	9809 (8.8)			
Former	534 (35.2)	3113 (37.3)	5203 (36.2)	2821 (37.1)	37 716 (33.7)	1100[c]	8	<.001
Never	663 (43.7)	3844 (46.0)	7526 (52.3)	3994 (52.5)	64 392 (57.5)			
Alcohol frequency								
Missing data *n*	3	7	4	4	61			
Daily/almost daily *n* (%)	299 (19.7)	1676 (20.1)	2792 (19.4)	1546 (20.3)	23 376 (20.9)			
3-4 times per week	258 (17.0)	1597 (19.1)	3112 (21.6)	1813 (23.9)	26 311 (23.5)			
1-2 times per week	344 (22.7)	1808 (21.7)	3572 (24.8)	1933 (25.4)	28 721 (25.7)	633.50[c]	20	<.001
1-3 times per month	191 (12.6)	1036 (12.4)	1911 (13.3)	924 (12.2)	12 182 (10.9)			
Special occasions only	228 (15.0)	1204 (14.4)	1900 (13.2)	861 (11.3)	12 579 (11.2)			
Never	198 (13.0)	1026 (12.3)	1095 (7.6)	526 (6.9)	8730 (7.8)			
On psychotropic medication								
Missing data *n*	24	95	190	118	1312			
Yes *n* (%)	464 (31.0)	2791 (33.8)	2550 (18.0)	587 (7.8)	3186 (2.9)	16 000[c]	4	<.001
Current depressive symptoms score (0-12)								
Missing data *n*	119	647	1051	529	9405			
Median (25^{th}, 75^{th} %ile)	3 (1, 5)	2 (1, 4)	2 (1, 4)	1 (0, 2)	1 (0, 2)	9478.18[d]	4	<.001

	Bipolar disorder *n* = 1521	Recurrent major depression (severe) *n* = 8354	Recurrent major depression (moderate) *n* = 14 386	Single episode major depression *n* = 7607	Control *n* = 111 960	Test statistic	df	*p* value
Neuroticism score (0-12)								
Missing data *n*	304	1372	2356	1107	18 575			
Mean (SD)	6.56 (3.58)	6.80 (3.39)	5.78 (3.18)	4.15 (2.99)	3.38 (2.93)	3769.96[b]	4, 120 109	<.001

SD, standard deviation.
a. No missing data.
b. One-way ANOVA.
c. Pearson χ^2 test.
d. Kruskal-Wallis test.

Table 2. Sample characteristics and cognitive performance in the wide and narrow control groups

	Wide control group[a] ***n* = 111 960**	**Narrow control group[b]** ***n* = 86 190**
Age[c] mean (SD), y	56.96 (8.20)	57.29 (8.12)
Female[c] *n* (%)	56 517 (50.5)	43 162 (50.1)
Townsend score quintile		
Missing data *n*	169	134
1 (least deprived) *n* (%)	19 071 (17.1)	15 029 (17.5)
2	22 721 (20.3)	17 764 (20.6)
3	23 238 (20.8)	17 990 (20.9)
4	25 016 (22.4)	18 970 (22.0)
5 (most deprived)	21 745 (19.5)	16 303 (18.9)
Educated to degree level		
Missing data *n*	1107	910
Yes *n* (%)	38 356 (34.6)	28 671 (33.6)
Smoking status		
Missing data *n*	43	34
Current *n* (%)	9809 (8.8)	7106 (8.3)
Former	37 716 (33.7)	28 878 (33.5)
Never	64 392 (57.5)	50 172 (58.2)
Alcohol frequency		
Missing data *n*	61	45
Daily/almost daily *n* (%)	23 376 (20.9)	18 083 (21.0)
3-4 times per week	26 311 (23.5)	20 410 (23.7)
1-2 times per week	28 721 (25.7)	22 068 (25.6)
1-3 times per month	12 182 (10.9)	9119 (10.6)
Special occasions	12 579 (11.2)	9662 (11.2)
Never	8730 (7.8)	6803 (7.9)
On psychotropic medication		
Missing data *n*	1312	1003
Yes *n* (%)	3186 (2.9)	2467 (2.9)
Current depressive symptoms score (0-12)		
Missing data *n*	9405	7464
Median (25^{th}, 75^{th} %ile)	1 (0, 2)	1 (0, 2)
Neuroticism score (0-12)		
Missing data *n*	18 575	14 563
Mean (SD)	3.38 (2.93)	3.25 (2.90)
Reasoning score (0-13)		
Missing data *n*	3835	3151
Mean (SD)	6.03 (2.16)	6.02 (2.16)
Reaction time (ms)		
Missing data *n*	1225	1012
Median (25^{th}, 75^{th} %ile)	543 (484, 621)	546 (485, 622)
Numeric memory score (2-12)		
Missing data[d] *n*	817	659
Mean (SD)	6.71 (1.34)	6.70 (1.33)
Pairs matching (errors)[c]		
Median (25^{th}, 75^{th} %ile)	3 (2, 5)	3 (2, 5)
Prospective memory		
Missing data *n*	568	483
Correct 1^{st} attempt *n* (%)	85 724 (77.0)	65 505 (76.4)

SD, standard deviation.

a. Full control group, including participants with no clinically significant mood disorder features, sub-threshold bipolar features or sub-threshold depressive features.

b. Subgroup restricted to participants with no clinically significant mood disorder features.
c. No missing data.
d. The Numeric Memory task was removed from the baseline battery before recruitment ended; missing data refers only to the period when this task was included in the battery.

Group differences were evident on the current depressive symptoms score, with the bipolar group having the highest median score and the single episode depression and control groups having the lowest. The severe recurrent depression group had a slightly higher mean neuroticism score than the bipolar group, with mean scores then reducing across the other groups.

3.2. Cognitive performance across groups

Table 3 shows the performance of each group on the five cognitive outcome measures. The single episode depression group performed best across all measures, followed by the moderate recurrent depression group and the control group. The severe recurrent depression and bipolar groups showed the poorest performance. Cognitive performance was very similar between the wide and narrow control groups (see Table 2).

Table 3. Cognitive performance, by lifetime mood disorder features

	Bipolar disorder *n* = 1521	**Recurrent major depression (severe) *n* = 8354**	**Recurrent major depression (moderate) *n* = 14 386**	**Single episode major depression *n* = 7607**	**Control *n* = 111 960**	**Test statistic**	**df**	***p* value**
Reasoning score (0-13)								
Missing data *n*	65	320	320	148	3835			
Mean (SD)	5.76 (2.24)	5.91 (2.22)	6.11 (2.08)	6.30 (2.08)	6.03 (2.16)	45.24	4, 139 135	<.001[a]
Reaction time (ms)								
Missing data *n*	21	101	108	38	1225			
Median (25^{th}, 75^{th} %ile)	547 (483, 620)	543 (484, 625)	540 (480, 614)	539 (480, 613)	543 (484, 621)	40.53	4	<.001[b]
Numeric memory score (2-12)								
Missing data[c] *n*	14	81	108	40	817			
Mean (SD)	6.62 (1.45)	6.63 (1.40)	6.71 (1.28)	6.78 (1.30)	6.71 (1.34)	4.26	4, 42 240	.002[a]
Pairs matching (errors)[d]								
Median (25^{th}, 75^{th} %ile)	3 (2, 6)	3 (2, 6)	3 (2, 5)	3 (2, 5)	3 (2, 5)	24.93	4	<.001[b]
Prospective memory								
Missing data *n*	11	48	49	16	568			
Correct 1^{st} attempt *n* (%)	1113 (73.7)	6252 (75.3)	11 381 (79.4)	6246 (82.3)	85 724 (77.0)	180.91	4	<.001[e]

SD, standard deviation.

a. One-way ANOVA.

b. Kruskal-Wallis test.

c. The Numeric Memory task was removed from the baseline battery before recruitment ended; missing data refers only to the period when this task was included in the battery.

d. No missing data.

e. Pearson χ^2 test.

