CME Background
Articles are selected for credit designation based on an assessment of the educational needs of CME participants, with the purpose of providing readers with a curriculum of CME articles on a variety of topics throughout each volume. Activities are planned using a process that links identified needs with desired results.

To obtain credit, read the article, correctly answer the questions in the Posttest, and complete the Evaluation. A $10 processing fee will apply.

CME Objective
After studying this article, you should be able to:

- Use measurement-based care for patients with depressive disorders treated with pharmacotherapy

Accreditation Statement
The CME Institute of Physicians Postgraduate Press, Inc., is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

ACCME ACCREDITED

Credit Designation
The CME Institute of Physicians Postgraduate Press, Inc., designates this journal-based CME activity for a maximum of 1 *AMA PRA Category 1 Credit*™. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

Note: The American Nurses Credentialing Center (ANCC) and the American Academy of Physician Assistants (AAPA) accept certificates of participation for educational activities certified for *AMA PRA Category 1 Credit*™ from organizations accredited by the ACCME.

Release, Expiration, and Review Dates
This educational activity was published in September 2021 and is eligible for *AMA PRA Category 1 Credit*™ through October 31, 2023. The latest review of this material was September 2021.

Financial Disclosure
All individuals in a position to influence the content of this activity were asked to complete a statement regarding all relevant personal financial relationships between themselves or their spouse/partner and any commercial interest. The CME Institute has resolved any conflicts of interest that were identified. In the past 3 years, Marlene P. Freeman, MD, Editor in Chief, has received research funding from JayMac and Sage; has been a member of the Independent Data Safety and Monitoring Committee for Janssen (Johnson & Johnson) and Novartis; and has served on advisory boards for Eliem and Sage. As an employee of Massachusetts General Hospital (MGH), Dr Freeman works with the MGH National Pregnancy Registry, which receives funding from Alkermes, Aurobindo, AuroMedics, Johnson & Johnson/Janssen, Otsuka, Sage, Sunovion, Supernus, and Teva, and works with the MGH Clinical Trials Network and Institute, which receives research funding from multiple pharmaceutical companies and the National Institute of Mental Health. No member of the CME Institute staff reported any relevant personal financial relationships. **Faculty financial disclosure appears at the end of the article.**

The Efficacy of Measurement-Based Care for Depressive Disorders: Systematic Review and Meta-Analysis of Randomized Controlled Trials

Maria Zhu, MSc[a]; Ran Ha Hong, BSc[a]; Tao Yang, MD, PhD[b]; Xiaorui Yang, MD[b]; Xing Wang, MD[b]; Jing Liu, MHA[a]; Jill K. Murphy, PhD[a]; Erin E. Michalak, PhD[a]; Zuowei Wang, MD, PhD[c]; Lakshmi N. Yatham, MBBS, MBA(Exec)[a]; Jun Chen, MD, PhD[b]; and Raymond W. Lam, MD[a,*]

ABSTRACT

Objective: To determine the efficacy of measurement-based care (MBC), defined as routinely administered outcome measures with practitioner and patient review to inform clinical decision-making, for adults with depressive disorders.

Data Sources: Embase, MEDLINE, PsycINFO, ClinicalTrials.gov, CNKI, and Wanfang Data were searched through July 1, 2020, using search terms for measurement-based care, depression, antidepressant or pharmacotherapy, and randomized controlled trials (RCTs), without language restriction.

Study Selection: Of 8,879 articles retrieved, 7 RCTs (2,019 participants) evaluating MBC for depressive disorders, all involving pharmacotherapy, were included.

Data Extraction: Two independent reviewers extracted data. The primary outcome was response rate (≥ 50% improvement from baseline to endpoint on a depression scale). Secondary clinical outcomes were remission rate (endpoint score in remission range), difference in endpoint severity, and medication adherence.

Results: Meta-analysis with random-effects models found no significant difference between MBC and comparison groups in response rates (3 studies; odds ratio [OR] = 1.66; 95% CI, 0.66–4.17; $P = .279$). MBC was associated with significantly greater remission rates (5 studies; OR = 1.83; 95% CI, 1.12–2.97; $P = .015$), lower endpoint severity (5 studies; standardized mean difference = 0.53; CI 0.06–0.99; $P = .026$), and greater medication adherence (3 studies; OR = 1.68; 95% CI, 1.22–2.30; $P = .001$).

