MS. HAILEY TREMAIN (Orcid ID : 0000-0001-6304-1940)

DR. JAN SCOTT (Orcid ID : 0000-0002-7203-8601)

PROF. GREG MURRAY (Orcid ID : 0000-0001-7208-5603)

Article type : Review

The influence of stage of illness on functional outcomes after psychological treatment in bipolar disorder: A systematic review

Hailey Tremain*[1]

Kathryn Fletcher[1]

Jan Scott[1, 2]

Carla McEnery[1, 4, 5]

Michael Berk[3, 4, 5, 6]

Greg Murray[1]

[1]Centre for Mental Health, Faculty of Health Arts and Design, Swinburne University, Melbourne Australia

[2]Academic Psychiatry, Institute of Neuroscience, Newcastle University, Newcastle UK

[3] IMPACT Strategic Research Centre, School of Medicine, Barwon Health, Deakin University, Geelong, VIC, Australia

[4] Orygen, The National Centre of Excellence in Youth Mental Health, Parkville Australia

[5] Centre for Youth Mental Health, The University of Melbourne, Melbourne
[6] The Department of Psychiatry and the Florey Institute for Neuroscience and Mental Health, the University of Melbourne, Parkville, VIC, Australia

* Correspondence, Hailey Tremain htremain@swin.edu.au, c/o Greg Murray PO Box 218 John St Hawthorn VIC, 3122 Australia

Abstract

Objectives: The aim of this study was to advance understanding of stage of illness in bipolar disorder (BD), by interrogating the literature for evidence of an influence of stage of illness on functional (i.e., non-symptom) outcomes following psychosocial intervention.

Methods: A systematic literature search following PRISMA guidelines was conducted to identify empirical studies of psychosocial interventions for established BD. To investigate stage as a predictor of three functional outcomes (general/social functioning, cognitive functioning, quality of life [QoL]), study samples were dichotomized into earlier and later stage using proxy measures identified in existing staging models. Findings were integrated using data-based convergent synthesis.

Results: 88 analyses from 62 studies were identified. Synthesis across studies suggested that psychosocial intervention was more likely to be effective for general functioning outcomes earlier in the course of established BD. No stage-related differences were found for cognitive or QoL outcomes. Exploratory investigations found some evidence of an interaction between specific intervention type and stage of illness in predicting outcomes.

Conclusions: A novel systematic review provided preliminary evidence that benefits for general/social functioning may be more pronounced in earlier versus later stages of established BD. The review also generated hypotheses about a potential three-way interaction, whereby specific psychosocial interventions may be best placed to target functional outcomes in earlier versus later stage BD. The strength of conclusions is limited by the overall low quality and significant heterogeneity of studies. Further research is urgently required to understand the impact of illness stage on the effectiveness of psychosocial interventions.

Key words: Bipolar Disorder, Psychotherapy, Social Functioning, Bipolar Disorder/psychology, Staging models, Treatment Outcomes

Data Availability: The data that support the findings of this study are available from the corresponding author upon reasonable request

1. Background

Clinical staging models revolutionised the conceptualisation and treatment of medical conditions such as cancer, and consequently, research has begun to explore their application to mental health conditions [1]. A growing body of research now suggests that, for at least a proportion of individuals, bipolar disorder (BD) can be characterized by progressive changes over time, with clinical and functional illness characteristics shifting across the course of illness [2-6]. Improved understanding of the nature and measurement of staging in BD has potential to support better tailoring of current interventions, and ultimately to develop new interventions targeting specific stages. Of particular relevance to this review, the staging heuristic may also help clinicians and consumers identify the most meaningful targets for interventions within a given stage. However, the literature on staging in BD is immature and limited in important ways. Below, we briefly review what is known about staging in BD to highlight the key gaps to be addressed within this systematic review.

1.1 Staging models in bipolar disorder

Clinical staging models categorise key features related to illness trajectory in BD within putative stages. There is as yet no consensus on the relative merits of transdiagnostic versus disorder-specific approaches to BD staging, but it has been proposed that transdiagnostic approaches may offer most utility in the at-risk and prodromal phases, while syndrome specific models are of most benefit once distinguishable syndromes emerge ('established BD' - the focus of this study)[7]. Two BD specific staging models have garnered most attention to date. A model proposed by Berk [8] uses number of mood episodes and hence recurrence as a proxy of stage, while a model proposed by Kapczinski [9] operationalises stage as reflecting functional impairment. Berk's and Kapczinski's models differ in their developmental starting point - Berk's model commences with an asymptomatic at-risk stage, while Kapczincski's commences with established BD (for more detail, see Appendix A).

A further important distinction pertinent to staging is that of disease progression versus disease extension. Scott and Henry suggest that while disease progression refers to worsening of the disorder itself, disease extension refers to the spread of the disorder to have wider reaching impacts on multiple outcomes (analogous to the spread of Hodgkin's disease to other body systems)[10]. The concept of disease extension, while currently underdeveloped in psychiatric research, may be informative for clinical decisions.

While no model of clinical staging in BD has been empirically validated, an International Society for Bipolar Disorders task force review concluded that there is sufficient evidence at this time supporting two stages of established BD, distinguishable based on number of episodes (NoE[11]), and functioning[12]. The task force also proposed that such a dichotomisation is likely to be of heuristic value in the treatment of BD, despite lack of validated boundaries or definitions at this time. Several studies have identified two broad post-onset groups based on features including functioning, NoE and cognitive impairments. [6,13,14]. For example, Rosa and colleagues [6] classified individuals with established BD into early versus late stages, finding the early group to have lower NoE, and less

cognitive and functional impairment. Similarly, Grande reported finding two clusters of participants, differing based on NoE and functioning[14].

1.2 Proxies of stage in established BD

Given the absence of robust evidence for biomarkers and biobehavioural boundaries of putative stages in BD, empirical research has relied on proxies to operationalise stage of illness. Most commonly, NoE has been used [15-19]. While a potentially useful and plausible proxy[1], NoE alone cannot fully capture the complexity of the proposed staging phenomenon [5,9,11], and in particular does not speak to the notion of disease extension. It has been proposed that analogues of physical disease progression in mental disorders may be functional impairment or disability, and cognitive functioning [3,6,9,20], while proxies such as medical comorbidities and some psychiatric comorbidities better reflect disease extension. Importantly though, as reviewed by Tremain, et al. [21] each of these plausible proxies has limitations, and may best be considered in combination to approximate progression via clinical staging. The present review employs those potential proxies with the most support to date (see Table 1), primarily reflecting progression.

1.3 The potential role of staging in psychosocial interventions

Prior research demonstrates that treatments for BD are more effective for some individuals than for others, and stage of illness is a potential moderator of response to treatment [18,22-27]. As summarised by Tremain, et al. [21], a number of studies concord with the staging hypothesis of BD, wherein psychological interventions may be generally less effective as BD advances. However, the focus to date has primarily been on symptom-focused outcomes.

Functional outcomes may have particular relevance to individuals later in the course of established BD [22]. Extending on this, certain interventions may be best positioned to target functional outcomes. For example, Torrent and colleagues compared functional remediation to treatment as usual (TAU) pharmacotherapy and psychoeducation (PE) for those with significant functional impairment and multiple previous mood episodes, mapping onto later stage [28,29]. Participants in the functional remediation group demonstrated greater improvements in functioning than both control groups at 12-month follow-up. These findings demonstrate the differential effectiveness of interventions for those considered at the later stage of BD for improving functioning.

1.4 The current Systematic Review

To our knowledge, only two reviews have examined plausible proxies of stage as moderators of psychosocial interventions for established BD, both using NoE as a proxy and focusing on symptom outcomes [18,30]. Conversely, several reports have reviewed the effectiveness of interventions

[1] We use the qualifier 'plausible' to remind the reader (and ourselves) that 'staging' remains a hypothetical construct in psychiatry. Any proxies are at best consistent with one or more of the various provisional conceptualisations of staging.

for the earliest stages of BD [31-33]. A comprehensive systematic review is required which attends to gaps in the literature through attention to (i) functional (non-symptom) outcomes, and ii) a broader investigation than previously conducted, via an expanded range of plausible proxies derived from empirical evidence and prominent theoretical models.

The present review aims to provide a synthesis of the evidence for stage-based moderation (as operationalised by plausible proxies of earlier versus later stage) of psychosocial interventions for functional outcomes in established BD. The core question of this review therefore is: Does the literature contain an evidentiary signal that stage of illness moderates functional outcomes in adjunctive psychosocial treatment of established BD? A subsidiary clinically-focused question was also explored: Is there evidence for an interaction between stage and intervention type, such that some psychosocial interventions are more beneficial for earlier versus later stage of established BD?

2. Method

The systematic review was registered with the International Prospective Register of Systematic Reviews (PROSPERO, registration number: CRD42016037868). A protocol describing the methods has been published [21]. It was conducted in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) recommendations for systematic reviews [34] and has been reported using PRISMA guidelines [35].

2.1 Inclusion Criteria

2.1.1 Population. Included studies were empirical, comprising participants aged 18 and over with a BD I, BD II, BD-NOS, not elsewhere classifiable, or other specified BD diagnosis (meeting internationally recognised diagnostic criteria, such as in the DSM-IV or 5, or ICD-10) with no restrictions based on characteristics such as psychosis, comorbidity or mood status at study entry.

2.1.2 Interventions. Studies reported the outcomes of any psychological or psychosocial intervention (any non-medical intervention intended to modify symptoms, behaviour, emotional state, or feelings) for BD. Definitions of interventions are provided in Appendix B.

2.1.3 Comparisons. Comparisons were conducted between, rather than within, studies and therefore the included studies did not require direct comparisons between individuals at different putative stages. Instead, samples were dichotomised into earlier versus later stages using plausible proxy measures of stage derived from an extensive review of published models and theories of staging and outlined in Table 1. For clarity, the reported mean, median, mode or proportion of individuals per category for plausible proxies was used to dichotomise samples. However, given that there are no established boundaries of putative stages of BD based on any of the plausible proxies, it is important to recognise that this dichotomisation was employed for heuristic purposes only; we do not intend to reify discrete and defined stages. We labelled these categories broadly as ‘earlier’ and ‘later’, reflecting a provisional, artificial categorisation of illness trajectory, intended to facilitate analyses rather than confirm the existence of such categories outside the present investigation. This dichotomy

was drawn from previous research and expert consensus on the status of the literature [6,12,14], and the cut-offs for later stage were derived either from specific investigations of these variables relevant to staging or previous studies which have indicated that these may be moderators of treatment effectiveness. However, we caution that the dichotomy was a heuristic device for the present study, not an empirical finding, and should not be reified.

[Insert Table 1 about here]

Where available, data about multiple proxies were integrated in assigning a likely 'stage', and if conflicting, hierarchically applied as ordered in Table 1. Of 26 studies with data for multiple proxies, these conflicted in four.

2.1.4 Outcomes. For the present review, 'functional' outcomes were; i) general/social functioning (including functional impairment, daily general, social or occupational functioning, disability or autonomy), ii) cognitive functioning or impairment and iii) QoL. Table 2 outlines the outcome measures to be included in this review.

[Insert Table 2 about here]

2.1.5 Study type. Treatment outcome studies were not restricted to RCTs, and included open pre-post trials and non-randomised controlled trials. Whilst broadening inclusion criteria invites the inclusion of studies with a higher risk of bias, this was addressed via implementation of risk of bias assessment and offset by gains including the implications of review findings for future research and clinical practice.

2.1 Exclusion criteria

Review studies (with no unique primary data) were excluded. Studies of broader clinical groups (e.g., psychosis, serious mental illness, depression) which did not report BD data separately were excluded. Studies with participants under 18 years old were excluded due to the additional complexity in diagnosis and treatment in this population, as well as potential divergent trajectory and phenomenology of early-onset or paediatric BD [93]. Case series and reports were excluded due to very high risk of bias in these designs. Grey literature was excluded as it was not expected to contribute to the fulfilment of the review aims.

2.2 Procedure

2.2.1 Search strategy. Scopus, PsycInfo, PubMed and Web of Science databases were searched to January 2019 to identify potentially eligible studies for inclusion. No date restrictions were applied, and search terms were intentionally broad. An example of the search strategy for Scopus is provided in the Appendix C, adjusted per database.

2.2.2 Screening procedures. Appendix D (Figure A1) demonstrates the PRISMA-P process for this review. After deduplication, both raters (HT and CMcE) screened all titles and abstracts against inclusion criteria for population, study type, interventions and outcomes. The full text of 183 studies was reviewed. Uncertainties regarding eligibility for inclusion were resolved via consensus.

2.2.3 Data extraction and quality evaluation. Data were extracted independently by both raters using a data extraction form developed for this review and piloted prior to extraction (Additional File 1). Risk of bias within randomised trials was assessed with the Cochrane Collaboration tool for assessing risk of bias [94], and within non-randomised studies using the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool [95]. To assess quality, GRADE methodology was applied across studies [96].

2.2.4 Data synthesis. Data were synthesised using narrative synthesis methods, as the broad aims and inclusion criteria precluded meaningful statistical data pooling. A data-based convergent synthesis design[97] focused on the interaction between interventions, stage of illness and outcomes. This method involves describing and synthesising results qualitatively, drawing out themes or categories. Results are reported at the analysis level: that is, per unique analysis as reported in the identified paper. This level of analysis was selected as different analyses from the same trial (e.g., analyses of different outcomes, subsamples or follow-up periods) may have generated divergent results. Multiple analyses may therefore result from the same trial and involve partially or completely overlapping samples. That is, the analyses are not necessarily independent: This complexity is noted where identified.

To aid synthesis, individual analyses were organised into a simple hierarchical grouping structure with outcome domain assessed (general/social functioning, QoL, cognitive functioning) at the highest level, dichotomised stage of the sample (earlier versus later) at the intermediate level, and finding (positive versus negative findings for the benefits of the tested intervention for that stage of BD on that outcome measure) at the lowest level. This logic is replicated in the structure of the Results section below. Within each of the resultant (3 x 2 x 2 = 12) subsections below, the strengths and weaknesses of individual studies and the set of studies is noted and a summary is presented. Where data were available, effect sizes were included (see Tables 3 and 4). Supplementary analyses were conducted taking into account the impact of sample size, risk of bias, and effect size on the results. Given overlaps between brands of psychosocial interventions for BD and their targets [98,99], as well as anticipated small number of studies examining each intervention brand, we chose not to examine different brands separately for the main analyses.

3. Results

The systematic review identified a total of 62 studies, generating 88 analyses for review. Summaries of identified analyses deemed to have samples in earlier (n = 37) and later (n = 51) stages of established BD, are presented in Tables 3 and 4, respectively.

[Insert Table 3 and Table 4 about here]

Given that we dichotomised samples into earlier and later stage using proxy measures, we have provided an example of the distribution of the most used proxy, NoE, in Figure 1. Note that

complete proxy information used for dichotomising samples' stage is reported in Table 3 and 4, and the criteria used are reported in Table 1.

[Insert Figure 1 about here]

As summarised in Figure 2, the number of analyses resulting in positive and negative (absence of significant positive) findings per dichotomised stage was the primary comparison made to explore stage-based moderation. Details of the analyses in each cell below can be found in supplementary materials.

[Insert Figure 2 about here]

3.1 Quality issues and risk of bias

The overall risk of bias for randomised studies was 'low' in 20 percent of studies, 'some concerns' in 41 percent and 'high' in 39 percent, with a Kappa value of 0.90, indicative of a 'near perfect' level of inter-rater agreement (see Appendix E for per domain ratings). For non-randomised studies, risk of bias ranged from moderate to severe. Regarding quality, GRADE ratings were very low for general/social functioning, QoL and cognitive functioning due to study type, risk of bias, imprecision, and publication bias criteria (Appendix F). Comparisons are limited by significant methodological heterogeneity; with broadly ranging interventions, outcome measures, measures of proxies, and study designs.

3.2 Functional outcome 1: General/Social Functioning

The majority of identified studies included one or more general or social functioning outcome measure (55 analyses from 48 trials).

3.2.1 Analyses finding benefits for general/social functioning in the earlier stage. From the identified studies dichotomised as investigating earlier stage samples, 20 analyses found benefits for general/social functioning. Some quality issues need to be considered before drawing conclusions, as can be seen in supplementary materials. Fourteen studies had small sample sizes and risk of bias was rated high in five analyses. However, five well-sized RCTs contribute to confidence in these findings.

3.2.2 Analyses finding no benefits for general/social functioning in the earlier stage.

From the identified studies dichotomised as investigating earlier stage samples, five analyses were identified which failed to demonstrate benefits for general/social functioning. Quality issues impact inferences from these analyses. Four analyses included small samples and just one was rated a low risk of bias; an RCT of moderate size. Therefore, caution is required when interpreting these findings.

