Supplementary Appendix

This supplementary material has been provided by the authors to give readers additional information about their work.

PRISMA Guidelines

eTable 1. Checklist Summarizing Compliance With PRISMA Guidelines

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	7
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	7
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7; 8; Table 1; Table 2; Table 3; Table 4; Table 5
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7; eTable2
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8; eFigure 2; eTable 3
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8; Table 1; Table 2; Table 3; Table 4; Table 5
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8

Search strings and Boolean algorithms used

eTable 2. Search strings used by PICOS framework category

eTable 2. Search algorithm for studies examining the efficacy and effectiveness of smartphone-based interventions in bipolar disorder by PICOS framework category[a]		
		Search words[b]
Population	Humans of any age with bipolar disorder	Bipolar disorder
Interventions	Smartphone-based interventions	Mobile, smartphone, app, application, mhealth, mobile-health
Comparisons	Control group without smartphone-based intervention or no comparison (observational studies)	[any]
Outcomes	**Primary**: Efficacy or effectiveness (mood episodes, psychiatric admissions, manic or depressive symptoms, perceived stress, functioning and quality of life) **Secondary**: user-engagement indicators	[any]
Study design	Randomized Controlled Trials and Observational Studies	[any]
[a] PICOS: population, interventions, comparisons, outcomes, study design. [b] "OR" terms		

Boolean algorithm used

Pubmed: "bipolar disorder" AND (smartphone OR mobile OR app OR application OR mHealth OR mobile-health)

- Results: 1226 entries retrieved
- Updated search on January 24, 2022: 126 new entries retrieved

Scopus: TITLE-ABS-KEY ("bipolar disorder") AND TITLE-ABS-KEY (mobile OR smartphone OR app OR application OR mhealth OR mobile-health) AND (LIMIT-TO (SRCTYPE , "j") OR LIMIT-TO (SRCTYPE , "p"))

- Results: 1594 entries retrieved

Embase:

Search item	Search algorithm	Filter	Total results
1	bipolar disorder.mp.	[mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword,	69429
2	(mobile or smartphone or app or application or mhealth or mobile-health).mp.		1251919

		floating subheading word, candidate term word]	
3	1 and 2		1410

- Results: 1410 entries retrieved

APA PsycINFO:

Search item	Search algorithm	Filter	Total results
1	bipolar disorder.mp.	[mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures, mesh]	41366
2	(mobile or smartphone or app or application or mhealth or mobile-health).mp.		154717
3	1 and 2		713