Conclusions: Although benefits for clinical response are unclear, MBC is effective in decreasing depression severity, promoting remission, and improving medication adherence in patients with depressive disorders treated with pharmacotherapy. The results are limited by the small number of included trials, high risk of bias, and significant study heterogeneity.

J Clin Psychiatry 2021;82(5):21r14034

To cite: Zhu M, Hong RH, Yang T, et al. The efficacy of measurement-based care for depressive disorders: a systematic review and meta-analysis of randomized controlled trials. *J Clin Psychiatry.* 2021;82(5):21r14034.

To share: https://doi.org/10.4088/JCP.21r14034

Clinical Points

- Measurement-based care (MBC) is an evidence-based practice that utilizes routine outcome monitoring to inform therapeutic decisions; however, MBC continues to be underutilized in clinical settings.
- For patients with depressive disorders treated with pharmacotherapy, MBC showed higher remission rates, reduced depression severity, and improved medication adherence compared to usual care.

[a]Department of Psychiatry, University of British Columbia, Vancouver, Canada
[b]Shanghai Mental Health Center and Shanghai Jiao Tong University School of Medicine, Shanghai, China
[c]Hongkou Mental Health Center, Shanghai, China
*Corresponding author: Raymond W. Lam, MD, 2255 Wesbrook Mall, Vancouver, BC, Canada V6T 2A1 (r.lam@ubc.ca).

Depressive disorders, including major depressive disorder (MDD), are common psychiatric disorders worldwide and a leading contributor to the global burden of disease.[1] Despite many evidence-based pharmacologic and psychosocial interventions, treatment outcomes for depression are relatively poor in clinical practice.[2–5] A key contributor to poor outcomes may be the unstructured approaches used by clinicians in assessing patient progress that make therapeutic effects difficult to quantify.[6] Clinicians may also have difficulty detecting deterioration of patient symptoms, leading to delay of needed treatment adjustments.[7] Thus, relying on clinical judgment alone to assess progress may lead to an inappropriate treatment regimen, poor treatment adherence, and failure to treat to full symptom remission.[7,8] Therefore, a standardized approach involving routine monitoring of patient outcomes may help to improve treatment adjustments and outcomes for MDD.

Measurement-based care (MBC) is an evidence-based practice that provides a systematic framework for routine outcome monitoring and has demonstrated benefit in treating a range of psychiatric disorders.[6,9,10] MBC includes (1) routine administration of validated rating scales, either by clinician-rated or patient-reported outcomes (PROs); (2) review of scores by practitioners and patients; and (3) using scores to inform shared clinical decision-making.[11,12] MBC has specifically been recommended for the management of MDD by the American Psychiatric Association (APA),[13] Canadian Network for Mood and Anxiety Treatments (CANMAT),[14] and UK National Institute for Health and Care Excellence (NICE)[15] clinical practice guidelines.

MBC offers several potential benefits for depression management. For example, MBC enables clinicians to individualize depression treatment based on up-to-date information about patient symptoms and severity. Quantifiable patient data can be readily incorporated into medication algorithms, facilitating standardized care.[16,17] MBC may also help to identify treatment nonresponders, detect residual symptoms, and increase treatment adherence by encouraging patient participation.[18–21] Although there are recognized barriers and facilitators to implementation and scalability of MBC,[22,23] its feasibility for depression treatment in clinical settings and propensity to improve patient outcomes were demonstrated in several large trials and projects.[11,24,25] Despite these benefits, MBC continues to be underutilized in clinical practice.[26–28]

A number of reviews and meta-analyses[6,9,29,30] have examined the efficacy of MBC for mental health outcomes, but these have focused on broad diagnoses and varied outcome measures. To our knowledge, there are no quantitative syntheses of the most rigorous studies investigating effects of MBC with depressive disorders as the primary diagnosis. Therefore, our aim was to conduct a systematic review and meta-analysis of randomized controlled trials (RCTs) to determine the efficacy of MBC for depression management. We sought to compare MBC to comparison interventions in adults with depressive disorders receiving antidepressant treatment, psychotherapy, or both for improvement on clinician-rated and PRO measures.