3.2.3 Analyses finding benefits for general/social functioning in the later stage. From the identified studies dichotomised as investigating later stage samples, 16 analyses demonstrated benefits for general/social functioning. Many had significant limitations. Seven included small samples and half were rate moderate or high risk of bias. However, three well-sized RCTs at low risk of bias met our criteria, strengthening inferences.

3.2.4 Analyses finding no benefits for general/social functioning in the later stage. From the identified studies dichotomised as investigating later stage samples, 13 analyses failed to demonstrate benefits for general/social functioning. The quality of these studies was generally adequate. Seven had some concerns regarding bias and small samples, however the remaining six were moderate to large sized RCTs, strengthening inferences.

3.2.5 Summary. In summary, in samples dichotomised as earlier stage, the majority demonstrated benefits for general/social functioning, and as seen in Table 3, effect sizes ranged from absent to large, with mean effect sizes in the moderate range (M = .53, SD = .55). Additionally, a similar number of analyses showed benefits for later stage participants, with effects ranging from absent to large and mean effect sizes in the small range (M = .40, SD = .39). Given that in the later stage (relative to earlier), comparatively more analyses failed to find benefits for improved general/social functioning and smaller effect sizes were observed, the present synthesis suggests that psychosocial interventions may be more consistently effective in the earlier than later stage. Considering the quality of studies from which these results are drawn, many were of low quality and high risk of bias per cell, however all cells include one or more moderate to large RCTs, lending weight to this inference.

3.3 **Functional outcome 2: Quality of life**

Twenty-three analyses (from 22 trials) reported on QoL outcomes.

3.3.1 Analyses finding benefits for quality of life in the earlier stage. From the identified studies dichotomised as investigating earlier stage samples, three analyses demonstrated benefits for QoL. Each of these studies suffered from considerable quality issues. Sample sizes were small-moderate, and risk of bias was at least moderate in each study, therefore inferences should be considered tentative.

3.3.2 Analyses finding no benefits for quality of life in the earlier stage. From the identified studies dichotomised as investigating earlier stage samples, five analyses failed to demonstrate benefits for QoL. All had small sample sizes, and two were at high risk of bias, raising concerns about these results.

3.3.3 Analyses finding benefits for quality of life in the later stage. From the identified studies dichotomised as investigating later stage samples, nine analyses demonstrated benefits for

QoL. Eight studies had at least some concerns regarding risk of bias, and three studies had small sample sizes. However, three moderately-sized and well-designed RCTs strengthen inferences.

3.3.4 Analyses finding no benefits for quality of life in the later stage. From the identified studies dichotomised as investigating later stage samples, six analyses failed to demonstrate benefits for QoL. This group of studies had some significant limitations: Five of these studies had small sample sizes, and at least some concerns regarding risk of bias.

3.3.5 Summary. Taken together, there is inconsistent evidence that psychosocial interventions may improve QoL for those the earlier stage of BD, with similar proportions of studies reporting positive and negative findings. More studies had later stage samples and, within these, a larger number of analyses demonstrated positive than negative findings. The mean effect sizes in both earlier (M = .24, SD = .32) and later stage (M =.38, SD = .18) were small. Therefore, our data do not indicate that psychosocial interventions are more effective in the earlier than later stage of BD for improving QoL. Additionally, overall QoL improved following psychosocial interventions in later stage samples in only a few studies, while only individual QoL sub-domains improved in others. The generally low quality of included studies must be considered in drawing conclusions.

3.4 Functional outcome 3: Cognitive functioning

Eleven analyses from seven trials included cognitive outcome measures.

3.4.1 Analyses finding benefits for cognitive functioning in the earlier stage. From the identified studies dichotomised as investigating earlier stage samples, three demonstrated benefits for cognitive functioning. Two studies were rated high risk of bias and had small sample sizes, while the third was a moderate size with some concerns regarding bias and as such inferences cannot be considered be reliable.

3.4.2 Analyses finding no benefits for cognitive functioning in the earlier stage. From the identified studies dichotomised as investigating earlier stage samples, two analyses failed to demonstrate benefits for cognitive functioning. Both were post-hoc analyses of (overlapping) small subsamples at high risk of bias, and as such inferences are tenuous.

3.4.3 Analyses finding benefits for cognitive functioning in the later stage. From the identified studies dichotomised as investigating later stage samples, six analyses demonstrated benefits for cognitive functioning in later stage. Five had small samples and just two were rated a low risk of bias. However, the inclusion of a reasonably sized RCT strengthens inferences.

3.4.4 Analyses finding no benefits for cognitive functioning in the later stage. From the identified studies dichotomised as investigating later stage samples, there were no analyses which failed to demonstrate benefits for cognitive functioning in later stage. However, most studies reported benefits for only some of the assessed cognitive domains.

3.4.5 Summary. Taken together, these results provide inconsistent support drawn from few studies for improvements to cognitive functioning following psychosocial interventions in earlier

stage BD, with a similar number of studies finding and failing to find benefits in earlier stage. Those in later stage BD, with cognitive impairments ranging from mild to severe, experienced improved cognitive functioning in specific domains in all included studies. Effect sizes were larger in the earlier (*M* = .59, *SD* = .29) than later (*M* =.32, *SD* = .16*)* stage, but these were driven by the very small number of analyses included. Therefore, these results do not support the hypothesis that psychosocial interventions are more effective for improving cognitive functioning in earlier than later stage. Our results suggest that those in both earlier and later stage BD may experience at least small improvements in cognitive functioning following psychosocial interventions, however benefits do not appear consistent or across all cognitive domains. Additionally, most studies were of low quality and high risk of bias, limiting our ability to draw firm conclusions.

4. Discussion

This first systematic review of the possible moderating effect of stage of illness on the impact of psychosocial treatments on functional outcomes in established BD identified a range of studies which provided insight into this hypothetical effect, and how it interacts with different outcomes. The three functional outcomes investigated here (social/general functioning QoL and cognitive functioning) were not equally prevalent in the 88 identified analyses (from 62 studies): the most commonly investigated outcome was general/social functioning while QoL and cognitive functioning were examined in only a handful of studies. A meaningful synthesis was achieved by organising the 88 analyses from 62 articles into a heuristic grid of outcome variable by stage by presence vs. absence of significant effects.

4.1 Core question: Stage-based moderation of functional outcomes

Our primary research question was: Is there any evidence that stage of illness moderates functional outcomes in adjunctive psychosocial treatment of established BD? Using the above approach, the core question of this study was provisionally answered in the positive: There was some evidence that benefits of psychosocial intervention for functional outcomes (specifically, general/social outcomes, accounting for the majority of analyses identified in the systematic review) were more pronounced in samples dichotomised as earlier versus later stages of established BD**.** The review also identified important qualifications to this conclusion, however, which are considered below before the study's limitations and future steps for this important research area are canvassed.

As a context for interpreting the present study's conclusions, it is worth noting that the reviewed literature contains substantial evidence that psychosocial interventions are generally beneficial for functional outcome variables in established BD (57 of 88 analyses found benefits). Turning then to the core moderation question of this review, we identified an evidentiary signal that psychosocial interventions are differentially effective at improving general/social functioning for samples dichotomised here to be in earlier versus later stages of BD. Specifically, consistent with

stage-moderation, analyses from multiple studies demonstrated that psychosocial interventions improved general/social functioning in both earlier and later stage BD, but the proportion of analyses finding benefits was higher in the earlier stage compared with the later stage studies and effect sizes were larger in the earlier stage.

The pattern of stage-moderation of general/social functioning outcomes was not repeated for the remaining two outcome variables. The proportion of studies to find benefits for QoL outcomes was greater in the group of identified studies focusing on later stage samples than the group focusing on earlier stage samples, and similarly our synthesis of analyses focusing on cognitive functioning suggested that psychosocial interventions are likely to be effective for improving at least some aspects of cognitive functioning in both the earlier and later stage, with no clear evidence for stage-based differences in benefits. Therefore, our data only support a role for stage-based moderation of general/social functioning.

4.1.1 Within-study moderation. Separate to the core between-study examination of the moderation hypothesis, the systematic review happened to identify a small number of studies that reported an analysis of moderation via variables we employed as plausible proxy measures of stage. Post hoc consideration of these incidental findings helps flesh out the core outcomes of the present systematic review. One study demonstrated that two plausible proxies of stage, anxiety disorder comorbidity and cognitive impairments, partially moderated the positive effects of cognitive remediation on psychosocial and occupational functioning [132]. Two additional studies found that controlling for proxies at baseline altered the significance of improvements to general functioning post-intervention [108,145,146]. However, other studies have found that improvements to QoL [119,129] or functioning [112] were not moderated by variables employed as proxies of stage within this review. While inconsistent in methods and results, these data provide additional support for the hypothesis that proxies of stage may influence the effectiveness of psychosocial interventions, and in parallel with the between-study analyses at the core of this review, support was strongest for functional outcomes.

4.1.2 Prior research findings supporting stage-based moderation. In line with our findings for general/social functioning, several prior studies considering only symptom-focused outcomes have demonstrated reduced effectiveness of psychosocial interventions in later stage samples (as defined by plausible proxies) [15,25,27,38,149]. Conversely, in the present systematic review, no superior effectiveness was evident for QoL or cognitive functioning in earlier, relative to later, stage. It is conceivable, therefore, that these outcomes continue to represent viable targets for psychosocial interventions into the later stage of illness, while symptoms and general/social functioning are most responsive to treatment early in the course of illness.

4.2 Subsidiary question: Interaction between stage and intervention type

Another possibility is that certain intervention types are effective for improving functional outcomes in the later stage of illness. Importantly, in interpreting any signal of differential

effectiveness, the key implication is not that psychosocial interventions are without hope in the later stage. Rather, the core clinical implication is that increased efforts may be needed to identify, tailor and refine interventions at different stages of illness, to maximise benefits. To this end, our subsidiary question was: Is there evidence for an interaction between stage and intervention type, such that some psychosocial intervention types are more beneficial in the earlier versus later stage of established BD?

4.2.1 General/social functioning. The review identified some interventions which appear to improve functional outcomes for those in the earlier stages, without consistent benefit in the later stages and additionally, some interventions which may have benefits into the later stage. For general/social functioning, our results suggest that some intervention types which were effective in the earlier stage, such as PE [103-105,116,125] and cognitive behavioural therapy (CBT) or cognitive therapy (CT) [108,110,120-124] were frequently found to have minimal benefits for those in the later stage [126,131,142,143,145-147,149]. While there are some inconsistencies, it appears that interventions specifically targeting general/social functioning, including metacognitive training [138], functional remediation [28,29,156] and less consistently, cognitive remediation [132] may be beneficial for improving general/social functioning in the later stage. Intensive, longer-term [133,135,140,148] [127,151,152,154] and combination interventions [130,136,137,153] also show promise for improving general/social functioning in later stage BD. Given that not all of these intervention types were trialled in earlier-stage, it is not possible to assess whether these interventions *become* or *remain* effective for improving general/social functioning in later stage.

4.2.2 Quality of Life. For QoL outcomes in the earlier stage, there was most support for the effectiveness of PE interventions [125,157], and some support for CBT [120]. These intervention types were less frequently effective in the later stage [131,142,143,149]. In contrast, other psychosocial interventions appear likely to contribute to improved QoL in later stage, including mindfulness [23], functional remediation [134] and intensive or combination interventions [127,129,130,154] but due to inconsistencies, more research is needed to draw conclusions as well as comparisons with earlier stage.

4.2.3 Cognitive functioning. For cognitive outcomes, a meaningful comparison of intervention types between stages is not possible as there were no negative trials in the later stage and few analyses overall. However, cognitive remediation and mindfulness-based cognitive therapy interventions demonstrated benefit in both earlier stage [107,118] and later stage [141,155], while functional remediation demonstrated benefits in the later stage [28,29], but not consistently in the earlier stage [102,158].

Taken together, the findings of studies included in this review suggest that specific interventions may best be positioned to target functional outcomes in earlier versus later stage. It is important to acknowledge that these conclusions are drawn from relatively few studies, with less systematic exploration possible, and should therefore be considered tentative. Nonetheless, these findings, if confirmed, have significant clinical utility.

4.2.4 Prior research supporting a stage by intervention interaction. Interestingly, results of previous studies examining symptom outcomes concord with these findings. Specifically, the intervention types in the current review (PE, CT and CBT) which showed little consistent benefit in later stage similarly have demonstrated reduced effectiveness in reducing symptoms in later stage BD [15,25,27,38,149]. An interesting prospect for future research therefore emerges; it may be that these types of interventions are effective earlier in the course of illness for both functional and symptom outcomes, while in the later stages a shift is required in both the type of outcome and the interventions best positioned to target these. An implication of this finding links to the primary aspiration for staging models; evaluating existing interventions, and ultimately developing new, stage-tailored, interventions.

4.3 Limitations

Key limitations of this review relate to the lack of a consensus staging model or established boundaries of stages in BD, based on any proxy measures of stage. In reviewing the identified studies, we have dichotomised samples into earlier versus later stages of illness based in proxies, and therefore conclusions should be considered in light of the significant potential for bias when i) applying categorisations which have not been adequately validated, ii) applying such categorisations post-hoc, with heterogeneous assessment and reporting thereof and iii) using per-study measures of central tendency to assign a single 'stage' to each sample. In some instances, this dichotomisation may not have accurately reflected sample characteristics (for example, we classified a sample described by authors as 'early/recent onset' as 'later'[159] and another with a 'severe or bad course' as 'earlier'[110,136]). Further, the dichotomised earlier/later stage model used here forgoes the possibility that subgroups, or those at the ends of the distribution (very highly recurrent or first episode) may behave differently. However, lack of validation regarding the role of these proxies in clinical staging does not preclude their potential importance as moderators of intervention effectiveness. A further issue is that we did not combine data statistically. We were limited to narrative synthesis methods, as preliminary scoping revealed significant heterogeneity in reporting of both proxy measures and outcomes, preventing meaningful statistical analyses to achieve the specific aims of this review.

Further, broader issues within the literature limit findings. Overall very low quality for each outcome significantly limits conclusions, with many studies at high risk of bias. More broadly, publication bias is likely; with fewer negative trials reaching publication [160], and multiple testing was evident, particularly for cognitive impairment. There may be systematic selection biases present; for example, trials targeting the outcomes in this review may be more likely to select samples with greater functional or cognitive impairment, while those with minimal impairments in these domains may also benefit from these interventions. Given that this review was interested in which interventions were and were not beneficial and for whom, unpublished studies or calculating fail-safe Ns may therefore

have contributed to our findings. Therefore, the conclusions offered here are not definitive, representing avenues for further exploration.

4.4 Future directions

Studies directly comparing the effectiveness of psychosocial interventions for functional outcomes for those at different stages of illness are required to answer this review's research questions definitively. Fewer intervention studies appear to prioritise functional outcomes in BD. Given that these are meaningful for those with BD [161,162] and may emerge as especially important in later illness stages [17,22], more emphasis on functional outcomes is encouraged. Next, planned moderation analyses are required to validate conclusions regarding the role of these proxies and implications for staging models. Examples of such analyses have been conducted within the Systematic Treatment Enhancement Program for BD (STEP-BD) and Parades samples, wherein NoE emerged as a moderator of the effectiveness of psychotherapy on symptom recovery and time to relapse, respectively [25,149]. Further, a key motivation for establishing staging models is the development of tailored interventions for specific stages; few studies have examined tailored psychosocial interventions, although this area of research appears to be growing [23,163]. Concurrently, staging models require further validation in BD. This includes the urgent need for improved consistency of measurement and reporting of potential stage proxies, such as NoE and functioning. Extrapolating from clinical samples, such as the treatment-seeking samples in the current review, may lead to distorted conclusions about illness progression; known as Berkson's bias [164]. Therefore, large scale epidemiological research is required. In addition, each of the plausible proxies of stage employed in this review has been assessed and applied inconsistently throughout the literature, and perhaps most problematically, NoE [11]; more consistent assessment and conceptualisation would guide research in this area. Finally, this review adopted a disorder specific conceptualisation of stage and it is not possible to infer whether our findings might extrapolate to transdiagnostic conceptualisations. Future studies should examine these relationships using a transdiagnostic model, given that the later stages of such models share characteristics with the employed model, such as impaired functioning.

4. 5 Conclusions

Results of this review suggest that psychosocial interventions are more effective for targeting general or social functioning in the earlier than later stage of BD. For both QoL and cognitive functioning, evidence for effective psychosocial interventions was inconsistent, with no support found for superior effectiveness in earlier stage. Further, the results of this review offer preliminary evidence in support of the hypothesis that specific types of psychosocial interventions are best positioned to target general/social functioning, QoL and cognitive functioning for individuals with different clinical experiences, classified as earlier and later stage within this review. Specific interventions which demonstrated promise for these outcomes in the later stage include functional or cognitive

remediation, mindfulness-based interventions and combined or intensive interventions. CT, CBT and PE appear to be less effective for these outcomes in the later than earlier stage. Additionally, some specific sub-domains of QoL and cognitive functioning may be more successfully targeted than others. However, this literature suffers from significant quality and bias issues and our conclusions must be considered tentative. In particular, we would not want our provisional findings of stage-related differences in benefits of some interventions to be interpreted fatalistically, and discourage attempts to improve these outcomes for all people with established BD. As noted, our findings should be considering preliminary and further research is required to validate these.