3.3. Association between cognitive performance and other variables

In the sample as a whole, statistically significant associations of generally small effect size were found between performance on cognitive tests and demographic, lifestyle and psychological variables (see Results S1 in the Supplementary Material for details).

3.4. Relationship between cognitive function and lifetime features of mood disorder

Due to the strong correlation between the current depressive symptoms score and the neuroticism score (ρ=0.534, p<0.001), only the former was included as a covariate in the regression analyses. The results of the unadjusted and fully adjusted regression models for the five cognitive measures are shown in Table 4. Table S1 in the Supplementary Material gives detailed results for each successive stage of the model adjustment. When the models were re-run using the narrow control group, the results were virtually identical with respect to the magnitude, direction and statistical significance of the coefficients, and so only the results using the wide control group are reported here. Model results using the raw and log-transformed reaction time data were very similar, and so the results using the raw data are presented for ease of interpretation.

Table 4. Regression models for the association between lifetime mood disorder features and cognitive performance

	Unadjusted model		Fully adjusted model[a]	
	Coefficient (95% CI)	***p* value**	**Coefficient (95% CI)**	***p* value**
Reasoning score[b]	n = 139 140		n = 125 789	
Single episode major depression	0.27 (0.23, 0.31)	<.001	0.26 (0.22, 0.31)	<.001
Recurrent major depression (moderate)	0.08 (0.05, 0.12)	<.001	0.22 (0.18, 0.26)	<.001
Recurrent major depression (severe)	-0.12 (-0.17, -0.07)	<.001	0.12 (0.07, 0.17)	<.001
Bipolar disorder	-0.27 (-0.38, -0.17)	<.001	0.04 (-0.06, 0.13)	.442
Reaction time (ms)[b]	n = 142 335		n = 128 104	
Single episode major depression	-6.59 (-9.41, -3.77)	<.001	-5.47 (-8.19, -2.75)	<.001
Recurrent major depression (moderate)	-4.96 (-7.11, -2.81)	<.001	-6.24 (-8.29, -4.20)	<.001
Recurrent major depression (severe)	4.12 (1.52, 6.72)	.002	-0.82 (-3.40, 1.75)	.530
Bipolar disorder	2.37 (-5.28, 10.02)	.543	1.70 (-5.11, 8.51)	.625
Numeric memory score[b]	n = 42 245[c]		n = 38 051[c]	
Single episode major depression	0.07 (0.01, 0.13)	.015	0.07 (0.02, 0.12)	.004
Recurrent major depression (moderate)	0.00 (-0.04, 0.04)	.868	0.06 (0.02, 0.11)	.004
Recurrent major depression (severe)	-0.08 (-0.14, -0.02)	.006	0.02 (-0.04, 0.08)	.505
Bipolar disorder	-0.10 (-0.26, 0.07)	.267	0.03 (-0.10, 0.16)	.606
	IRR (95% CI)	***p* value**	**IRR (95% CI)**	***p* value**
Pairs matching (errors)[d]	n = 143 828		n = 129 229	
Single episode major depression	0.97 (0.95, 0.98)	<.001	0.98 (0.96, 0.99)	.009
Recurrent major depression (moderate)	0.98 (0.97, 0.99)	.023	0.99 (0.98, 1.01)	.428
Recurrent major depression (severe)	1.04 (1.02, 1.06)	<.001	1.03 (1.02, 1.05)	<.001
Bipolar disorder	1.04 (0.99, 1.08)	.085	1.04 (0.99, 1.08)	.095
	OR (95% CI)	***p* value**	**OR (95% CI)**	***p* value**
Prospective memory (correct 1st attempt)[e]	n = 143 136		n = 128 727	
Control group (reference)	1	-	1	-
Single episode major depression	1.39 (1.31, 1.48)	<.001	1.37 (1.28, 1.47)	<.001
Recurrent major depression (moderate)	1.15 (1.10, 1.20)	<.001	1.25 (1.19, 1.31)	<.001
Recurrent major depression (severe)	0.91 (0.87, 0.96)	<.001	1.11 (1.05, 1.18)	.001
Bipolar disorder	0.84 (0.75, 0.94)	.003	1.05 (0.92, 1.19)	.501

CI, confidence interval; IRR, incidence rate ratio; OR, odds ratio.

a. Full model adjusted for age, gender, current smoker, alcohol daily/almost daily, on psychotropic medication, current depressive symptoms score, has a degree, Townsend score.

b. Linear regression with bootstrapped standard errors; omitted reference group was the control group.

c. The Numeric Memory task was removed from the baseline battery before recruitment ended, yielding smaller possible sample sizes than for the other cognitive measures.

d. Negative binomial regression; omitted reference group was the control group.

e. Logistic regression.

3.4.1. Reasoning

The fully adjusted model indicated that all three major depression groups were outperforming controls, and the bipolar group was no longer significantly different from controls. The magnitude of the effect of group membership on performance compared to controls in the fully adjusted model was small, being no larger than an increase of a quarter of a point on a 0-13 point scale, seen in the single episode depression group.

3.4.2. Reaction time

In the fully adjusted model, the single episode and moderate recurrent depression groups were significantly faster than controls and the other groups were not significantly different from controls. Effect sizes were small, with the single episode and moderate recurrent depression groups being approximately 6ms faster than the control group in the fully adjusted model.

3.4.3. Numeric memory

The fully adjusted model suggested that performance was significantly better in the single episode and moderate recurrent depression groups, although this effect was small, representing a gain of less than a tenth of a point on a scale which ranges from 2 to 12. Performance in the other groups was no different from controls.

3.4.4. Pairs matching

The fully adjusted model indicated fewer errors in the single episode depression group and more errors in the severe recurrent depression group compared to controls, but the other groups were not significantly different from controls. Magnitude of effect of group membership was small (IRR of errors in the severe recurrent depression group versus controls = 1.03; 95% CI 1.02, 1.05). In contrast to the other cognitive measures, adjusting for confounders did not lead to notable changes in the effect sizes across successive models for each mood group on this measure.

3.4.5. Prospective memory

Following full adjustment, performance was significantly better than controls in the three major depression groups, and the bipolar group was not significantly different. Similar to the findings for the other cognitive measures, the magnitude of effect of group membership was modest (OR for correct response in single episode depression group versus controls = 1.37; 95% CI 1.28, 1.47).

3.5. Interaction between lifetime mood disorder status and psychotropic medication

A significant interaction was found between mood disorder group and current use of psychotropic medication in predicting cognitive performance. Table 5 shows the results of the fully adjusted regression models, stratified by use of psychotropic medication. For each cognitive measure, the overall better performance observed for the single episode depression group was driven by those not taking psychotropic medication and was absent in participants who reported being on psychotropic medication. Furthermore, although the bipolar group was not significantly different from controls on any cognitive measure in the overall analyses, the stratified models revealed that those taking psychotropic medication performed significantly worse than controls on two measures (reaction time and pairs matching). In the case of reaction time, participants with bipolar features who were on psychotropic medication were 24ms slower than controls, whereas those not on medication were 7ms faster. The pattern across most cognitive measures was that use of psychotropic medication attenuated any cognitive advantage in the milder depression groups, or worsened any disadvantage in the severe depression and bipolar groups; there was also greater variance in the groups taking psychotropic medication, which may partly be due to smaller sample sizes. An exception to this pattern of performance was the reasoning measure, on which the use of psychotropic medication slightly strengthened the better performance of the moderate and severe recurrent depression groups.