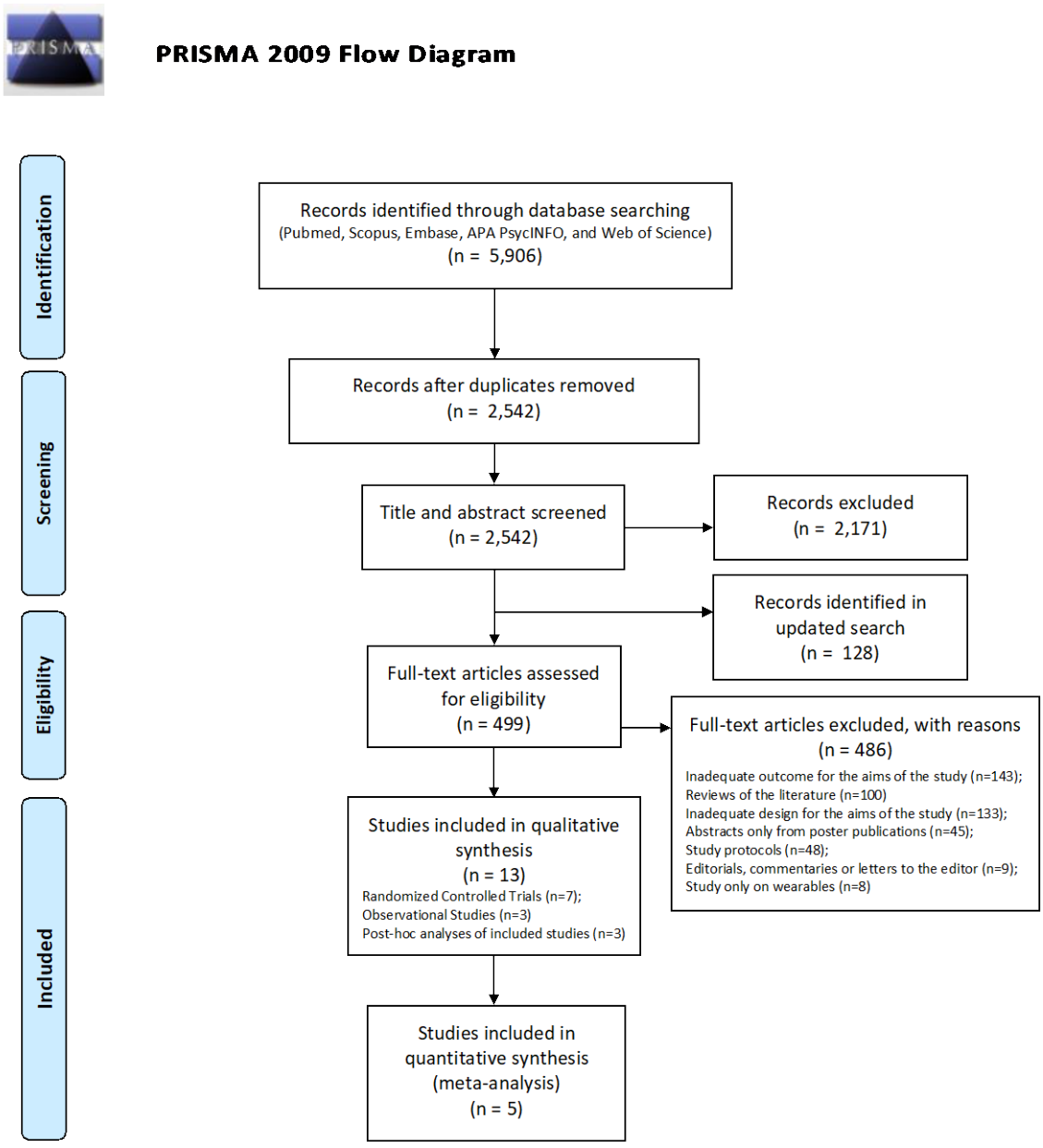
- Results: 713 entries retrieved

Web of Science: ((TI=(bipolar disorder)) OR AB=(bipolar disorder)) AND ((TI=(mobile OR smartphone OR app OR application OR mhealth OR mobile-health)) OR AB=(mobile OR smartphone OR app OR application OR mhealth OR mobile-health))

- Results: 963 entries retrieved

PRISMA Flow Diagram

eFigure 1. PRISMA flow diagram showing the search, article selection, and extraction process for the search regarding the efficacy and effectiveness of smartphone-based interventions in bipolar disorder.



Quality assessment of the included studies

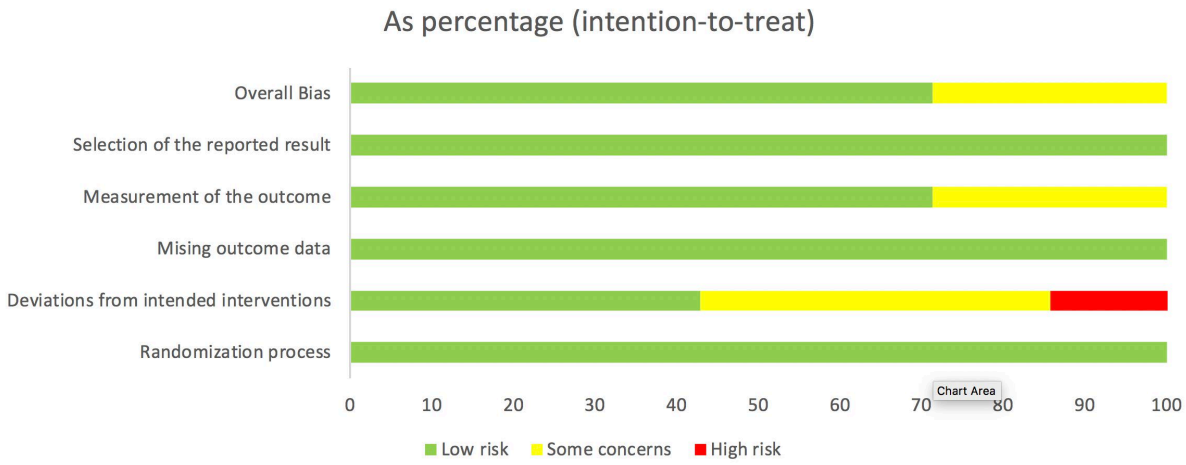
eFigure 2: Risk of bias for included Randomized Controlled Trials assessed using the Cochrane Collaboration's Risk of Bias tool 2. (A) per individual studies; (B) as percentage.

