Table A.1: Hierarchical exclusion criteria

Order	Exclusion criteria	Explanation
1	Publication type	Studies excluded under this category: dissertations, theses, reviews, case studies, discussion articles, summaries, theoretical and policy papers.
2	Design	Studies using qualitative design were excluded
3	Diagnosis	Studies were excluded if they defined remission criteria or recovery from other mental health problems, substance misuse, addiction or eating disorders. Additional explanation for full-text screening: BD was not verified based on DSM or ICD criteria.
4	Recovery	Studies were excluded if solely focused on clinical recovery through symptoms remission and relapse prevention. Additional explanation for full-text screening: Recovery (other than clinical or symptomatic) definition was provided and operationalised as an outcome measure
5	Age	All participants must have been 16 years old or older at the time of inclusion. Additional explanation for full-text screening: No minimum age reported (unless directly referenced to primary source which provides this data)
6	Availability	No English abstract available. Additional explanation for full-text screening: No English full-text available
7	Prediction	Full-text screening only: studies were excluded if they did not investigate any predictors of recovery (including prevalence studies and papers comparing recovery across mental health diagnoses).

Table A.2 Study characteristics, methods and analysis

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
P1	Cross-sectional study	**Demographic factors:** age, gender, education, and employment status **Clinical factors:** years since diagnosis, medication, depressive and manic symptoms, recent hypomania and depression relevant experiences **Psychosocial factors:** positive and negative self-dispositional appraisals, normalising appraisals of depression and mania relevant experiences, illness perception, and positive and negative appraisals of internal states	1) Exploring bivariate associations between demographic and clinical factors and personal recovery: independent t-tests were used to check if recovery differed by categorical variables; Pearson's correlation was used to test associations between recovery and continuous variables. 2) Pearson's correlations were conducted to test associations between personal recovery and psychosocial factors. 3) To control for potential confounding effects, demographic and clinical variables that showed significant bivariate associations with recovery (step one) and psychosocial factors (step two) were included in a hierarchical multiple regression analysis with personal recovery as dependent variable. 4) Statistical significance set as $p < .05$. To control for Type I error while not compromising power and reducing the likelihood of Type II error, the sequential Holm-Bonferroni correction was applied for each set of hypothesis-driven tests.	Significant negative associations: • **Depressive symptoms** showed negative bivariate association with personal recovery (r = -.62, $p < .05$) and remained significant in the hierarchical multiple regression models- Step 1: [standardised β = -.77, β = -26.91, *95% CI* (-34.06; -19.78), *SE* = 3.59, $p < .001$] and Step 2: [standardised β = -.53, β = -18.45, *95% CI* (-25.66; -11.24), *SE* = 3.62, $p < .001$] • **Negative illness model** showed negative bivariate association with personal recovery (r = -.59, $p < .05$) and remained significant in the hierarchical multiple regression model- Step 2: [standardised β = -.38, β = -13.49, *95% CI* (-20.08; -6.89), *SE* = 3.31, $p < .001$]. • **Recent depression relevant experiences** (r = -.32, $p < .05$) showed negative bivariate association with recovery. • **Negative self-dispositional appraisals** (r = -.39, $p < .05$), **endorsement of positive and negative appraisals of internal states** (r = -.44, $p < .05$) showed negative correlations with personal recovery (but not significant in multiple regression, see below). • BIPQ dimensions representing negative beliefs about mood swings were negatively correlated with recovery, including **consequence** (r = -.39, p <.001); **identity** (r = -.31, p = .004); **emotional response- concern** (r = -.53; $p < .001$); **emotional response- emotion** (r = -.35; p = .001) and **self-blame** (r = -.27; p = .011) No association:

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				• **Age** *(r* = -.04) and **years since diagnosis** (*r* = .18) did not show any significant association with personal recovery. • Participants' recovery did not differ significantly by **gender** (*t* = -0.5, ns) or current **medication** use (*t* = 0.98, ns), or across **education** level (*F* = 1.1, ns). • **Manic symptoms** (*r* = .12, ns) and **recent mania relevant experiences** (*r* = .05, ns) were not associated with personal recovery. • Recovery did not show association with **positive self-dispositional appraisals** (*r* = -.13). • **Normalising appraisals of hypomania** relevant experiences did not show significant associations with personal recovery when examined in the hierarchical multiple regression model- Step 2: [standardised β = .09, β = 6.07, *95% CI* (-5.95; 18.09), *SE* = 6.04, *p* = .318]. • **Normalising and self-dispositional appraisals of depression** relevant experiences did not show significant associations with personal recovery when examined in the hierarchical multiple regression. Normalising appraisals- Step 2: [standardised β = .04, β = 2.68, *95% CI* (-9.73; 15.09), *SE* = 6.23, *p* = .669]; Negative self-dispositional appraisals- Step 2: [standardised β = -.10, β = -5.51, *95% CI* (-15.35; 4.34), *SE* = 4.95, *p* = .269]. • **Endorsement of positive and negative appraisals of internal states** did not show significant associations with personal recovery when examined in the hierarchical multiple regression model- Step 2: [standardised β = -.05, β = 1.60, *95% CI* (-6.40; 3.22), *SE* = 2.42, *p* = .512]. • Some of the BIPQ dimensions were not associated with personal recovery, including **timeline** (*r* = -.17; *p* = .119),

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				illness comprehensibility (r = .20; p = .071) and **cause internal** (r = -.08, p = .448). Significant positive associations: • **Employed** participants had significantly higher personal recovery compered to unemployed participants [M = 2381.0 (SD = 428.3) vs M = 2076.7 (SD = 514.9), t = -2.98, p < .05] and employment status showed positive association with personal recovery in the hierarchical multiple regression models- Step 1: [standardised β = .34, β = 335.93, *95% CI* (185.31; 486.56), SE = 75.73, p < .001] ; Step 2: Step 2: [standardised β = .39, β = 386.44, *95% CI* (252.41; 520.47), SE = 67.32, p < .001]. • Recovery showed positive bivariate associations with **normalising appraisals** of both depression (r = .28, p < .05) and mania (r = .25, p < .05) relevant experiences. • **Recent depression relevant experiences** showed positive association with personal recovery in the hierarchical multiple regression models- Step 1: [standardised β = .21, β = 29.80, *95% CI* (0.34; 59.27), SE = 14.81, p = .05] ; Step 2: [standardised β = .30, β = 42.47, *95% CI* (15.63; 69.31), SE = 13.48, p = .002] • BIPQ dimensions representing positive beliefs about mood swings were positively correlated with recovery, including **personal control (**r = .38; p < .001**), treatment control (**r = .23; p = .036), and **personal effort (**r = .25; p = .019**).**
P2	Prospective study with 6 month FU period	**Demographic factors:** age and gender **Clinical factors:** manic (activation) and depressive	1) Correlations to explore bivariate associations. 2) T-tests to compare mean scores between time point 1 and 2.	Significant negative associations: • Personal recovery showed negative cross-sectional correlations with **depressive symptoms** (r = -.56, $p \leq .01$), **perceived conflict** (psychopathology; r = -.29, $p \leq .01$) and

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		symptoms, well-being and perceived conflict in internal states; time since last episode **Psychosocial factors:** resilience domains (self-management of BD, turning point, self-care, self-confidence, and interpersonal support), psychosocial functioning (functional impairment) and quality of life	3) Path analysis (multivariate approach) to test associations between the clinical and psychosocial variables and personal recovery. Due to the lack of adequate distribution of the variables, Satorra-Bentler maximum likelihood method was used. 4) The goodness of fit of the model was evaluated using the CFI and NNFI, the RMSEA and the SRMR. 5) All missing values were imputed using the expectation-maximization imputation algorithm with SPSS.	**psychosocial functional impairment** (r = -.65, $p \leq .01$) at BL. • BL personal recovery showed significant negative associations with FU scores on **depression** (r =- .31, $p \leq .05$); **perceived conflict** (psychopathology; r = -.37, p ≤ .01) and **psychosocial functional impairment** (r = -.39, p ≤ .01). • FU personal recovery showed significant negative associations with BL scores on **depression** (r =- .33, $p \leq .05$) and **psychosocial functional impairment** (r = -.50, p ≤ .01). No association: • BL personal recovery did not show significant cross-sectional association with **manic symptoms** (activation; r = -.17, *ns*) nor with the FU score on the **activation** subscale (r = -.12, ns). • BL personal recovery was not associated with the FU scores on the **turning point** (r = -.14, ns) and **interpersonal support** (r = .18, ns) resilience subscales. • FU personal recovery was not associated with the BL scores on the **turning point** (r = .02, ns) and **interpersonal support** (r = .15, ns) resilience subscales and not with BL **activation** scores (r = -.10, ns) or BL **perceived conflict**/psychopathology scores (r = -.21, ns). • **Age, gender and time since last episode** were added to the FU multivariate path analysis, but did not show any significance and were subsequently removed (no statistics reported). • **Turning point and interpersonal support** did not remain significant in the multivariate baseline or follow-up path analyses. **Self-management** did not remain significant in the multivariate follow-up path analysis.

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				Significant positive associations: • BL personal recovery showed positive cross-sectional correlation with the total score of **resilience** (r = .66, $p \leq .01$) and its subscales: **self-management** (r = .68, $p \leq .01$), **turning-point** (r = .2, $p \leq .05$), **self-care** (r = .62, $p \leq .01$), **self-confidence** (r = .58, $p \leq .01$) and **interpersonal support** (r = .39, $p \leq .01$); and with **quality of life** (r = .72, $p \leq .01$), and **wellbeing** (r = .65, $p \leq .01$). • **BL personal recovery** showed significant positive correlations with the **FU resilience** total score (r = .32, $p \leq .05$),as well as some of the FU resilience subscale scores: **self-management** (r = .39, $p \leq .01$), **self-care** (r = .33, $p \leq .01$), and **self-confidence** (r = .51, $p \leq .01$); with FU **personal recovery** (r = .58, $p \leq .01$), with FU **quality of life** (r = .38, $p \leq .01$), and FU **wellbeing** (r = .29, $p \leq .05$)/ • FU personal recovery showed significant positive associations with BL **resilience** (r = .45, $p \leq .01$) and with some of the BL resilience subscales: **self-management** (r = .47, $p \leq .01$), **self-care** (r = .53, $p \leq .01$) and **self-confidence** (r = .30, $p \leq .05$), as well as with BL **quality of life** (r = .46, $p \leq .01$), and BL **wellbeing** (r = .46, $p \leq .01$). • Some of the resilience factors remained significant in the cross-sectional multivariate path analysis: **self-management** (r = .39, $p < .001$), **self-care** (r = .20, $p < .05$), and **self-confidence** (r = .25, $p < .05$). • FU multivariate path analysis showed that BL **self-confidence** predicted a significant increase (direct effect) in FU personal recovery (*unstandardized coefficient* = 0.42, *SE* = 0.06, $p < .001$), as well as BL personal recovery did in FU personal recovery (*unstandardized coefficient* = 0.23, *SE* = 0.08, $p < .05$). An indirect effect of BL **self-care** on FU

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				personal recovery via **self-confidence** was identified (z = 3.47, p < .001). (Please note that the authors also indicate an indirect effect of BL interpersonal support on FU personal recovery as significant, however the reported p-value is .15, therefore we do not consider this significant).
P3	Prospective study with 6 month FU period (personal recovery related results are **cross-sectional only**)	**Psychosocial factors:** resilience (subscales: self-management, turning point, self-care, self-confidence, interpersonal support)	Data analysis relevant to personal recovery: 1) **Concurrent validation of the Resilience Questionnaire for BD (RBD):** correlation with personal recovery. 2) **Known-group validity of the RBD**: • BD participants were grouped based on the personal recovery scores: above the percentile 75 (BRQ ≥ 277) were labelled as "recovered", while patients with scores below this percentile were labelled as "not recovered". • Cohen´s d values were calculated to indicate the magnitude of the differences between means of each group on the RBD. • Analysis of variance (ANOVA) with Tukey post hoc test for multiple comparisons and nonparametric Welch post	Significant negative association: None reported No association: • **Turning point** resilience subscale did not show significant bivariate correlation with personal recovery (no statistics reported) and individuals in the recovered group did not have significantly different score on this subscale compared to the not recovered group and control groups: [Not recovered: M = 19.77 (SD = 3.81); recovered: M = 20.32 (SD = 4.82); CG: M = 19.11, SD = 3.61) Significant positive associations: • All of the **resilience factors/subscales** (except turning point) showed positive bivariate associations with personal recovery (.23 < r >.74, p < .05). • Recovered group scored significantly higher on the RBD **resilience** total score and factors/subscales (except turning point) [M = 100.07 (SD = 10.16) vs. M = 85.33 (SD = 13.58), p < .001]. Post hoc Tukey analysis revealed that most of the differences were between not recovered and recovered patients, and between not recovered and the CG. Cohen´s d ranged from -0.71 for the interpersonal support factor to -1.42 for the self-care factor. • Resilience standardised mean scores on both **resilience** measures were higher in the recovered BD group than the not recovered or CG. Mean standardized scores (and SD) in the

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
			hoc test (as applicable) were used to compare scores on resilience measures (RS-25 and RBD) of the recovered, not recovered, and general population CG. • Items were adapted for the general population group so that references to BD were substituted by 'the personal problem'. • To facilitate interpretation of the comparison between these two measures, RBD and RS-25 scores were standardized to range from 0 to 100 (exclusively for the ANOVA).	global indexes of the RBD versus RS-25 for recovered patients were 83.77 (11.04) versus 85.62 (9.36), respectively; for the comparison group from the general population, they were 75.72 (11.23) versus 79.80 (13.74), respectively; and for not recovered patients, they were 67.74 (14.76) versus 61.83 (17.72), respectively, and all these mean group differences were statistically significant ($p < .001$).
P4	Cross-sectional study	**Psychosocial factors:** functional impairment	Data analysis relevant to personal recovery: 1) Kolmogorov-Smirnov Lilliefors test was used to explore the normality of distributions of the measures for the BD group. 2) Concurrent validity of the WSAS was measured through correlations. As BD data from the WSAS did not follow the normal distribution ($p = 0.02$),	Significant negative associations: None reported. No association: None reported. Significant positive associations: • Non-parametric correlations showed that the deficit in **psychosocial functioning** correlated negatively personal recovery (*rho* = 0.61; $p < .001$).