METHODS

Search Strategy and Selection Criteria

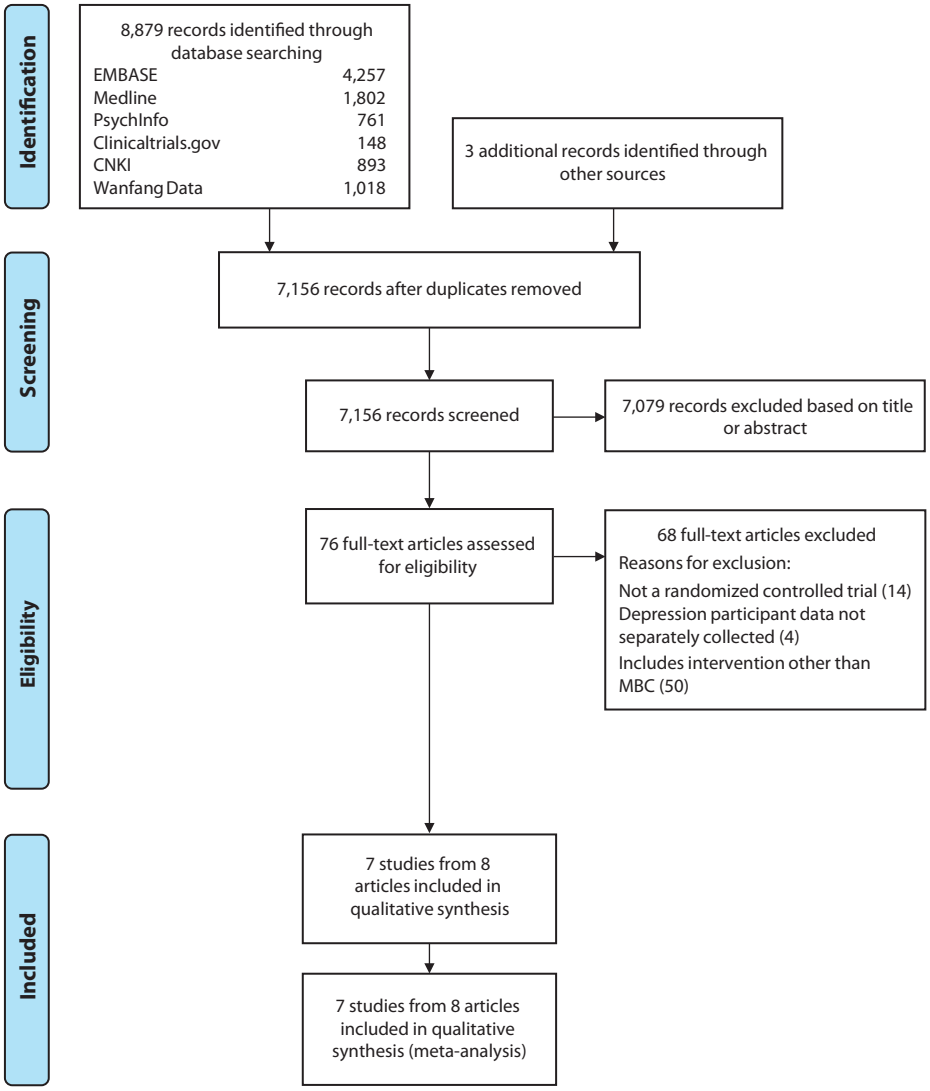
The systematic review was registered with PROSPERO (https://www.crd.york.ac.uk/prospero/), number CRD#42019147274. The electronic databases Embase (OVID), MEDLINE (OVID), PsycINFO, ClinicalTrials.gov, CNKI, and Wanfang Data were searched from their inception until July 1, 2020. The search was conducted by combining search terms from 4 categories: (1) measurement-based care, (2) depressive disorder, (3) antidepressant OR psychotherapy, (4) randomized controlled trial (see Supplementary Appendix 1 for complete search strategy used for each database). Medical subject heading (MeSH) terms for these 4 categories of search terms were included when available. Reference mining was performed by searching through bibliographies of relevant articles to identify any studies that were missed through the database search.

Studies were included if they were (1) RCTs involving (2) adults ≥18 years of age (3) currently diagnosed with a depressive disorder based on validated criteria (ie, *DSM-IV*, *DSM-5*, or *ICD-10*) and were given (4) MBC as an intervention. To be considered MBC, the intervention had to include a routinely administered validated symptom, outcome, or process measure that involved practitioner review of data and use of data to inform clinical decisions. There was no restriction on language.