Acknowledgements:

MB is supported by a NHMRC Senior Principal Research Fellowship (1059660) HT and CMcE are supported Australian Government Research Training Program Scholarships. GM receives funding from the National Health and Medical Research Council for a related project, grant number APP1102097. JS was a visiting fellow with Swinburne University. These funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Hickie IB, Scott J, McGorry PD. Clinical staging for mental disorders: A new development in diagnostic practice in mental health. *Med J Aust.* 2013;198(9):461-462.
2. Duffy A, Horrocks J, Doucette S, Keown-Stoneman C, McCloskey S, Grof P. The developmental trajectory of bipolar disorder. *Br J Psychiatry.* 2014;204(2):122-128.
3. Berk M, Berk L, Dodd S, et al. Stage managing bipolar disorder. *Bipolar Disord.* 2014;16(5):471-477.
4. Post RM, Fleming J, Kapczinski F. Neurobiological correlates of illness progression in the recurrent affective disorders. *J Psychiatr Res.* 2012;46(5):561-573.
5. Vieta E, Reinares M, Rosa AR. Staging bipolar disorder. *Neurotox Res.* 2011;19(2):279-285.
6. Rosa AR, Magalhaes PV, Czepielewski L, et al. Clinical staging in bipolar disorder: Focus on cognition and functioning. *J Clin Psychiatry.* 2014;75(5):e450-456.
7. Berk M, Post R, Ratheesh A, et al. Staging in bipolar disorder: From theoretical framework to clinical utility. *World psychiatry : official journal of the World Psychiatric Association (WPA).* 2017;16(3):236-244.

8. Berk M, Conus P, Lucas N, et al. Setting the stage: From prodrome to treatment resistance in bipolar disorder. *Bipolar Disord.* 2007;9(7):671-678.
9. Kapczinski F, Dias VV, Kauer-Sant'Anna M, et al. Clinical implications of a staging model for bipolar disorders. *Expert Rev Neurother.* 2009;9(7):957-966.
10. Scott J, Henry C. The staging model (or evolutionary stages) applied to bipolar disorder. *L'Encephale.* 2018;44(6S):S30-S33.
11. Tremain H, Fletcher K, Murray G. Number of episodes in bipolar disorder: The case for more thoughtful conceptualization and measurement. *Bipolar Disord.* 2019.
12. Kapczinski F, Magalhaes PV, Balanza-Martinez V, et al. Staging systems in bipolar disorder: An international society for bipolar disorders task force report. *Acta Psychiatr Scand.* 2014;130(5):354-363.
13. Deckersbach T, Peters AT, Sylvia LG, et al. A cluster analytic approach to identifying predictors and moderators of psychosocial treatment for bipolar depression: Results from step-bd. *J Affect Disord.* 2016;203:152-157.
14. Grande I, Magalhaes PV, Chendo I, et al. Staging bipolar disorder: Clinical, biochemical, and functional correlates. *Acta Psychiatr Scand.* 2014;129(6):437-444.
15. Colom F, Reinares M, Pacchiarotti I, et al. Has number of previous episodes any effect on response to group psychoeducation in bipolar patients? A 5-year follow-up post hoc analysis. *Acta Neuropsychiatr.* 2010;22(2):50-53.
16. Passos IC, Mwangi B, Vieta E, Berk M, Kapczinski F. Areas of controversy in neuroprogression in bipolar disorder. *Acta Psychiatr Scand.* 2016;134(2):91-103.
17. Berk M, Brnabic A, Dodd S, et al. Does stage of illness impact treatment response in bipolar disorder? Empirical treatment data and their implication for the staging model and early intervention. *Bipolar Disord.* 2011;13(1):87-98.
18. Joyce K, Thompson A, Marwaha S. Is treatment for bipolar disorder more effective earlier in illness course? A comprehensive literature review. *Int J Bipolar Disord.* 2016;4(1):19.
19. Lavagnino L, Cao B, Mwangi B, et al. Changes in the corpus callosum in women with late-stage bipolar disorder. *Acta Psychiatr Scand.* 2015;131(6):458-464.
20. Reinares M, Papachristou E, Harvey P, et al. Towards a clinical staging for bipolar disorder: Defining patient subtypes based on functional outcome. *J Affect Disord.* 2013;144(1-2):65-71.
21. Tremain H, Fletcher K, Scott J, McEnery C, Berk M, Murray G. Does stage of illness influence recovery-focused outcomes after psychological treatment in bipolar disorder? A systematic review protocol. *Systematic reviews.* 2019;8(1):125.
22. Berk M, Berk L, Udina M, et al. Palliative models of care for later stages of mental disorder: Maximizing recovery, maintaining hope, and building morale. *Aust N Z J Psychiatry.* 2012;46(2):92-99.

23. Murray G, Leitan ND, Berk M, et al. Online mindfulness-based intervention for late-stage bipolar disorder: Pilot evidence for feasibility and effectiveness. *J Affect Disord.* 2015;178:46-51.
24. Miziou S, Tsitsipa E, Moysidou S, et al. Psychosocial treatment and interventions for bipolar disorder: A systematic review. *Ann Gen Psychiatry.* 2015;14:19.
25. Peters A, Sylvia LG, Magalhaes PV, et al. Age at onset, course of illness and response to psychotherapy in bipolar disorder: Results from the systematic treatment enhancement program for bipolar disorder (step-bd). *Psychol Med.* 2014;44(16):3455-3467.
26. Meyer TD, Hautzinger M. Cognitive behaviour therapy and supportive therapy for bipolar disorders: Relapse rates for treatment period and 2-year follow-up. *Psychol Med.* 2012;42(7):1429-1439.
27. Scott J, Paykel E, Morriss R, et al. Cognitive-behavioural therapy for severe and recurrent bipolar disorders: Randomised controlled trial. *Br J Psychiatry.* 2006;188(4):313-320.
28. Bonnin CM, Torrent C, Arango C, et al. Functional remediation in bipolar disorder: 1-year follow-up of neurocognitive and functional outcome. *Br J Psychiatry.* 2016;208(1):87-93.
29. Torrent C, Bonnin Cdel M, Martinez-Aran A, et al. Efficacy of functional remediation in bipolar disorder: A multicenter randomized controlled study. *Am J Psychiatry.* 2013;170(8):852-859.
30. Lam DH, Burbeck R, Wright K, Pilling S. Psychological therapies in bipolar disorder: The effect of illness history on relapse prevention - a systematic review.(report). *Bipolar Disord.* 2009;11(5):474.
31. Perich T, Mitchell PB. Psychological interventions for young people at risk for bipolar disorder: A systematic review. *J Affect Disord.* 2019.
32. Vieta E, Salagre E, Grande I, et al. Early intervention in bipolar disorder. *Am J Psychiatry.* 2018;175(5):411-426.
33. Vallarino M, Henry C, Etain B, et al. An evidence map of psychosocial interventions for the earliest stages of bipolar disorder. *The Lancet Psychiatry.* 2015;2(6):548-563.
34. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015: Elaboration and explanation. *BMJ.* 2015;350(jan02 1):g7647.
35. Liberati A, Altman DG, Tetzlaff J, et al. The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1-e34.
36. Rosa AR, Reinares M, Michalak EE, et al. Functional impairment and disability across mood states in bipolar disorder. *Value Health.* 2010;13(8):984-988.
37. Park DY, Do D, Chang L, et al. Episode accumulation associated with hastened recurrence and delayed recovery in bipolar disorder. *J Affect Disord.* 2018;227:657-664.

38. Reinares M, Colom F, Rosa AR, et al. The impact of staging bipolar disorder on treatment outcome of family psychoeducation. *J Affect Disord.* 2010;123(1-3):81-86.
39. Bonnin CM, Martinez-Aran A, Reinares M, et al. Thresholds for severity, remission and recovery using the functioning assessment short test (fast) in bipolar disorder. *J Affect Disord.* 2018;240:57-62.
40. Judd LL, Schettler PJ, Solomon DA, et al. Psychosocial disability and work role function compared across the long-term course of bipolar i, bipolar ii and unipolar major depressive disorders. *J Affect Disord.* 2008;108(1-2):49-58.
41. van der Voort TY, Seldenrijk A, van Meijel B, et al. Functional versus syndromal recovery in patients with major depressive disorder and bipolar disorder. *J Clin Psychiatry.* 2015;76(6):e809-814.
42. Weinstock LM, Miller IW. Psychosocial predictors of mood symptoms 1 year after acute phase treatment of bipolar i disorder. *Compr Psychiatry.* 2010;51(5):497-503.
43. Nasrallah H, Morosini P, Gagnon DD. Reliability, validity and ability to detect change of the personal and social performance scale in patients with stable schizophrenia. *Psychiatry Res.* 2008;161(2):213-224.
44. Rybarczyk B. Social and occupational functioning assessment scale (sofas). *Encyclopedia of clinical neuropsychology.* 2011:2313-2313.
45. Leon AC, Olfson M, Portera L, Farber L, Sheehan DV. Assessing psychiatric impairment in primary care with the sheehan disability scale. *Int J Psychiatry Med.* 1997;27(2):93-105.
46. Bosc M, Dubini A, Polin V. Development and validation of a social functioning scale, the social adaptation self-evaluation scale. *Eur Neuropsychopharmacol.* 1997;7 Suppl 1(1):S57-70; discussion S71-53.
47. Martino DJ, Igoa A, Marengo E, Scapola M, Strejilevich SA. Longitudinal relationship between clinical course and neurocognitive impairments in bipolar disorder. *J Affect Disord.* 2018;225:250-255.
48. Malhi GS, Bargh DM, Cashman E, Frye MA, Gitlin M. The clinical management of bipolar disorder complexity using a stratified model. *Bipolar Disord.* 2012;14 Suppl 2:66-89.
49. Levy B, Medina AM, Weiss RD. Cognitive and psychosocial functioning in bipolar disorder with and without psychosis during early remission from an acute mood episode: A comparative longitudinal study. *Compr Psychiatry.* 2013;54(6):618-626.
50. Jones SH, Thornicroft G, Coffey M, Dunn G. A brief mental health outcome scale-reliability and validity of the global assessment of functioning (gaf). *Br J Psychiatry.* 1995;166(5):654-659.
51. Rosa AR, Sanchez-Moreno J, Martinez-Aran A, et al. Validity and reliability of the functioning assessment short test (fast) in bipolar disorder. *Clin Pract Epidemiol Ment Health.* 2007;3(1):5.

52. Leon A, Solomon D, Mueller T, Turvey C, Endicott J, Keller M. The range of impaired functioning tool (life–rift): A brief measure of functional impairment. *Psychol Med.* 1999;29(4):869-878.
53. Organization WH. World health organization disabilty assessment schedule: Whodas ii. Phase 2 field trials. Health services research. 2000.
54. Arbuckle R, Frye MA, Brecher M, et al. The psychometric validation of the sheehan disability scale (sds) in patients with bipolar disorder. *Psychiatry Res.* 2009;165(1-2):163-174.
55. Birchwood M, Smith J, Cochrane R, Wetton S, Copestake S. The social functioning scale the development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. *The British Journal of Psychiatry.* 1990;157(6):853-859.
56. Altshuler L, Mintz J, Leight K. The life functioning questionnaire (lfq): A brief, gender-neutral scale assessing functional outcome. *Psychiatry Res.* 2002;112(2):161-182.
57. Patterson TL, Goldman S, McKibbin CL, Hughs T, Jeste DV. Ucsd performance-based skills assessment: Development of a new measure of everyday functioning for severely mentally ill adults. *Schizophr Bull.* 2001;27(2):235-245.
58. Aydemır Ö, Eren İ, Savaş H, et al. Development of a questionnaire to assess inter-episode functioning in bipolar disorder: Bipolar disorder functioning questionnaire. *Turkish Journal of Psychiatry.* 2007;18(4).
59. Burdick KE, Ketter TA, Goldberg JF, Calabrese JR. Assessing cognitive function in bipolar disorder: Challenges and recommendations for clinical trial design. *J Clin Psychiatry.* 2015;76(3):e342-350.
60. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12(3):189-198.
61. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The mini-cog: A cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *Int J Geriatr Psychiatry.* 2000;15(11):1021-1027.
62. Buschke H, Kuslansky G, Katz M, et al. Screening for dementia with the memory impairment screen. *Neurology.* 1999;52(2):231-238.
63. Brodaty H, Pond D, Kemp NM, et al. The gpcog: A new screening test for dementia designed for general practice. *J Am Geriatr Soc.* 2002;50(3):530-534.
64. Shulman KI. Clock-drawing: Is it the ideal cognitive screening test? *Int J Geriatr Psychiatry.* 2000;15(6):548-561.
65. Nasreddine ZS, Phillips NA, Bedirian V, et al. The montreal cognitive assessment, moca: A brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* 2005;53(4):695-699.

66. Wechsler D. Wechsler adult intelligence scale–fourth edition (wais–iv). *San Antonio, TX: NCS Pearson.* 2008;22:498.
67. Wechsler D. Wechsler memory scale iv (wms-iv). In: New York, NY: Psychological Corporation; 2009.
68. Sheslow D, Adams W. Wide range assessment of memory and learning–second edition (wraml2). *Wide range Inc, Delaware.* 2003.
69. Woods SP, Delis DC, Scott JC, Kramer JH, Holdnack JA. The california verbal learning test–second edition: Test-retest reliability, practice effects, and reliable change indices for the standard and alternate forms. *Arch Clin Neuropsychol.* 2006;21(5):413-420.
70. Benedict RH, Schretlen D, Groninger L, Brandt J. Hopkins verbal learning test–revised: Normative data and analysis of inter-form and test-retest reliability. *The Clinical Neuropsychologist.* 1998;12(1):43-55.
71. Benedict RH, Schretlen D, Groninger L, Dobraski M, Shpritz B. Revision of the brief visuospatial memory test: Studies of normal performance, reliability, and validity. *Psychol Assess.* 1996;8(2):145.
72. Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique.(les problems.). *Arch Psychol (Geneve).* 1941.
73. Gordon NG. The trail making test in neuropsychological diagnosis. *J Clin Psychol.* 1972;28(2):167-169.
74. Heaton RK. Wisconsin card sorting test manual; revised and expanded. *Psychological Assessment Resources.* 1981:5-57.
75. Delis DC, Kramer JH, Kaplan E, Holdnack J. Reliability and validity of the delis-kaplan executive function system: An update. *J Int Neuropsychol Soc.* 2004;10(2):301-303.
76. Goodglass H, Kaplan E, Weintraub S. *Boston naming test.* Lea & Febiger; 1983.
77. Ruff RM, Light RH, Parker SB, Levin HS. Benton controlled oral word association test: Reliability and updated norms. *Arch Clin Neuropsychol.* 1996;11(4):329-338.
78. Kern RS, Gold JM, Dickinson D, et al. The mccb impairment profile for schizophrenia outpatients: Results from the matrics psychometric and standardization study. *Schizophr Res.* 2011;126(1-3):124-131.
79. WHOQOL Group. The world health organization quality of life assessment (whoqol): Position paper from the world health organization. *Soc Sci Med.* 1995;41(10):1403-1409.
80. Michalak EE, Murray G, Collaborative RTtSPIiBD. Development of the qol.Bd: A disorder-specific scale to assess quality of life in bipolar disorder. *Bipolar Disord.* 2010;12(7):727-740.
81. Rabin R, Charro Fd. Eq-sd: A measure of health status from the euroqol group. *Ann Med.* 2001;33(5):337-343.