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
			researchers performed correlations using the non-parametric coefficient of Spearman's *rho* between the WSAS and BRQ.	
P5	Prospective study with 6 month FU period (relevant outcome assessments at BL only-**cross-sectional data**)	**1) SFR components:** Quality of life, including psychological wellbeing, self-esteem, family relationships, relationship with friends, resilience, physical well-being, autonomy	1) Pearson's correlation coefficients (*r*) were used to investigate the relationships between quality of life and recovery measures.	Significant negative associations: None reported. No association: • The following subscales of quality of life were not significantly associated with the recovery subscales: **family relationships** (with personal confidence and hope: $r = .13$, ns.; with willingness to ask for help: $r = .13$, ns; with goal and success orientation: $r = .05$, ns; with no domination by symptoms: $r = .13$, ns.), **relationships with friends** (with goal and success orientation: $r = .10$, ns; with no domination by symptoms: $r = .08$, ns.). Significant positive associations: • The following subscales/total score of quality of life showed significant positive associations with the recovery subscales: **psychological wellbeing** (with personal confidence and hope: $r = .35$, $p < .01$; with willingness to ask for help: $r = .29$, $p < .01$; with goal and success orientation: $r = .15$, $p < .05$; with reliance on others: $r = .22$, $p < .01$; with no domination by symptoms: $r = .23$, $p < .01$), **self-esteem** (with personal confidence and hope: $r = .57$, $p < .01$; with willingness to ask for help: $r = .38$, $p < .01$; with goal and success orientation: $r = .26$, $p < .01$; with reliance on others: $r = .34$, $p < .01$; with no domination by symptoms: $r = .46$, $p < .01$), **family relationships** (with reliance on others: $r = .33$, $p < .01$), relationship with friends (with personal confidence and hope: $r = .19$, $p < .01$; with willingness to ask for help: $r = .19$, $p < .01$; with reliance on others: $r = .45$, $p < .01$;),

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				resilience (with personal confidence and hope*: r* = .47, *p* < .01; with willingness to ask for help*: r* = .36, *p* < .01; with goal and success orientation: *r* = .43, *p* < .01; with reliance on others: *r* = .20, *p* < .01; with no domination by symptoms: *r* = .30, *p* < .01), **physical well-being** (with personal confidence and hope: *r* = .44, *p* < .01; with willingness to ask for help: *r* = .15, *p* < .05; with goal and success orientation: *r* = .26, *p* < .01; with reliance on others: *r* = .26, *p* < .01; with no domination by symptoms: *r* = .39, *p* < .01), **autonomy** (with personal confidence and hope: *r* = .46, *p* < .01; with willingness to ask for help: *r* = .31, *p* < .01; with goal and success orientation: *r* = .25, *p* < .01; with reliance on others: *r* = .35, *p* < .01; with no domination by symptoms*: r* = .23, *p* < .01), **sentimental life** (with personal confidence and hope: r = .30, *p* < .01; with willingness to ask for help: *r* = .16, *p* < .05; with goal and success orientation: *r* = .16, *p* < .05; with reliance on others: *r* = .24, *p* < .01; with no domination by symptoms: *r* = .24, p<.01), **total score** (with personal confidence and hope: *r* = .57, *p* < .01; with willingness to ask for help: *r* = .39, *p* < .01; with goal and success orientation: *r* = .31, *p* < .01; with reliance on others: *r* = .48, *p* < .01; with no domination by symptoms: *r* = 0.42, *p* < .01).
P6	Cross-sectional study	**1) Demographic factors:** age, gender, marital status, education, employment status, religion, family type (nuclear/extended), locality (rural/urban). **2) Clinical factors:** Age of onset, illness duration, remission duration,	1) Associations were studied by using Pearson's correlation coefficient and Spearman rank correlations. 2) Comparisons were done by using t-test, Chi-square test, and FET. 3) Significance was set at two-tailed values at 0.05.	Significant negative associations: • Higher levels of residual **depressive symptoms** were associated with significantly lower level of recovery in all the domains of recovery: **original 5 recovery factors** (personal confidence and hope: *r* = -.256, *p* < .001; willingness to ask help: *r* = -.274, *p* < .001; goal and success orientation: *r* = -.197, *p* = .007; reliance on others: *r* = -.247, *p* = .001; no domination of symptoms: *r* = -.215, *p* = .003) **current study recovery factors** (defeated/overcome the illness: *r* = -.231, *p* = .002; personal confidence and hope: *r* = -.251, *p* = .001;

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		number of episodes (total), number of hospital appointments in last 3 months, depressive and manic symptoms.		seeking and relying on social support: r = -.269, p < .001; awareness and control over the illness: r = -.241, p = .001; goal and success orientation: r = -.227, p = .002). No association: • Participants on **paid jobs** did not differ in other domains of recovery (statistics not reported). None of the other **demographic or clinical factors** were associated with recovery (statistics not reported). Significant positive associations: • Participants, who were on **paid jobs** reported higher level of recovery in the domain of "willingness to ask for help" (t = 2.08; p =.039).
P7	Cross-sectional study	**1) Demographic factors:** age, gender, marital status, education, employment status, religion, income of the patient, family type (nuclear/extended), locality (rural/urban). **2) Clinical factors:** Age of onset, illness duration, remission duration, number of episodes (total), number of hospitalisations (lifetime and in past 6 months), depressive and manic symptoms. **3) Psychosocial factors:** internalized stigma	1) Comparisons using t-test. 2) Correlations were studied using Pearson's correlation coefficient. 3) Multiple regression analysis was used to study the predictors of recovery.	Significant negative associations: • **Depressive symptoms** correlated negatively with all of the recovery domains and (remained significant predictor in the regression model): with the **original 5 recovery factors** (personal confidence: r = -.326, $p \leq .001$; willingness to ask help: r = -.353, $p \leq$ .001; goal orientation: r = -.256, $p \leq$ .001; reliance on others: r = -.306, $p \leq .001$; not dominated by symptoms: r = -.385, $p \leq .001$; Total score: r = -.325, $p \leq$ .001) and the **current study recovery factors** (defeated/overcome the illness: r = -.231, p < .01; personal confidence and hope: r = -.251, $p \leq .001$; seeking and relying on social support: r = -.269, $p \leq .001$; awareness and control over the illness: r = -.241, $p \leq .001$; goal and success orientation: r = -.227, p < .01, total score: r = -.341, $p \leq$ .001). • **Internalised stigma** was negatively associated with each domain of recovery (total score without stigma resistance reported here, subscale associations also presented in the paper): with the **original 5 recovery factors** (personal confidence: r = -.593, $p \leq .001$; willingness to ask help: r = -

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		(alienation, stereotype endorsement, perceived discrimination, social withdrawal, and stigma resistance), religious coping (positive and negative), religiosity (organisational, non-organisational, intrinsic), religiousness (involvement, influence, hope). **4) SFR components:** global functioning		.491, $p \leq .001$; goal orientation: $r = -.462$, $p \leq .001$; reliance on others: $r = -.504$, $p \leq .001$; not dominated by symptoms: $r = -.551$, $p \leq .001$; Total score: $r = -.576$, $p \leq .001$) and the **current study recovery factors** (defeated/overcome the illness: $r = -.602$, $p \leq .001$; personal confidence and hope*:* $r = -.566$, $p \leq .001$; seeking and relying on social support: $r = -.524$, $p \leq .001$; awareness and control over the illness: $r = -.557$, $p \leq .001$; goal and success orientation: $r = -.506$, $p \leq .001$; total score: $r = -.581$, $p \leq .001$). Subscales of discrimination experience, stereotype endorsement and alienation remained significant in the regression model. • The **absence of stigma** in all the domains was associated with significantly higher recovery [recovery total (24 items): $t = 6.598$, $p < .001$; recovery total (41 items): $t = 6.593$, $p < .001$]. No association: • There was no significant correlation between recovery scores and **age, gender, education, marital status, family type and locality** (statistics not reported). • There was no association between recovery and **manic symptoms, number of episodes, age of onset, illness or remission duration and number of hospitalizations** (statistics not reported). • **Employment status and income** did not correlate with other domains of recovery (statistics not reported). • **Positive religious coping and religiosity** did not correlate with other domains of recovery (no statistics reported)- original recovery factors: willingness to ask for help; goal orientation; not dominated by symptoms; and current study recovery factors: goal and success orientation • **Religiousness and negative religious coping** did not correlate with recovery (no statistics reported). • **Non-organisational/private religiosity** did not correlate with the following recovery factors: Original recovery factors: personal confidence; willingness to ask for help;

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				reliance on others; not dominated by symptoms; recovery total score. Current study recovery factors: defeated/overcome illness; personal confidence & hope; seeking and relying on social support; awareness and control over illness; total score (stats not reported). Significant positive associations: • Those who were on **paid employment** experienced high level of recovery in the domain of 'willingness to ask for help' (t = -2.079, p < .05). • Participants who **were earning more** reported higher level of recovery in the domain of 'goal orientation' (t = -2.225, p < .05) and 'not dominated by symptoms' (t = -2.387, p < .05). • **Functioning** correlated positively with all the recovery domains (and remained significant in the regression model); with the **original 5 recovery factors** (personal confidence*: r* = .450, $p \leq .001$; willingness to ask help: r = .445, $p \leq .001$; goal orientation: r = .435, $p \leq .001$; reliance on others: r = .480, $p \leq .001$; not dominated by symptoms: r = .426, $p \leq$ .001; total score: r = .484, $p \leq .001$) and the **current study recovery factors** (defeated/overcome the illness: r = .440, $p \leq .001$; personal confidence and hope: r = .497, $p \leq .001$; seeking and relying on social support: r =.497, $p \leq .001$; awareness and control over the illness: r = .450, $p \leq .001$; goal and success orientation: r = .483, p < .01, total score: r = .497, $p \leq .001$). • **Stigma resistance** (reverse coded) was positively associated all domains of recovery: with **the original 5 recovery factors** (personal confidence: r = -.259, $p \leq .001$; willingness to ask help: r = -.351, $p \leq .001$; goal orientation: r = -.171, $p \leq .001$; reliance on others: r = -.277, $p \leq .001$; not dominated by symptoms: r = -.286, $p \leq .001$; total score: r = -.282, $p \leq$ .001) and the **current study recovery factors** (defeated/overcome the illness: r = -.287, $p \leq .001$; personal confidence and hope: r = -.239, $p \leq .001$; seeking and relying on social support: r = -.329, $p \leq .001$; awareness and control

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				over the illness: r = -.339, $p \leq .001$; goal and success orientation: r = -.218, $p \leq .001$; total score: r = -.299, $p \leq .001$). • **Positive religious coping** showed positive associations with some of the recovery domains: from the **original 5 recovery factors** (personal confidence: r = .203, $p < .01$; reliance on others: r = .169, $p < .05$; total score: r = .172, $p < .05$) and from the **current study recovery factors** (defeated/overcome the illness: r = .165, $p < .05$; personal confidence and hope: r = .162, $p < .05$; seeking and relying on social support: r = .158, $p < .05$; awareness and control over the illness*:* r = .184, $p < .05$; total score: r =.168, $p < .05$). • **Non-organisational religiosity** showed positive association with some domains of recovery (goal orientation r = .144, $p <.05$ and goal and success orientation r = .149, $p < .05$).
P8	Cross-sectional study	**1) Clinical factors:** Observer rated and self-reported depressive and manic symptom. **2) Psychosocial factors:** Observer rated and self-reported growth **3) SFR component:** Observer rated and self-reported functioning	1) Cross-sectional analysis (correlation) of relationships between recovery scores and the self-reported and observer rated measures. 2) To more rigorously assess the unique associations between measures of symptoms and function and BRQ scores those measures which were significantly associated with BRQ were entered together into a series of regression analyses to explore the variance accounted for by each, one exploring the variance explained by symptom measures and a second, exploring the variance explained by measures of growth and functioning.	Significant negative associations: • Recovery was negatively associated with both observer rated (**depressive symptomatology**: r = -.495, $p < .01$, depressive mood item separately: r = -.456, $p < .01$) and self-reported depression (r = -.665, $p < .01$); with observer rated specific elevated mood items (r = -.304, $p < .05$) and with **bipolar symptomatology** (internal states; activation: r = -.289, $p < .05$, depression: r = -.459, $p < .01$, perceived conflict: r = -.448, $p < .01$). • **Self-reported depression** remained significant and predicted recovery in the regression model including symptom measures (standardised β = -.503, t = -3.096, $p < .01$) and in the regression model including both symptom and other measures (standardised β = -.401, t = -3.097, $p <.001$). No association: • Recovery was not associated with **manic symptoms** total score (observer rated; r = -.144, ns.) and **physical**

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
			Significant predictors from these initial analyses were then entered into a final regression analysis to explore the specific measures that uniquely predicted recovery.	**functioning** (r = .058, ns.). **Observer or self-report manic symptoms** did not remain in the regression analyses, and overall and mental functioning did not remain in the combined regression adjusting for symptom and other measures (statistics not reported). Significant positive associations: • Recovery was positively associated with **post-traumatic growth** (r = .591, p < .01), with **overall functioning** (r = .489, p < .01), with self-reported **well-being** measures (positive well-being: r = .549 and **internal state/symptomatic well-being**: r = .525, p < .01) and **mental functioning** (r = .561, p < .01). • **Internal state/symptomatic well-being** (standardised β = .423, t = 3.234, p < .01) remained significant and predicted recovery in the regression model including symptom measures • **Overall functioning** (standardised β = .221, t = 2.028, p < .047) **post-traumatic growth** (standardised β = .448, t = 4.708, p < .001) and **mental functioning** (standardised β = .310, t = 2.805, p < .005) remained significant and predicted recovery in the regression model including functioning and growth measures only. • **Post traumatic growth** (standardised β = .363, t = 4.114, p < .001) and **well-being** (symptomatic/internal state; standardised β = .199, t = 2.173, p < .05) remained significant and predicted recovery in the regression model including both symptom and other measures. • PTGI items independently were also positively correlated with BRQ total score (data not extracted, as PTGI items have not been validated at item level).
P9	Pilot Randomised Control Trial (6	**1) Psychosocial factors:** Recovery focused cognitive-behavioural	1) All therapy effects were estimated using a random-effects (random intercepts) model, assuming that the	Significant negative associations: None reported. No association: None reported

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
	and 12 months FU assessment)	therapy (EG: therapy, CG: TAU)	effects were the same for each FU time (having first checked that there was no significant therapy by FU time interaction). 2) The BL value of the relevant outcome measure was used as a covariate. The intention-to-treat principle was followed throughout. 3) Missing data were assumed to be missing at random (ignorable) and automatically allowed for in fitting the random-effects or analysis of covariance models.	Significant positive associations: • The recovery score was higher in the **recovery-focused CBT** group at FU than the TAU group [310.87, *95% CI* 75.00–546.74 (*SE* = 120.34), p = .010, d = 0.62] with no interaction between this effect and FU assessment point (6 or 12 month).
P10	Cross-sectional study	**1) Demographic factors:** gender, age, education, marital status, number of children, employment status, religion, family monthly income. **2) Clinical factors:** age of onset, number of life time hospitalisation, longest hospitalisation, lifetime alcohol and substance use, lifetime binge drinking, manic and depressive symptoms. **3) PR components:** recovery elements	1) To explore the four stages of recovery (operationalized as the four ranges of the total score on the SRS) bivariate analyses were used, including cross-tabulations, FET and ANOVA. For ANOVA, Bonferroni test (equal variances assumed) and Tamhane's T2 (equal variances not assumed) were also conducted as post hoc analyses. 2) Decision tree analysis (also known as recursive partition analysis) was conducted to identify the variables associated with each of the four stages of recovery. In each split of the decision tree, the classification accuracy of the partition is indicated by the G^2 and LogWorth statistics which	Significant negative associations: • In participants under **age** 45 an earlier **age of onset** (under 22) was associated with more advanced recovery (G^2 = 43.22, *LogWorth* = 1.14). No association: • There were no significant demographic differences across the four stages of recovery using bivariate analyses: **gender** (*FET*, p = .247), **age** (F = 1.348, ns.), **education** (*FET*, p = .524), **marital status** (*FET*, p = .082), **number of children** (F = 0.667, ns.), **employment status** (*FET*, p = .072), **religion** (*FET*, p = .971), **family monthly income** ($\chi 2$, p = .293) • There were no significant clinical differences across the four stages of recovery using bivariate analyses: **age of onset** (*Welch's ANOVA* = 0.517, ns), number of lifetime **hospitalisations** (F = 0.534, ns.) and longest hospitalisation (F = 0.551, ns.), life time **binge drinking** (*FET*, p = .407),