Studies were excluded if the participants had other comorbidities as a primary diagnosis or were children or adolescents. Review articles, abstracts, commentaries, case reports, and case series were also excluded.

Two independent reviewers (M.Z., X.Y.) screened titles and abstracts of retrieved articles for inclusion. Full-text reviews were subsequently conducted for potentially eligible articles. Discrepancies were resolved by consensus or through consultation with an independent third reviewer (R.W.L.) when consensus could not be reached.

Figure 1. PRISMA Flow Diagram Showing Number of Database Search Results and Article Selection



Data Extraction

Two independent reviewers extracted data using a data extraction form designed for the study. Any discrepancies were resolved through consensus or consultation with a third reviewer (R.W.L.). Study authors were contacted if additional information was required. The following data were extracted: study setting, study duration, study design, participant characteristics (eg, age, sex, primary diagnosis), outcome measures and scores at baseline and post-intervention, and details of the interventions and comparison conditions.

Data Analysis

The primary outcome was efficacy based on clinical response rate (50% or greater improvement from baseline score at end of treatment), which was originally defined from a clinician-administered depression rating scale. However, after the literature search, we found that only 2 studies reported clinician-rated scales; hence, we expanded our definition to include either clinician-rated or patient-rated depression scales. Secondary outcomes included (1) clinical response rate on patient-rated depression measures, (2) clinical remission rate (endpoint score in the remission range, eg, 17-item Hamilton Depression Rating Scale [HDRS] score ≤ 7) by either clinician- or patient-rated depression measures, (3) depression severity at endpoint (endpoint scores on a clinician- or patient-rated depression rating scale) or change scores (endpoint scores minus the baseline scores) if endpoint scores were not available, (4) change in measures of quality of life or functioning, (5) attrition rate (dropouts for any reason), and (6) medication adherence.

All outcomes were analyzed with the intent-to-treat (ITT) samples if available. To protect for inflation of effect size for studies with more than one intervention arm, the results of active arms were pooled as one intervention, as recommended by the Cochrane Handbook for Systematic Reviews of

Interventions.[31] Risk of bias was assessed using version 2 of the Cochrane Risk of Bias tool for RCTs.[32] Meta-analysis was performed using the Comprehensive Meta-Analysis Version 2.0 software (Biostat, USA). Categorical outcomes such as response and remission rates were analyzed using odds ratios (ORs). Continuous outcomes were analyzed using the standardized mean difference (SMD) as a measure of effect size. Outcome data were taken at the end of treatment for each study unless otherwise specified.

A random-effects model was used to account for expected study heterogeneity, including variations in depression severity, the type of depression (eg, major depression, chronic depression, treatment-resistant depression), study duration, clinician-rated versus self-rated measures, and how MBC was delivered. Statistical heterogeneity was assessed using Q χ^2 statistics and I^2; the I^2 can be interpreted as 50%–70% indicating substantial heterogeneity and 75%–100% indicating considerable heterogeneity.[33] Publication bias was examined using funnel plots of outcomes plotted against their standard error (with asymmetric distribution of data points suggesting bias), the Rosenthal fail-safe *N* (the number of unidentified null studies that would need to exist to change the result),[34] and the Egger regression intercept (a statistical test for asymmetry in the funnel plot).[35] If bias was suggested, the trim-and-fill procedure with a random-effects model was used to impute potential missing studies.[36]

RESULTS

Study Selection

The literature search and selection are outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (Figure 1). The search yielded 8,879 articles and 3 that were identified through other sources, such as hand-searching. After duplicate removal, 7,156 remaining articles were assessed through title and abstract review and 76 articles were assessed for eligibility through full-text review. The final 8 articles yielded outcome data from 7 studies for the meta-analysis. Two articles, Chang et al[19] and Yeung et al,[37] reported different outcomes from a single study, the Clinical Outcomes in Measurement-based Treatment (COMET) trial. For the meta-analysis, clinical response, remission, and endpoint PHQ-9 scores were retrieved from Yeung et al,[37] while medication adherence data was taken from Chang et al.[19] Additionally, the COMET trial used a quasi-randomized cluster design in which physician sites were alternately assigned to study arms. The remaining 6 studies were parallel-group RCTs. Therefore, data from a total of 7 studies were included in the meta-analysis.