Table 5. Regression models for the association between lifetime mood disorder features and cognitive performance, stratified by use of psychotropic medication

	On psychotropic medication[a]		Not on psychotropic medication[a]	
	Coefficient (95% CI)	***p* value**	**Coefficient (95% CI)**	***p* value**
Reasoning score[b]	n = 8272		n = 117 517	
Single episode major depression	0.05 (-0.14, 0.25)	.582	0.28 (0.23, 0.34)	<.001
Recurrent major depression (moderate)	0.29 (0.19, 0.40)	<.001	0.21 (0.17, 0.25)	<.001
Recurrent major depression (severe)	0.20 (0.08, 0.32)	.001	0.09 (0.03, 0.16)	.005
Bipolar disorder	0.17 (-0.04, 0.38)	.110	-0.02 (-0.14, 0.11)	.814
Reaction time (ms)[b]	n = 8494		n = 119 610	
Single episode major depression	-3.35 (-14.96, 8.26)	.572	-5.56 (-8.17, -2.96)	<.001
Recurrent major depression (moderate)	-6.17 (-12.90, 0.57)	.073	-5.62 (-8.70, -2.54)	<.001
Recurrent major depression (severe)	6.93 (-0.25, 14.11)	.058	-2.92 (-6.81, 0.97)	.141
Bipolar disorder	23.72 (11.20, 36.23)	<.001	-6.82 (-12.98, -0.66)	.030
Numeric memory score[b]	n = 2636[c]		n = 35 415[c]	
Single episode major depression	0.03 (-0.19, 0.25)	.786	0.07 (0.02, 0.13)	.013
Recurrent major depression (moderate)	0.14 (-0.01, 0.30)	.069	0.05 (0.01, 0.09)	.024
Recurrent major depression (severe)	0.12 (-0.04, 0.27)	.132	0.00 (-0.07, 0.06)	.885
Bipolar disorder	0.00 (-0.34, 0.35)	.983	0.07 (-0.06, 0.21)	.307
	IRR (95% CI)	***p* value**	**IRR (95% CI)**	***p* value**
Pairs matching (errors)[d]	n = 8607		n = 120 622	
Single episode major depression	1.01 (0.94, 1.09)	.734	0.97 (0.95, 0.99)	.005
Recurrent major depression (moderate)	0.99 (0.95, 1.04)	.715	0.00 (0.98, 1.01)	.683
Recurrent major depression (severe)	1.07 (1.02, 1.11)	.003	1.03 (1.00, 1.05)	.025
Bipolar disorder	1.13 (1.04, 1.22)	.003	0.00 (0.95, 1.05)	.916
	OR (95% CI)	***p* value**	**OR (95% CI)**	***p* value**
Prospective memory (correct 1st attempt)[e]	n = 8551		n = 120 176	
Control group (reference)	1	-	1	-
Single episode major depression	1.13 (0.90, 1.41)	.296	1.40 (1.30, 1.50)	<.001
Recurrent major depression (moderate)	1.21 (1.06, 1.38)	.004	1.25 (1.18, 1.31)	<.001
Recurrent major depression (severe)	1.01 (0.89, 1.15)	.898	1.13 (1.05, 1.21)	.001
Bipolar disorder	0.85 (0.68, 1.07)	.177	1.11 (0.95, 1.31)	.183

CI, confidence interval; IRR, incidence rate ratio; OR, odds ratio.

a. Full model adjusted for age, gender, current smoker, alcohol daily/almost daily, current depressive symptoms score, has a degree, Townsend score.

b. Linear regression with bootstrapped standard errors; omitted reference group was the control group.

c. The Numeric Memory task was removed from the baseline battery before recruitment ended, yielding smaller possible sample sizes than for the other cognitive measures.

d. Negative binomial regression; omitted reference group was the control group.

e. Logistic regression.

3.6. Effect of missing data

Results across all stages of model adjustment were similar when the analyses were re-run including only those participants with complete data on all variables required for the final model (Model 5), indicating that the differences between initial and final models reported above were not simply driven by participants who contributed data to the unadjusted models but were missing from the fully adjusted analyses.

4. Discussion

These findings describe associations of cognitive performance with mood disorder features in a non-clinical population and over a range of disorders and severity, and may be contrasted with studies focussing on clinical populations. These data suggest that for mild mood disorder there is no detectable cognitive impairment at the group level, other than that associated with psychotropic medication.

Across all cognitive measures, the reduced performance compared to controls observed in the participants with a lifetime history of bipolar or severe recurrent depression features was attenuated or reversed after adjustment for confounders. Much of the attenuation followed adjustment for psychotropic medication use, alcohol consumption and smoking, as well as severity of current depressive symptoms. A different pattern of performance was evident for the groups with a lifetime history of single episode or moderate recurrent depression features: these participants slightly outperformed the controls on most cognitive measures, and this effect was stable or strengthened after adjustment for confounders. An interaction effect was observed between lifetime mood disorder features and use of psychotropic medication, such that current use of psychotropic medication attenuated any cognitive advantage in the single episode and moderate recurrent depression groups, or worsened any disadvantage in those with severe recurrent

368 depression or bipolar features. Effect sizes in the adjusted regression models were small and unlikely
369 to be clinically significant.
370
371 It is likely that mood disorder features are highly heterogeneous in the general population, covering
372 a broad range of symptom experience and disability. A substantial proportion will not come to the
373 attention of mental health services and will not be included in clinic-based studies of cognition. It
374 may therefore be the case that much clinic-based research is not representative of the abilities and
375 functioning of a large sector of the general population who are living with, or have past experiences
376 of, mild to moderate features of mood disorder. Variation in cognitive performance within clinic-
377 based samples also merits greater attention, with two recent studies showing that almost one-third
378 of affectively stable bipolar disorder patients were indistinguishable from healthy controls on
379 neurocognitive measures (20, 21). It is therefore apparent that the cognitive phenotype associated
380 with mood disorder is diverse, both in clinical settings and in the broader population. It may be that,
381 even after adjustment for confounding factors, summary measures of group-level performance on
382 cognitive tests mask wide variation across participants. Some sub-groups of patients will show
383 prominent cognitive impairment as part of the depression or bipolar syndrome, while others are
384 cognitively resilient. We therefore do not believe that cognitive impairment should be dismissed as a
385 key feature of affective disorders, despite our study findings, but rather suggest that careful
386 consideration be given to issues of heterogeneity and confounding when considering the presence
387 of cognitive dysfunction in mood disorder populations.
388
389 Several explanations may be offered for the slightly better performance found in the milder
390 depression groups relative to controls. The apparent cognitive advantage in these groups may be
391 artefactual. Possible explanations include differential selection, differential recall or residual
392 confounding. Differential selection might operate through invitees with mild mood disorder being
393 more likely to respond than those of comparable ability without mood disorder. Differential recall
394 bias might be operating through cognitive performance being associated with greater ability and/or

motivation to recall or report an episode of mood disorder, or with different interpretation of the questionnaire leading to classification bias (22). Alternatively, higher cognitive ability might be associated with differential help-seeking or clinical course among people who have experienced depression: if higher ability were associated with reduced likelihood of psychiatric healthcare input then those who did experience recurrent depression would have been classified as moderate rather than severe by our criteria; or if higher ability were protective in some way against recurrence of depression then such people would be over-represented in the single episode group. Finally, despite the wide range of variables analysed here, as with all observational studies, there is a possibility of residual confounding.