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		(respect, hope, self-directed empowerment, meaningful role, asset and strength base, social role), organisational climate (recovery-enhancing environment)	are analogous to the fitness index (least residual) in a regression. To avoid overfitting the model and to validate the results, the decision tree analysis software randomly selected a percentage of the sample as a training set (72%) and the remainder as a validation set (28%). Separate analyses were run on these subsets. Adjusted receiver operating characteristic (ROC) curves depicting the rates of correct classification and misclassification were calculated on this final model.	life time **substance use** (*FET*, $p = 1.00$), depressive ($F = 1.129$, ns.) and manic ($F=1.852$, ns.) **symptoms.** • There were no other significant differences across the four stages of recovery using bivariate analyses: asset and strength-based recovery element ($F = 2.086$, ns.), social role recovery element ($F = 2.636$, ns.), recovery enhancing environment ($F = 1.789$, ns.) Significant positive associations: • **Respect, hope and self-directed empowerment** ($F = 6.720$, $p < .001$) and **meaningful role** ($F = 3.658$, $p < .05$) recovery elements were more important to individuals in more advanced stages of recovery in bivariate analyses. The former was the strongest differentiator of recovery stages ($G^2 = 113.99$, LogW*orth* = 1.56); the latter was important in differentiating recovery in individuals with later age of onset ($G^2 = 20.59$, *LogWorth* = 0.66) in decision tree analysis. • **Age** was the second differentiator in decision tree (participants over 45 were more likely to be in more advanced recovery ($G^2 = 43.22$, *LogWorth* = 1.14). • In participants over 45 life time **binge drinking** was associated with better recovery ($G^2 = 26.40$, *LogWorth* =1.19)
S1	Prospective cohort study (1 year post hospitalisation FU period- data collection 5 times, every 10 weeks)	**1) Demographic factors**: age, gender, ethnicity. **2) Clinical factors:** substance use	1) General linear mixed-effects models using restricted maximum likelihood estimation were constructed predicting functioning measures from time and time-varying substance use variables. 2) Diagnostic differences in these relationships were also investigated by examining diagnosis by substance use interactions.	Significant negative associations: • Interaction between gender and diagnostic groups with regard to alcohol use and functional recovery: **men with BD who used alcohol** exhibited poor functioning compared to women: $F(2, 2872) = 5.64$, $p = .004$. No association: • No interaction effect between **cannabis and gender** on functional recovery in bipolar subsample.

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
			3) Exploratory analyses were conducted to examine the degree to which gender moderated these relationships. 4) All conditional growth models included age, race and gender, as well as initial levels of the outcome variable that was under study (e.g. BL functioning).	• No associations reported between **age, ethnicity** and recovery. Significant positive associations: None reported.
S2	Prospective cohort study (6 months FU period-data collection 6 times, monthly for outcome)	**1) Clinical factors:** substance abuse, treatment/medication adherence (prescribed medications included-valproate and lithium).	The cumulative probabilities of outcomes between adherence/non-adherence, substance abuse/no substance abuse compared using log-rank test at a significance level of $p < .05$.	Significant negative associations: • **Substance abuse** associated with longer time to functional recovery based on LIFE-RIFT (log rank: $\chi = 4.36$, $p = .037$). No association: • **Treatment adherence and substance use** was not associated with recovery based on GAF scores. Significant positive associations: • Full **treatment adherence** shortened time to functional recovery based on LIFE-RIFT (log rank: $\chi = 4.5$, $df = 1$, $p = .03$).
S3	Prospective cohort study (6 months FU period: 3 assessments (BL, 1 month and 6 months)	**1) Demographic factors**: age, gender, marital status and employment status. **2) Clinical factors:** Family psychiatric history, psychiatric comorbidity, polarity of first episode, lifetime psychotic	1) Preliminary Pearson bivariate correlation between predictors and outcome at 6 month. 2) Bivariate association with qualitative variables explored using Mann-Whitney U test. 3) All associated (showed at least trend) variables from preliminary	Significant negative associations: • Recovered participants were **younger** ($p = .03$), had **lower BMI** ($p = .005$), had **fewer** number of total **episodes** ($p = .02$), shorter illness duration (chronicity) ($p < .001$) compared to non-recovered participants. • Correlation results: **age** ($r = .21$; $p = .01$), **years of illness** ($r = .22$; $p = .006$); total number of **episodes** ($r = .19$; $p = .02$), number of **depressive episodes** ($r = .24$; $p = .005$), number of days of **hospitalisation** between BL and at 6 month FU ($r = .26$; $p = .004$).

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		symptom, rapid cycling, age of onset, number and type of episodes, number of suicide attempts, number of hospital admissions, cannabis consumption, hours of sleep at BL.	analysis and literature underwent stepwise multiple regression.	• Best regression model (Adjusted R^2 = .22; *df* = 6, *F* = 3.95; *p* = .002) included 5 variables, 3 were significant: number of previous **depressive episodes** (β = 3.25; *t* = 3.23; *p* =.002), presence of **psychotic symptoms** during index episode (β = 7.007; *t* = 2.2; *p* = .031) and **BMI** (β = 0.62; *t* = 2.09; *p* = .041) No association: • Recovered participants did not differ significantly in **age at onset** (*p* = .47), presence of **psychotic symptoms** during the index manic episode (*p* = .26) or days of **hospitalisation** (*p* = .39), no difference reported in **gender, marital status** or **employment status** (no statistic reported). • Mann-Whitney U-test results: **psychiatric comorbidity** (*p* =.26), presence of **mixed symptoms** (*p* =.15), **family history of affective disorders** (*p* =.61), previous **suicide** attempts (*p* =.42), **cannabis** consumption at BL (*p* =.31), presence of **psychotic symptoms** during index episode (*p* =.059) were not associated with recovery. • Regression model analysis: **number of days hospitalised** between BL and FU1 (β = -0.133; *t* = -0.75; *p* = .45), **years of illness** (β = -0.16; *t* =-0.92; *p* = .45) and **hours of sleep** at BL (β = -1.12; *t* =-1.31; *p* =.194) • No analytic statistics reported for: **rapid cycling, number of manic episodes, lifetime psychotic symptoms, and polarity of first episode.** Significant positive associations: None reported.
S4	Cross-sectional study	**Demographic factors:** age, gender and years of education **Clinical factors:** diagnosis subtype, number and type of episodes,	1) Global FAST score was calculated and used to categorise participants as functionally remitted or impaired. 2) Descriptive analyses of the two groups were performed using Chi-square tests for categorical variables	Significant negative associations: • Functionally impaired group presented higher **depressive symptoms** [*M* = 5.1 (*SD* = 2.9) vs. *M* = 2.1 (*SD* = 2.2) *t* = -11.11, *p* < .001], more **depressive episodes** [*M* = 6.7 (*SD* = 10.8) vs. *M* = 4.6 (*SD* = 8.3) *t* = -2.17, *p* = .03], worse **chronicity**- years of illness [*M* = 17.5 (*SD* = 11.1) vs. *M* =

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		chronicity (illness duration in years), number of hospitalizations, history of psychosis, history of rapid cycling and family affective psychiatric history, depressive and manic symptoms **Psychosocial factors:** estimated IQ, processing speed, working memory, verbal memory, executive functions, visual learning and memory, and attention	and t-tests for continuous variables (using z scores for psychosocial factors). 3) Multivariate logistic regression model was performed. Logistic regression was used to estimate the effects of the risk factors associated with functional impairment. Variables were selected for inclusion in logistic regression when significance at $p < .05$ in the univariate analysis was met. All analyses were two-tailed with alpha set at $p < .05$.	14.2 (*SD* = 10.3) *t* = -3.1, *p* = .02], and more previous **hospitalizations** [*M* = 2.1 (*SD* = 2.4) vs. *M* = 1.6 (*SD* = 1.7) *t* = -2.36, *p* = .02], were **older** [*M* = 44.5 (*SD* = 10.1) vs. M = 38.4 (*SD* = 11.1) *t* = -5.85, *p* <.001] and had worse **global functioning**, which was the basis for categorisation [*M* = 32.6 (*SD* = 8.9) vs. *M* = 10.5 (*SD* = 6.6) *t* = -28.5, *p* < .001]. • The regression analysis showed that individuals with higher rates of **depressive symptoms** [β = 0.39, *Wald* = 32.56, *OR* = 1.48, *CI 95%* (1.29-1.7), *p* < .01] and history of **psychotic symptoms** [β = 1.07, *Wald* = 4.77, *OR* = 2.91 *CI 95%* (1.11-7.54), *p* = .03] were less likely to achieve functional recovery. No association: • The functionally remitted and impaired groups did not differ in years of **education** [*M* = 14.1 (*SD* = 3.9) vs. *M* = 14.3 (*SD* = 3.3) *t* = 1.38, *p* = .16], **estimated IQ** [*M* = 108.4 (*SD* = 8.8) vs. *M* = 108.6 (*SD* = 9.7) *t* = -0.16, *p* = .52], **manic symptoms** [*M* = 1.2 (*SD* = 1.7) vs. *M* = 1.4 (*SD* = 1.8) *t* = -1.29, *p* = .19], number of **manic episodes** [*M* = 2.1 (*SD* = 2.6) vs. *M* = 2.4 (*SD* = 3.1) *t* = -1.10, *p* = .27], **BD diagnosis** [*n* = 173 (79.7%) vs *n* = 142 (72.1%), X^2 = 3.31, *p* = .08), lifetime **rapid cycling** [*n* = 15 (9%) vs *n* = 26 (15.9%), X^2 = 3.5, *p* = .07), lifetime **psychotic symptoms** [*n* = 142 (66%) vs *n* = 131 (67.2%), X^2 = 0.06, *p* = .83), and **family affective psychiatric history** [*n* = 109 (66.9%) vs *n* = 118 (71.5%), X^2 = 0.83, *p* = .40). • The functionally remitted group did not differ from the functionally impaired group in **executive functions**- as measured on SCWT interference test [*M* = 52.6 (*SD* = 6.4) vs. *M* = 52.3 (*SD* = 7.6), *t* = 0.47, *p* = .63] and **verbal**

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				memory- as measured on CVLT delay free recall test [*M* = 12.6 (*SD* = 2.8) vs. *M* = 11.8 (*SD* = 7.3), *t* = 1.4, *p* = .15]
				Significant positive associations: • Functionally remitted group was more prevalent in **female gender** [*n* = 116 (52.5%) vs *n* = 80 (40.6%), X^2 = 5.9, *p* = .01). • Functionally remitted group had significantly better cognitive functioning in **processing speed** [*M* = 104.3 (*SD* = 17.3) vs *M* = 100.7 (*SD* = 14.3), *t* = 2.2, *p* = .02], in **working memory** [*M* = 100.5 (*SD* = 12.9) vs *M* = 95.6 (*SD* = 14.9), *t* = 3.5, *p* < .01], in **executive functions**-as assessed by WCST categories [*M* = 5.4 (*SD* = 1.4) vs *M* = 4.5 (*SD* = 1.9), *t* = 5.1, *p* < .01], WCST preservative errors [*M* = 11.8 (*SD* = 10.6) vs *M* = 18.1 (*SD* = 14.9), *t* = - 4.8, *p* < .01], further examination of executive functions- phonemic fluency [*M* = 36.0 (*SD* =9.7) vs *M* = 33.7 (*SD* = 10.2), *t* = 2.3, *p* = .02] and animal naming [*M* =20.2 (*SD* = 4.5) vs *M* = 18.3 (*SD* = 5.6), *t* = 3.9, *p* < .01], in **attention**- as assessed by TMT-A [*M* = 27.8 (*SD* = 9.5) vs *M* = 37.6 (*SD* = 18.1), *t* = -6.8, *p* < .01] and TMT-B [*M* = 70.7 (*SD* = 37.6) vs *M* = 110.6 (*SD* = 75.8), *t* = -6.7, *p* < .01], in **verbal memory** assessed as CVLT total words [*M* = 56.5 (*SD* = 10.8) vs *M* = 50.9 (*SD* = 13.7), *t* = 4.6, *p* < .01], CVLT short-free recall [*M* = 12.1 (*SD* = 2.9) vs *M* = 11.1 (*SD* = 6.1), *t* = 2.2, *p* = .02], CVLT short-cued recall [*M* = 13.0 (*SD* = 2.5) vs *M* = 11.8 (*SD* = 3.2), *t* = 4.2, *p* < .01] and CVLT delay cued recall [*M* = 13.2 (*SD* = 2.4) vs *M* = 12.1 (*SD* = 3.1), *t* = 4.1, *p* < .01], and in **visual learning and memory** [*M* = 19.2 (*SD* = 4.8) vs *M* = 17.3 (*SD* = 5.3), *t* = 3.7, *p* < .01]. • The regression analysis showed that individuals with better **executive functions**-as measured on WCST number of