82. van Nieuwenhuizen C, Schene AH, Koeter MW, Huxley PJ. The lancashire quality of life profile: Modification and psychometric evaluation. *Soc Psychiatry Psychiatr Epidemiol.* 2001;36(1):36-44.
83. Priebe S, Huxley P, Knight S, Evans S. Application and results of the manchester short assessment of quality of life (mansa). *Int J Soc Psychiatry.* 1999;45(1):7-12.
84. Russo J, Roy-Byrne P, Reeder D, et al. Longitudinal assessment of quality of life in acute psychiatric inpatients: Reliability and validity. *J Nerv Ment Dis.* 1997;185(3):166-175.
85. Keller MB, Lavori PW, Friedman B, et al. The longitudinal interval follow-up evaluation. A comprehensive method for assessing outcome in prospective longitudinal studies. *Arch Gen Psychiatry.* 1987;44(6):540-548.
86. Ware Jr JE, Kosinski M, Keller SD. A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Med Care.* 1996:220-233.
87. Ware JE, Sherbourne CD, Davies AR. Developing and testing the mos 20-item short-form health survey. 1992.
88. McHorney CA, Ware JE, Jr., Lu JF, Sherbourne CD. The mos 36-item short-form health survey (sf-36): Iii. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care.* 1994;32(1):40-66.
89. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of life enjoyment and satisfaction questionnaire: A new measure. *Psychopharmacol Bull.* 1993;29(2):321-326.
90. Hunt SM, McKenna SP. The qlds: A scale for the measurement of quality of life in depression. *Health Policy.* 1992;22(3):307-319.
91. Ferrans CE, Powers MJ. Quality of life index: Development and psychometric properties. *ANS Adv Nurs Sci.* 1985;8(1):15-24.
92. Group TW. The world health organization quality of life assessment (whoqol): Development and general psychometric properties. *Soc Sci Med.* 1998;46(12):1569-1585.
93. Faedda GL, Baldessarini RJ, Glovinsky IP, Austin NB. Pediatric bipolar disorder: Phenomenology and course of illness. *Bipolar Disord.* 2004;6(4):305-313.
94. Higgins JP, Altman DG, Gotzsche PC, et al. The cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ.* 2011;343(oct18 2):d5928.
95. Sterne JA, Hernan MA, Reeves BC, et al. Robins-i: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ.* 2016;355:i4919.
96. Guyatt GH, Oxman AD, Vist GE, et al. Grade: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ.* 2008;336(7650):924-926.
97. Hong QN, Pluye P, Bujold M, Wassef M. Convergent and sequential synthesis designs: Implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Systematic reviews.* 2017;6(1):61.

98. Fletcher K, Murray G. Beyond pharmacotherapy for bipolar disorder: The role of adjunctive psychological treatments. *Psychiatric Annals.* 2019;49(9):399-404.
99. Geddes JR, Miklowitz DJ. Treatment of bipolar disorder. *Lancet (London, England).* 2013;381(9878):1672-1682.
100. Bonnin CM, Reinares M, Martinez-Aran A, et al. Effects of functional remediation on neurocognitively impaired bipolar patients: Enhancement of verbal memory. *Psychol Med.* 2016;46(2):291-301.
101. Sanchez-Moreno J, Bonnin C, Gonzalez-Pinto A, et al. Do patients with bipolar disorder and subsyndromal symptoms benefit from functional remediation? A 12-month follow-up study. *Eur Neuropsychopharmacol.* 2017;27(4):350-359.
102. Sole B, Bonnin CM, Mayoral M, et al. Functional remediation for patients with bipolar ii disorder: Improvement of functioning and subsyndromal symptoms. *Eur Neuropsychopharmacol.* 2015;25(2):257-264.
103. Chen R, Zhu X, Capitao LP, et al. Psychoeducation for psychiatric inpatients following remission of a manic episode in bipolar i disorder: A randomized controlled trial. *Bipolar Disord.* 2019;21(1):76-85.
104. Clarkin JF, Carpenter D, Hull J, Wilner P, Glick I. Effects of psychoeducational intervention for married patients with bipolar disorder and their spouses. *Psychiatr Serv.* 1998;49(4):531-533.
105. Colom F, Vieta E, Sanchez-Moreno J, et al. Psychoeducation for bipolar ii disorder: An exploratory, 5-year outcome subanalysis. *J Affect Disord.* 2009;112(1-3):30-35.
106. Deckersbach T, Holzel BK, Eisner LR, et al. Mindfulness-based cognitive therapy for nonremitted patients with bipolar disorder. *CNS Neurosci Ther.* 2012;18(2):133-141.
107. Stange JP, Eisner LR, Holzel BK, et al. Mindfulness-based cognitive therapy for bipolar disorder: Effects on cognitive functioning. *J Psychiatr Pract.* 2011;17(6):410-419.
108. Díaz-Zuluaga AM, Vargas C, Duica K, et al. Impact of a multimodal intervention on the psychological profile of schizophrenic and bipolar i patients: A study of prisma programme. *Revista Colombiana de Psiquiatría (English ed).* 2017;46(2):56-64.
109. Frank E, Kupfer DJ, Thase ME, et al. Two-year outcomes for interpersonal and social rhythm therapy in individuals with bipolar i disorder. *Arch Gen Psychiatry.* 2005;62(9):996-1004.
110. Gonzalez Isasi A, Echeburua E, Mosquera F, Ibanez B, Aizpuru F, Gonzalez-Pinto A. Long-term efficacy of a psychological intervention program for patients with refractory bipolar disorder: A pilot study. *Psychiatry Res.* 2010;176(2-3):161-165.
111. Harvey AG, Soehner AM, Kaplan KA, et al. Treating insomnia improves mood state, sleep, and functioning in bipolar disorder: A pilot randomized controlled trial. *J Consult Clin Psychol.* 2015;83(3):564-577.

112. Hawke LD, Velyvis V, Parikh SV. Bipolar disorder with comorbid anxiety disorders: Impact of comorbidity on treatment outcome in cognitive-behavioral therapy and psychoeducation. *Int J Bipolar Disord.* 2013;1(1):15.
113. Kilbourne AM, Post EP, Nossek A, Drill L, Cooley S, Bauer MS. Improving medical and psychiatric outcomes among individuals with bipolar disorder: A randomized controlled trial. *Psychiatr Serv.* 2008;59(7):760-768.
114. Kilbourne AM, Goodrich DE, Lai Z, Clogston J, Waxmonsky J, Bauer MS. Life goals collaborative care for patients with bipolar disorder and cardiovascular disease risk. *Psychiatr Serv.* 2012;63(12):1234-1238.
115. Kilbourne AM, Goodrich DE, Lai Z, et al. Randomized controlled trial to assess reduction of cardiovascular disease risk in patients with bipolar disorder: The self-management addressing heart risk trial (smahrt). *J Clin Psychiatry.* 2013;74(7):e655-662.
116. Kurdal E, Tanriverdi D, Savas HA. The effect of psychoeducation on the functioning level of patients with bipolar disorder. *West J Nurs Res.* 2014;36(3):312-328.
117. Lahera G, Benito A, Montes JM, Fernandez-Liria A, Olbert CM, Penn DL. Social cognition and interaction training (scit) for outpatients with bipolar disorder. *J Affect Disord.* 2013;146(1):132-136.
118. Lewandowski KE, Sperry SH, Cohen BM, et al. Treatment to enhance cognition in bipolar disorder (trec-bd): Efficacy of a randomized controlled trial of cognitive remediation versus active control. *J Clin Psychiatry.* 2017;78(9):e1242-e1249.
119. Nagy N, Sabry W, Khalifa D, Hashem R, Zahran N, Khalil AH. Relapse rate and outcome correlates in egyptian patients with bipolar disorder treated with behavioural family psychoeducation. *Middle East Current Psychiatry.* 2015;22(3):121-131.
120. Patelis-Siotis I, Young LT, Robb JC, et al. Group cognitive behavioral therapy for bipolar disorder: A feasibility and effectiveness study. *J Affect Disord.* 2001;65(2):145-153.
121. Scott J, Garland A, Moorhead S. A pilot study of cognitive therapy in bipolar disorders. *Psychol Med.* 2001;31(3):459-467.
122. Sylvia LG, Nierenberg AA, Stange JP, Peckham AD, Deckersbach T. Development of an integrated psychosocial treatment to address the medical burden associated with bipolar disorder. *J Psychiatr Pract.* 2011;17(3):224-232.
123. Sylvia LG, Salcedo S, Bernstein EE, Baek JH, Nierenberg AA, Deckersbach T. Nutrition, exercise, and wellness treatment in bipolar disorder: Proof of concept for a consolidated intervention. *Int J Bipolar Disord.* 2013;1(1):24.
124. Wenze SJ, Gaudiano BA, Weinstock LM, Tezanos KM, Miller IW. Adjunctive psychosocial intervention following hospital discharge for patients with bipolar disorder and comorbid substance use: A pilot randomized controlled trial. *Psychiatry Res.* 2015;228(3):516-525.

125. Zaki N, Awaad MI, Elbatrawy AN, Elmissiry MA, Zahran NS. Effectiveness of a behavioral family psychoeducational program for egyptian patients with bipolar disorder and their caregivers. *Middle East Current Psychiatry.* 2014;21(4):212-221.

126. Ball JR, Mitchell PB, Corry JC, Skillecorn A, Smith M, Malhi GS. A randomized controlled trial of cognitive therapy for bipolar disorder: Focus on long-term change.[cme]. *The Journal of clinical psychiatry.* 2006;67(2):1,478-286.

127. Bauer MS, McBride L, Williford WO, et al. Collaborative care for bipolar disorder: Part ii. Impact on clinical outcome, function, and costs. *Psychiatr Serv.* 2006;57(7):937-945.

128. Ellard KK, Bernstein EE, Hearing C, et al. Transdiagnostic treatment of bipolar disorder and comorbid anxiety using the unified protocol for emotional disorders: A pilot feasibility and acceptability trial. *J Affect Disord.* 2017;219:209-221.

129. Kilbourne AM, Biswas K, Pirraglia PA, Sajatovic M, Williford WO, Bauer MS. Is the collaborative chronic care model effective for patients with bipolar disorder and co-occurring conditions? *J Affect Disord.* 2009;112(1-3):256-261.

130. Castle D, Berk M, Berk L, Lauder S, Chamberlain J, Gilbert M. Pilot of group intervention for bipolar disorder. *Int J Psychiatry Clin Pract.* 2007;11(4):279-284.

131. de Barros Pellegrinelli K, de OCLF, Silval KI, et al. Efficacy of psychoeducation on symptomatic and functional recovery in bipolar disorder. *Acta Psychiatr Scand.* 2013;127(2):153-158.

132. Deckersbach T, Nierenberg AA, Kessler R, et al. Research: Cognitive rehabilitation for bipolar disorder: An open trial for employed patients with residual depressive symptoms. *CNS Neurosci Ther.* 2010;16(5):298-307.

133. Deckersbach T, Peters AT, Shea C, et al. Memory performance predicts response to psychotherapy for depression in bipolar disorder: A pilot randomized controlled trial with exploratory functional magnetic resonance imaging. *J Affect Disord.* 2018;229:342-350.

134. Demant KM, Vinberg M, Kessing LV, Miskowiak KW. Effects of short-term cognitive remediation on cognitive dysfunction in partially or fully remitted individuals with bipolar disorder: Results of a randomised controlled trial. *PLoS One.* 2015;10(6):e0127955.

135. Fiorillo A, Del Vecchio V, Luciano M, et al. Efficacy of psychoeducational family intervention for bipolar i disorder: A controlled, multicentric, real-world study. *J Affect Disord.* 2015;172:291-299.

136. Gonzalez Isasi A, Echeburua E, Liminana JM, Gonzalez-Pinto A. How effective is a psychological intervention program for patients with refractory bipolar disorder? A randomized controlled trial. *J Affect Disord.* 2010;126(1-2):80-87.

137. Gonzalez Isasi A, Echeburua E, Liminana JM, Gonzalez-Pinto A. Psychoeducation and cognitive-behavioral therapy for patients with refractory bipolar disorder: A 5-year controlled clinical trial. *Eur Psychiatry.* 2014;29(3):134-141.

138. Haffner P, Quinlivan E, Fiebig J, et al. Improving functional outcome in bipolar disorder: A pilot study on metacognitive training. *Clin Psychol Psychother.* 2018;25(1):50-58.
139. Hidalgo-Mazzei D, Reinares M, Mateu A, et al. Opensimple: A real-world implementation feasibility study of a smartphone-based psychoeducation programme for bipolar disorder. *J Affect Disord.* 2018;241:436-445.
140. Hoberg AA, Ponto J, Nelson PJ, Frye MA. Group interpersonal and social rhythm therapy for bipolar depression. *Perspect Psychiatr Care.* 2013;49(4):226-234.
141. Ives-Deliperi VL, Howells F, Stein DJ, Meintjes EM, Horn N. The effects of mindfulness-based cognitive therapy in patients with bipolar disorder: A controlled functional mri investigation. *J Affect Disord.* 2013;150(3):1152-1157.
142. Jones SH, Smith G, Mulligan LD, et al. Recovery-focused cognitive-behavioural therapy for recent-onset bipolar disorder: Randomised controlled pilot trial. *Br J Psychiatry.* 2015;206(1):58-66.
143. Jones SH, Knowles D, Tyler E, et al. The feasibility and acceptability of a novel anxiety in bipolar disorder intervention compared to treatment as usual: A randomized controlled trial. *Depress Anxiety.* 2018;35(10):953-965.
144. Lam DH, Bright J, Jones S, et al. Cognitive therapy for bipolar illness—a pilot study of relapse prevention. *Cognit Ther Res.* 2000;24(5):503-520.
145. Lam DH, Hayward P, Watkins ER, Wright K, Sham P. Relapse prevention in patients with bipolar disorder: Cognitive therapy outcome after 2 years. *Am J Psychiatry.* 2005;162(2):324-329.
146. Lam DH, Watkins ER, Hayward P, et al. A randomized controlled study of cognitive therapy for relapse prevention for bipolar affective disorder: Outcome of the first year. *Arch Gen Psychiatry.* 2003;60(2):145-152.
147. Lenz G, Berg A, Breit-Gabauer B, et al. Kognitiv-psychoedukative therapie im vergleich zu bibliotherapie bei bipolarer störung: Eine kontrollierte gruppentherapiestudie. *Verhaltenstherapie.* 2016;26(2):92-98.
148. Miklowitz DJ, Otto MW, Frank E, et al. Psychosocial treatments for bipolar depression: A 1-year randomized trial from the systematic treatment enhancement program. *Arch Gen Psychiatry.* 2007;64(4):419-426.
149. Morriss R, Lobban F, Riste L, et al. Clinical effectiveness and acceptability of structured group psychoeducation versus optimised unstructured peer support for patients with remitted bipolar disorder (parades): A pragmatic, multicentre, observer-blind, randomised controlled superiority trial. *Lancet Psychiatry.* 2016;3(11):1029-1038.
150. Sajatovic M, Davies MA, Ganocy SJ, et al. A comparison of the life goals program and treatment as usual for individuals with bipolar disorder. *Psychiatr Serv.* 2009;60(9):1182-1189.

151. Schottle D, Schimmelmann BG, Karow A, et al. Effectiveness of integrated care including therapeutic assertive community treatment in severe schizophrenia spectrum and bipolar i disorders: The 24-month follow-up access ii study. *J Clin Psychiatry.* 2014;75(12):1371-1379.

152. Schottle D, Schimmelmann BG, Ruppelt F, et al. Effectiveness of integrated care including therapeutic assertive community treatment in severe schizophrenia-spectrum and bipolar i disorders: Four-year follow-up of the access ii study. *PLoS One.* 2018;13(2):e0192929.

153. Todd NJ, Jones SH, Hart A, Lobban FA. A web-based self-management intervention for bipolar disorder 'living with bipolar': A feasibility randomised controlled trial. *J Affect Disord.* 2014;169:21-29.

154. van der Voort TY, van Meijel B, Hoogendoorn AW, Goossens PJ, Beekman AT, Kupka RW. Collaborative care for patients with bipolar disorder: Effects on functioning and quality of life. *J Affect Disord.* 2015;179:14-22.

155. Veeh J, Kopf J, Kittel-Schneider S, Deckert J, Reif A. Cognitive remediation for bipolar patients with objective cognitive impairment: A naturalistic study. *Int J Bipolar Disord.* 2017;5(1):8.

156. Zyto S, Jabben N, Schulte PF, Regeer BJ, Kupka RW. A pilot study of a combined group and individual functional remediation program for patients with bipolar i disorder. *J Affect Disord.* 2016;194:9-15.

157. Smith DJ, Griffiths E, Poole R, et al. Beating bipolar: Exploratory trial of a novel internet-based psychoeducational treatment for bipolar disorder. *Bipolar Disord.* 2011;13(5-6):571-577.

158. Sanchez-Moreno J, Martinez-Aran A, Vieta E. Treatment of functional impairment in patients with bipolar disorder. *Curr Psychiatry Rep.* 2017;19(1):3.

159. Jones S, Mulligan L, Law H, et al. A randomised controlled trial of recovery focused cbt for individuals with early bipolar disorder. *BMC Psychiatry.* 2012;12.

160. Driessen E, Hollon SD, Bockting CL, Cuijpers P, Turner EH. Does publication bias inflate the apparent efficacy of psychological treatment for major depressive disorder? A systematic review and meta-analysis of us national institutes of health-funded trials. *PLoS One.* 2015;10(9):e0137864.