Study Characteristics

Table 1 shows the main characteristics of the 7 included RCTs. All the included studies involved pharmacotherapy management. Identified studies of psychotherapy were excluded because they were not RCTs or did not involve depressive disorders separately from other psychiatric diagnoses. Individual study sample sizes ranged from 108 to 915 participants with a total of 2,019 participants in the 7 studies. The participants in all studies were diagnosed with a depressive disorder or a major depressive episode based on *DSM-IV*, *DSM-5*, or *ICD-10* criteria, with severity of depression ranging from mild to severe. Three studies were conducted in outpatient psychiatry settings, 2 studies in outpatient primary care settings, and 2 studies in inpatient psychiatry settings.

The studies used various outcome measures for depression severity. Clinician-rated depression measures included the HDRS and Bech-Rafaelsen Melancholia Scale (BRMS). Patient-rated depression measures included the Patient Health Questionnaire-9 (PHQ-9), and Beck Depression Inventory-II (BDI-II).

The routinely administered outcome measure for MBC varied across all studies and included the QIDS-SR,[38] MADRS-S,[39] PHQ-9,[37,40,41] 21-item HDRS,[42] and the BRMS.[25] In keeping with the criteria for MBC, all studies involved physician review of the measures to inform clinical decisions. Physicians in 3 studies used algorithms to further guide clinical decision-making,[25,38,42] while the physicians in the other 4 studies were not guided by algorithms. The physicians assigned to the comparison groups were able to freely choose the treatments for the patients without routine depression measures as feedback. Three studies used antidepressant monotherapy as the chosen form of treatment, whereas the Adli et al[42] and Bauer et al[25] studies used a complex stepwise approach that included antidepressant monotherapy, monoamine oxidase inhibitors, lithium, and triiodothyronine augmentation, sleep deprivation, and electroconvulsive therapy. None of the studies explicitly reported psychotherapy as a form of treatment. Frequency of MBC measures ranged from once every week to once a month. The study durations ranged from 3 months to 12 months.

Risk of Bias Assessment

Figure 2 shows the summary of the risk of bias assessment from the Cochrane Risk of Bias tool. A study was judged to be high risk if greater than one domain received an assessment of high risk of bias. Two studies were assessed to have an overall low risk of bias,[38,40] whereas the remaining 5 studies were judged to have a high risk of bias.[25,37,39,41,42] For the randomization process, Yeung et al[37] used a pseudorandomized cluster design, and the randomization process was unclear for Chen et al.[41] Four studies either lacked or did not report blinding of participants, physicians, and assessors, and were assessed to have either uncertain or high risk of bias due to potential deviations from intended interventions. Two studies had missing outcome data due to incomplete data on dropouts; ie, the COMET study[19,37] and the Zhao et al study[40] excluded data from dropouts in their final analyses. Five studies had either high or uncertain risk of bias for measurement of the outcome. None of the studies had high risk of bias in the selection of reported outcomes.