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				categories [β = -0.35, *Wald* = 4.17, *OR* = 0.7, *CI 95%* (0.5-0.98), p = .04], **working memory** [β = -0.04, *Wald* = 7.57*, *OR* = 0.95, *CI 95%* (0.93-0.98), p < .01], and **verbal memory**-as measured on CVLT short-cued recall [β = -0.38, *Wald* = 5.52, *OR* = 0.68 *CI 95%* (0.49-0.93), p = .02] were more likely to achieve functional recovery.
S5	Prospective cohort study (12 months FU period: 4 assessments at BL, time of stabilisation, 6 months and 12 months	**1) Clinical factors:** Age at admission symptomatic remission, negative symptoms, family history of schizophrenia and/or affective disorder, duration of untreated psychosis (DUP) symptoms prior to admission, duration of untreated mania (DUM) symptoms, alcohol use and illicit drug use. **2) SFR components:** Functional recovery at 6 months (as predictor at 12 months FU).	1) Comparisons between patients who had and had not recovered function were conducted using the non-parametric Mann–Whitney U-test. 2) Backward stepwise logistic regressions based on the *Wald* statistic were conducted to determine which factors significantly predicted dichotomous outcome variables (presence or absence of functional recovery after 12 months). OR and 95% CI were calculated for the identified predictors. The capacity of the model to correctly distinguish between patients with different outcome was explored with the Hosmer and Lemeshow test. The level of variance explained by the model was assessed by the Nagelkerke R^2.	Significant negative associations: • Mann-Whitney U-test: Non-recovered participants had significantly higher scores of **negative symptoms** (p < .01)-except alogia. • Final model: $\chi^2(4) = 28.96$, $p < 0.01$; Hosmer and Lemeshow test: $\chi^2(8) = 13.49$, $p > .05$; included 4 variables- 1 showed significantly negative association with functional recovery: **illicit drug use:** β = 1.79, z = 4.98, *OR* = 19.21 CI 95% (1.43, 257.23). No association: • **DUP and DUM** (no statistics reported) and **family history of affective disorders** [β=3.08, z=3.73, *OR* = 21.12 CI 95% (0.96, 466.84), significance not reported]. • Patients who had not recovered function at 12 months did not have significantly higher **alogia** scores than those that had recovered function (on the SANS). Significant positive associations: • Functional recovery at 12 months was associated with **functional recovery at 6 months:** $\chi^2(1) = 11.53$, $p < .05$; and remission of symptoms at 6 months: $\chi^2(1) = 4.88$, $p < .05$. • Final model ($\chi^2(4) = 28.96$, $p < 0.01$; Hosmer and Lemeshow test: $\chi^2(8) = 13.49$, $p > 0.05$) included 4 variables- 2 showed significantly positive association: **age:** β= -.037, z = 5.48, *OR* = 0.69 CI 95% (0.50, 0.94) p < .05; and achieving **functional**

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				recovery at 6 month: β = 5.72, z = 7.89, OR = 305.81, CI 95 % (5.65, 257.23), $p < 0.01$.
S6	Randomised control trial (12 month FU period, after 2 months of therapy: 5 assessment points: BL, after 8 sessions (at weeks 4, after 16 sessions (at week 8), 6 and 12 months after the end of the treatment)	**1) Psychosocial factors:** Psychoeducation treatment: comparison of EG (pharmacological treatment and psychoeducation) and CG: (pharmacological treatment and placebo intervention-relaxation).	1) Categorical variables were compared using Pearson's chi-squared test, continuous variables were compared using the t-test. 2) Groups were compared at the five time-points using two-way ANOVA for repetitive measurements. Inter- and intragroup comparisons were also performed. Significance was set at p = .05 for all comparisons.	Significant negative associations: • The scores on the **environmental domain** (WHOQOL-BREF) suggested a worsening over time (p = .025) in both groups. No association: • The means for the **social component** of the Social Adjustment Scale were stable over time (p =.114, ES = 0.078) with no difference between groups (p =.416, ES = 0.036). • **Functioning** levels (GAF) did not change over time (p =.097, ES = 0.089) in either group (p =.586, ES = 0.027). Significant positive associations: None reported
S7	Prospective cohort study (12 months FU period: outcome assessments at 2, 6 and 12 months after discharge.)	**1) Demographic factors**: age, race, sex, SES **2) Clinical factors:** number of episodes, presence of personality disorder, treatment compliance	1) Kaplan-Meier survival curves were used to estimate the probability of recovery. The log-rank test determined differences between groups. 2) Logistic regression analysis were performed to determine whether personality disorder were associated with functional recovery controlling for demographic and clinical variables. 3) Chi-square analysis was performed on the first episode sub-group to	Significant negative associations: • Patients with **personality disorder** and BD were significantly less likely to recover from a manic episode one year after hospitalisation (χ^2 = 6.6, df =1, p =.01). No association: • **Age, race, sex, number of manic or mixed episodes and treatment compliance** were not associated with functional recovery (no statistics reported). • First episode sub-group: no association between **personality disorder** and functional recovery. Significant positive associations: None reported

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
			determine whether personality disorder was associated with functional recovery.	
S8	Retrospective cohort study (FU period not specified- 3 assessment points: premorbid highest functioning, worst ever functioning and current functioning)	**1) Demographic factors:** sex **2) Clinical factors:** illness onset, duration of illness, gene CACNA1C **3) SFR component:** premorbid functioning	1) All p-values reported are two-sided. 2) Linear regression residuals at all three time points were jointly analysed by non-parametric longitudinal rank-sum test. Analysis adjusted for age and illness onset or duration- separately for males and females. 3) Sex-stratified analyses also considered the recovery phenotype (GAF3 minus GAF2).Latter was adjusted for sex, duration of illness, and premorbid GAF. A non-parametric maximum test (nparcom) was used for analysing recovery in males and females, which accounts for unknown genetic mode of inheritance. These tests are robust when used on non-normally distributed variables.	Significant negative associations: None reported. No association: • Regression detected no **sex CACNA1C interaction** in the BD sample (p = .870). Also found when additionally adjusting GAF scores for diagnostic subcategory and when only analysing the largest diagnostic subgroup (i.e.BD-I). • No statistics reported on the association of **illness onset, duration of illness, and premorbid functioning** (adjusted for in regression) with recovery. Significant positive associations: None reported
S9	Prospective cohort study (12 months FU period- 2 assessments: BL and 12 months after hospital discharge)	**1) Clinical factors:** Depressive, manic and psychotic symptoms (assessed both at BL and FU), lifetime alcohol and drug dependence, presence of lithium or/and benzodiazepine treatment.	1) Logistic regression: dependent variable was the MSIF global at 12 months after hospital discharge. Independent variables were the neurocognitive factors. Five covariates were used: depressive and manic symptom scores at BL and FU (at the same time as the outcome), and time between BL and FU. Each	Significant negative associations: • **Manic symptoms** at FU were significantly associated with functional recovery [p = -.0007, OR = 0.86, 95% CI (0.79; 0.94)]- cross-sectional finding. • **Psychotic symptoms** at FU were associated with worse functional recovery (statistics not reported)- cross-sectional finding. • **Lifetime alcohol and drug dependence** was significantly associated with recovery (statistics not reported).

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		2) Psychosocial factors: 1) Neurocognitive factors: attention, working memory, ideational fluency, verbal knowledge, non-verbal functions and learning, and time between BL and FU assessments	neurocognitive factor was examined independently (in each case including all five covariates). 2) Additional logistic regressions were completed with psychosis symptoms, lifetime alcohol and drug dependence, presence of lithium or/and benzodiazepine treatment as covariates.	No association: • BL **manic and depressive symptoms** and FU depressive symptoms were not associated with functional recovery (no statistics reported). • Presence of **lithium or/and benzodiazepine treatment** was not associated with functional recovery (no statistics reported). • Neither **working memory** nor **learning** showed any relationship with 12-month functional recovery. • Trend level associations were observed for **verbal knowledge** and **non-verbal functions** (no statistics reported) Significant positive associations: • **Attention** [*Wald* χ^2 = 4.256, *p* = .039, *OR* = 1.87, 95% *CI* (1.032; 3.397)] and **Ideational Fluency** [*Wald* χ^2 = 3.927, *p* = .048, *OR* = 1.62, 95% *CI* (1.005; 2.601)] were associated with recovery at FU.
S10	Prospective study (12 months FU period)- data combined from two studies: a randomised-withdrawal study and open label maintenance study	**Clinical factors:** 52 weeks of aripiprazole maintenance treatment (AOM 400 = Aripiprazole 400mg once monthly)	1) Functioning total and domain scores were summarized at BL and at 52 weeks of the respective maintenance phases using mean and SD for a) all participants included in the analysis and b) those participants who met criteria for functional recovery. 2) Between-group differences were derived from an ANOVA model with treatment and region as BL factors.	Significant negative associations: None reported No association: • During the maintenance phase of the randomised withdrawal study: 30.2% of participants (35/116) receiving AOM 400 and 24.8% of participants (28/113) receiving placebo achieved recovery. Recovery rates were not significantly different between AOM 400 and placebo groups (*p* = .394). • In the randomised-withdrawal study, there were no significant differences between the AOM 400 and CG in any of the functioning domains at 52 weeks in the total study sample (except in interpersonal relationship functioning): **Autonomy** [BL: *M* = 1.47 (*SD* = 2.07) in AOM 400 group

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				vs. *M* = 1.35 (*SD* = 2.10) in CG; FU: *M* = 1.72 (*SD* = 2.49) in AOM 400 group vs. *M* = 2.43 (*SD* = 3.19) in CG, *LS M* = -0.67, *95% CI* (-1.35, 0.020), *p* = .055]; **Occupational functioning**: [BL: *M* = 5.73 (*SD* = 4.91) in AOM 400 group vs. *M* = 5.69 (*SD* = 5.35) in CG; FU: *M* = 5.29 (*SD* = 4.91) in AOM 400 group vs. *M* = 6.39 (*SD* = 5.54) in CG, *LS M* = -0.82, *95% CI* (-1.88, 0.24), *p* = .128]; **Cognitive functioning**: [BL: *M* = 3.28 (*SD* = 3.38) in AOM 400 group vs. *M* = 2.91 (*SD* = 3.18) in CG; FU: *M* = 3.64 (*SD* = 3.62) in AOM 400 group vs. *M* = 4.21 (*SD* = 4.06) in CG, *LS M* = -0.56, *95% CI* (-1.44, 0.32), *p* = .212]; **Financial issues**: [BL: *M* = 1.20 (*SD* = 1.66) in AOM 400 group vs. *M* = 1.17 (*SD* = 1.48) in CG; FU: *M* = 1.16 (*SD* = 1.53) in AOM 400 group vs. *M* = 1.57 (*SD* = 1.72) in CG, *LS M* = -0.35, *95% CI* (-0.74, 0.04), *p* = .075]; **Leisure time functioning**: [BL: *M* = 1.40 (*SD* = 1.57) in AOM 400 group vs. *M* = 1.18 (*SD* = 1.28) in CG; FU: *M* = 1.62 (*SD* = 1.65) in AOM 400 group vs. *M* = 1.81 (*SD* = 1.76) in CG, *LS M* = -0.19, *95% CI* (-0.58, 0.20), *p* = .339]; • Or in the recovery subgroup: **Autonomy** [BL: *M* = 0.29 (*SD* = 0.63) in AOM 400 group vs. *M* = 0.04 (*SD* = 0.19) in CG; FU: *M* = 0.17 (*SD* = 0.45) in AOM 400 group vs. *M* = 0.11 (*SD* = 0.42) in CG, *LS M* = 0.01, *95% CI* (-0.23, 0.24), *p* = .958]; **Occupational functioning**: [BL: *M* = 2.26 (*SD* = 2.74) in AOM 400 group vs. *M* = 1.56 (*SD* = 1.89) in CG; FU: *M* = 1.14 (*SD* = 1.65) in AOM 400 group vs. *M* = 0.93 (*SD* = 1.59) in CG, *LS M* = 0.24, *95% CI* (-0.63, 1.11), *p* = .582]; **Cognitive functioning**: [BL: *M* = 0.97 (*SD* = 1.57) in AOM 400 group vs. *M* = 0.96 (*SD* = 1.29) in CG; FU: *M* = 0.66 (*SD* = 1.21) in AOM 400 group vs. *M* = 0.29 (*SD* = 0.66) in CG, *LS M* = 0.32, *95% CI* (-0.12, 0.76), *p* = .155]; **Financial issues**: [BL: *M* = 0.32 (*SD* = 0.68) in AOM 400 group vs. *M*

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				= 0.57 (*SD* = 1.23) in CG; FU: *M* = 0.29 (*SD* = 0.83) in AOM 400 group vs. *M* = 0.43 (*SD* = 0.88) in CG, *LS M* = -0.06, *95% CI* (-0.49, 0.36), *p* = .777]; **Leisure time functioning**: [BL: *M* = 0.47 (*SD* = 1.08) in AOM 400 group vs. *M* = 0.50 (*SD* = 0.64) in CG; FU: *M* = 0.49 (*SD* = 0.92) in AOM 400 group vs. *M* = 0.54 (*SD* = 0.69) in CG, *LS M* = 0.00, *95% CI* (-0.38, 0.38), *p* = .992]; **Interpersonal relationship functioning**: [BL: *M* = 1.15 (*SD* = 1.62) in AOM 400 group vs. *M* = 0.71 (*SD* = 1.01) in CG; FU: *M* = 0.77 (*SD* = 1.26) in AOM 400 group vs. *M* = 0.46 (*SD* = 0.79) in CG, *LS M* = 0.17, *95% CI* (-0.37, 0.70), *p* = .537]. Significant positive associations: • In the randomised-withdrawal study, functional recovery total scores were generally maintained in the group of participants who received AOM 400 [*M* = 15.92 (*SD* = 13.19) at BL; *M* = 16.59 (*SD* = 13.98) at FU] and were worsened in the CG [*M* = 14.82 (*SD* = 12.12) at BL *M* = 20.91 (*SD* = 16.87) at FU. At 52 weeks: [AOM 400 vs CG mean treatment effect = - 3.98, *95% CI* (-7.52; -0.44) *p* = .028]. Participants who met criteria for functional recovery, FAST total scores improved from *M* = 5.47 (*SD* = 5.50) at BL to *M* = 3.51 (*SD* = 3.62) at FU in the AOM 400 group and from *M* = 4.44 (*SD* = 4.23) at BL to *M* = 2.75 (*SD* = 2.86) at FU in the placebo group. • In the randomised-withdrawal study, **interpersonal functioning** worsened more in the total population CG at 52 weeks compared to the AOM 400 group [At BL: *M* = 2.84 (*SD* = 3.5) in AOM 400 group vs. *M* = 2.45 (*SD* = 2.87) in CG; at FU: *M* = 3.15 (*SD* = 3.82) in AOM 400 group vs. *M* = 4.50 (*SD* = 4.67) in CG, *LS mean* = -1.32; *95% CI* (-2.32, -0.31), *p* = .011]

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				• In the open-label study, de novo participants significantly improved from BL to the end of the 4-12 weeks stabilisation phase (*M* = 17.90 (*SD* = 13.51) vs. *M* = 14.02 (*SD* = 12.02), *p* < .001, one-sided Z test).
S11	Prospective study with 48 months FU period	**Clinical factors:** Number of affective episodes, time: FU period	1) Due to skewed distribution of scores on continuous variables nonparametric tests were used. 2) Differences at BL between participants with more or less than five previous affective episodes were analysed with the Mann–Whitney test for continuous variables and chi-squared test for categorical variables. 3) Differences between Time 1 and Time 2 were analysed as two related samples with the Wilcoxon signed-rank test for ordinal/continuous variables and McNemar or Marginal Homogeneity Test for categorical variables.	Significant negative associations: • At BL participants with more previous **affective episodes** self-reported worse levels of functional recovery than patients with fewer episodes (56.7% vs. 85.0%, $\chi 2$ = 4.42, *p* = .035). No association: • Improvements in functioning (including recovery rates) **between BL and FU** did not seem to be influenced by the number of affective episodes during the FU period. Significant positive associations: • At FU, participants showed a better level of functional recovery compared to BL (57.42% vs. 70.4%, *p* = .039).
S12	Prospective cohort study (6 months FU period: 3 assessments (BL, 1 month and 6 months)	**1) Clinical factors:** Presence of mixed symptoms during current manic episode	1) Comparison between manic patients with and without mixed features, using descriptive statistics, independent samples t-test or chi-square, depending on the nature of the variables. All the analyses were two-tailed with alpha set at $p < 0.05$.	Significant negative associations: None reported No association: • No differences were found between groups (with and without **mixed features**) in functional recovery using either BL functioning total score (*t* = 0.69, *p* =.492) or at FU (*t* = 1.73; *p* = .085) or comparing the proportion of participants who achieved functional recovery. Significant positive associations: None reported