Use of psychotropic medication was associated with worse cognitive performance, even after adjustment for other variables. This is congruent with previous research (5, 11-13). It may be that this relationship is indirect, with use of psychotropic medication being a proxy for mood disorder severity. If a direct relationship exists, it is unclear what the mechanism might be; certain medications might have sedative or anticholinergic effects, or may contribute to adiposity and associated cardiometabolic problems which in turn increase the risk of cognitive impairment.

The unique strength of this study is that it has assessed the relationship between lifetime features of mood disorder and cognitive performance in a large population sample. The use of standardised procedures and outcome measures across study sites represents a significant advantage over previous studies that have relied on sample aggregation or meta-analytic methods to achieve statistical power. Unlike other studies in this field, it has been possible to take into account a wide range of potential confounding variables, also measured in a standard way. Robust multivariate analysis methods were used, including testing for interactions. The classification of participants across the severity spectrum from single depressive episode to recurrent depression and bipolar disorder has allowed us to delineate differential effects between groups and to investigate 'dose-response' associations between mood disorder severity and cognitive performance.

Limitations in the data included lack of access to clinical records including healthcare service use, and the brevity and self-report nature of mood measurement. We have previously described the validity of the mood disorder criteria used (16), however, and scores on measures such as the neuroticism scale successfully differentiated the various groups in accordance with what would be expected from the clinical literature. The cognitive assessment battery was brief in comparison to a typical clinical neuropsychological assessment; it did not include classic tests of learning and delayed memory; and it used bespoke tasks which are not directly comparable to standard clinical measures routinely used in clinical neuropsychology practice and research. The pattern of findings across the five different cognitive measures was consistent, however. The presence of missing data on the covariates reduced the sample size at each successive stage of regression model adjustment, although results were similar when analyses were repeated using only participants who provided complete data on all variables. The sample was aged between 40 and 69 years and so findings may not be extrapolated beyond this age range. The associations were based on cross-sectional data and so cannot address questions of causality and temporal relationships, but they provide a useful basis for future longitudinal analyses.

Future prospective work in UK Biobank will use repeat assessments and linked routine healthcare data to investigate a range of long-term outcomes. It will be possible to investigate whether baseline measures of cognition and other characteristics predict incident mood disorder in the control participants, or recurrence of mood disorder in those who have had a single episode of depression. There is potential for a range of studies of cognitive function and mood disorder incorporating other ongoing UK Biobank assessments including neuroimaging, genotyping and biochemical assays. In particular, we will be able to examine the role of genetic markers alongside the covariates analysed in this study, to elucidate mediating and moderating effects of genetic and environmental factors in order to understand the wide variation in cognitive functioning among people with mood disorders. The present study contributes important knowledge about the multifactorial nature of mood

disorder-related cognitive function, and lays a foundation for future research which will increase our understanding in this area.

Acknowledgements

We are grateful to Dr Jason Gill, University of Glasgow, for contributions to discussions of this study. This research was conducted using the UK Biobank resource. UK Biobank was established by the Wellcome Trust medical charity, Medical Research Council, Department of Health, Scottish Government and Northwest Regional Development Agency. UK Biobank has also had funding from the Welsh Assembly Government and the British Heart Foundation. Data collection was funded by UK Biobank. The funders had no role in the design or analysis of this study, decision to publish, or preparation of the manuscript.

Figure legend

Figure 1. Mood status criteria

Criteria for hierarchical classification of the sample according to self-reported lifetime features of major depression or bipolar disorder

References

1. Austin MP, Mitchell P, Goodwin GM. Cognitive deficits in depression - Possible implications for functional neuropathology. Br J Psychiatry. 2001;178:200-6.

2. McClintock SM, Husain MM, Greer TL, Cullum CM. Association Between Depression Severity and Neurocognitive Function in Major Depressive Disorder: A Review and Synthesis. Neuropsychology. 2010;24(1):9-34.

3. Lee RSC, Hermens DF, Porter MA, Redoblado-Hodge MA. A meta-analysis of cognitive deficits in first-episode Major Depressive Disorder. J Affect Disord. 2012;140(2):113-24.

4. Bora E, Harrison BJ, Yucel M, Pantelis C. Cognitive impairment in euthymic major depressive disorder: a meta-analysis. Psychol Med. 2013;43(10):2017-26.

5. Bourne C, Aydemir O, Balanza-Martinez V, Bora E, Brissos S, Cavanagh JTO, et al. Neuropsychological testing of cognitive impairment in euthymic bipolar disorder: an individual patient data meta-analysis. Acta Psychiatr Scand. 2013;128(3):149-62.

6. Samame C, Martino DJ, Strejilevich SA. A quantitative review of neurocognition in euthymic late-life bipolar disorder. Bipolar Disord. 2013;15(6):633-44.

7. Porter RJ, Bowie CR, Jordan J, Malhi GS. Cognitive remediation as a treatment for major depression: A rationale, review of evidence and recommendations for future research. Aust N Z J Psych. 2013;47(12):1165-75.

8. Zihl J, Reppermund S, Thum S, Unger K. Neuropsychological profiles in MCI and in depression: Differential cognitive dysfunction patterns or similar final common pathway disorder? J Psychiatr Res. 2010;44(10):647-54.

9. Kessing LV. Cognitive impairment in the euthymic phase of affective disorder. Psychol Med. 1998;28(5):1027-38.

10. Vythilingam M, Vermetten E, Anderson GM, Luckenbaugh D, Anderson ER, Snow J, et al. Hippocampal volume, memory, and cortisol status in major depressive disorder: Effects of treatment. Biol Psychiatry. 2004;56(2):101-12.

11. Herzallah MM, Moustafa AA, Natsheh JY, Danoun OA, Simon JR, Tayem YI, et al. Depression impairs learning, whereas the selective serotonin reuptake inhibitor, paroxetine, impairs generalization in patients with major depressive disorder. J Affect Disord. 2013;151(2):484-92.

12. Culang ME, Sneed JR, Keilp JG, Rutherford BR, Pelton GH, Devanand DP, et al. Change in Cognitive Functioning Following Acute Antidepressant Treatment in Late-Life Depression. Am J Geriatr Psychiatr. 2009;17(10):881-8.

13. Palsson E, Figueras C, Johansson AGM, Ekman CJ, Hultman B, Ostlind J, et al. Neurocognitive function in bipolar disorder: a comparison between bipolar I and II disorder and matched controls. BMC Psychiatry. 2013;13:9.

14. Townsend P. Deprivation. J Soc Policy. 1987;16:125-46.

15. First MB, Spitzer RL, Gibbon M, Williams JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. New York, NY: Biometrics Research, New York State Psychiatric Institute; 2002.

16. Smith DJ, Nicholl BI, Cullen B, Martin D, Ul-Haq Z, Evans J, et al. Prevalence and Characteristics of Probable Major Depression and Bipolar Disorder within UK Biobank: Cross-Sectional Study of 172,751 Participants. PLoS One. 2013;8(11):7.

17. Eysenck SBG, Eysenck HJ, Barrett P. A revised version of the Psychoticism scale. Pers Individ Differ. 1985;6(1):21-9.

18. StataCorp. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP; 2011.

19. Manolio TA, Weis BK, Cowie CC, Hoover RN, Hudson K, Kramer BS, et al. New Models for Large Prospective Studies: Is There a Better Way? Am J Epidemiol. 2012;175(9):859-66.

20. Burdick KE, Russo M, Frangou S, Mahon K, Braga RJ, Shanahan M, et al. Empirical evidence for discrete neurocognitive subgroups in bipolar disorder: clinical implications. Psychol Med. 2014;44(14):3083-96.