Table 1. Characteristics of Included Studies

Authors	Depression Type	Diagnostic Criteria	Study Design	Country	Setting	MBC Criteria: Symptom, Outcome, or Process Measure	MBC Criteria: Routinely Administered Measure	MBC Criteria: Physician Review of Data	MBC Criteria: Physician Use of Data to Inform Treatment	Comparison Condition	Participants Receiving MBC, n	Participants Receiving Comparison Condition, n	Treatment(s) Received	Duration of Intervention	Depression Outcome Instrument
Adli et al[42]	Major depressive episode	*ICD-10*	RCT	Germany	Inpatient psychiatry	21-item HDRS	Once every 2 wk ± 3 d	Yes	Yes: treatment based on algorithm	Free selection of treatment by physician based on clinical judgment alone	266	84	Antidepressant monotherapy, MAOI, lithium, T_3, electroconvulsive therapy	16 wk	21-item HDRS
Bauer et al[25]	Major depressive episode	*ICD-10*	RCT	Germany	Inpatient psychiatry	BRMS	Once every 2 wk	Yes	Yes: treatment based on algorithm	Free selection of treatment by physician based on clinical judgment alone. BRMS was administered but results were not presented to physicians	74	74	Antidepressant monotherapy, MAOI, lithium, T_3, electroconvulsive therapy	12 wk	BRMS
Chang et al[19]	Major depressive disorder	*DSM-IV*	Cluster RCT	United States	Outpatient primary care	PHQ-9, PGI-S	Once a mo for 6 mo	Yes	Yes: choice of treatment by physician	Free selection of treatment by physician based on clinical judgment alone. PHQ-9 administered at 3 and 6 mo, but faxed to the physicians only at 6 mo	380	284	Antidepressants	6 mo	Same study as Yeung et al[37]—medication adherence as outcome
Chen et al[41]	Depressive disorder	*DSM-5*	RCT	China	Outpatient psychiatry	PHQ-9	Unspecified	Unspecified	Unspecified	Free selection of treatment by physician based on clinical judgment alone	54	54	Tricyclic antidepressants	Unspecified	PHQ-9
Guo et al[38]	Major depressive disorder	*DSM-IV*	RCT	China	Outpatient psychiatry	QIDS-SR, FIBSER	Baseline and at 2, 4, 8, 12, and 24 wk	Yes	Yes: treatment based on algorithm	Free selection of treatment by physician based on clinical judgment alone	61	59	Open-label paroxetine (20–60 mg/d) or open-label mirtazapine (15–45 mg/d)	24 wk	17-item HDRS
Wikberg et al[39]	Depressive disorder	*DSM-IV*	RCT	Sweden	Outpatient primary care	MADRS-S	4 visits over 3 mo	Yes	Yes: choice of treatment by physician	Free selection of treatment by physician based on clinical judgment alone	125	133	Antidepressants	3 mo of intervention, follow-up at 3, 6 and 12 mo	BDI-II
Yeung et al[37]	Major depressive disorder	*DSM-IV*	Cluster RCT	United States	Outpatient primary care	PHQ-9, PGI-S	Once a mo for 6 mo	Yes	Yes: choice of treatment by physician	Free selection of treatment by physician based on clinical judgment alone. PHQ-9 administered at 3 and 6 mo, but faxed to the physicians only at 6 mo	364	278	Antidepressants	6 mo	PHQ-9
Zhao et al[40]	Depressive disorder	*DSM-5*	RCT	China	Outpatient psychiatry	PHQ-9	Once every wk for 12 mo	Yes	Yes: choice of treatment by physician	Free selection of treatment by physician based on clinical judgment alone	60	60	Antidepressants (SSRIs and SNRIs)	12 mo	17-item HDRS

Abbreviations: BDI-II = Beck Depression Inventory II; BRMS = Bech-Rafaelsen Melancholia Scale; DSM = *Diagnostic and Statistical Manual of Mental Disorders*; FIBSER = Frequency, Intensity, and Burden of Side Effects Rating; HDRS = Hamilton Depression Rating Scale; *ICD* = *International Classification of Diseases*; MADRS-S = Montgomery-Asberg Depression Rating Scale–Self Assessment; MAOI = monomamine oxidase inhibitor; MBC = measurement-based care; PGI-S = Patient Global Impression–Severity; PHQ-9 = Personal Health Questionnaire-9; PRIME-MD = Primary Care Evaluation of Mental Disorders; QIDS-SR = Quick Inventory of Depressive Symptomatology–Self-Report; RCT = randomized controlled trial; SNRI = serotonin-norepinephrine reuptake inhibitor; SSRI = selective serotonin reuptake inhibitor; T_3 = triiodothyronine.

Figure 2. Cochrane Risk-of-Bias Assessment for Included Studies

First author, year	Randomization Process	Deviations from interventions	Missing outcome data	Measurement of the outcome	Selection of reported outcomes	Overall bias	Notes
Adli, 2017[42]	+	−	+	−	+	−	
Bauer, 2009[25]	+	−	+	−	+	−	
Chen, 2016[41]	?	−	+	−	+	−	
Guo, 2015[38]	+	+	+	+	+	+	
Wikberg, 2017[39]	+	−	+	−	+	−	
Yeung, 2012[37]	?	+	−	?	+	−	Cluster randomized trial
Zhao, 2015[40]	+	+	−	+	+	+	

+ Low risk of bias
? Some concerns
− High risk of bias

Meta-Analysis

Efficacy outcomes. The original primary outcome was clinical response rate using clinician-rated depression scales. As only 2 studies reported clinician-rated scales, we expanded our definition to include either clinician-rated or patient-rated depression scales. Overall, 3 studies (1,112 participants) provided clinical response rates, with the Guo et al[38] and Adli et al[42] studies using versions of the clinician-rated HDRS and the Yeung et al[37] study using the patient-rated PHQ-9. The Adli et al study[42] had 3 intervention arms that were consistent with MBC, so the data from these arms were pooled into 1 intervention group.