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
S13	Prospective cohort study (8 months FU period: assessments at BL, 1, 4 and 8 months-maximum)	**1) Demographic factors:** age, sex, ethnicity, years of education, and highest employment level (past 5 years) and SES (based on education and employment). **2) Clinical factors:** depressive and manic symptoms, symptomatic recovery, age at onset, presence of psychosis, index episode duration and polarity (mixed/manic), history of untreated affective episode, current alcohol and cannabis use disorder, pharmacological treatment compliance, non-pharmacologic mental health contacts per month. **3) SFR components**: BL functioning in 4 areas: role performance, recreational enjoyment, interpersonal relationship, sexual activity.	1) Differences in the timing and rates of recovery of the four areas of function were compared using the Kaplan-Meier survival curves and the two-tailed log-rank statistic. 2) Associations among the areas of function were determined using Spearman correlations. 3) Logistic regression techniques were employed to identify specific variables that predicted recovery in each of the four major areas of function. In this analysis, age, sex, and socioeconomic status (``forced variables") were included in all regression models. The BL rating in each of the four major areas of function was included in each model. 4) Other potential outcome predictors were examined for inclusion in the final logistic regression models using stepwise selection. In this stepwise selection process, additional variables were entered into the model if they were associated with recovery at a $p < .2$. They were retained in the model only if the association with recovery persisted at a $p < .05$ after adjusting for the forced variables and BL ratings.	Significant negative associations: • Patients with **index episodes** longer than 2 months exhibited poorer BL interpersonal relationships ratings (i.e., best score in the previous 5 years) compared to the remaining participants ($t = 1.9$, $df = 40$, $p = .065$). • Participants who failed to achieve recovery of sexual activity were more likely to exhibit **mood incongruent psychosis** at the index assessment ($\chi^2 = 6.4$, $df = 1$, $p = .01$). No association: • None of the four areas of function were significantly correlated with each other at BL (maximum $r < 0.25$, $p > .07$). The times to achieve recovery of the areas did not correlate (maximum $r < 0.18$, $p > .2$). • Recovery of **role performance** was not associated with the examined predictors (except age at onset and SES). • Recovery of **interpersonal relationships** was not associated with the examined predictors (except duration of index episode and symptomatic recovery). • Recovery of **recreational enjoyment and sexual activity** were not associated with any of the predictors (statistics not reported) in the regression models. Significant positive associations: • **Age of onset:** Patients whose bipolar illness began prior to age 20 years were less likely to achieve recovery of role performance compared to those whose illness began later (adjusted *Wald* $\chi^2 = 4.6$, $df = 1$, $p = .03$). • Higher **socioeconomic status** (SES) was associated with a greater likelihood of recovery of role performance in this statistical model (adjusted *Wald* $\chi^2 = 5.2$, $df = 1$, $p = .02$) of achieving a good outcome (adjusted *Wald* $\chi^2 = 6.6$, $df = 1$, $p = .01$). • Recovery of interpersonal relationships was more likely for patients with **index episodes** longer than 2 months than

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
			Demographic and clinical variables were examined in this manner.	those with shorter index episode duration (adjusted *Wald* χ^2 = 7.3, *df* = 1, *p* = .007). • Recovery of interpersonal relationships was also significantly more likely for patients who achieved **symptomatic recovery** during FU than those who did not (adjusted *Wald* χ^2 = 4.4, *df* = 1, *p* = .035).
S14	Cross-sectional study	**1) Demographic factors:** gender, age, education, parents' education, employment status, marital status, ethnicity. **2) Clinical factors:** age at onset, subtype of BD, illness duration, co-morbid illnesses (medical and psychiatric-Axis I), history of psychosis, rapid cycling, number of episodes/year, number of suicide attempts and hospitalisations, current symptoms, time since last episode (months), number of psychotropic medications (with/without antidepressants). **3) Psychosocial factors:** estimated premorbid IQ, executive functions, attention, concentration,	1) Two-sample t-tests or *WRS* tests compared group means of continuous variables. Chi-square ($\chi 2$) or FET compared proportions. 2) To explore factors associated with social-functional recovery, variables with at least suggestive differences ($p <$ .15) between socially recovered and unrecovered patients based on univariate descriptive statistics were entered into a multiple logistic regression model using stepwise selection method. Statistical significance required a two-sided $p <$ .05.	Significant negative associations: • **Age:** recovered subjects were significantly younger (*t* = 2.99, *p* = .004), • Socially unrecovered participants had more **depressive symptoms** (*WRS* = 747, *p* =.002) and had been **ill longer** (*WRS* = 834, *p* =.04), and received more **psychotropic medication** (*WRS* =814, *p* =.02) than the socially recovered participants • Selection by stepwise inclusion of potential factors found two factors to be significantly and independently associated with social-functional recovery: younger **age** [Adjusted *OR* = 0.93; 95% *CI* (0.89; 0.98,), *p* =.005] and lower **current depression** scores [Adjusted *OR* = 0.82; 95% *CI* (0.69; 0.97), *p* = .020]. No association: • The recovered and unrecovered subgroups had similar previous highest levels of **social functioning** (*WRS* = 1023, *p* = .66), were similar in **sex-distribution** (χ^2 = 0.15 *p* =.70), ethnicity (*FET*, *p* = .74), years of **education** (*t* = 0.29, *p* = .77), **parental education** (*WRS* = 1105, *p* = .12; *WRS* = 1090, *p* = .12), **employment** (χ^2 = 2.52, *p* = .11), and **marital status** (*FET*, *p* =.29). • The recovered and unrecovered subgroups were similar in **estimated IQ** (*WRS* = 1095, *p* =.17), **attention, concentration**, and **mental tracking** (*t* =0.24, *p* = .81), **verbal learning** and **memory** (*t* = −0.49, *p* = .62), and **executive functions** (*WRS* = 1114, *p* =.11).

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		mental tracking, verbal learning and memory.		• Recovered vs. non-recovered were similar in **onset age** (*WRS* = 991 p = .99); **BD-subtypes** (χ^2 = 0.71, p = .40) prevalence of **co-morbid psychiatric illnesses**, including **substance use** disorder (χ^2 = 0.01, p = .90) or **medical illnesses** (χ^2 = 0.12, p = .73); past **psychosis** (χ^2 = 0.48, p = .49) and **rapid cycling** (χ^2 = 0.10, p = .75); annual rates of lifetime major depressive (*WRS* = 1039, p = .52) or manic/hypomanic **episodes** (*WRS* = 1000, p = .90) or total mood episodes (*WRS* = 1015, p =.74); number of **suicide attempts** (*WRS* = 951 p = .56); number of **hospitalizations** (*WRS* = 1011, p =.78), and proportions taking **antidepressants** (with antidepressant χ^2 = 1.39, p = .24), current **manic symptoms** (*WRS* = 921 p = .35), and **time since last episode** (*WRS* = 1086, p = .14) • Factors entered into the regression model that were non-significant: **months since last major episode** (p = .611); **co-morbid psychiatric illness** (p = .704); **executive function** (p = .571); **BD diagnostic type** – I vs II (p =.724) Significant positive associations: none reported
S15	Prospective cohort study (max. 9 months FU period-monthly assessments until functional recovery achieved)	**1) Clinical factors:** Depressive and manic symptoms **2) Psychosocial factors**: Acute stress- stressful life events in the past 3 months.	1) One-way ANOVAS with planned contrasts compared concurrent vs. delayed functional recovery groups and delayed versus non-recovered groups, on depressive and manic symptoms prior to each content domain functional recovery assessment. 2) Logistic regression analyses were used to test the contribution of recent stressors to functional outcome status (concurrent with clinical recovery vs. delayed), controlling for depression and mania residual scores in the month before functional recovery. Four	Significant negative associations: • Delayed recovery of work/school functioning was significantly associated with presence of one or more **stressors** in the prior 3 months [$\beta(SE)$ = 2.07 (0.73), *Wald* = 7.98, *OR* = 7.93, *95% CI* (1.89; 33.3), p = .005]. Similarly in the friendship domain and family domain, presence of a stressor significantly predicted delayed functional recovery: friendship: [$\beta(SE)$ = 2.08 (0.87), *Wald* = 5.76, OR=7.99 *95% CI* (1.46; 43.65), p = .02; family: $\beta(SE)$ = 2.34 (0.96), *Wald* = 5.98, *OR* = 10.37, *95% CI* (1.59; 67.7), p = .01. • Recovery of home duties functioning was related to higher **depressive symptom** scores among those in the delayed recovery group (statistics not reported). • Not recovered participants (in family, home duties and work/school domains) had significantly higher depressive

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
			separate regressions were conducted, one for each role domain. 3) Kaplan–Meier survival analyses were conducted to evaluate the time to achieve functional recovery in each of the four domains, as a function of presence/ absence of recent stressors.	symptoms compared to the concurrent recovered group ($p <$ 0.01). • The not recovered group in work/school domain had significantly higher depressive symptoms compared to the delayed recovery group ($p <$. 01). • Not recovered participants (in family, friends, home duties and work/school domains) had significantly higher **manic symptom** manic than the concurrent recovered group ($p <$ 0.01). • The not recovered group in home duties and work/school domains had significantly higher manic symptoms ($p <$ .01) and had significantly higher stress levels prior to family domain assessment compared to the delayed recovery group ($p <$ .01). • Participants who did no experience **stressful life events** had quicker recovery in the work/school domain (*log-rank* = 12.99, $p <$ .001), in the friend domain (*log-rank* = 11.56, $p <$ .001), in the family domain (*log-rank* = 10.58, $p <$ .001) compared to participants who experienced a stressful life event. No association: • There was no association between delayed recovery and **stress** occurrence in the home duties domain [*OR* = 2.84, *95% CI* (0.57–14.09)]. • **Symptoms** were generally not significant predictors of concurrent versus delayed recovery, except home duties (statistic not reported). • Recovered and not recovered participants in friends domain did not differ significantly in depressive symptoms (statistic not reported). • The delayed recovery group was similar to the not recovered group in depressive and manic symptomatology (statistic not reported).

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				• Not recovered participants were similar in experienced **stress** to the delayed functional recovery on three of the domains, friends, work/school and home duties (no statistics reported) • In the home duties domain, participants who did not experience a **stressful life event** had similar time to recovery (*log-rank* = 0.35, ns.) compared to participants who experienced a stressful life event. Significant positive associations: None reported
O1	Prospective cohort study- maximum 9 months FU period (until occupational recovery achieved): BL and monthly FU assessments of mood and occupational functioning (operationalised as occupational recovery) and neurocognitive assessments every 3 months.	**1) Demographic factors:** age, education, ethnicity, gender, and marital status. **2) Clinical factors:** age of onset, depressive and manic symptomology, number of depressive and manic episodes and therapy/ medication usage, being in therapy at the time of the assessment. **3) Psychosocial factors**: Episodic memory, visual scanning, working memory/attention, executive function, speed of processing. BL neurocognitive function and change/improvement in neuro-cognition over time.	1) In order to identify potential confounders, first associations between BL recovery and individual demographic and course of illness measures were examined using two sample t-tests or χ^2 tests. 2) Multiple logistic regression was used to determine the joint contributions of the neurocognitive domain scores to the prediction of functional recovery, adjusting for key demographic and clinical covariates as identified in the preliminary analyses. • Model 1 evaluated the relationship between neurocognitive performance and occupational recovery at BL; • Model 2 analysed BL neurocognitive scores as predictors of occupational recovery at three month; • Model 3 used change scores in neurocognitive domains from	Significant negative associations: • **Age** (*OR* = .33; *p* < .01) and BL **depressive symptoms** (*OR* = 0.95; *p* < .01) predicted BL occupational recovery (negative associations). • Age predicted occupational recovery at 3 month (*OR* = .61, *p* = .013-when adjusted for BL neurocognitive factors; *OR* = .99, *p* = .02- when adjusted for changes scores in neuro-cognitive factors and depressive symptoms). No association: • There were no significant differences between recovered and unrecovered individuals in demographic factors: **age** (*p* = .24) **education** (*p* = .47), **ethnicity** (*p* = .54), **gender** (*p* = .97) and **marital status** (*p* = .86) at BL. • There were no significant differences between recovered and unrecovered individuals at BL in clinical factors: prior manic (*p* = .25) or depressive **episodes** (*p* = .17), manic (*p* = .29) and depressive **symptoms** (*p* = .06) **age of onset** (*p* = .46) or **medication** usage (*p* > .15) or being in **therapy** (*p* = .77). • BL **executive functions** did not predict BL (*OR* = 1.59, *p* = .08) or FU (*OR* = 1.82, *p* = .17) occupational recovery. • BL psychosocial factors did not predict occupational recovery at 3 months: **episodic memory** (*OR* = 1.89, *p* = .081), **visual scanning**: (*OR* = 1.14, *p* = .66), **working memory/attention** (*OR* = 1.62, *p* = .20), **speed of processing** (*OR* = 1.5, *p* = .11).