161. Maassen EF, Regeer BJ, Regeer EJ, Bunders JFG, Kupka RW. The challenges of living with bipolar disorder: A qualitative study of the implications for health care and research. *Int J Bipolar Disord.* 2018;6(1):23.

162. Murray G, Leitan ND, Thomas N, et al. Towards recovery-oriented psychosocial interventions for bipolar disorder: Quality of life outcomes, stage-sensitive treatments, and mindfulness mechanisms. *Clin Psychol Rev.* 2017;52:148-163.

163. Fletcher K, Foley F, Thomas N, et al. Web-based intervention to improve quality of life in late stage bipolar disorder (orbit): Randomised controlled trial protocol. *BMC Psychiatry.* 2018;18(1):221.

164. Peters AT, West AE, Eisner L, Baek J, Deckersbach T. The burden of repeated mood episodes in bipolar i disorder: Results from the national epidemiological survey on alcohol and related conditions. *J Nerv Ment Dis.* 2016;204(2):87-94.

Table 1

Plausible Proxy Variables and Later Stage Cut-offs

Proxy variables	Later stage cut-off for this review
Number of mood episodes [15-19].	>10 episodes [3,15,36,37].
General or social functioning/Disability [6,9,20,38].	Established cut-offs or qualitative descriptors for later stage where cut-offs are unavailable (moderate to severe functional impairments): • FAST: >32 [6,36,39] • GAF: <60 [39] • LIFE-RIFT: >11 [40]. • SFQ: <120, based on descriptors • WHODAS 2.0 12 item: Summary score of 24 or > 25 standardised score represents moderate overall impairment [41] • WHODAS 2.0 36 item: Summary score of 72 or > 25 standardised score represents moderate overall impairment [41] • UCLA SAS >3 [42] • PSP <60 [43] • SOFAS <60 [44] • SDS scale scores >5 or total >15 [45] • SASS <25 [46].
Cognitive functioning/cognitive impairment/neurocognitive deficit [3,6,9,20].	Marked impairment > 2 SD from the mean (considered in the context of other proxies) or qualitative moderate to severe impairment [6,47].

Psychiatric comorbidities [5,9].	No cut-off (considered in the context of other proxies where possible); presence of anxiety disorders or substance use disorders [13,48].

Note. GAF = Global Assessment of Functioning; FAST = Functioning Assessment Short Test; LIFE-RIFT = Longitudinal Interval Follow-up Evaluation-Range of Impaired Functioning Tool; WHODAS 2.0 36 = World Health Organization Disability Assessment Schedule 2.0, 36 item; WHODAS 2.0 12 = World Health Organization Disability Assessment Schedule 2.0, 12 item; SOFAS = Social and Occupational Functioning Assessment Scale; SDS = Sheehan Disability Scale;; SFQ = Social Functioning Questionnaire; SPS = Social Performance Scale; SAS = Social Adjustment Scale; UCLA SAS = UCLA Social Attainment Scale; PSP = Personal and Social Functioning Scale; SASS = Social Adaptation Self-Evaluation Scale.
This table provides 'later' stage cut-offs, while the 'earlier' stage therefore includes individuals with diagnosable BD not surpassing any of the cut-off criteria.

Table 2

Outcome Measurement

Outcome measure instruments
1. General/social functioning: A multidimensional construct encompassing an individual's capacity for independent living, occupational and educational achievement, interpersonal relationships and recreation[49]. Measures will include any validated self-report or clinician-rated, real-world functioning or performance-based measure, including: • Global Assessment of Functioning (GAF) [50]• Functioning Assessment Short Test (FAST)[51] •The clinician-rated Longitudinal Interval Follow-up Evaluation-Range of Impaired Functioning Tool (LIFE-RIFT) [52]• World Health Organization Disability Assessment Schedule (WHODAS-2) [53]• Sheehan Disability Scale (SDS) [54]• Social Functioning Scale (SFS) [55]• Life Functioning Questionnaire (LFQ)[56] • USCD Performance-Based Skills Assessment [57]• Bipolar Disorder Functioning Questionnaire[58]• Social and Occupational Functioning Assessment Scale (SOFAS)[44]

2. Cognitive functioning

   An individual's range of abilities relating to cognition, memory and learning, including attention, executive function, verbal learning and memory, verbal fluency, processing speed, working memory, visual learning and memory, psychomotor speed, visuo-spatial ability)[59]. Measures will include any validated measure or screening tool assessing cognitive function or impairment overall or by domain including:

   • Mini Mental State Exam[60] • Mini-Cog[61] • Memory Impairment Screen (MIS)[62] • General Practitioner Assessment of Cognition (GPCOG)[63] • The clock drawing test (CDT) [64]• Montreal Cognitive Assessment (MoCA) [65]• FAST, cognitive domain [51]•Weschler Adult Intelligence Scale (WAIS-IV) or its subtests [66] • Wechsler Memory Scale (WMS-IV) or its subtests [67]• Wide Range Assessment of Memory and Learning [68]• California Verbal Learning Test [69]• Hopkins Verbal Learning Test-Revised [70]• Brief Visuospatial Memory Test-Revised [71]• Rey-Osterrieth Complex Figure Test [72]• Trail Making Test Part A or Part B [73]• Wisconsin Card Sorting Test [74]• Delis-Kaplan Executive Function System [75]• Boston Naming Test [76]• Controlled Oral Word Association [77] •MATRICS Consensus Cognitive Battery (MCCB) [78]

3. Quality of life

   While quality of life is used in varying ways, the World Health Organisation defines this as ''an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns''[79]. Measures will include any validated self-report or clinician-rated, general or disorder specific measure of quality of life, including:

   •brief Quality of Life in Bipolar Disorder questionnaire (QoL.BD)[80] • EuroQoL /EuroQoL 5 dimension (EQ-5D) [81]• Lancashire Quality of Life Profile [82]/ Manchester Short Assessment of Quality of Life (MANSA) [83]• Lehman Quality of Life Interview[84]• Longitudinal Interval Follow-up Evaluation [85]• Medical Outcomes Study (MOS) Short Form 12[86]•MOS Short Form 20 [87]• MOS Short Form 36 [88]•Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q)[89]• Quality of Life in Depression Scale[90] •Quality of Life Index [91]• World Health Organization Quality of Life Assessment [92]

Table 3

Summary of Studies Meeting Earlier Stage Criteria

Study	Sample	Intervention (control)	Stage proxy and information	Stage	Outcome (measures)	Findings	Notes

Bonnin, et al.[100]	Subsample of 188 participants with neurocognitive impairment	FR (TAU, PE)	**Functioning:** Moderate to severe impairment (FAST) **30.68 (10.09)** **NoE:** Mean in each group above 10 Bonnin NoE 7.01 (7.33) **Cognitive impairment:** (2 SDs below norm) in 1 or more domains	Earlier	Functioning (FAST) Cognitive functioning (comprehensive neuropsychological battery)	Improved verbal memory relative to TAU (ES .2) Improved functioning relative to both controls	Six-month follow-up subsample
Sanchez-Moreno, et al.[101]	Subsample of 99 participants with residual symptoms	FR (TAU, PE)	**Functioning:** Moderate to severe impairment (FAST) **31.39 (9.51)**	Earlier	Functioning (FAST)	Significant improvements in functioning in functional remediation group Superior to TAU (ES .78 6 mo†; .71 12 mo†) and to PE (ES .45 6 mo†; .32† 12mo).	12-month follow-up, subsample

			NoE: Mean of less than 10 in each group **6.63(6.41)**		Cognitive functioning (comprehensive neuropsychological battery)	No significant changes on any of the cognitive variables	
			Functioning: Moderate to severe impairment (FAST) 29.76 (9.69)		Functioning (FAST)	Significant overall improvements in functioning only in the functional remediation group.	
Sole, et al. [102]	Subsample of 53 participants with BD II	FR (TAU, PE)	**NoE:** Mean of less than 10 in each group **9.11 (8.32)**	Earlier	Cognitive functioning (comprehensive neuropsychological battery)	No group demonstrated superiority regarding cognitive improvements, although some domains significantly improved over time	Subsample

Chen, et al. [103]	140 inpatients with BD and current manic episode (remitting)	Group PE (group discussions)	**NoE:** Median of 4 (intervention) 3 (control) Median of 3.5 (50% of sample had 3.5 or less) **Functioning:** Mild impairment (WHODAS 2.0 36) 47.73 (11.82)	Earlier	Functioning (WHODAS 2.0 36)	WHODAS scores were significantly improved, relative to control (ES .21†).	Controlled for baseline functioning Manic episode
Clarkin, et al. [104]	42 with BD diagnosis	Marital PE (TAU)	**Functioning:** GAF in the mild-moderate impairment range **64.53 (No SD)**	Earlier	Functioning (GAF)	GAF significantly improved in the treatment group relative to control	
Colom, et al. [105]	20 euthymic participants with BD II	Group PE (Unstructured support group)	**Functioning: M**ild difficulty but generally functioning well ~ **60 (No SD)**	Earlier	Functioning (SOFAS)	SOFAS significantly improved at 2 and 5 years relative to control. (5 year only, ES = 1.715)	
Deckersbach, et al. [106]	12 adults with BD and subthreshold residual depressive symptoms	Mindfulness based cognitive therapy (none)	**Functioning:** Mild impairment (LIFE-RIFT) 10.2 (5.27)	Earlier	Functioning (LIFE-RIFT)	Functioning improved from pre- to posttreatment and to follow-up	

Stange, et al. [107]	8 adults with BD and subthreshold residual depressive symptoms	Mindfulness based cognitive therapy (none)	**Cognitive functioning:** Substantial impairment (1.6 SDs below normative mean = BRIEF 66.38 (15.63)	Earlier	Cognitive functioning (BRIEF, FrSBe)	Self-reported cognitive functioning improved across the domains of: Initiate, Working Memory, Apathy and Executive Functioning with large effect sizes post intervention and at follow-up	Post-hoc subsample analysis
Díaz-Zuluaga, et al. [108]	198 outpatients with BD and no psychiatric comorbidities	Multimodal intervention including PE and CBT (TAU)	**NoE:** fewer than 10 **3(4.47)** **Functioning: GAF aligne**d with no impairment 80 (23) **Comorbidity:** None	Earlier	Functioning (FAST)	Significant functional improvement when baseline characteristics including functioning were not controlled	Evidence of moderation: no significant improvements when baseline characteristics including functioning controlled

Frank et al. [109]	125 participants with BD 1, current mood episode and at least 2 prior episodes	Interpersonal and Social Rhythm therapy, then either maintenance IPSRT or Clinical Management with PE (Clinical management then maintenance)	**NoE:** Median of 7 (50% had 7 or fewer) **Functioning:** Severe impairments	Earlier	Functioning (Occupational: UCLA SAS)	Both groups' occupational functioning significantly improved over time, but the IPSRT group improved faster	Acute episode remitting may have improved occupational functioning Active control
Gonzalez Isasi, et al. [110]	20 euthymic participants with BD and prior 'severe' or 'bad' course	CBT (TAU)	**Functioning:** Mild impairments 71.13 (13.3)	Earlier	Functioning (GAF)	Functioning improved relative to control, becoming significant at 12-month follow-up	Functioning aligned with 'earlier' despite inclusion criteria of severe/ bad course.

Harvey, et al.[111]	58 euthymic participants with BD I and insomnia	CBT (PE)	**Functioning:** Minimal impairment 4.03 (2.46)	Earlier	Functioning (SDS) Quality of Life (Q-LES-Q)	Both groups experienced improved functioning no significant differences between groups (between ES .44 in favour of PE post treatment and .15 in favour of CBT at follow up) Quality of life did not significantly improve in either group (between ES .04 in favour of PE post treatment and .11 in favour of CBT at follow-up).	Active control
Hawke, et al.[112]	204 euthymic adults with BD and at least 2 episodes within 3 years	CBT (PE)	**Functioning:** Mild-moderate impairment **64.82 (10.94)** **Comorbidity:** With (42%) versus without comorbid anxiety disorder	Earlier	Functioning (GAF)	Both intervention groups experienced improved functioning, and this was not moderated by presence of anxiety disorders	

Kilbourne, et al. [113]	58 with BD and cardiovascular disease risk	Collaborative Care (TAU)	**Functioning:** Mild to moderate impairment **17.9 (9.9)**	Earlier	Functioning (WHODAS 2.0 12 item) Quality of Life (SF-12)	No significant differences between the two groups for overall functioning (ES .18) or overall quality of life. Significant differences physical health–related quality of life between groups (ES .32). Non-sig mental health QoL = ES between .2
Kilbourne, et al. [114]	68 with BD and cardiovascular disease risk	Collaborative Care: Lifegoals (TAU plus newsletters)	**Functioning:** Mild to moderate impairment **18.8 (8.7)**	Earlier	Functioning (WHODAS 2.0 12 item) Quality of Life (SF-12)	No significant differences in functioning (ES .20) or quality of life - mental ES .01, physical ES .12
Kilbourne, et al. [115]	118 with BD and cardiovascular disease risk	Collaborative Care: Lifegoals (TAU plus newsletters)	**Functioning:** Mild to moderate impairment **17.26 (9.64)**	Earlier	Functioning (WHODAS 2.0 12 item) Quality of Life (SF-12)	No significant differences in functioning (ES.14) or quality of life (mental ES -.04; physical ES -.07)

Kurdal, et al.[116]	80 euthymic individuals with BD	PE (waitlist TAU)	**NoE**: 58.75% had ≤4 episodes	Earlier	Functioning (Bipolar Disorder Functioning Questionnaire)	Significant improvement in functioning across the domains of emotional functioning (within ES .81†; ES .91†), intellectual functioning (ES 1.67†), feelings of stigmatization (ES 1.75†), social withdrawal (ES 1.56†), household relations (ES 1.75†), relations with friends (ES .77†), participating in social activities (ES .81†), daily and recreational activities (ES .70†), taking initiative and self-sufficiency (ES .80), and occupation (ES .48†) relative to control. No significant improvement in the sexual functioning subscale.	BDFQ does not have available norms/cut-offs Unbalanced NoE in controls versus experimental group, limiting conclusions

Lahera, et al. [117]	37 euthymic adults with BD	Social cognition and interaction training (TAU)	**NoE:** Less than 10 NoE **7.79 (3.25)** **Functioning:** At least mild impairments **GAF 78.78 (15.30) FAST 40.43 (16.36)**	Earlier	Functioning (GAF; FAST)	No significant improvements in functioning following the intervention GAF ES .01, FAST ES -.43 (attributable to baseline differences between groups)	Functional measures conflicting at baseline: FAST in moderate-severe range, GAF in mild impairment range
Lewandowski, et al. [118]	72 adults with BD and a history of psychosis	Computer-based cognitive remediation (computer-based control)	**Cognitive functioning:** Less than .5 SD below the mean on the MATRICS 45.33 (9.77) T 50, SD 10	Earlier	Cognitive functioning (MATRICS)	Significant effects over control in several cognitive domains including processing speed (ES .42) visual learning and memory (ES .92), and the cognitive composite (ES .80).	
Meyer, et al. [26]	76 euthymic adults with BD	CBT (supportive psychotherapy)	**NoE:** Median of 6 **(50% with 6 or less)** Functioning: mild-moderate impairment GAF 69.6 (12.61) **T 50, SD 10**	Earlier	Functioning (GAF)	Significant improvements over time but no significant differences between groups - between ES - .20 in favour of supportive	Active control

			Cognitive impairment: Significant impairment exclusion criteria **58.25 (6.81)**				
Nagy, et al. [119]	111 with BD and their caregivers	Family PE (supportive psychotherapy)	**NoE:** Not provided, but moderation analyses conducted	No information	Functioning (SFQ) Quality of Life (W-QLI)	PE obtained greater percentage of change in scores on SFQ (between ES .52) and W-QLI (between ES .55), but not response rate in comparison with the control group. These results were not moderated by NoE.	Active control Evidence of moderation: results were not moderated by NoE.
Patelis-Siotis, et al. [120]	49 euthymic or mildly depressed adults with BD	Group CBT (none)	**Functioning:** Mild-moderate impairment **GAF 62.20 (8.13)**	Earlier	Functioning (GAF) Quality of Life (SF-36)	Significant improvements in QoL; specifically, vitality and emotional role and functioning were observed post-intervention.	