21. Martino DJ, Strejilevich SA, Marengo E, Ibanez A, Scapola M, Igoa A. Toward the identification of neurocognitive subtypes in euthymic patients with bipolar disorder. J Affect Disord. 2014;167:118-24.

22. Deary IJ, Watson R, Booth T, Gale CR. Does Cognitive Ability Influence Responses to the Warwick-Edinburgh Mental Well-Being Scale? Psychol Assess. 2013;25(2):313-8.

Criteria for lifetime experience of features of bipolar disorder

Bipolar disorder, type I
Ever 'manic or hyper' for 2 days OR ever 'irritable/argumentative' for 2 days; plus at least 3 features from 'more active', 'more talkative', 'needed less sleep' and 'more creative/more ideas'; plus duration of a week or more; plus 'needed treatment or caused problems at work'.

Bipolar disorder, type II
Ever 'manic or hyper' for 2 days OR ever 'irritable/argumentative' for 2 days; plus at least 3 features from 'more active', 'more talkative', 'needed less sleep' and 'more creative/more ideas'; plus duration of a week or more.

Criteria for lifetime experience of features of major depression

Probable recurrent major depression (severe)
Ever depressed/down for a whole week; plus at least two weeks duration; plus at least two episodes; plus ever seen a psychiatrist for 'nerves, anxiety, tension, depression' OR ever anhedonic (unenthusiasm/uninterest) for a whole week; plus at least two weeks duration; plus at least two episodes; plus ever seen a psychiatrist for 'nerves, anxiety, tension, depression'.

Probable recurrent major depression (moderate)
Ever depressed/down for a whole week; plus at least two weeks duration; plus at least two episodes; plus ever seen a GP (but not a psychiatrist) for 'nerves, anxiety, tension, depression' OR ever anhedonic (unenthusiasm/uninterest) for a whole week; plus at least two weeks duration; plus at least two episodes; plus ever seen a GP (but not a psychiatrist) for 'nerves, anxiety, tension, depression'.

Single probable episode of major depression
Ever depressed/down for a whole week; plus at least two weeks duration; plus only one episode; plus ever seen a GP or a psychiatrist for 'nerves, anxiety, tension, depression' OR ever anhedonic (unenthusiasm/uninterest) for a whole week; plus at least two weeks duration; plus only one episode; plus ever seen a GP or a psychiatrist for 'nerves, anxiety, tension, depression'.

Appendix S1. List of neurological conditions recorded in the UK Biobank dataset (self-reported by participants; from data fields 6150, 20001 and 20002).

Brain cancer/primary malignant tumour
Brain haemorrhage
Brain/intracranial abscess
Cerebral aneurysm
Cerebral palsy
Chronic/degenerative neurological problem
Dementia/Alzheimer's disease/cognitive impairment
Encephalitis
Epilepsy
Head injury
Infection of nervous system
Ischaemic stroke
Meningeal cancer/malignant meningioma
Meningioma (benign)
Meningitis
Motor neurone disease
Multiple sclerosis
Neurological injury/trauma
Neuroma (benign)
Other demyelinating condition
Other neurological problem
Parkinson's disease
Spina bifida
Stroke
Subarachnoid haemorrhage
Subdural haematoma
Transient ischaemic attack

Appendix S2. List of psychotropic medications recorded in the UK Biobank dataset (self-reported by participants; from data field 20003).

Mood stabiliser	**Anti-depressant (SSRI)**	**Anti-depressant (other)**	**Anti-psychotic (traditional)**	**Anti-psychotic (2nd generation)**	**Sedative/hypnotic**
lithium product Priadel (lithium) Camcolit (lithium) sodium valproate Epilim (sodium valproate) Depakote (semisodium valproate) valproic acid carbamazepine product carbamazepine Tegretol (carbamazepine) Teril (carbamazepine) Teril retard (carbamazepine) Timonil retard (carbamazepine) Epimaz (carbamazepine)	paroxetine Seroxat (paroxetine) fluoxetine Prozac (fluoxetine) citalopram Cipramil (citalopram) escitalopram Cipralex (escitalopram) sertraline Lustral (sertraline) fluvoxamine	mirtazapine Zispin (mirtazapine) duloxetine Cymbalta (duloxetine) Yentreve (duloxetine) venlafaxine Efexor (venlafaxine) amitriptyline Elavil (amitriptyline) Tryptizol (amitriptyline) Lentizol (amitriptyline) amitriptyline+perphenazine Triptafen (amitriptyline+perphenazine) amitriptyline+chlordiazepoxide Limbitrol 10 (amitriptyline+chlordiazepoxide) Limbitrol-5 (amitriptyline+chlordiazepoxide) phenelzine maoi - phenelzine Nardil (phenelzine) moclobemide Manerix (moclobemide) imipramine Tofranil (imipramine) trimipramine Surmontil (trimipramine) dothiepin dosulepin Prothiaden (dosulepin) Thaden (dosulepin) clomipramine Anafranil (clomipramine) lofepramine Gamanil (lofepramine) Lomont (lofepramine) mianserin Bolvidon (mianserin) Norval (mianserin)	chlorpromazine cpz - chlorpromazine Largactil (chlorpromazine) haloperidol Haldol (haloperidol) Serenace (haloperidol) fluphenazine decanoate fluphenazine Modecate (fluphenazine) Moditen tablet (fluphenazine) Moditen enanthate (fluphenazine) flupentixol Flupenthixol (flupentixol) Depixol (flupentixol) Fluanxol (flupentixol) zuclopenthixol Clopixol (zuclopenthixol) loxapine Loxapac (loxapine) droperidol Droleptan (droperidol) trifluoperazine Stelazine (trifluoperazine) thioridazine Melleril (thioridazine)	quetiapine Seroquel (quetiapine) risperidone Risperdal (risperidone) olanzapine Zyprexa (olanzapine) aripiprazole Abilify (aripiprazole) amisulpride Solian (amisulpride) clozapine Clozaril (clozapine)	diazepam diazepam product Valium tablet (diazepam) Valium syrup (diazepam) Valium supp (diazepam) temazepam Normison (temazepam) Euhypnos (temazepam) zopiclone Zimovane (zopiclone) zaleplon Sonata (zaleplon) zolpidem Stilnoct (zolpidem) nitrazepam Mogadon (nitrazepam) Nitrados (nitrazepam) Remnos (nitrazepam) Somnite (nitrazepam) Noctesed (nitrazepam) Surem (nitrazepam) Unisomnia (nitrazepam) flunitrazepam Rohypnol (flunitrazepam) triazolam Halcion (triazolam)

Methods S1. Detailed description of the cognitive assessments.

Five cognitive measures were administered via computerised touchscreen interface:

- Reasoning

Thirteen questions were presented sequentially via touchscreen on a self-paced basis with an overall time limit of two minutes. Responses were selected from a multiple-choice array. Any questions not attempted during the two-minute time limit were scored as zero. The score for analysis was an unweighted total from 0 to 13 (UK Biobank data field 20016, known as the 'fluid intelligence' test), with higher scores indicating better performance.

- Reaction time

This test was based on a 'Snap'-style computer game, in which participants were asked to press a button with their dominant hand as quickly as possible each time a matching pair of symbols was presented on-screen. Twelve pairs of symbols were presented in total. The score for analysis was the mean time (in milliseconds) to press the button, derived from all trials in which a matching pair occurred (UK Biobank data field 20023). Higher scores indicate slower (i.e. worse) performance.