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
			Time 1 to Time 2 as predictor variables to assess whether improvement in neurocognitive function between BL and three months was associated with three-month occupational recovery. Overall performance of the models was examined using the area under the receiver operating characteristic curve (AUC) which is a plot of the false positive rate versus the false negative rate. 3) Age and subsyndromal symptoms of depression were included in all logistic regression models. All analyses were two- tailed with alpha set at $p < .05$. 4) The stability of the third logistic regression model was assessed using a bootstrap re-sampling procedure.	• BL **depressive symptoms** (*OR* = .098, *p* = .55) or changes in depressive symptoms between BL and FU (*OR* = 0.93, *p* = .96) did not predict occupational recovery at 3 months. • Changes in **speed of processing** between BL and FU did not predict occupational recovery at 3 months (*OR* = 3.78, *p* = .06). • The unrecovered and recovered group at 3 months did not differ significantly in their psychosocial change scores (effect sizes for group difference in change score-Cohen's d): **episodic memory** (*d* = .0.80, *p* < .1), visual scanning (*d* = 0.05, ns), **executive functions** (*d* = 0.49, ns.), **speed of processing** (*d* = 0.19, ns.) Significant positive associations: • BL **episodic memory** (*OR* = 1.55, *p* = .018), **visual scanning** (*OR* = 2.21 *p* = .006), **working memory/attention** (*OR* = 2.49, *p* < .01) and **speed of processing** (*OR* = 2.62, *p* < .01) predicted BL occupational recovery. • Changes in psychosocial factors between BL and FU predicted occupational recovery at 3 months: **episodic memory** (*OR* > 10, *p* < .01), **visual scanning** (*OR* = 5.25, *p* < .01), **working memory/attention** (*OR* > 10, *p* < .01), **executive functions** (*OR* > 10, *p* < .01). • The recovered and unrecovered group differed significantly in their **attention/working memory** change score between BL and FU (*d* =1.05, *p* < .05).
O2	Cross-sectional study	**1) Demographic factors:** age, ethnicity, marital status, gender, education (years) **2) Clinical factors:** Presence of Personality Disorder (PD-categorical	1) Nonparametric (χ^2 with FET) and parametric methods (Student t test) were used to compare variables as appropriate. 2) Multiple linear and logistic regression analyses were conducted to examine the effects of PDs/traits and	Significant negative associations: • Participants with a greater number of **maladaptive PD traits** relative to those with fewer traits were more likely to be classified in the poor work functioning group (*t* = 2.50, *p* = .016). • Participants with poorer work functioning had a significantly greater number of prior **hospitalizations** (*t* = 2.07, *p* = .044),

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		or trait scores), age of onset, number of hospitalisation, other psychiatric comorbidities (substance abuse & anxiety).	other clinical variables on work, residential, and social/leisure outcomes.	a higher level of residual **manic symptoms** (t = 2.18, p = .034). • Residential role recovery showed negative association with manic (r = .39, p = .005) and depressive (r = .30, p = .035) **symptoms.** • **Depressive symptoms** remained significant predictor of residential recovery in the regression model (t=2.58, p=.013) No association: • No significant differences reported between poor work functioning and good work functioning group: **age, ethnicity, marital status, gender, education, residual depressive symptoms, age of onset, or other psychiatric comorbidities,** including **substance abuse** (χ^2 = 4.13, p = .073). • PD traits (*Wald* χ^2=2.73, p = .098), number of **hospitalisatio**ns, and residual **manic symptoms** did not remain independent significant predictors of occupational recovery in the regression model. • No associations reported between **PD traits** (r =.26, p = .066), **ethnicity, gender, marital status, age of onset,** number of **hospitalisation**, **psychiatric comorbidities** and residential role recovery. Significant positive associations: • Residential role recovery showed positive association with **age** (r = –.40, p = .004) (older individuals) and **education** (r = –.38, p = .006) (higher education levels) were more likely to achieve residential role recovery. • **Age** was a significant contributor to residential role recovery in regression model (t = 3.18, p = .003).

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
O3	Prospective cohort study (6 months FU-outcome data collected at hospital discharge (BL) and at 6 month (FU).	**1) Demographic factors**: age, sex and race. **2) Clinical factors:** manic and depressive symptomology. **3) Psychosocial factors:** Personality factors: Novelty Seeking, Harm Avoidance and Reward Dependence (dimensional scores)	1) Categorical variables were compared by a two-tailed FET. Two-class comparisons were made with the *WRS* test. For these comparisons, the analysis proceeded into three steps to control for multiple comparisons. The initial analysis compared the three outcome measures (syndromic remission at discharge, syndromic and functional recovery at 6 months) with the three dimensional scores. To control for Type 1 error, a Bonferroni correction to the standard $a = 0.05$ was used, resulting in a corrected ($\alpha = 0.0055$ as the significance level for these comparisons. The second step in the analysis involved determining which of the corresponding sub-dimensional scores contributed to any significant differences noted in the dimensional score analysis. Since this analysis was dependent upon results from the dimensional score analysis, an $\alpha = 0.05$ was used for the significance level. Finally, for comparisons between outcome measures and sub-dimensional scores that did not exhibit significant differences on the corresponding dimensional scores (maximum of 36 comparisons), a	Significant negative associations: • **Novelty seeking** (impulsiveness and disorderliness sub-dimensions) at discharge was significantly higher in the functionally not recovered participants ($z = 3.0$, $p = .003$), with most of this variance reflecting differences in the sub-dimensional scores "impulsiveness" ($z = 2.5$, $p = .01$) and "disorderliness" ($z = 2.2$, $p = .02$). Six patients (22.2%) had Novelty-Seeking scores > 20, and five of these patients (83.3%) failed to achieve functional recovery (*FET*, $p = .0004$). • Novelty seeking remained significant in logistic regression (*OR* = 2.9; CI=1.1-8.0, $p = .04$) No association: • There were no association between functional recovery and syndromic recurrence (*FET*, $p = 0.1$) or syndromic recovery. • There were no significant differences in manic or depressive **symptomatology** between patients who did and did not functionally recover ($z = 1.0$). • There were no significant differences in **sex or race** between patients who did and did not attain functional recovery. • **Age, sex, race** and manic and depressive **symptomatology** were not associated with a risk if failure to achieve functional recovery in the regression model. • **Harm Avoidance and Reward Dependence** did not associate significantly with recovery (statistics not reported). Significant positive associations: None reported.

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
			Bonferroni-corrected a = 0.001 was used to determine significance. 2) Correlations were made using the Pearson *r* statistic. 3) Logistic regression was performed for dimensional scores demonstrating significant associations with outcome variables from the previous analysis, controlled for confounding factors and calculated *OR* with *95% CI*. α = 0.05 was used to determine significance for the logistic regression analysis.	
O4	Prospective cohort study (6 months FU period- outcome collected at 6 months after discharge)	**1) Demographic variables:** age, marital status, race and gender. **2) Clinical factors:** psychiatric (Axis-I diagnosis) or medical comorbidity.	1) *OR* between discrete variables and outcome measures were obtained. 2) Non-paired *t* tests were used to compare continuous variables. 3) To simultaneously estimate the effects of risk factors (discrete and continuous variables) and to control for confounding factors, logistic regression models were fitted and adjusted OR (*ORa*) with 95% CIs were obtained. 4) Survival analysis curves using the Kaplan-Meier method were used to estimate time to recovery and time to recurrence.	Significant negative associations: • **Males** were less likely to recover functionally at 6 months [*ORa* = 4.9; *95%, CI* (- 1.4; 19.4); χ^2 = 5.9; *p* = .01) after controlling for age. No association: • No association reported between **age, marital status, race and psychiatric or medical comorbidity** in the bipolar cohort. Significant positive associations: none reported

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
O5	Prospective cohort study (2-4 years FU period: outcome assessments at: 6, 12, 24, 36 and 48 months)	**1) Demographic factors:** Age, sex, marital status and race. **2) Clinical factors:** Episode type (manic/mixed), psychotic features, prior major depressive episodes, comorbidities (psychiatric-Axis I diagnosis- and medical), alcohol and drug abuse, BL symptomology (depression, mania and psychosis), length of index hospitalisation, pharmacological treatment. **3) SFR components:** BL global functioning.	1) Rates of recovery or new episodes among recovered patients were compared in subgroups of interest by using contingency tables (chi-square) or *FET* if cells held <10 subjects (with $df = 1$, unless stated otherwise). 2) Mann-Whitney (U) rank methods compared distributions of continuous variables in subgroups. Group recovery and recurrence latencies were compared by Kaplan-Meier life table survival analyses, tested with Mantel- Cox log-rank (*chi-square*) tests. Variables with preliminary bivariate associations ($p \leq .10$) with recovery or recurrence were included in multivariate analyses. 3) Multiple logistic regression models (for categorical functional recovery) evaluated candidate variables for independent association with outcomes. For both types of models, we computed robust *SE*s or associated 95% CIs. Explanatory variables with adjusted odds ratios (for logistic regression) different from 1.0 ($p < .05$) were retained for final multivariate regression models. 4) Times to recovery (and 95% CI) in survival analyses were estimated as	Significant negative associations: • Preliminary bivariate analyses for likelihood of achieving functional recovery at 2 years found the following factors: shorter length of index **hospitalization** ($\chi^2 = 9.34$, $p = .002$). • Shorter initial hospitalization (*OR* = 2.82, 95% CI = 1.36–5.88; $p = .006$) was associated with functional recovery at 2 years in logistic multivariate regression. No association: • Having below- versus above-median BL **depression** ratings was weakly related to functional recovery ($\chi^2 = 2.37$, $p = .12$). • **Gender**-women and men did not differ in likelihood of functional recovery ($\chi^2 = 0.09$, $p = .76$), and there was no correlation of presence/absence of mood-incongruent **psychotic features** with functional recovery ($\chi^2 = 0.08$, $p = .78$). • **Ethnicity and marital status** did not remain significant in the logistic multivariate regression (statistics not reported). • No association reported with **initial episode type** (mixed/manic), prior major **depressive episodes,** medical and psychiatric **comorbidities, alcohol and drug abuse**, BL **manic symptoms**, **pharmacological treatment** or BL **global functioning** (statistics not reported). Significant positive associations: • Preliminary bivariate analyses for likelihood of achieving functional recovery at 2 years found the following factors: older **age** (≥30 years) at entry ($\chi^2 = 12.0$, $p = .001$), Caucasian versus other **race** ($\chi^2 = 6.69$, $p = 0.01$); being **married** ($\chi^2 = 4.64$, $p = .03$). • Being **older** than 30 (*OR* = 3.28, 95% CI = 1.58–6.82; $p = .001$) was associated with functional recovery at 2 years in logistic multivariate regression.

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
			weeks by which 50% of subjects (or 25%, if <50% by 2 years) reached recovery. 5) Correlations were determined by linear regression (*r*) or Spearman nonparametric rank (*rs*) methods. Statistical significance required two-tailed $p < .05$.	
O6	Cross-sectional study	**1) Demographic factors:** gender, age, education, parents' education, marital status, and ethnicity. **2) Clinical factors:** onset age, subtype of BD, illness duration, medical and psychiatric comorbidities (Axis-I), history of psychosis, rapid cycling, number of episodes/year, number of suicide attempts and hospitalisations, current symptoms (depressive and hypomanic), time since last episode (month), number of psychotropic treatment (with/without antidepressants). **3) Psychosocial factors:** executive functioning,	1) Chi-square (χ^2) or *FET* test was used to compare proportions. Two-sample t-test or *WRS* was used to compare group means of continuous variables. 2) Cognitive scores of functionally recovered and unrecovered patients were compared using multiple linear regression with cognitive z-scores as the dependent variable, recovery status as the independent variable, and residual mood symptoms and education as covariates. 3) To explore factors associated with recovery, variables with at least suggestive differences ($p < .15$) between recovered and unrecovered patients, based on bivariate descriptive statistics, were entered into a multiple logistic regression model using backward, forward, and stepwise selection methods. To that end, 10 covariates considered for the logistic	Significant negative associations: • **Illness duration** was a significant independent predictor of recovery in multiple regression model [*OR* = 0.95, *95% CI* (0.91; 0.997), $p = .037$]. No association: • Recovered participants did not differ significantly from unrecovered participants in the following demographic variables: **gender** ($\chi^2 = 2.6$; $p = .11$), **age** ($t = 0.5$, $p = .61$), **estimated IQ** (*WRS* = 1028, $p = .17$), and **parental education** (father: *WRS* = 1007.5, $p = .26$; mother: *WRS* = 977, $p = .36$). • Recovered participants did not differ significantly from unrecovered participants in the following clinical variables: **age of onset** (W*RS* = 997,$p = .34$), **type of BD** ($\chi^2 = 1$, $p = 0.32$), **illness duration** (*WRS* = 812, $p = .14$), **comorbidity** (psychiatric-including substance use: $\chi^2 = 0.6$, $p = .45$; medical $\chi 2 = 0.7$, $p = .41$), history of **psychosis** ($\chi^2 = 0.6$, $p = .45$), **rapid cycling** ($\chi^2 = 0.0$, $p = .96$), number of **episodes** [(hypo)mania: *WRS* = 932, $p = .92$; depression: *WRS* = 900, $p = .75$; and total: *WRS* = 914, $p = .90$], **suicide attempts** (*WRS* = 897, $p = .68$), number of **hospitalisations** (*WRS* = 895, $p = .70$), current **symptomatology** [depressive: *WRS* = 783, $p = .07$, and (hypo)manic symptoms: *WRS* = 909, $p = .84$], **time since a last major mood episode** recurrence (*WRS* = 1046, *p*

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		attention concentration, mental tracking, verbal learning and memory, and estimated premorbid IQ	regression model were education, marital status, race, MADRS score, time since the last major mood episode, sex, being treated with or without an antidepressant, number of current psychotropic medications, illness duration, and executive function (FAS z-score: Controlled Oral Word Association Test). 4) Statistical significance required a two-sided $p \leq .05$.	= .07), number of current **psychotropic medications** (*WRS* = 810, *p* = .13), and number of participants treated with **antidepressants** (χ^2 = 2.5, *p* = .11). • Recovered participants did not differ significantly from unrecovered participants in the following psychosocial factors: **executive functions** measured on Letter-number sequence [unadjusted *ES* = -0.06, *p* = .81; adjusted (for symptoms and education) difference in *z*-scores = 0.19 CI = -0.70–0.31, *p* = .45]; on FAS (adjusted difference in *z*-scores = 0.42; CI = -0.17–1.01; p = .16) and on TMT-B (unadjusted *ES* = -0.23, *p* = .38; adjusted difference in *z*-scores = 0.10, CI = -0.95–1.15, *p* = .85); **Attention, concentration and mental tracking** measured on Digit span test (unadjusted *ES* = 0.27, *p* = .28; adjusted difference in *z*-scores = 0.02, CI = -0.45–0.48, *p* = .94) and on TMT-A (*ES* = 0.32, *p* = .20; adjusted difference in *z*-scores = 0.14, CI = - 0.36–0.64, *p* = .59), in **verbal learning** and **memory** measured on RAVLT trials I-V (*ES* = 0.28, *p* = .27; adjusted difference in *z*-scores = 0.19, CI = -0.50–0.88, *p* = .58), RAVLT immediate recall (*ES* = 0.21, *p* = .40; adjusted difference in *z*-scores = 0.12, CI = -0.55–0.80, *p* = .72) and on RAVLT delayed recall (*ES* = 0.13, *p* = .60; adjusted difference in *z*-scores = 0.05; CI = -0.61-0.72, *p* = .87) in **estimated premorbid IQ** as measured on vocabulary test (adjusted difference in *z*-scores = 0.17; CI = -0.39–0.73, *p* = .55). • After adjusting for residual mood **symptoms** and **education** in multiple linear regressions, differences in cognitive performance between the functionally recovered and unrecovered patients were no longer statistically significant. • **Ethnicity**, **time since last episodes, gender**, being treated with **antidepressants**, number of current **psychotropic medication**, **executive functions** (no statistic reported) and **depressive symptoms** (*p* = .349), **comorbid psychiatric disorder** (*p* = .543), and **BD subtype** (*p* = .411) did not remain significant in the regression model. When **time since last episode and depressive symptoms** were adjusted for in

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
				the regression model the significance level of **marital status** became insignificant (p = .06). Significant positive associations: • **Employment status** – recovered group more likely to be employed (χ^2 = 23.5; p < .0001) • **Education, marital status** (married) and **ethnicity** (Caucasian) showed positive association with recovery when recovered and unrecovered participants were compared: Unrecovered patients had fewer years of education (t = -3.4 p = .001), were less likely to be married (χ^2 = 5.7 p = .02), and were more often African American than Caucasian (FET = 9.0, p = .03). • **Education** [OR = 1.45, *95% CI* (1.11; 1.90), p = .006] and **marital status** (OR = 4.27, 95% CI (1.03; 17.68), p = .045] were significant predictors of recovery in the regression model adjusted for **comorbidities, BD subtype, illness duration** (see under negative significant association) and **depressive symptoms.** • Unrecovered patients performed significantly less well than recovered patients in **executive functions** as measured on FAS (unadjusted ES = 0.54, p =.03) and had poorer **estimated premorbid IQ** as measured on vocabulary test (unadjusted ES = 0.47, p = .05).
M1	Prospective cohort study with 36 month FU period (relevant results reported at 36 month- **cross-sectional data**)	**1) Demographic factors:** employment (competitive employment) and residential status (independent housing) **1) Clinical factors:** Global psychiatric	1) The relationships among six major outcomes were assessed with simple bivariate (Pearson Product Moment) correlations at 36 months.	Significant negative associations: • Levels of **symptomatology** showed negative correlation with social-functional recovery (operationalised as quality of life/overall life satisfaction-higher score indicates higher satisfaction) r = -.34; p < .05. No association: • Levels of symptomatology was not associated with occupational and residential recovery (residential recovery r = .03, ns; occupational recovery r = -.13, ns.)