Scott, et al. [121]	42 with BD in any mood state, with ≥1 episode in the last 2 years	Cognitive therapy (waitlist)	**Functioning:** Mild-moderate impairments **70.65 (11.96)**	Earlier	Functioning (GAF)	Significant improvement in functioning pre-post and relative to control (ES .57†)
Smith, et al. [12]	50 euthymic adults with BD	Online PE (TAU)	**Functioning:** Mild-moderate impairments GAF 70.95 (14.35) FAST 24.12 (15.87)	Earlier	Functioning (GAF; FAST) Quality of Life WHOQOL–BREF)	No significant improvements in overall QoL (.04) or functioning (GAF and FAST .26). However, the domain of psychological QoL improved relative to control (ES .23†)
Sylvia, et al. [122]	6 participants with BD in any mood state	Lifestyle-focused CBT (none)	**Functioning:** Mild impairment **12 (3.1)**	Earlier	Functioning (LIFE-RIFT) Quality of Life (Quality of Life scale)	Functioning improved post intervention No improvements in quality of life were observed post-intervention
Sylvia, et al. [123]	5 overweight or obese participants with BD	Lifestyle-focused CBT (none)	**Functioning:** Mild impairment **(11.8 3.4)**	Earlier	Functioning (LIFE-RIFT)	Functioning improved post intervention

Wenze, et al.[124]	30 adult inpatients and outpatients with BD and comorbid SUD	Multimodal CBT-based with ACT and family therapy (enhanced TAU)	**Functioning:** Mild-moderate impairments **18.83 (9.33)**	Earlier	Functioning (WHODAS 2.0 12 item)	Small (ES .12) significant improvements in functioning relative to control
Zaki, et al.[125]	111 participants with BD in any mood state and caregivers	Family PE (Supportive Psychotherapy)	**NoE:** < 10 **5.85 (3.08)** **Functioning:** Mild to moderate impairment **SFQ 107.57 (10.27)**	Earlier	Functioning (SFQ) Quality of Life (W-QLI)	Both groups showed improved functioning and QoL. Significantly greater improvements in social functioning (ES .93†) and QoL (ES 1.03†) in the PE group.

BD = Bipolar Disorder; TAU = Treatment as Usual; ED = effect size; SD = Standard Deviation; SUD = Substance Use Disorder; NoE = Number of Episodes; QoL = Quality of Life; FR = Functional Remediation; CBT = Cognitive behavioural therapy; DBT = Dialectical Behaviour therapy; PE = Psychoeducation; FFT = Family focused therapy; IPSRT = Interpersonal and Social Rhythm Therapy; CT = Cognitive Therapy; MBCT = Mindfulness-based Cognitive Therapy; GAF = Global Assessment of Functioning; MCCB = MATRICS Consensus Cognitive Battery (MCCB); FAST = Functioning Assessment Short Test; LIFE-RIFT = Longitudinal Interval Follow-up Evaluation-Range of Impaired Functioning Tool; WHODAS 2.0 36 = World Health Organization Disability Assessment Schedule 2.0, 36 item; WHODAS 2.0 12 = World Health Organization Disability Assessment Schedule 2.0, 12 item; SOFAS = Social and Occupational Functioning Assessment Scale; SDS = Sheehan Disability Scale; IS = Inadaptation Scale; SF12 = MOS Short Form 12; SF36 = MOS Short Form 36; Q-LES-Q = Quality of Life Enjoyment and Satisfaction Questionnaire; QoL.BD = brief Quality of Life in Bipolar Disorder questionnaire; SFQ = Social Functioning Questionnaire; W-QLI = Wisconsin Quality of Life Index; SPS = Social Performance Scale; SAS = Social Adjustment Scale; UCLA SAS = UCLA Social Attainment Scale; PSP = Personal and Social Functioning Scale; WHO-QOL-BREF = World Health Organization Quality of Life Assessment; SASS = Social Adaptation Self-Evaluation Scale; BRIEF = Behavior Rating Inventory of Executive Function; FrSBe = Frontal Systems Behavior Scale; RAVLT = the Rey Auditory Verbal Learning Test; CFQ = Cognitive Failures Questionnaire. Effect sizes are as reported. Where these were not reported data were available, these were calculated per relevant contrast using Cohen's *d*, Hedge's *g* or Glass' Δ as appropriate, denoted by †. Note effect sizes are inflated with small samples.

Table 4

Summary of Studies Meeting Later Stage Criteria

Study	Sample	Intervention (control)	Stage proxy and information	Stage	Outcome (measures)	Findings	Notes
Ball, et al.[126]	52 individuals with BD I or II on medication	Cognitive Therapy (TAU)	**NoE:** Mean lifetime NoE 10 episodes **14.11 (10.23)** **Functioning:** GAF score in the MILD range at baseline **77.58 (13.12)**	Later	Functioning (GAF, SPS, SAS)	No significant improvements in global or social functioning post-treatment (GAF between ES .43) or at follow-up (GAF between ES.25, SAS .17, SPS ;02) relative to control	Differences in functional ratings between measures at baseline of comparable magnitude in same direction (no change score available due to SD)

Bauer, et al. [127]	306 veterans with BD	Bipolar Disorders Program: Intensive Chronic Collaborative Care (TAU)	**Comorbidity:** Lifetime SUD 72%, lifetime Anxiety Disorder 42% **Functional impairment:** Social role dysfunction scores in severe to very severe range **Ms and SDs not provided.**	Later	Functioning (SAS) Quality of Life (SF-36)	Improved mental QoL over time, relative to control, (ES .27) Improved role functioning over time, relative to control (ES overall .30, parental 1.04, extended family.32)
Ellard, [128]	29 euthymic with BD and at least one anxiety disorder	CBT – the Unified Protocol for Emotional Disorders (TAU)	**Functioning: M**oderate impairment **11.99 (2.81)**	Later	Functioning (LIFE-RIFT)	No significant improvements in functioning relative to TAU. (ITT between groups ES 0.48 completers 0.86)

Kilbourne, et al. [129]	290 Veterans	Bipolar Disorders Program: Intensive Chronic Collaborative Care (TAU)	**Comorbidity:** At least one comorbid anxiety disorder	Later	Quality of Life (SF-36)	Effectiveness of the intervention for improving quality of life was not moderating by current comorbid anxiety disorders, substance use disorder, or psychosis	
Bonnin, et al. [28] Torrent, et al. [29]	239 Euthymic participants with BD, moderate to severe functional impairment and no current psychiatric comorbidity (172 at 1 year follow-up)	FR (TAU, PE)	**Functioning:** Moderate to severe impairment (FAST) **29.95 (9.84) Both** **NoE:** Mean in each group above 10 **11.59 (12.33) Both**	Later	Functioning (FAST) Cognitive functioning (comprehensive neuropsychological battery)	FR group had significantly improved functioning at post-intervention and follow-up FR was superior to TAU post-intervention (ES .3) and to both PE (ES .16†) and TAU (ES .18) at follow-up. Only significant at the domain level for autonomy (relative to PE PE ES .22†, TAU ES .26†) at follow-up. Improved verbal memory at f/up (within ES .25†) and compared with	Trial and follow-up – full sample

						TAU (ES .17†) but no other cognitive domain.	
Castle, et al. [130]	17 euthymic participants with BD	Group intervention including CBT, PE and DBT elements (supportive phone calls)	**Functioning:** Moderate to severe impairment **58.65 (10.94)**	Later	Functioning (GAF) Quality of Life (WHO-QOL-Bref)	Functioning improved in the treatment group relative to TAU (ES 1.00†), only the social relationships domain of quality of life improved significantly relative to control.	
de Barros Pellegrinelli, et al. [131]	55 euthymic individuals with BD aged 18+, receiving TAU pharmacotherapy	PE (group control)	**NoE**: 63% of sample ≥10, 80% ≥6 **GAF not reported** **Comorbidity:** Axis I comorbidity present in 28% of the sample	Later	Functioning (GAF) Quality of Life (WHO-QOL)	No differences in QoL or functioning (ES SAS .08 GAF .1) between groups, but QoL worsened pre-post	About 60% of patients in the sample had more than 10 or countless mood episodes, and 80% had six or more

			Functioning: Mild psychosocial impairment, moderate to severe work impairment **11.29 (3.16)**			
Deckersbach, et al. [132]	18 euthymic adults with BD	Cognitive remediation (none)	**Cognitive impairment:** Below average performance in the areas of executive function, attention and memory **Comorbidity:** Rates not specified	Later	Cognitive functioning (various assessments) Functioning (LIFE-RIFT)	Improved occupational functioning pre-post and follow-up Improved psychosocial functional pre-post and follow-up.	Evidence of moderation: Improvements in occupational functioning were partially moderated by improved executive functioning and history of comorbid anxiety disorders
Deckersbach, et al. [133]	32 participants with BD, current	CBT (supportive psychotherapy)	**Functioning:** Moderate impairment 14.6 (3.7)	Later	Functioning (LIFE-RIFT)	No significant difference from control; both groups experienced improved functioning	Improved functioning may be linked to significant

	MDE		**Cognitive functioning:** 1 – 1.5 SDs below the mean across a number of tasks			improvement in depression in both groups Active control
Demant, et al.[134]	46 participants with BD and subjective cognitive difficulties	Group cognitive remediation (TAU)	**NoE:** Mean over 10 **21.89 (21.54)** **Functioning:** Moderate impairment (FAST) **27.70 (10.50)** **Cognitive functioning:** Subjective cognitive difficulties (CPFQ)	Later	Functioning (FAST), Quality of Life (WHOQOL) Cognitive functioning: (RAVLT, CFQ)	No significant improvement in objective cognitive functioning (9 different tests). FAST (ES b/w .0042), or overall QoL (ES b/w.47) at follow-up. Superior improvement to TAU in the psychological domain of QoL (ES .78†). At follow-up, subjective mental acuity (ES *partial Eta squared* = 0.19) and some aspects of verbal fluency (up to ES b/w .26) had improved

Fiorillo, et al.[135]	137 participants with BD, any mood state and at least one episode within 3 years	PE (TAU)	**Functioning:** Poor adjustment **(obvious dysfunction = 2, serious dysfunction = 3) 2.9 (.95)**	Later	Functioning (WHODAS)	Significant improvements in functioning scores over time relative to control (ES .52†)	Real world trial

Gonzalez Isasi, et al. [136] Gonzalez Isasi, et al. [137]	40 euthymic participants with BD and ≥2 episodes in previous year or severe difficulties in functioning	CBT with PE (TAU)	**NoE:** Mean over 10 **10.75 (no SD) range 2 - 18, 45% had severe inadaptation scores**	Later	Functioning (IS)	Functioning significantly improved compared with control (ES = 1.8) and these improvements were consistent across 5 years (ES = 2.72).	This is potentially an overlapping sample with Gonzalez Isasi, et al. [50]
Haffner, et al. [138]	34 euthymic BD with some functional impairment	Metacognitive training (none)	**NoE:** Mean of 22.1 (12.6) **Functioning:** all had at least mild functional impairments **Mdn 24.5 (IQR 18.0)**	Later	Functioning (FAST) Quality of Life (WHO-QOL)	Overall functioning improved significantly post-treatment and in the domains of autonomy, occupational, cognitive, and interpersonal functioning. Those with worse functioning improved more. QoL did not improve significantly	

Hidalgo-Mazzei, et al. [139]	201 adults with a diagnosis of BD	PE smartphone app (none)	**NoE:** Mean above 10 **13.1 (5.9)** **Comorbidity:** Above 60%	Later	Quality of Life (SF-36)	Of the completers, quality of life improved in 6 of 8 domains: general health perceptions, mental health, physical role functioning, emotional role functioning, social functioning, and bodily pain. (ES within from .40 -.78)	Only 51% completed and outcomes may be influenced by dropouts
Hoberg, et al. [140]	9 adults with current depression	Intensive short-term group IPSRT (none)	**Functioning:** Moderate – severe impairment 19.88 (5.57)	Later	Functioning (SDS)	Functioning was improved following the intervention and reached significance at 12-week follow-up	
Ives-Deliperi, et al. [141]	23 euthymic adults with BD	Mindfulness-based cognitive therapy (waitlist, healthy controls)	**Cognitive impairment:** Significantly lower scores across several domains (executive functioning including working memory and inhibition) - can't	Later	Cognitive functioning (neuropsychological battery)	Significantly improved performance in neuropsychological tasks measuring working memory (ES .38) spatial memory (ES .60), and verbal fluency (ES.35), following MBCT relative to waitlist.	

			assess against norms,				
Jones, et al.[142]	67 euthymic individuals with BD diagnosis within 5 years	Functional CBT (TAU)	**NoE:** 58% with ≥6 **Functioning:** Mean score in the 'mild manifest problems' range **70.21 (17.45)**	Later	Functioning (PSP) Quality of life (QoL.BD)	Non-significant trends towards improved QoL (b/w ES d=0.47) and functioning (b/w ES d=0.35) relative to TAU	Despite being aimed at recent onset BD, proportion of individuals with ≥6 NoE, and manifest functional issues are more likely to reflect later stage
Jones, et al.[143]	72 euthymic adults with BD and current anxiety	CBT targeted for BD and anxiety (TAU)	**NoE: 93% had 8 or more;** 70% had more than 20 episodes **Functioning:** Mild impairment **72.48 (11.23)** **Comorbidity:** Over 85 had a	Later	Functioning (PSP) Quality of Life (QoL.BD)	No significant improvements in functioning (between ES 0.089) or quality of life (Between ES 0.158) compared with TAU	

			current anxiety disorder				
Lam, et al. [144]	25 euthymic participants with BD I (and at least 2 episodes within 2 years or 3 within 5 years)	Cognitive therapy (minimum psychiatric care)	**NoE:** Mean >10 **17.28 (11.32)**	Later	Functioning (SPS)	Significant improvement in SPS scores at 6 (ES 1.04†) and 12 months (ES 1.04†) relative to control	
Lam, et al. [145], Lam, et al. [146]	103 euthymic participants with BD I (and at least 2 episodes within 2 years or 3 within 5 years)	Cognitive therapy (minimum psychiatric care)	**NoE:** Mean >10 **10.92 (8.40)**	Later	Functioning (SPS)	No significant differences in SPS scores between groups post-intervention (ES .26) or at 6-month follow-up (ES .18) When NoE was controlled for, CT was superior to control at 24-month follow-up (ES .73†), but not at 18 (ES = 0) or 30-month (ES.36) follow-up.	Evidence of moderation: When NoE was controlled for, CT was superior to control at 24-month follow-up (ES .73†),

Lenz, et al. [147]	100 euthymic participants with at least 2 episodes of illness in the previous 3 years or at least 3 episodes in 5 years	Group Cognitive-psychoeducational Therapy (bibliotherapy book and three group sessions)	**Functioning:** moderate impairment, 76.6% were impaired on at least one functioning scale. **Work: 5.53 (3.17), Leisure 5.31 (3.08)** **Comorbidity:** Over 40% had at least one psychiatric comorbidity	Later	Functioning (SDS)	While functioning improved in both groups, no benefit was observed over control (ES between .19, .03, .30) for the three SDS scales	Active control
Miklowitz, et al. [148]	152 outpatients with BD, current MDE	Intensive intervention CBT, IPSRT, FFT (Brief psychoeducational Collaborative Care)	**NoE**: 68% had > 10 manic and 69% > 10 depressed episodes **No M/SD** **Functioning:** moderate	Later	Functioning (LIFE-RIFT)	Patients in intensive psychotherapy had better total functioning (ES .34) relationship functioning (ES .37) , and life satisfaction scores over 9 months than patients in collaborative care. These remained even after pre-treatment functioning scores were covaried. No effects of psychosocial	

			impairment **51.79 (7.68)**			intervention were observed on work/role functioning	
Morriss, et al. [149]	304 euthymic individuals with BD with at least 1 episode within 24 months	PE (group peer support)	**NoE:** 87% had 8 or more episodes, 56% had 20 or more	Later	Functioning (SAS and SOFAS) Quality of Life (SF-12)	No benefit of PE relative to control (peer support) for improved functioning as assessed with the SAS (ES .26) and SOFAS (ES .16), nor QoL (ES physical .18, mental .14)	Active control
Murray, et al. [23]	26 individuals with 6+ mood episodes	Online mindfulness intervention (none)	**NoE:** Median of 16 episodes **R (6-248)**	Later	Quality of Life (QoL.BD)	Improvements pre-post in QoL. Symptom change was not significant	
Sajatovic, et al. [150]	164 outpatients with BD –	Collaborative Care – Lifegoals	**Functioning:** Moderate impairment **57.35**	Later	Functioning (GAF)	No significant differences in functioning between groups post-intervention or at follow-up (ES .06)	

	any mood state	(TAU)	**(12.21)**				
Schottle, et al.[151]	23 adults with BD I and a history of psychosis and 'severe' illness features	Assertive long-term community management including psychotherapy - ACCESS II (none)	**Functioning:** **Severe impairment 38 (12.8)** **Comorbidity:** 70% with psychiatric comorbidity	Later	Functioning (GAF) Quality of life (Q-LES-Q 18)	Functioning and QoL improved significantly over time	24-month follow-up
Schottle, et al.[152]	23 adults with BD I and a history of psychosis and 'severe' illness features	Assertive long-term community management including psychotherapy - ACCESS II (none)	**Functioning:** Severe impairment **38 (12.8)** **Comorbidity:** 70% with psychiatric comorbidity	Later	Functioning (GAF) Quality of life (Q-LES-Q 18)	Functioning and QoL improved significantly over time.	48-month Follow-up

Todd, et al. [153]	122 adults with BD	Recovery informed online CBT with PE (waitlist)	**NoE:** 56% had 12 or more depressed and 74% has 7 or more depressed eps. **36% had 12 or more manic, 54% had 7 or more manic Functioning:** Mild impairment **34.97 (8.70)**	Later	Functioning (SASS) Quality of Life (QoL.BD, WHOQOL–BREF)	Functioning significantly improved relative to control (ES .54) QoL (ES .50) QoL.BD scores became nonsignificant, as did some WHO-QOL domains, when controlling for baseline characteristics (not specified) – but psychological (ES .50) and physical (ES .60) quality of life remained significant.
van der Voort, et al. [154]	138 euthymic participants with BD and 'unstable' illness course	Intensive Collaborative Care with PE and problem-solving therapy	**NoE:** Most had ≥5 manic and ≥ depressed episodes**. 64% had experienced 5 or more manic, 36% 10 or more. 71% had experienced 5 or depressed, 47% had experienced**	Later	Functioning (FAST)	Functioning significantly improved at 12 months (ES .30) and specifically in the autonomy (ES .50), cognitive (ES .40) and leisure (ES .40) domains. No significant improvements to QoL except physical health QoL (ES .40)

			10 or more depressed.			
			Functioning: Moderate		Quality of Life WHOQOL–BREF	
Veeh, et al. [155]	26 euthymic participants with BD and objective cognitive impairments	Cognitive remediation (TAU)	**Cognitive functioning:** Below average performance relative to normative data Ms and SDs for overall cog not supplied.	Later	Functioning (Mini- ICF-App Social Functioning Scale) Quality of Life (WHOQOL-Bref) Cognitive functioning	No significant improvement in functioning (ES between .2) or QoL (ES between .001) , Significant improvement of cognitive performance after CR relative to control was observed in working memory (ES .76), and problem solving (ES .80).

			in at least two of seven test measures		(neuropsychological battery)	
Zyto, et al.[156]	12 euthymic participants with BD and at least moderate functional impairment	Functional remediation (none)	**NoE:** Mean of 10 **(3.28)** **Functioning:** Moderate impairments 29.3 (6.1) **Cognitive functioning:** Some cognitive impairments	Later	Functioning (FAST)	Significant improvements in global functioning and specifically, within the domains of autonomy and occupational functioning.