With a random-effects model, there was no significant difference in clinical response between MBC and comparison conditions (OR 1.66; 95% CI, 0.66–4.17; $P = .279$) (Figure 3A). There was significant heterogeneity between these trials, with Q statistic of 15.02 and I^2 of 86.7% (degrees of freedom [df] = 2; $P = .001$). The funnel plot of standard errors by effect size estimates was broadly symmetrical (Supplementary Figure 1), the fail-safe N was 4, and the Egger intercept had a 1-tailed P value of .45, suggesting there is little probability of publication bias (see Supplementary Appendix 2). The unadjusted rates for clinical response were 67.6% and 62.7%, respectively, for active MBC versus comparison conditions.

Secondary efficacy outcomes included clinical remission rate and endpoint depression severity. Five studies (1,518 participants) reported remission rates. Three studies used clinician-rated scales, and 2 studies used patient-rated scales. For clinician-rated scales, remission was defined as 17-item HDRS score ≤ 7,[38] 21-item HDRS score ≤ 9,[42] and BRMS score ≤ 7.[25] For patient-rated scales, remission was defined as BDI-II score ≤ 13 at 3-month follow-up[39] and PHQ-9 score < 5 at 6 months.[37] MBC was associated with a significantly higher clinical remission rate compared to the comparison condition (OR 1.83; 95% CI, 1.12–2.97; $P = .015$) (Figure 3B). There was significant heterogeneity between these trials, with Q statistic of 16.14 and I^2 of 75.2% ($df = 4$, $P = .003$). The funnel plot of standard errors by effect size estimates was broadly symmetrical (Supplementary Figure 2), the fail-safe N was 23, and the Egger intercept had a 1-tailed P value of .08, suggesting a low probability of publication bias (see Supplementary Appendix 2). The unadjusted rates for clinical remission were 52.8% for MBC and 43.0% for comparison conditions.

Difference in endpoint depression scores were reported for 5 studies (1,248 participants). For endpoint depression scores, 3 studies used the PHQ-9,[37,40,41] 1 study used the 17-item HDRS,[38] and 1 study used the BDI-II.[39] Overall, the MBC condition had significantly lower endpoint depression scores versus the comparison condition (SMD 0.53; 95% CI, 0.06–0.99; $P = .026$), with the SMD representing a medium effect size (Figure 3C). There was significant heterogeneity between these trials, with Q statistic of 53.55 and I^2 of 92.5% ($df = 4$; $P < .001$). The funnel plot of standard errors by effect size estimates showed some asymmetry (Supplementary Figure 3) and, although the fail-safe N was 45, the Egger intercept was 8.10 (95% CI, 2.26–13.95, $t_3 = 4.4$, 1-tailed $P = .01$), suggesting probability of publication bias (see Supplementary Appendix 2). The trim-and-fill procedure, however, suggested that there were no missing studies, and the random-effects model did not change the SMD point estimate.

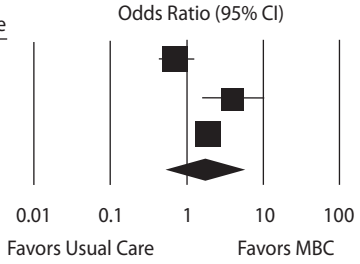
Because all 3 efficacy outcomes showed significant statistical heterogeneity, we conducted post hoc sensitivity analyses to explore potential sources of heterogeneity, segregating by methodology variables that might differ between studies. For each outcome, we examined (1) studies that used antidepressants only (ie, excluding those with multitreatment algorithms), (2) studies conducted in psychiatric settings (ie, excluding those in primary care settings), and (3) studies conducted outside of China (ie, because of potential differences in health care systems) (Supplementary Table 1). For clinical response, the studies using antidepressants only did not have significant heterogeneity (Q=1.73, $df = 1$, $P = .19$, I^2=42.2); the OR was 2.48 (95% CI, 1.36–4.51; $P = .003$) in favor of MBC. For clinical remission, the 4 studies conducted outside of China did not show significant heterogeneity (Q=2.28, $df = 3$, $P = .52$, I^2=0.00); the OR was 1.42 (95% CI, 1.11–1.80; $P = .005$) in favor of MBC. All of the remaining analyses continued to show significant heterogeneity in the results.