A

Unique ID	Study ID	Experimental	Comparator	Outcome	Weight	D1	D2	D3	D4	D5	Overall
Faurholt-Jepsen 2020	RADMIS	app	TAU	Readmissions and Change in Clinical Scales	1	+	+	+	+	+	+
Faurholt-Jepsen 2019	MONARCA II	app	TAU	Change in Clinical Scales	1	+	+	+	+	+	+
Faurholt-Jepsen 2015	MONARCA I	app	TAU	Change in Clinical Scales	1	+	+	+	+	+	+
Depp 2019	MOBIT	app	SM-app or TAU	Change in Clinical Scales	1	+	!	+	!	+	!
Depp 2015	PRISM	app	Paper-pencil monitoring	Change in Clinical Scales	1	+	!	+	+	+	+
Ben-Zeev 2018	FOCUS	app	Group intervention	Change in Clinical Scales	1	+	!	+	+	+	+
Ben-Zeev 2021	CORE	app	Waiting list	Change in Self-reported Clinical Scales	1	+	-	+	!	+	!

+ Low risk
! Some concerns
- High risk

B



Risk of bias domains:

D1: Randomisation process
D2: Deviations from the intended interventions
D3: Missing outcome data
D4: Measurement of the outcome
D5: Selection of the reported result

eTable 3. Risk of bias for included observational studies assessed using the National Institutes of Health (NIH) quality assessment tool for before-after (Pre-Post) study with no control group

Major Components	Hidalgo-Mazzei et al.; 2016	Hidalgo-Mazzei et al.; 2018	Ryan et al.; 2021
1. Was the study question or objective clearly stated?	Yes	Yes	Yes
2. Were eligibility/selection criteria for the study population prespecified and clearly described?	Yes	Yes	Yes
3. Were the participants in the study representative of those who would be eligible for the test/service/intervention in the general or clinical population of interest?	Yes	Cannot Determine	Yes
4. Were all eligible participants that met the prespecified entry criteria enrolled?	No	Yes	No
5. Was the sample size sufficiently large to provide confidence in the findings?	Yes	Yes	Cannot Determine
6. Was the test/service/intervention clearly described and delivered consistently across the study population?	Yes	Yes	Yes
7. Were the outcome measures prespecified, clearly defined, valid, reliable, and assessed consistently across all study participants?	Yes	Yes	Yes
8. Were the people assessing the outcomes blinded to the participants' exposures/interventions?	No	Not Applicable	Not Applicable
9. Was the loss to follow-up after baseline 20% or less? Were those lost to follow-up accounted for in the analysis?	No	No	No
10. Did the statistical methods examine changes in outcome measures from before to after the intervention? Were statistical tests done that provided p values for the pre-to-post changes?	Yes	Yes	Yes
11. Were outcome measures of interest taken multiple times before the intervention and multiple times after the intervention (i.e., did they use an interrupted time-series design)?	No	No	No
12. If the intervention was conducted at a group level (e.g., a whole hospital, a community, etc.) did the statistical analysis take into account the use of individual-level data to determine effects at the group level?	Not Applicable	Not Applicable	Not Applicable
Quality Rating	Fair	Fair	Fair
Abbreviations: CD: cannot determine; NA: not applicable; NR: not reported.			