BD = Bipolar Disorder; TAU = Treatment as Usual; ED = effect size; SD = Standard Deviation; SUD = Substance Use Disorder; NoE = Number of Episodes; QoL = Quality of Life; FR = Functional Remediation; CBT = Cognitive behavioural therapy; DBT = Dialectical Behaviour therapy; PE = Psychoeducation; FFT = Family focused therapy; IPSRT = Interpersonal and Social Rhythm Therapy; CT = Cognitive Therapy; MBCT = Mindfulness-based Cognitive Therapy; GAF = Global Assessment of Functioning; MCCB = MATRICS Consensus Cognitive Battery (MCCB); FAST = Functioning Assessment Short Test; LIFE-RIFT = Longitudinal Interval Follow-up Evaluation-Range of Impaired Functioning Tool; WHODAS 2.0 36 = World Health Organization Disability Assessment Schedule 2.0, 36 item; WHODAS 2.0 12 = World Health Organization Disability Assessment Schedule 2.0, 12 item; SOFAS = Social and Occupational Functioning Assessment Scale; SDS = Sheehan Disability Scale; IS = Inadaptation Scale; SF12 = MOS Short Form 12; SF36 = MOS Short Form 36; Q-LES-Q = Quality of Life Enjoyment and Satisfaction Questionnaire; QoL.BD = brief Quality of Life in Bipolar Disorder questionnaire; SFQ = Social Functioning Questionnaire; W-QLI = Wisconsin Quality of Life Index; SPS = Social Performance Scale; SAS = Social Adjustment Scale; UCLA SAS = UCLA Social Attainment Scale; PSP = Personal and Social Functioning Scale; WHO-QOL-

BREF = World Health Organization Quality of Life Assessment; SASS = Social Adaptation Self-Evaluation Scale; BRIEF = Behavior Rating Inventory of Executive Function; FrSBe = Frontal Systems Behavior Scale; RAVLT = the Rey Auditory Verbal Learning Test; CFQ = Cognitive Failures Questionnaire. Effect sizes are as reported. Where these were not reported data were available, these were calculated per relevant contrast using Cohen's *d*, Hedge's *g* or Glass' Δ as appropriate, denoted by †. Note effect sizes are inflated with small samples

Appendices

Appendix A

Dominant Proposed Disorder-specific Staging Models in BD

Stage	Berk et al. model	Kapczinski et al. model	Proposed features	Current review
Increased Risk	Stage 0	Latent phase	Well. Familial risk is the only confirmed risk factor. Some support for subthreshold sleep issues [1], while other risk factors such as trauma and substance use remain unconfirmed [2]	Not considered
Mild/non-specific symptoms	Stage 1a	Latent phase	Emergence of generic or subthreshold symptoms, mood dysregulation, anxiety disorders [1,3] sleep disorders or ADHD [1]	
Prodromal features (Ultra-high risk)	Stage 1b	Latent phase	While often indistinct from the prodrome of other disorders, symptoms of either pole (dysthymia, cyclothymia) may emerge [2,4]	
First episode	Stage 2	I	First episode (diagnosis at first elevated episode despite initial episodes most commonly depressive)[1,5]	Aligned with Earlier in the current review
Subthreshold recurrence	Stage 3a	I	Symptom recurrence without full threshold episode [2]	
First relapse	Stage 3b	II	Relapse (most often depressive) likely to be related to discernible life events[6], inter-episode symptoms related to comorbidities emerge [7]	
Multiple relapses	Stage 3c	III	Some discernible neurocognitive and functional impairment, worsening inter-episode functioning, residual symptoms [8,9]	Aligned with Later in the current review
Persistent unremitting illness	Stage 4	III-IV	Significant neurocognitive and functional impairment, ultimately unable to live independently, episodes	

	more severe and autonomous, treatment resistance[9]	

Appendix B

Characteristics of Psychosocial Interventions

Brand	General contents	Specific refinements within included studies
Cognitive behavioural therapy (CBT)	An intervention focused on the reciprocal relationship between thoughts, behaviours and emotions. Involves identifying and modifying dysfunctional and negative thoughts, underlying maladaptive assumptions and beliefs, as well as implementing behavioural changes.	RfCBT; focus eliciting and targeting client focussed recovery goals [10]. Several trials implemented transdiagnostic approaches [11,12]
Cognitive therapy	An intervention focused on the role of cognitions in emotional and behavioural difficulties. Shifting cognition is seen as the main mechanism by which changes occur.	
Interpersonal and social rhythm therapy	An adaptation of interpersonal therapy (IPT) for bipolar disorder, incorporating emphasis the relationship between disrupted circadian and social rhythms and bipolar episodes	
Psychoeducation	A structured educational intervention to help individuals experiencing bipolar disorders to become experts in their own	Psychoeducation was delivered individually or in group settings, in

	condition to improve medication adherence, mood stability and self-management. Psychoeducation broadly includes provision of information about the nature of the illness, its treatments, and key coping strategies to the patient and family.	person or via digital health platforms
Family-focused therapy	A psychoeducation-based intervention delivered to individuals and their families. FFT focuses on communication styles between patients and their families or spousal relationships, with the goal of improving relationship functioning as well as disorder-specific risks, and problem-solving skills in the family.	
Mindfulness-based cognitive therapy (MBCT)	An integrated approach combining cognitive therapy with mindfulness, drawn from mindfulness-based stress reduction. MBCT differs from CBT as it does not emphasize changing the content of thoughts but changing one's awareness of and relationship to thoughts, feelings, and sensations; observing and accepting these without judgment.	
Metacognitive therapy	Metacognition is the part of cognition that controls mental processes and thinking. This approach focuses on metacognitive beliefs that maintain psychological symptoms and how to modify these.	
Functional remediation	An approach which focuses on the functional and cognitive impairments common within bipolar disorder; aimed at enhancing functioning within one's daily life. This is a structured approach involving neurocognitive training, psychoeducation on cognition-related issues, and problem-	

	solving within an ecological framework; aimed at the application of specific skills within one's routine.	
Cognitive remediation/rehabilitation	An intervention which targets cognitive functioning. Interventions typically involve psychoeducation about cognitive dysfunction in BD, identification of practical compensatory and adaptive strategies and neurocognitive training involving a range of computer-based or pencil and paper tasks.	
Social Cognition and Interaction Training	A manualised group intervention targeting emotion perception, attributional style, and theory of mind abilities.	
Dialectical Behavioural therapy	A therapy aimed at changing patterns of unhelpful thinking. This therapy combines CBT techniques with distress tolerance, emotion regulation, acceptance, and mindfulness strategies.	
Collaborative care	A community based, multidisciplinary approach typically coordinated by a case manager who coordinates the care including; medical input, medication management, psychotherapy, vocational and social support.	Widely varying applications and parameters, from brief psychoeducational to long-term, spanning 3 years.

Appendix C

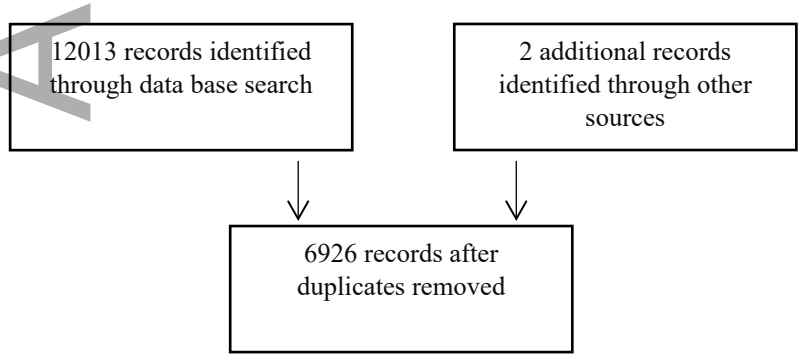
Scopus Search Terms

Search terms
TITLE-ABS-KEY (bipolar OR mania OR manic* OR "manic depress*") **AND TITLE-ABS-KEY** ("psychological therap*" OR psychotherap* OR psychosocial OR "self-management" OR "Psychological intervention" OR cbt OR "cognitive therapy" OR

"behavior therapy" OR "behavioral therapy" OR "cognitive behavioural" OR "cognitive behavior" OR mindfulness OR mbct OR "collaborative care" OR ipsrt OR "social rhythm" OR interpersonal OR "Acceptance and Commitment Therapy" OR "ACT" OR dbt OR "dialectical behaviour therapy" OR psychoeducation OR "family therapy" OR "family focused" OR carer OR "functional remediation" OR "cognitive remediation" OR "schema therapy") **AND TITLE-ABS-KEY** (stage* OR staging OR predict* OR mediat* OR moderat* OR neurocognit* OR cognit* OR comorbid* OR episodes OR "number of episodes" OR function* OR trajectory OR "illness history" OR "illness characteristics" OR "course of illness") **AND TITLE-ABS-KEY** (recovery OR "patient-reported outcomes" OR "quality of life" OR function* OR employment OR autonomy OR wellbeing OR social OR disability OR meaning OR vocation* OR cognit* OR acceptance OR self-compassion)

Appendix D

PRISMA flow chart



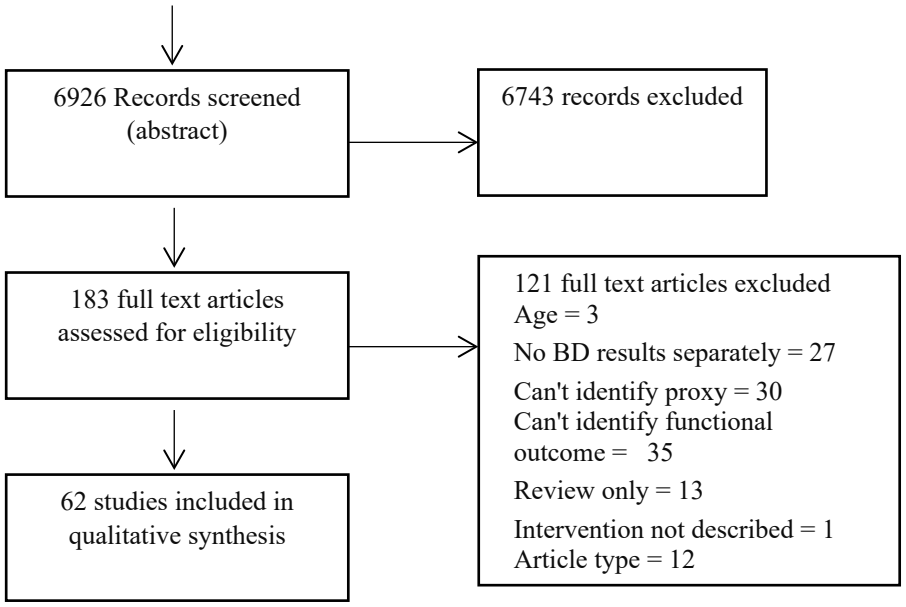
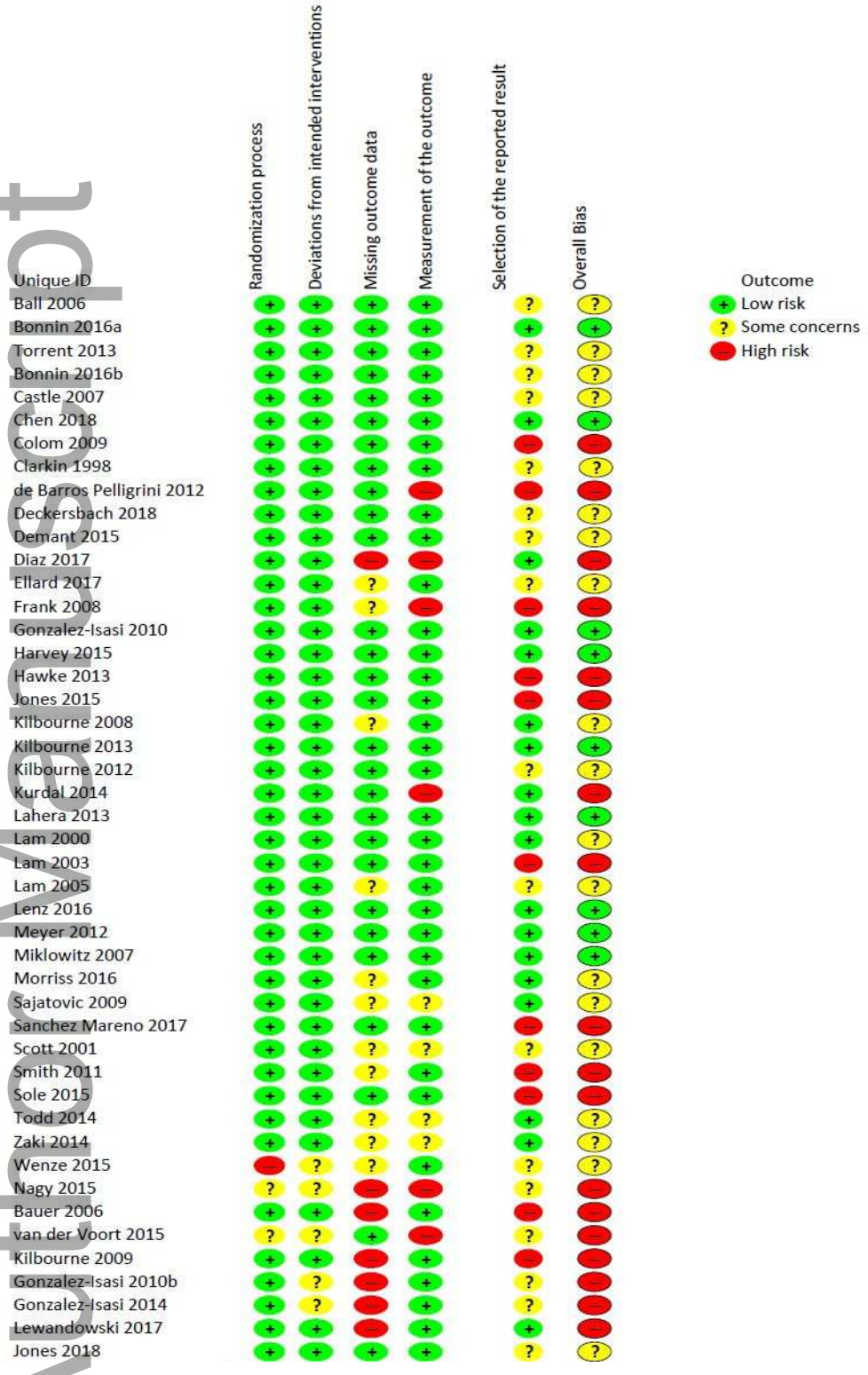
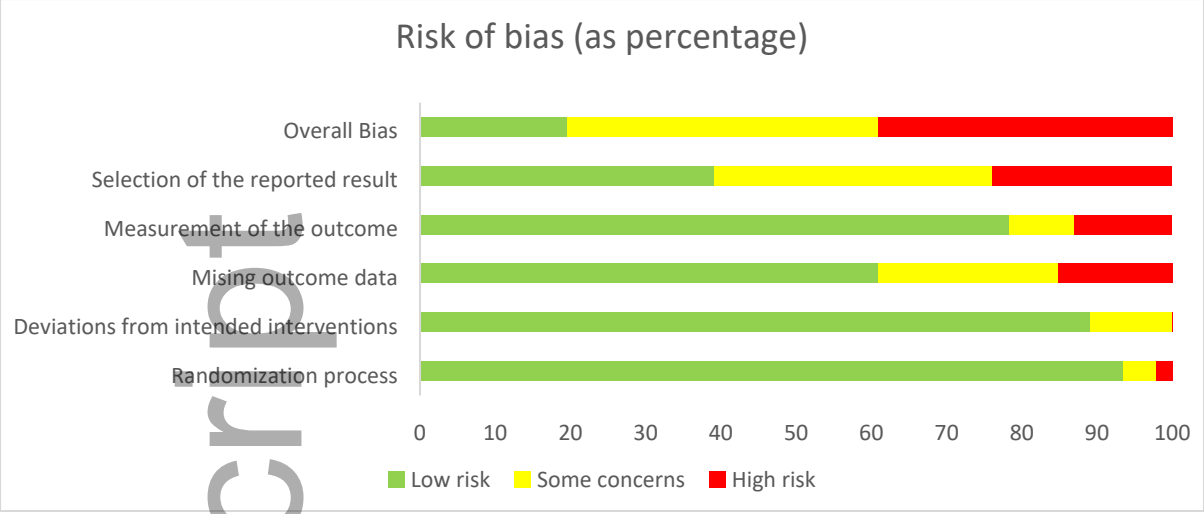