- Numeric memory

A string of numbers was presented on-screen, and after a brief delay participants were asked to enter it from memory, in reverse order, via a numeric keypad. Each string was presented on screen for a period of 2000ms, plus an additional 500ms multiplied by the string length. A delay of 3000ms occurred between clearing the screen and activating the response keypad. All participants began with a string length of two, and successive strings increased by one, up to a maximum string length of 12. The test was discontinued after five successive incorrect responses at a string length of two, or after two successive incorrect responses at string lengths of three or more. The score for analysis was the maximum string length recalled correctly (UK Biobank data field 4282), with higher scores indicating better performance. This task was phased out before recruitment ended for reasons of time, yielding a smaller sample size than for the other cognitive measures.

- Pairs matching

Pairs of symbols were presented on-screen in a random array. Participants were asked to memorise the position of as many matching pairs as possible. The cards were then turned face down on the screen and participants were asked to touch as many pairs as possible in the fewest tries. The score for analysis was the number of errors made while attempting to select the pairs, with a higher score indicating worse performance. Two trials of this task were administered, one with three pairs of symbols and one with six pairs. Because of a ceiling effect on the three-pair trial, only the score on the six-pair trial of the test was analysed in the present study (UK Biobank data field 399.0.2).

- Prospective memory

The following instruction appeared on the touchscreen: "At the end of the games we will show you four coloured symbols and ask you to touch the blue square. However, to test your memory, we want you to actually touch the orange circle instead". After a delay during which participants underwent the other cognitive tasks described above, a screen appeared showing four coloured shapes with the instruction to touch the blue square. If the participant touched the orange circle, their response was recorded as 'correct on first attempt'. If they touched the blue square, they were given a prompt on-screen to try to recall what the original instruction was, and asked to respond again. If they correctly selected the orange circle after receiving this prompt, their response was recorded as 'correct on second attempt'. All other responses were recorded as incorrect. For the present analyses, data were dichotomised as either 'correct on first attempt' or not (derived from UK Biobank data field 20018).

Methods S2. Statistical analysis.

Townsend index scores were categorised into quintiles based on frequency to facilitate comparisons across groups on the descriptive and unadjusted analyses; quintile 1 represents the least deprived and quintile 5 the most deprived areas. The characteristics and cognitive performance of each group were summarised using percentages, means and medians, as appropriate, and were compared using one-way ANOVA or the Kruskal-Wallis test for continuous data, and the Pearson χ^2 test for categorical data. Bivariate associations between cognitive scores and other characteristics were tested using Pearson χ^2 tests, Pearson or Spearman correlation tests, *t*-tests, or Mann-Whitney *U*-tests, depending on the type and distribution of the variables; standardised effect sizes are reported as Pearson's *r*, Spearman's ρ, risk ratios (RR), or η^2. A natural log transformation was applied to normalise the reaction time data, but where results were similar between raw and transformed data, the former are reported for ease of interpretation.

Separate univariate regression models were constructed to investigate the relationship between mood disorder category and each of the five cognitive measures, followed by a series of multivariate models with additional covariates added sequentially. In each regression model, the dependent variable was the cognitive measure. The independent variable was mood disorder category, which was entered as a categorical variable, with the control group as the reference value. The type of regression model used depended on the nature and distribution of the cognitive data and the need to address any violation of model assumptions. For reasoning, reaction time and numeric memory, linear regression models were used, with bootstrapped standard errors to minimise the effect of non-normal residuals and heteroscedasticity; results are presented as unstandardised coefficients with 95% confidence intervals (CI). The distribution of the pairs matching data was found to be significantly positively skewed and leptokurtic, but a log transformation was not advisable due to the presence of zero-values in the data. Instead, a negative binomial regression model was used (results presented as incidence rate ratios [IRR] with 95% CI). Performance on the prospective memory test was coded dichotomously, with 1 representing 'correct on first attempt' and 0 representing other outcomes (including correct on second attempt and incorrect), and a logistic regression model was applied; results are presented as odds ratios (OR) with 95% CI.

The first multivariate model (Model 1) adjusted for age and gender. Model 2 also included current smoking status (smoker versus non), alcohol use (daily/almost daily versus other) and psychotropic medication (yes/no). Current depressive symptoms score was added in Model 3. Models 4i and 4ii added degree (yes/no) or Townsend score, respectively, to Model 3; these were added separately at this stage in order to allow separate investigation of their effects, because they are often conflated. Model 5 was the fully adjusted model and included all the covariates in Model 3 plus degree and Townsend score.

The presence of statistically significant interactions between the mood disorder category variable and other covariates (age group [decade], gender, degree, psychotropic medication) was tested by applying the likelihood ratio test to model estimations with and without an interaction term. Where this indicated presence of a significant interaction across the cognitive variables, stratified analyses were then carried out to investigate the differential effect of the covariate on the regression results.

Analyses were repeated using a narrower control group comprising only participants with no mood disorder, for the purpose of comparison. Repeat analyses were also conducted which were restricted to only those participants who provided complete data on all the covariates (and therefore could be entered in to the final model), in order to explore the potential effect of missing data across successive adjusted models.

Results S1. Association between cognitive performance and other variables.

- Age

Older age was associated with worse performance on all cognitive tests: reasoning r=-0.051; reaction time ρ=0.321; numeric memory r=-0.118; pairs matching ρ=0.145; prospective memory r=-0.103 (all p<0.001).

- Gender

Men scored slightly better than women on all tests (reasoning r=0.058; reaction time r=0.085; numeric memory r=0.088; prospective memory risk ratio [RR]=1.02; all p<0.001) except the pairs matching task, on which women performed marginally better (r=0.008, p=0.004).

- Deprivation

Slightly better performance was seen on all tests in participants with less deprived Townsend scores: reasoning ρ=-0.121; reaction time ρ=0.042; numeric memory ρ=-0.063; pairs matching ρ=0.011; prospective memory r=-0.097 (all p<0.001).

- Education

Having a degree was associated with better cognitive performance: reasoning r=0.307; reaction time r=0.089; numeric memory r=0.169; pairs matching r=0.045; prospective memory RR=1.11 (all p<0.001).

- Smoking

There were differences in performance (p<0.001) associated with smoking status on all tests except numeric memory. Current smokers showed relatively lower performance on reasoning (mean [SD]: current 5.67 [2.17], former 6.11 [2.11], never 6.07 [2.17]) and prospective memory (% correct on first attempt: current 74.9, former 78.5, never 77.1) but relatively better performance on pairs matching (error score mean [SD]: current 3.87 [3.26], former 4.04 [3.25], never 4.04 [3.41]). Former smokers had the slowest reaction time (median [25^{th}, 75^{th} percentile]: current 543 [484, 622], former 547 [488, 621], never 542 [481, 620]).

- Alcohol

Participants who never consume alcohol obtained the poorest scores on all tests (all p<0.001):

  - Reasoning η^2 0.034;
  - Reaction time in ms, median (25^{th}, 75^{th} percentile): Never 570 (504, 660); special occasions 559 (496, 640); 1-3 times p/month 539 (480, 617); 1-2 times p/week 539 (480, 617); 3-4 times p/week 535 (480, 610); daily/almost daily 539 (481, 613);
  - Numeric memory η^2 0.015;
  - Pairs matching errors, median (25^{th}, 75^{th} percentile): Never 4 (2, 6); special occasions 3 (2, 6); each other group 3 (2, 5);
  - Prospective memory % correct on first attempt: Never 64.4; special occasions 69.7; 1-3 times p/month 78.3; 1-2 times p/week 77.4; 3-4 times p/week 81.2; daily/almost daily 81.8.