Study keys	Study design	Variables examined in association with recovery	Data analysis	Key findings
		symptomology and substance abuse 2) **SFR components:**: Regular contact with peers who are not substance abusers and quality of life		• Levels of symptomatology was not associated with social-functional recovery (operationalised as frequency of social contact with non-abusers; r = -.11, ns). • **Substance abuse** was not associated with occupation residential (occupational recovery: r = .08, ns; residential recovery: r = -.09, ns.) or social-functional recovery (regular contact with non-abusers r = .11, ns; quality of life: r = .15, ns.). • Occupational (r = 0.3) and residential recovery (r = 0.13) was not associated with social-functional recovery (operationalised as quality of life/life satisfaction; ns.). • Social recovery (contact with non-abusers) was not associated with quality of life/general life satisfaction (r = .23, ns) Significant positive associations: • Occupational and residential recovery were associated positively with each other (r = .32, p < .05) and with social-functional recovery (operationalised as frequency of social contact with non-abusers) occupational recovery: r = .32, p < .05; residential recovery: r = .29, p < .05)
M2	Cross-sectional study	**1) PR**	1) Correlations were used to test the relationship between personal and occupational and residential recovery.	Significant negative association: none reported No association: • No association was found between occupational and residential and personal recovery (statistics not reported). Significant positive association: none reported

Abbreviations: ANOVA: Analysis of Variance; AOM-400: Aripiprazole 400 mg once monthly; BD: Bipolar Disorder (BD-I: Type-I; BD-II: Type-II); BL: Baseline Assessment; BMI: Body Mass Index; CBT: Cognitive Behavioural Therapy; CFI: Comparative Fit Index; CG: Control Group; CI: Confidence Intervals; d=Cohen's d (effect size); df: Degrees of freedom; EG: Experimental Group; ES: Effect Size; F: F-statistics; FET: Fisher's exact test; FU: Follow-up assessment; M= mean; NNFI: Non-normed fit index; ns: not significant; OR: Odds ration; PD: Personality Disorder; PR: Personal Recovery; r: Pearson's correlation coefficient; RMSEA: Root Mean Square Error of

Approximation; SD: Standard Deviation; SE: Standard Error; SES: Social Economic Status; SFR: Social-functional Recovery; SRMR: Standardised Root Mean Square Residual; t: T-test; TAU: Treatment as usual; WRS: Wilcoxon Rank Sum;

Measures: BIPQ: Brief Illness Perception Questionnaire (Broadbent et al., 2006; Lobban et al., 2013); BRQ**:** Bipolar Recovery Questionnaire (Jones et al., 2013); CVLT: California Verbal Learning Test (Delis et al., 1987); GAF: Global Assessment of Functioning (American Psychiatric Association, 1987, 2000, 2003); FAS: Controlled Oral Word Association test with three word-naming components (trials F, A, and S; Benton & Hamsher,1978); FAST: Functioning Assessment Short Test (Rosa et al., 2007); LIFE-RIFT: Longitudinal Interval Follow-Up Evaluation-Range Impaired Functioning Tool (Leon et al., 2000); MADRS: Montgomery-Asberg Depression Rating Scale (Montgomery & Asberg, 1979); MSIF: Multidimensional Scale of Independent Functioning (Jaeger, Berns, & Czobor, 2003); PTGI: Post Traumatic Growth Inventory (Tedeschi & Calhoun, 1996); RBD: Resilience Questionnaire for BD (Echezarrage et al., 2017); RAVLT = Rey Auditory Verbal Learning Test (Lezak, Howieson, Loring, 2004); RS-25: The Resilience Scale-25 (Spanish version; Las Hayas et al., 2014), SCWT: Stroop Color-Word Interference Test (Golden, 1978); SRS: Stages of Recovery Scale (Song & Hsu, 2011); TMT: Trail Making Test (Reitan, 1958); WHOQOL-BREF: Quality of Life Scale of the World Health Organisation Quality of Life Assessment-shorter version; (Fleck et al., 2000); WCST: Wisconsin Card Sorting Test (Heaton, 1981); WSAS: Work and Social Adjustment Scale (Mundt et al., 2002)

* There is a discrepancy in the S4 paper – in the text, the authors reported *Wald* = 7.52 and in the table it is reported as *Wald* = 7.57

Table A.3: Demographic characteristics

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
P1	BD-I, BD-II, cyclothymia and BD-NOS	DSM-IV-TR SCID	*N* = 87 (BD-1: *n* = 55 (63.2%), BD-II: *n* = 29 (33.3%), cyclothymia: *n* = 2 (2.3%), BD-NOS *n* = 1 (1.1%)	1) Aged >18 years; 2) UK-based; 3) Self-reported diagnosis of BD 4) a research diagnosis of BD, confirmed via SCID	1) Closing the browser while completing the survey (non-completers) 2) Participants who answered every catch item in the survey incorrectly	*M* = 44.46 (*SD* = 12.16)
P2	BD-I, II, Cyclothymic/others, or NOS	DSM-IV Clinician confirmed	BL: *N* = 125 FU: *n* = 60	1) Clinician-confirmed diagnosis of BD-I, II, Cyclothymic/others, or NOS according to DSM-IV criteria 2) Aged 18–65 years old 3) Sufficiently fluent in Spanish to be able to complete the measures 4) Providing voluntary informed consent.	1) Clinically serious multiorganic disorder, acute psychosis, or cerebral organic deterioration that would prevent the person from completing the questionnaires	BL age: *M* = 46.13 (*SD* = 10.89); BL Age of onset: *M* = 29.46 (*SD* = 10.79) FU age: *M* = 45.13 (*SD* = 11.06) FU age of onset: *M* = 30.34 (*SD* = 10.51)
P3	BD and general population sample without BD	DSM-IV No information on method	At BL: BD group*: n* =125 CG: *n* = 107 For ANOVA group comparison: BD recovered: *n* = 28 BD not recovered: *n* = 83 CG: *n* = 71	1) Confirmed diagnosis of BD according DSM-IV (for BD group only) For both BD and control groups: 1) Age 18–65 years; 2) Sufficient fluency in Spanish for completing the battery of tests 3) Informed consent for voluntary participation after	For both BD and control groups: 1) Clinically serious multi-organic disorder, acute psychosis, or cerebral organic deterioration that would prevent the participant from completing the questionnaires	BD group BL: *M* = 46.13 (*SD* = 10.89); CG BL: *M* = 35.42 (*SD* = 10.61) Participants in the CG were younger than BD participants (*t* = -7.56, *p* = <.05)

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
			At FU: BD group: *n* = 63 CG: *n* = 54	being personally informed by his/her therapist		Age of BD onset: *M* = 29.46 (*SD* = 10.79) BD group FU: *M* = 45.13 (*SD* = 11.06) CG age at FU not reported.
P4	BD and general population sample without BD	DSM-IV Therapist confirmed	BD group: *n* = 120 CG: *n* = 97	1) Confirmed diagnosis of BD according DSM-IV (for BD group only) For both BD and control groups: 1) Age 18–65 years; 2) Sufficient fluency in Spanish for completing the battery of tests 3) Informed consent for voluntary participation after being personally informed by his/her therapist	For both BD and control groups: 1) Clinically serious multi-organic disorder, acute psychosis, or organic cerebral deterioration that would prevent them from completing the tests	BD group: *M* = 45.83 (*SD* = 10.76) CG: *M* = 35.25 (*SD* = 10.49) Participants in the CG groups were significantly younger (*t* [215] = 7.28, *p* <0.001), and the effect size was medium (*r* = 0.44).
P5	Homeless individuals with BD	DSM-IV-TR Psychiatrist confirmed	*N* = 216	1. Age over 18 years; 2. Absolutely homeless or precariously housed; 3. Diagnosis of BD by a psychiatrist based on the DSM-IV-TR; 4. Ability to speak French.	1. Reduced capacity to provide consent	*M* = 39.7 (*SD* = 9.3)
P6	BD- in remission	ICD-10 No information on method	*N* = 185	1. Diagnosis of BD as per ICD -10 criteria;	1. Patients with comorbid intellectual disability,	*M* = 40.5 (*SD* = 11.26) *Range*:19-63

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
				2. Aged between 18 and 65 years; 3. Have an illness of at least 1 year. 4. Currently in euthymic state YMRS and HDRS scores of <7.	organic brain disease, and chronic physical illnesses.	
P7	BD- in remission	DSM-IV No information on method	*N* = 185	1. Diagnosis of BD as per the DSM-IV, 2. Aged between 18 and 65 years 3. Currently in euthymic state (<7 on YMRS and HDRS).	1. Participants with comorbid intellectual disability.	*M* = 40.5 (*SD* = 11.26) *Range*:19-63
P8	BD Type I, II-clinically stable	DSM-IV SCID	*N* = 60	1.Verified diagnosis (SCID-DSM-IV) 2. Aged 18-65 years old. 3. Sufficient fluent in English.	1. Current acute episode of major depression or mania (or experienced in a month prior to assessment).	*M* = 42.37 (*SD* = 11.42) *Range*: 19-63
P9	BD Type I and II-clinically stable	DSM-IV SCID	*N* = 67 EG: *n* = 33 CG: *n* = 34 *n* = 45 (at 12 months FU) EG: *n* = 22 CG: *n* = 23	1. DSM-IV diagnosis of primary BD with onset in past 5 years. 2. Sufficient understanding of written and spoken English in order to provide consent, engage with interviews and use the intervention; 3. Aged between 18 and 65 years.	1. Manic, hypomanic, depressed or mixed episode currently or in the past 4 weeks.	EG: *M* = 38.3 (*SD* = 12.8) CG: *M* = 39.9 (*SD* = 10.4)
P10	BD Type I or II-in remissison	DSM-IV-TR SCID	*N* = 75	1. Aged between 18 and 65	1. Hospitalisation in the previous 6 months.	*M* = 45.25 (*SD* = 9.73)

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
			n = 54 in decision tree analysis (training set)	2. Being ethnic Chinese and able to communicate. 3. Being in clinical remission for at least 6 months (HAM-D and YMRS < 7).		
S1	BD- 1 year post hospitalisation	DSM-III-R criteria checklist	*N* = 137	1.Ability to read and comprehend English, 2. Aged between 18 and 40 years. 3. Having a medical chart diagnosis, 4. Being at risk for future violence.	1.Hospitalization for more than 145 days, 2. Being under commitment for more than 21 days.	*M* = 29.68 (*SD* = 6.18)
S2	BD Type I, admitted for first manic episode	DSM-III M.I.N.I	*N* = 13 at BL *n* = 10 at 6 months	1. Adults (18-65 years old) 2. Meeting DSM-IV criteria for a current manic episode.	1. Affective episode resulting from unstable medical or neurological disorder or acute substance intoxication or withdrawal (determined by symptom resolution in 72 hours).	*M* = 26.7 (*SD* = 9.9) *Range*:18-53
S3	BD Type I, current manic episode	DSM-IV-TR No information on method	*N* = 169	1.Diagnostic criteria: DSM-IV Bipolar I diagnosis with an index/current manic episode 2. Manic symptom score (YMRS)≥15 3. 18 years or older 4.inpatient or outpatient treatment of the current episode	Not meeting the inclusion criteria.	*M* = 42.5 (*SD* = 12.7)

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
S4	Difficult-to-treat patients with BD in euthymia	DSM-IV-TR SCID	*N* = 420 Functionally remitted: *n* = 221 impaired: *n* = 199	1) Verified BD diagnosis (DSM-IV-TR) 2) In euthymia, defined as YMRS < 6 and HAM-D < 8 3) aged between 18 and 70 years old	1) Current diagnosis of substance abuse or dependence 2) History of mental retardation or any clinical condition that could interfere in the interview 3) Estimated IQ lower than 85	Functionally remitted: *M* = 38.4 (*SD* = 11.1) Functionally impaired: *M* = 44.5 (*SD* = 10.1) The impaired group was significantly older (*t* = -5.85; *p* < .001)
S5	BD Type I-First episode psychotic mania patients	DSM-III-R RPMIP	*N* = 87 at BL *n* = 56 at 6 month FU (46 in regression model) *n* = 49 at 12 months FU (43 in regression model)	1. Age of onset of first psychotic episode to be between 16 and 45 years. 2. Meeting DSM-III-R criteria for a manic episode with psychotic features in the context of a BD. 3. Being a resident of the catchment area (western suburbs of Melbourne). 4. Sufficient command of English. 5. To be able to provide written consent form.	1. Psychotic episode caused by substance abuse, withdrawal (symptoms resolving within the expected period of acute intoxication or withdrawal) or medical illness (determined by medical evaluation). 2. IQ below 70. 3. Previous psychiatric admission, previous substantial antipsychotic or mood stabiliser treatment (>6 month).	*M* = 22.1 (*SD* = 3.5) Age at onset of psychotic symptoms: *M* = 22.1 (*SD* = 3.6)
S6	BD Type I and Type II in remission	DSM-IV-TR Psychiatrist/Psychologist verified	*N* = 55 EG: *n* = 32 CG: *n* = 23	1. Verified diagnosis of BD I or II 2. Age between 18-65 3. To be in remission for a minimum of 1 month=	1. Diagnosis of personality disorder, schizophrenia or other psychotic conditions. 2. Organic mental disorders, deafness, mental retardation.	*M* = 43.58 (*SD* = 11.34) Age at onset: *M* = 24.61 (*SD* = 12.68)