Figure A1. **PRISMA Flow chart from search results to qualitative synthesis**

Appendix E

Risk of Bias Assessment





Appendix F

GRADE Rating

GRADE criteria	Rating (circle one)	Footnotes (explain reasons for down- or upgrading)	Quality of the evidence (Circle one)
Outcome: General functioning			
Study design	RCT (starts as high quality) **Non-RCT (starts as**	Most are RCTS	⊕⊕⊕⊕ High
Risk of Bias *(use the Cochrane Risk of Bias tables and figures)*	No **serious (-1)** **very** serious	Serious risk	⊕⊕⊕□ Moderate
Inconsistency	**No** serious (-1) very	Some inconsistency but not serious	⊕⊕□□ Low
Indirectness	**No** serious (-1) very	Adequate directness	
Imprecision	No **serious (-** **1) very**	Issues with sample sizes	⊕□□□ **Very Low**
Publication Bias	**Undetected** Strongly suspected (-1)	Many negative trials found but some risk	
Other (upgrading factors, circle all that apply)	Large effect (+1 or +2) Dose response (+1 or +2) No Plausible	No upgrade factors	

GRADE criteria	Rating (circle)	Footnotes (explain reasons for down- or)	Quality of the evidence (Circle one)
Outcome: Cognitive functioning			
Study design	RCT (starts as high quality) **Non-RCT (starts as**	Many are not RCTS	⊕⊕⊕⊕ High
Risk of Bias *(use the Cochrane Risk of Bias tables and figures)*	No **serious (-1)** **very** serious	Serious risk	⊕⊕⊕□ Moderate
Inconsistency	**No** serious (-1) very	Some inconsistency but not serious	⊕⊕□□ Low
Indirectness	**No** serious (-1) very	Adequate directness	
Imprecision	No **serious (-** **1) very**	Issues with sample sizes	⊕□□□ **Very Low**
Publication Bias	Undetected **Strongly suspected (-1)**	No negative trials found	
Other (upgrading factors, circle all that apply)	Large effect (+1 or +2) Dose response (+1 or +2) No Plausible	No upgrade factors	

GRADE criteria	Rating (circle)	Footnotes (explain reasons for down- or)	Quality of the evidence (Circle one)
Outcome: Quality of life			
Study design	RCT (starts as high quality) **Non-RCT (starts as**	Many are not RCTS	⊕⊕⊕⊕ High
Risk of Bias *(use the Cochrane Risk of Bias tables and figures)*	No **serious (-1)** **very** serious	Serious risk	⊕⊕⊕□ Moderate
Inconsistency	**No** serious (-1) very	Some inconsistency but not serious	⊕⊕□□ Low
Indirectness	**No** serious (-1) very	Adequate directness	⊕□□□ **Very Low**
Imprecision	No **serious (-** **1) very**	Issues with sample sizes	
Publication Bias	Undetected **Strongly suspected (-1)**	Some risk, some negative trials published	
Other (upgrading factors, circle all that apply)	Large effect (+1 or +2) Dose response (+1 or +2) No Plausible	No upgrade factors	

References

1. Duffy A, Goodday S, Keown-Stoneman C, Grof P. The emergent course of bipolar disorder: Observations over two decades from the canadian high-risk offspring cohort. *Am J Psychiatry.* 2018:appiajp201818040461.

2. Berk M, Conus P, Lucas N, et al. Setting the stage: From prodrome to treatment resistance in bipolar disorder. *Bipolar Disord.* 2007;9(7):671-678.
3. Duffy A, Horrocks J, Doucette S, Keown-Stoneman C, McCloskey S, Grof P. The developmental trajectory of bipolar disorder. *Br J Psychiatry.* 2014;204(2):122-128.
4. Skjelstad DV, Malt UF, Holte A. Symptoms and signs of the initial prodrome of bipolar disorder: A systematic review. *J Affect Disord.* 2010;126(1-2):1-13.
5. Mesman E, Nolen WA, Reichart CG, Wals M, Hillegers MH. The dutch bipolar offspring study: 12-year follow-up. *Am J Psychiatry.* 2013;170(5):542-549.
6. Post RM. Kindling and sensitization as models for affective episode recurrence, cyclicity, and tolerance phenomena. *Neurosci Biobehav Rev.* 2007;31(6):858-873.
7. Kapczinski F, Dias VV, Kauer-Sant'Anna M, et al. Clinical implications of a staging model for bipolar disorders. *Expert Rev Neurother.* 2009;9(7):957-966.
8. Robinson LJ, Ferrier IN. Evolution of cognitive impairment in bipolar disorder: A systematic review of cross-sectional evidence. *Bipolar Disord.* 2006;8(2):103-116.
9. Rosa AR, Magalhaes PV, Czepielewski L, et al. Clinical staging in bipolar disorder: Focus on cognition and functioning. *J Clin Psychiatry.* 2014;75(5):e450-456.
10. Jones SH, Smith G, Mulligan LD, et al. Recovery-focused cognitive-behavioural therapy for recent-onset bipolar disorder: Randomised controlled pilot trial. *Br J Psychiatry.* 2015;206(1):58-66.
11. Jones SH, Knowles D, Tyler E, et al. The feasibility and acceptability of a novel anxiety in bipolar disorder intervention compared to treatment as usual: A randomized controlled trial. *Depress Anxiety.* 2018;35(10):953-965.
12. Ellard KK, Bernstein EE, Hearing C, et al. Transdiagnostic treatment of bipolar disorder and comorbid anxiety using the unified protocol for emotional disorders: A pilot feasibility and acceptability trial. *J Affect Disord.* 2017;219:209-221.

ESTIMATED DISTRIBUTIONS OF NUMBER OF EPISODES BY STUDY (EARLIER VERSUS LATER GROUPS)

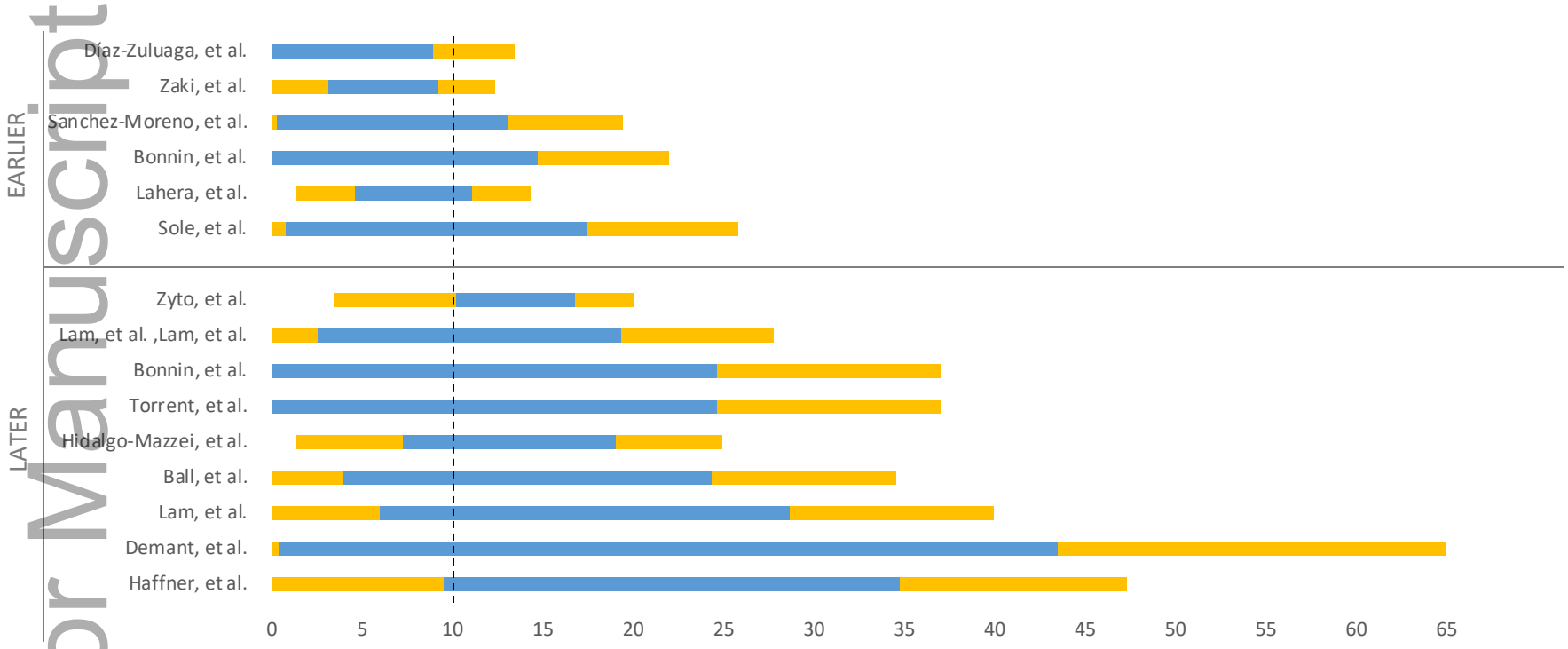


Figure 1. Distribution of number of episodes in samples dichotomised to the earlier versus later stage. Distribution is based on means and standards deviations for number of episodes wherever supplied, assuming a normal distribution, and is therefore an approximation only.

NB. Blue bar represents 1 standard deviation +/- the mean, yellow bar represents 2 standard deviations +/- the mean. The dotted line represents the cut-off of a mean of 10 episodes use to form stage categories. We present the distribution of number of episodes as an example of a proxy distribution in the current analyses. Samples were dichotomised based on additional proxies of functioning, comorbidity and cognitive impairment, however where available number of episodes was used and hierarchically prioritised.

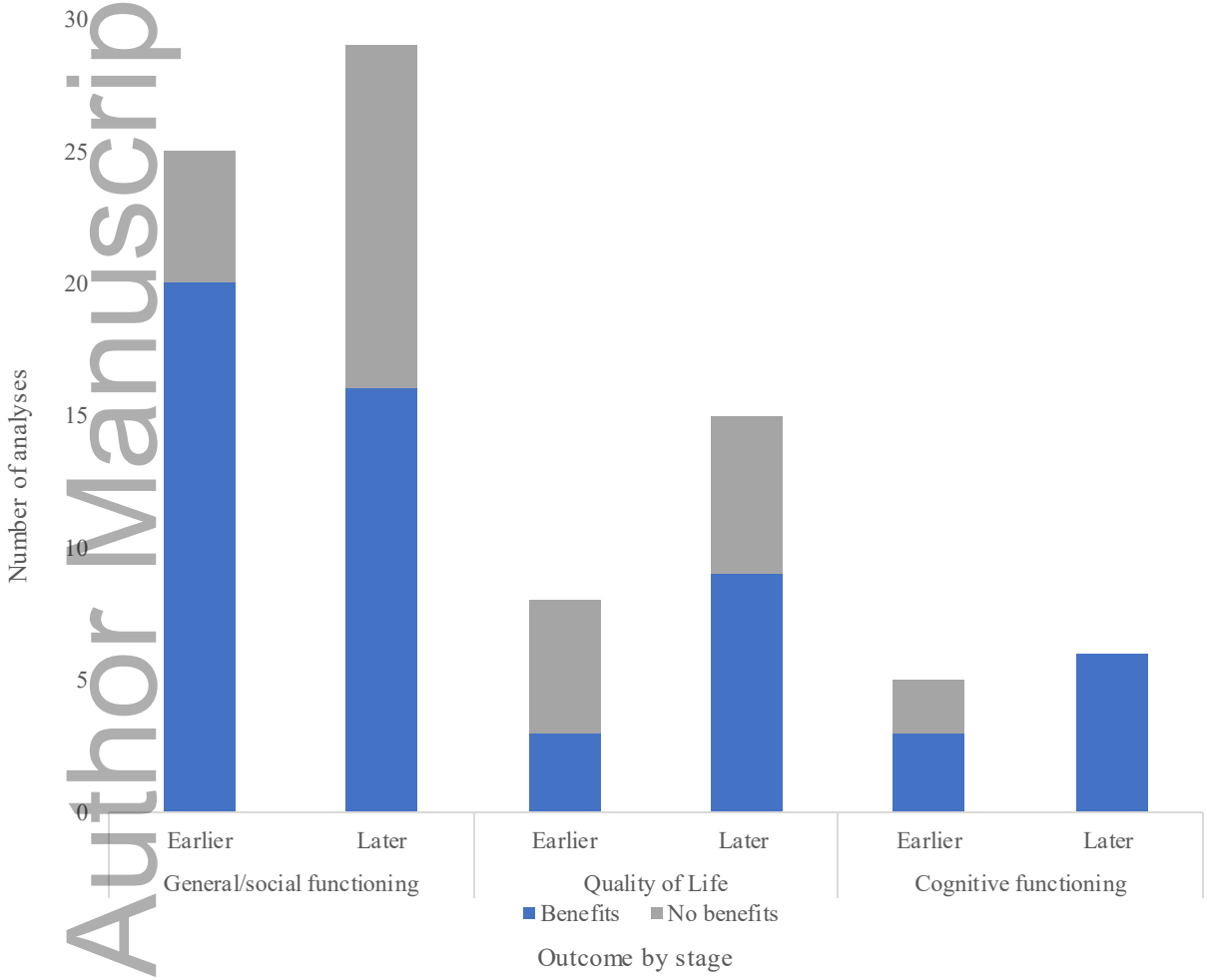


Figure 2. Number of analyses with and without benefits for general/social functioning, quality of life and cognitive functioning in each dichotomised stage.

Box 1

Highlights

- It has been hypothesised that stage of illness may moderate treatment outcomes in BD
- Evidence was found for stage moderation of general/social outcomes, but not QoL or cognitive outcomes: with general/social outcomes more responsive to psychosocial intervention earlier in the course of illness
- Some evidence was found of an interaction between specific intervention type and stage of illness in predicting outcomes
- Methodological heterogeneity and quality issues limit conclusions, however the most significant hindrance is the field's current lack of consensus regarding staging definitions and applications

Author/s:
Tremain, H;Fletcher, K;Scott, J;McEnery, C;Berk, M;Murray, G

Title:
The influence of stage of illness on functional outcomes after psychological treatment in bipolar disorder: A systematic review

Date:
2020-07-27

Citation:
Tremain, H., Fletcher, K., Scott, J., McEnery, C., Berk, M. & Murray, G. (2020). The influence of stage of illness on functional outcomes after psychological treatment in bipolar disorder: A systematic review. BIPOLAR DISORDERS, 22 (7), pp.666-692. https://doi.org/10.1111/bdi.12974.

Persistent Link:
http://hdl.handle.net/11343/276063