- Psychotropic medication

Consumption of any psychotropic medication was weakly associated with poorer performance: reasoning r=-0.038; reaction time r=0.036; numeric memory r=-0.041; pairs matching r=0.016; prospective memory RR=0.93 (all p<0.001).

- Current depressive symptoms

Higher current depressive symptom scores were associated with slightly worse performance on all tests except pairs matching: reasoning ρ=-0.056; reaction time r=0.016; numeric memory ρ=-0.034; prospective memory r=-0.037 (all p<0.001).

- Neuroticism

Higher neuroticism scores were similarly associated with slightly poorer performance: reasoning r=-0.059; reaction time ρ=0.011; numeric memory r=-0.043; pairs matching ρ=0.017; prospective memory r=-0.036 (all p<0.001).

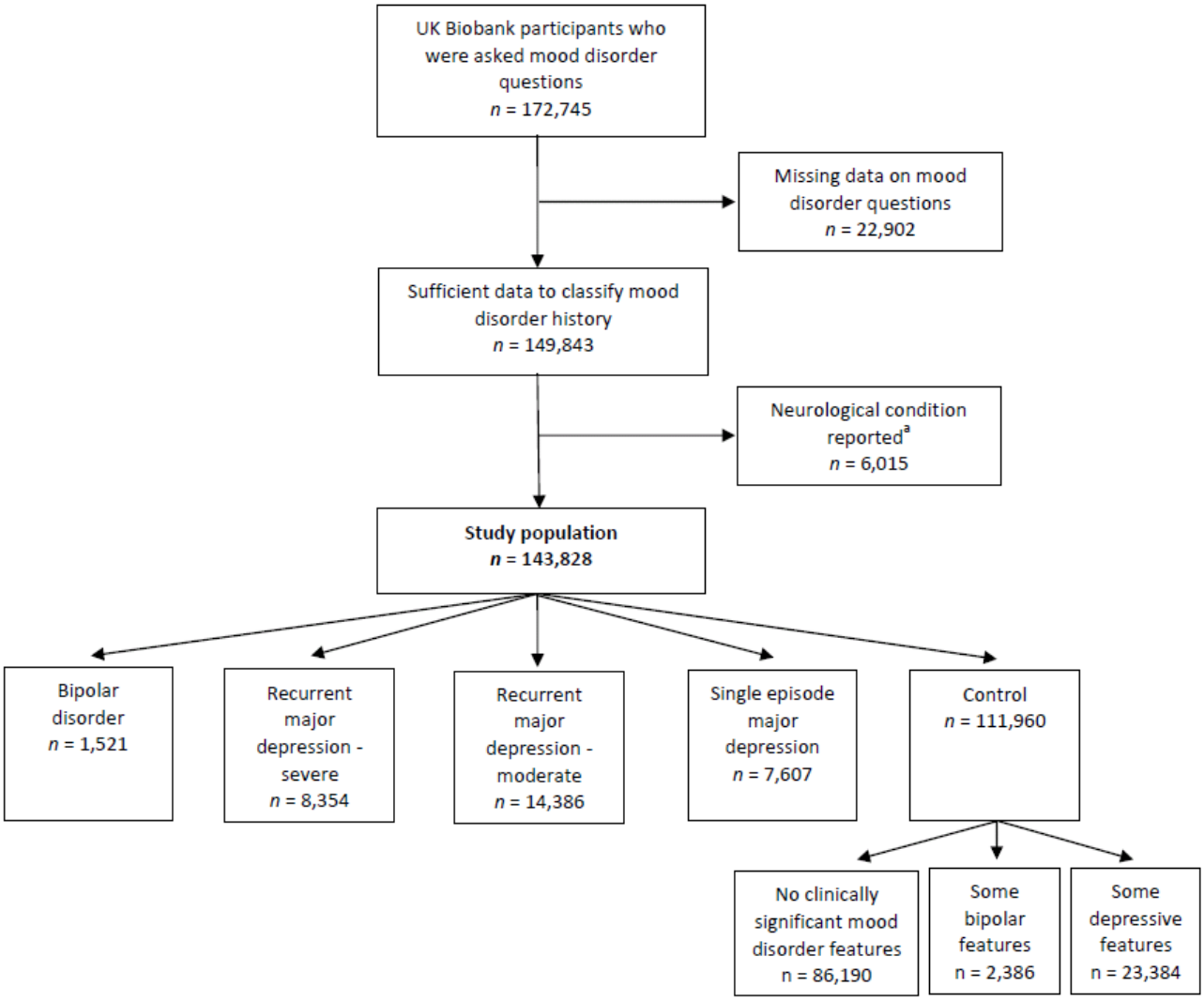

Figure S1. Sample composition.

a. Participants self-reported one or more neurological conditions which can potentially affect cognitive performance (see Appendix S1 for full list).

Table S1. Regression models for the association between lifetime mood disorder features and cognitive performance.

	Unadjusted		Model 1[a]		Model 2[b]		Model 3[c]		Model 4i[d]		Model 4ii[e]		Model 5[f]	
	Coefficient (95% CI)	***p***	**Coefficient (95% CI)**	***p***	**Coefficient (95% CI)**	***p***	**Coefficient (95% CI)**	***p***	**Coefficient (95% CI)**	***p***	**Coefficient (95% CI)**	***p***	**Coefficient (95% CI)**	***p***
Reasoning score[g]	n = 139 140		n = 139 140		n = 137 377		n = 126 796		n = 126 000		n = 126 584		n = 125 789	
Single episode major depression	0.27 (0.23, 0.31)	<.001	0.30 (0.25, 0.34)	<.001	0.31 (0.27, 0.36)	<.001	0.31 (0.26, 0.35)	<.001	0.27 (0.23, 0.32)	<.001	0.30 (0.25, 0.35)	<.001	0.26 (0.22, 0.31)	<.001
Recurrent major depression (moderate)	0.08 (0.05, 0.12)	<.001	0.11 (0.08, 0.14)	<.001	0.16 (0.12, 0.20)	<.001	0.26 (0.22, 0.29)	<.001	0.22 (0.19, 0.26)	<.001	0.25 (0.21, 0.30)	<.001	0.22 (0.18, 0.26)	<.001
Recurrent major depression (severe)	-0.12 (-0.17, -0.07)	<.001	-0.12 (-0.17, -0.06)	<.001	0.01 (-0.04, 0.05)	.758	0.18 (0.12, 0.23)	<.001	0.09 (0.03, 0.14)	.002	0.21 (0.15, 0.27)	<.001	0.12 (0.07, 0.17)	<.001
Bipolar disorder	-0.27 -0.38, -0.17)	<.001	-0.32 (-0.41, -0.22)	<.001	-0.18 (-0.32, -0.04)	.012	0.04 (-0.08, 0.17)	.488	-0.03 (-0.13, 0.08)	.625	0.11 (-0.02, 0.23)	.087	0.04 (-0.06, 0.13)	.442
Reaction time (ms)[g]	n = 142 335		n = 142 335		n = 140 501		n = 129 242		n = 128 320		n = 129 025		n = 128 104	
Single episode major depression	-6.59 (-9.41, -3.77)	<.001	-6.21 (-8.56, -3.87)	<.001	-6.80 (-9.18, -4.42)	<.001	-6.09 (-8.94, -3.23)	<.001	-5.76 (-8.82, -2.69)	<.001	-5.78 (-8.28, -3.28)	<.001	-5.47 (-8.19, -2.75)	<.001
Recurrent major depression (moderate)	-4.96 (-7.11, -2.81)	<.001	-1.85 (-3.95, 0.25)	.084	-3.92 (-6.10, -1.73)	<.001	-6.70 (-9.00, -4.41)	<.001	-6.28 (-8.29, -4.26)	<.001	-6.64 (-8.62, -4.66)	<.001	-6.24 (-8.29, -4.20)	<.001
Recurrent major depression (severe)	4.12 (1.52, 6.72)	.002	8.48 (5.13, 11.83)	<.001	3.55 (0.39, 6.71)	.028	-0.51 (-3.08, 2.07)	.700	0.32 (-2.78, 3.41)	.842	-1.64 (-4.23, 0.96)	.216	-0.82 (-3.39, 1.75)	.530