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
				HAMD score ≤7 and YMRS ≤6.	3. Psychoactive substance dependence.	EG: *M* = 43.43 (*SD* = 11.14), CG: *M* = 43.74 (*SD* = 11.55)
S7	BD Type I- current episode mania or mixed	DSM-III-R SCID	*N* = 56 Personality disorder group: *n* = 27; No personality disorder group: *n* = 29 (*n* = 52 for survival analysis *n* = 42 for logistic regression)	1. Hospitalisation and meeting BD diagnosis (manic or mixed). 2. Age between 18 and 65. 3. Ability to communicate in English 4. Residents within the Cincinnati metropolitan area. 5. Providing written informed consent. 6. Participating in SCID interview- patient and personality disorders version.	1. Psychiatric symptoms resulted entirely from acute alcohol and drug intoxication, withdrawal or acute medical illness. Determined by medical examination and rapid symptom resolution after the medical event.	Personality disorder group *M* = 34 (*SD* = 12) No Personality disorder group: *M* = 31 (*SD* = 13).
S8	BD Type I, II, NOS	DSM-IV-TR Axis I Disorders SCID	*N* = 516 *n* = 443 BD-I *n* = 71 BD-II *n* = 2 BD-NOS	1. Adult inpatients aged between 17 and 80 years 2. Verified SCID diagnosis of BD according to DSM-IV criteria and a minimum illness duration of 6months	1. Not meeting the inclusion criteria.	*Range:* 17-80
S9	BD Type I, II, NOS	DSM-IV SCID	*N* = 78 *n* = 66 BD-I *n* = 4 BD-II *n* = 8 BD-NOS *n* = 29 euthymic *n* = 8 depressive episode	1. Diagnosis of BD I or II (NOS) 2. English as primary language. 3. Age between 18 and 59.	1. Co-occurring medical condition that may cause or contribute to disability. 2. BD-NOS superimposed upon another Axis-I diagnosis.	*M* = 35.8 (*SD* = 10.23)

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
			n = 5 manic episode others were sub-syndromal		3. Positive toxicology screening for substance abuse at the time of neurocognitive assessment	
S10	BD-I	DSM-III M.I.N.I.	Randomized-withdrawal study: *N* = 229 EG: *n* = 116; EG recovery group: *n* = 35 CG: *n* = 113 CG recovery group: *n* = 28 Open-label maintenance study: *N* = 402 New participants: *n* = 321 New participants recovery group: *n* = 116 Roll-over participants: *n* = 81 Roll-over participants recovery group: *n* = 35	1) Clinical diagnosis of BD-I (DSM) verified by MINI 2) Experienced ≥1 previous manic or mixed episode with manic symptoms of sufficient severity to require hospitalization, treatment with a mood stabilizer, or treatment with an antipsychotic agent 3) Age 18–65 years Combined sample of a two trials: Randomized-withdrawal study: 4a) YMRS score >20 Open-label maintenance study: 4b) no YMRS criterion 5) New participants meeting criterion 2 or participants who had completed the maintenance phase of the randomized-withdrawal study (EG or CG) without recurrence of a mood episode and meeting criterion 2.	Combined sample of a two trials: Randomized-withdrawal study: 1) Participants with a mixed or depressive episode Open-label maintenance study: 1) Participants with a depressive episode	Randomized-withdrawal study: *M* = 40.6 (*SD* = 11) Age at first manic episode: *M* = 25.0 (*SD* = 10.1) Open-label maintenance study: *M* = 41.1 (*SD* = 11.8) Age at first BP-I diagnosis: *M* = 29.1 (*SD* = 11.7)*

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
S11	BD-I (46.3%), BD-II (53.7%)	DSM-IV SCID	BL and FU: *N* = 55 Participants with more (*n* = 33) or less (*n* = 22) than five previous affective episodes.	1) Diagnosis of BD-I or BD-II according to DSM-IV 2) Age between 18 and 65 years, 3) A FU period of more than 48 uninterrupted months in the Program 4) Euthymic (defined by HMD ≤9 and YMRS ≤8) for at least 8 weeks both at BL and FU.	1) History of substance abuse/dependence, mental retardation, neurological disease, or any unstable clinical condition that could affect functional outcome.	BL: *M* = 43.64 (*SD* = 12.62; *median* = 44; *range* = 43)
S12	BD Type I with index/ current manic episode	DSM-IV-TR Psychiatrist verified	*N* = 169 *n* = 46 mania with mixed features *n* = 123 mania without mixed feature	1.Diagnostic criteria: DSM-IV Bipolar I diagnosis with an index/current manic episode 2. Manic symptom score (YMRS)≥15 3. 18 years or older 4. Inpatient or outpatient treatment of the current episode.	Not meeting the inclusion criteria.	Mania without mixed features: *M* = 41.85 (*SD* = 12.66) Mania with mixed features: *M* = 44.35 (*SD* = 13.07)
S13	BD Type I-first hospitalisation for manic or mixed episode	DSM-IV SCID-P	*N* = 42 Good outcome group: *n* = 20 Poor outcome group: *n* = 22	1. DSM-IV criteria for BD. 2. Aged between 16-45 years. 3. No prior psychiatric hospitalisations. 4. Last than 1 month of prior psychotropic medication use. 5. English speaking 6. Living within 50 miles of Cincinnati metropolitan region.	1. Psychiatric symptoms are secondary to acute medical illness, determined by medical examination. 2. Symptoms result from acute intoxication or withdrawal, determined by symptom resolution within the expected period. 3. Mental retardation (IQ < 70).	Good outcome group: *M* = 25 (*SD* = 7) Poor outcome group: *M* = 27 (*SD* = 7).

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
				7. Provision of written informed consent (including parental consent if under 18).		
S14	BD Type I or II- clinically stable	DSM-IV-TR SCID	*N* = 65 (BDI: *n* = 42, BDII: *n* = 23) Recovered group: *n* = 30, Unrecovered group: *n* = 35.	1. Male or female outpatients; 2. Age 18–65 years; 3. English as primary language; SCID-supported DSM-IV diagnosis of type I or II BD; 4. Having no history of hospitalization within the past 3 months 5. Currently clinically stable, supported by MADRS scores ≤14 (no more than mildly depressed) and MRS scores ≤11 (no more than mildly hypomanic) at the time of assessment.	1. Meeting DSM-IV criteria for a substance use disorder within 30 days. 2. Given a schizoaffective diagnosis within the past year; 3. Pregnancy; 4. Severe and unstable medical condition; 5. Neuropsychiatric illnesses associated with cognitive impairment; 6. Previous brain injury or severe cerebral trauma; 7. Any history of electroconvulsive treatment; 8 IQ <70	*M* = 40.1 (*SD* = 13.2) Recovered group *M* = 35.1 (*SD* = 12.3), Unrecovered group: *M* = 44.4 *SD* = 12.6).
S15	BD Type I- with (hypo)manic episode in past 3 months, but achieved clinical recovery	DSM-IV SCID	*N* = 65 *n* = 35 recovered all 4 domains *n* = 62 recovered at least one domain: (n=54 recovery in the friends domain;	1. Confirmed diagnosis of BD-I 2. Mania or hypomania within 3 months of study enrolment. 3. Treatment for the index manic or hypomanic episode with a mood stabilizer or combination of mood stabilizers such as lithium, divalproex sodium,	1. Significant alcohol or substance abuse or dependence within the past 3 months; 2. Rapid cycling within the year prior to enrolment or prior to the index episode; 3. Organic mood disorder.	*M* = 36.8 (*SD* = 11.3) *Range*: 18-63

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
			n = 50 recovery in the family domain; *n* = 53 recovery in home duties; *n* = 50 recovery in the work/school domain.	carbamazepine or a second generation antipsychotic. 4. Worked in the year prior to the index manic episode, with work defined broadly to include a variety of full-time equivalent "primary life roles" such as work for pay, student status, and homemaker role. 5. Maintained ≥80% medication adherence at each study visit. 6. Achieved clinical recovery YMRS score < 7) during the first phase (6 months FU) and maintained symptomatic recovery for 6 weeks.		
O1	BD Type I, with manic episode in past 6 months, but achieved clinical recovery	DSM-IV- SCID	*N* = 79 n = 45 occupationally recovered at BL *n* = 34 occupationally unrecovered at BL *n* = 25 participants at 3 months FU, who were unrecovered at BL	1. Age of 18-65 years 2. DSM-IV diagnosis of BD-I 3. Achieved symptomatic recovery by the 6 moths FU assessment (Phase 1)- eligible to take part in the 9 months long second phase. 4. Had a manic episode in past 6 months. 5. History having worked in year prior to manic episode.	1. Significant alcohol or substance use disorder (abuse ⁄dependence) within the past three months. 2. Rapid cycling within the year prior to the manic episode. 3. Organic mood disorder (e.g., head trauma or cerebrovascular accident preceding their first manic episode).	Recovered group: *M* = 35.02 (*SD* = 11.53); Unrecovered group: *M* = 38.18. (*SD* = 11.83)

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
					4. Not being able to attain symptomatic recovery after six months of registration in the study. 5. Meeting criteria for a depressive episode at 6 months FU. 6. Developed substance use disorders at any point during the FU period.	
O2	BD Type I-clinically stable 1 year after acute episode	DSM-IV SCID- I/P	*N* = 51 *n* = 21 good occupational group *n* = 30 poor occupational group	1. Hospitalisation for an acute affective episode. 2. Aged between 18 and 59. 3. Mild range symptoms (defined as 17 on the first 17 items of the HRS-D and 15 on the first 10 items of the Clinician-ARSM).	1. Mental retardation, neurological disease or serious medical illness. 2. Lack of fluency in English.	*M* = 35.47 (*SD* = 11.39)
O3	BD Type I –first hospitalisation with mania	DSM-III-R SCID	*N* = 27	1. First psychiatric hospitalisation with psychotic or manic symptoms. 2. Minimum age of 18. 3. Ability to communicate in English. 4. Provision of informed consent. 5. Completion of SCID for DSM-III-R	1. Patients with BD in the depressed or mixed states.	*M* = 32.2 (*SD* = 14.1)

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
O4	BD Type I – first hospitalisation with psychosis	DSM-III-R SCID-P	*N* = 60	1. DSM-III verified diagnosis 2. First lifetime admission for a psychotic disorder. 3. Aged ≥ 18. 4. Providing informed consent.	1. Organic psychosis 2. Dementia 3. IQ < 70	*M* = 31 (*SD* = 12.4)
O5	BD Type I-first manic or mixed episodes	DSM-III –R (SCID-P) DSM-IV (diagnosis updated)	*N* = 166 *n* = 151 at 24 months FU	1. Aged between 18 and 75. 2. DSM criteria for mixed or manic episode. 3. Provided written informed consent.	1. Current substance withdrawal. 2. Delirium 3. Previous psychiatric hospitalisation, unless for detoxification only. 4. Documented IQ < 70. 5. Ill for more than 1 year. 6. Previous treatment with antipsychotic or mood stabilizer for more than 3 months total.	*M* = 32.5 (*SD* = 13.7) *Range:* 18-72, *Mdn* = 28
O6	BD Type I or II-clinically stable	DSM-IV SCID	*N* = 65 (64.6% BDI; 35.4% BDII) Recovered group: *n* = 28, Unrecovered group: *n* = 37	1. Male or female outpatients; 2. Age 18–65 years; 3. English as primary language; 4. SCID verified DSM-IV diagnosis of type I or II BPD; 5. Having no history of hospitalization within the past 3 months. 6. Currently clinically stable, supported by MADRS scores ≤14 (no more than mildly	1. Meeting DSM-IV criteria for a substance use disorder within 30 days. 2. Meeting DSM-IV criteria for schizoaffective diagnosis within the past year; 3. Pregnancy; 4. Unstable medical condition; 5. Neuropsychiatric illnesses associated with cognitive impairment;	*M* = 40.1 (*SD* = 13.2) Recovered group: *M* = 39.1 (*SD* = 14.3) Unrecovered group: *M* = 40.8 (*SD* = 12.5)

Study key	Diagnosis of participants	Diagnosis verification	Sample size used to evaluate recovery	Inclusion criteria	Exclusion criteria	Reported details on BD sample age (in years)
				depressed) and MRS scores ≤11 at the time of assessment.	6. Previous brain injury or severe cerebral trauma; 7. Any history of electroconvulsive treatment (ECT); 8. IQ <70	
M1	Individuals with co-occurring bipolar and substance use disorders	DSM-III-R SCID	*N* = 51	1) Diagnosis of BD 2) Active SUD diagnosis (abuse or dependence of alcohol and/or other drugs) according to DSM-III-R criteria within the past 6 months. 3) Aged between 18-60 4) Absence of additional medical conditions or mental retardation 5) Willingness to provide written informed consent (substituted from Drake et al., 1998)	Not meeting the inclusion criteria.	*M* = 37.5 (*SD* = 9.6)
M2	BD- in remission	DSM-IV-TR SCID	*N* = 75	1. Aged between 18 and 65 2. Diagnosis of BD. 3. Being ethnic Chinese and able to communicate in Cantonese. 4. Being in clinical remission for at least 6 months (HAM-D and YMRS ≤ 7).	1. Hospitalisation in the previous 6 months.	*M* = 45.25 (*SD* = 9.73)

Abbreviations: BD-I/II: Bipolar Disorder Type I/II; BD-NOS: BD not otherwise specified; BL: baseline; CG: Control Group; DSM: Diagnostic and Statistical Manual of Mental Disorders: DSM-III-R: 3rd edition revised (American Psychiatric Association, 1987), DSM-IV-TR: 4th edition text revised (American Psychiatric Association, 2000); EG: Experimental Group; FU: follow-up; ICD-10: International Classification of Diseases-10; *M*: mean, *Mdn*: median, *N*: number (sample size); *p:* significance value; *r*: Pearson's correlation coefficient; SCID: Structured Clinical Interview for DSM; *SD*: Standard deviation; *t*: t-test

Measures: ARSM: Administered Rating Scale for Mania (Altman, 2994); DSM-III-R criteria checklist (J'anca A. & Helzer, 1990); MADRS: Montgomery-Asberg Depression Rating Scale (Montgomery & Asberg, 1979); M.I.N.I: The Mini-International Neuropsychiatric Interview (Sheehan et al., 1998), DSM-III-R SCID (Spitzer, First, Gibbon, & Williams, 1990); DSM-IV SCID-I/P: Axis I Disorders – Patient Edition (First, Spitzer, Gibbon, & Williams, 1995) DSM-IV SCID-TR (First et al., 2002); ; HAM-D/HDRS: Hamilton Depression Rating Scale (Hamilton, 1960);; MRS: Mania Rating Scale (Endicott & Spitzer, 1978); PST: Present State Examination; YMRS: Young Mania Rating Scale (Young, Biggs, Ziegler, & Meyer, 1978); RMIP (McCorry, Copolov, Singh, 1990; McCorry, Singh, Copolov, Kaplan, Dossetor, van Riel, 1990)

*Age reported for larger sample sizes (Randomised withdrawal study: *N* = 266; Open-label maintenance study: *N* = 464) rather than the sample included in the present post-hoc analysis