	Unadjusted		Model 1[a]		Model 2[b]		Model 3[c]		Model 4i[d]		Model 4ii[e]		Model 5[f]	
Bipolar disorder	2.37 (-5.28, 10.02)	.543	13.93 (7.17, 20.70)	<.001	8.67 (1.91, 15.44)	.012	3.48 (-2.53, 9.49)	.257	4.06 (-0.83, 8.95)	.104	1.15 (-5.43, 7.72)	.733	1.70 (-5.11, 8.51)	.625
Numeric memory score[g]	n = 42 245[h]		n = 42 245[h]		n = 41 580[h]		n = 38 426[h]		n = 38 151[h]		n = 38 325[h]		n = 38 051[h]	
Single episode major depression	0.07 (0.01, 0.13)	.015	0.08 (0.03, 0.13)	.001	0.09 (0.03, 0.16)	.003	0.09 (0.03, 0.15)	.002	0.07 (0.02, 0.13)	.009	0.09 (0.04, 0.14)	<.001	0.07 (0.02, 0.12)	.004
Recurrent major depression (moderate)	0.00 (-0.04, 0.04)	.868	0.01 (-0.03, 0.05)	.644	0.04 (-0.00, 0.08)	.054	0.07 (0.02, 0.12)	.003	0.06 (0.02, 0.10)	.006	0.08 (0.04, 0.11)	<.001	0.06 (0.02, 0.11)	.004
Recurrent major depression (severe)	-0.08 (-0.14, -0.02)	.006	-0.09 (-0.15, -0.03)	.002	-0.03 (-0.10, 0.04)	.364	0.04 (-0.02, 0.11)	.175	0.01 (-0.04, 0.06)	.687	0.05 (-0.02, 0.12)	.126	0.02 (-0.04, 0.08)	.505
Bipolar disorder	-0.10 (-0.26, 0.07)	.267	-0.15 (-0.29, -0.01)	.035	-0.06 (-0.19, 0.07)	.354	0.04 (-0.12, 0.21)	.596	0.01 (-0.14, 0.16)	.860	0.07 (-0.08, 0.22)	.383	0.03 (-0.10, 0.16)	.606
	IRR (95% CI)	***p***	**IRR (95% CI)**	***p***	**IRR (95% CI)**	***p***	**IRR (95% CI)**	***p***	**IRR (95% CI)**	***p***	**IRR (95% CI)**	***p***	**IRR (95% CI)**	***p***
Pairs matching (errors)[i]	n = 143 828		n = 143 828		n = 141 964		n = 130 413		n = 129 446		n = 130 195		n = 129 229	
Single episode major depression	0.97 (0.95, 0.98)	<.001	0.98 (0.96, 0.99)	.008	0.97 (0.95, 0.99)	.002	0.97 (0.95, 0.99)	.004	0.97 (0.96, 0.99)	.007	0.97 (0.96, 0.99)	.005	0.98 (0.96, 0.99)	.009
Recurrent major depression (moderate)	0.98 (0.97, 0.99)	.023	1.01 (0.99, 1.02)	.452	1.00 (0.99, 1.02)	.852	0.99 (0.98, 1.01)	.259	0.99 (0.98, 1.01)	.385	0.99 (0.98, 1.01)	.301	0.99 (0.98, 1.01)	.428

	Unadjusted		Model 1[a]		Model 2[b]		Model 3[c]		Model 4i[d]		Model 4ii[e]		Model 5[f]	
Recurrent major depression (severe)	1.04 (1.02, 1.06)	<.001	1.06 (1.04, 1.08)	<.001	1.05 (1.03, 1.07)	<.001	1.03 (1.01, 1.05)	.001	1.04 (1.02, 1.06)	<.001	1.03 (1.01, 1.05)	.002	1.03 (1.02, 1.05)	<.001
Bipolar disorder	1.04 (0.99, 1.08)	.085	1.08 (1.04, 1.12)	<.001	1.07 (1.03, 1.11)	.001	1.04 (0.99, 1.08)	.076	1.04 (1.00, 1.09)	.045	1.03 (0.99, 1.07)	.148	1.04 (0.99, 1.08)	.095
	OR (95% CI)	***p***	**OR (95% CI)**	***p***	**OR (95% CI)**	***p***	**OR (95% CI)**	***p***	**OR (95% CI)**	***p***	**OR (95% CI)**	***p***	**OR (95% CI)**	***p***
Prospective memory (correct 1st attempt)[j]	n = 143 136		n = 143 136		n = 141 288		n = 129 885		n = 128 944		n = 129 667		n = 128 727	
Control group (reference)	1	-	1	-	1	-	1	-	1	-	1	-	1	-
Single episode major depression	1.39 (1.31, 1.48)	<.001	1.38 (1.30, 1.47)	<.001	1.41 (1.32, 1.50)	<.001	1.41 (1.32, 1.50)	<.001	1.38 (1.29, 1.47)	<.001	1.40 (1.31, 1.49)	<.001	1.37 (1.28, 1.47)	<.001
Recurrent major depression (moderate)	1.15 (1.10, 1.20)	<.001	1.12 (1.07, 1.17)	<.001	1.17 (1.12, 1.22)	<.001	1.27 (1.21, 1.33)	<.001	1.25 (1.19, 1.31)	<.001	1.26 (1.20, 1.32)	<.001	1.25 (1.19, 1.31)	<.001
Recurrent major depression (severe)	0.91 (0.87, 0.96)	<.001	0.88 (0.83, 0.93)	<.001	0.97 (0.92, 1.02)	.261	1.12 (1.05, 1.18)	<.001	1.08 (1.02, 1.15)	.012	1.15 (1.08, 1.22)	<.001	1.11 (1.05, 1.18)	.001
Bipolar disorder	0.84 (0.75, 0.94)	.003	0.77 (0.68, 0.86)	<.001	0.86 (0.76, 0.96)	.010	1.01 (0.89, 1.14)	.915	0.98 (0.86, 1.12)	.781	1.07 (0.94, 1.22)	.293	1.05 (0.92, 1.19)	.501

CI, confidence interval; IRR, incidence rate ratio; OR, odds ratio.
a. Adjusted for age, gender.
b. Adjusted for age, gender, current smoker, alcohol daily/almost daily, on psychotropic medication.
c. Adjusted for age, gender, current smoker, alcohol daily/almost daily, on psychotropic medication, current depressive symptoms score.
d. Adjusted for age, gender, current smoker, alcohol daily/almost daily, on psychotropic medication, current depressive symptoms score, has a degree.
e. Adjusted for age, gender, current smoker, alcohol daily/almost daily, on psychotropic medication, current depressive symptoms score, Townsend score.
f. Adjusted for age, gender, current smoker, alcohol daily/almost daily, on psychotropic medication, current depressive symptoms score, has a degree, Townsend score.

g. Linear regression with bootstrapped standard errors; omitted reference group was the control group.
h. The Numeric Memory task was removed from the baseline battery before recruitment ended, yielding smaller possible sample sizes than for the other cognitive measures.
i. Negative binomial regression; omitted reference group was the control group.
j. Logistic regression.