Figure 3. Forest Plots Displaying Meta-Analyses of (A) Odds Ratios for Clinical Response for MBC Versus Usual Care, (B) Odds Ratios for Clinical Remission for MBC Versus Usual Care, and (C) Standardized Mean Differences in Endpoint Scores for MBC Versus Usual Care

A. Clinical Response

	MBC		Usual Care				
Study (first author, year)	Events	Total	Events	Total	Weight	Odds Ratio (95% CI)	*P* Value
Adli, 2017[42]	170	266	61	84	34.8%	0.67 (0.39–1.15)	.143
Guo, 2015[38]	53	61	37	59	28.5%	3.94 (1.58–9.80)	.003
Yeung, 2012[37]	NA	81	NA	42	36.8%	2.02 (1.36–3.01)	.001
Total (random)	NA	408	NA	185		1.66 (0.66–4.17)	.279

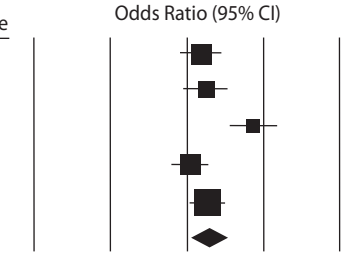
Heterogeneity $Q = 15.02$, $P = .001$, $I^2 = 86.7\%$



B. Clinical Remission

	MBC		Usual Care				
Study (first author, year)	Events	Total	Events	Total	Weight	Odds Ratio (95% CI)	*P* Value
Adli, 2017[42]	154	266	42	84	22.3%	1.38 (0.84–2.24)	.205
Bauer, 2009[25]	40	74	29	74	12.6%	1.83 (0.95–3.51)	.071
Guo, 2015[38]	45	61	17	59	8.4%	6.95 (3.12–15.49)	.000
Wikberg, 2017[39]	61	125	63	133	22.6%	1.06 (0.65–1.73)	.818
Yeung, 2012[37]	NA	81	NA	42	34.1%	1.59 (1.07–2.37)	.022
Total (Random)	NA	607	NA	392		1.83 (1.12–2.97)	.015

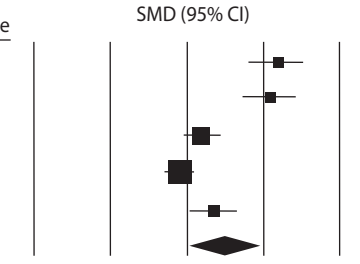
Heterogeneity $Q = 16.14$, $P = .003$, $I^2 = 75.2\%$



C. Endpoint Scores

	MBC		Usual Care				
Study (first author, year)	Mean (SD)	N	Mean (SD)	N	Weight	SMD (95% CI)	*P* value
Chen, 2016[41]	5.3 (1.8)	54	7.6 (2.1)	54	7.6%	1.18 (0.77–1.59)	.000
Guo, 2015[38]	4.8 (3.6)	61	8.6 (3.6)	59	8.8%	1.06 (0.67–1.44)	.000
Wikberg, 2017[39]	15.6 (11.2)	125	18.1 (11.6)	133	21.3%	0.22 (-0.03–0.46)	.079
Yeung, 2012[37]	7.6 (5.5)	364	7.2 (5.4)	278	52.4%	-0.07 (-0.23–0.08)	.358
Zhao, 2015[40]	5.3 (5.8)	60	7.3 (4.8)	60	9.8%	0.36 (0.00–0.72)	.049
Total (random)	NA	664	NA	584		0.53 (0.06–0.99)	.026

Heterogeneity $Q = 53.55$, $P = .000$, $I^2 = 92.5\%$



Abbreviations: MBC = measurement-based care, NA = not available, SMD = standardized mean difference.

Attrition rate and adherence to medication. Data on attrition rates were available for all 7 studies (1,746 participants). There was no significant difference in attrition rates between MBC and comparison conditions (OR 1.32; 95% CI, 0.61–2.87; $P = .49$) (Figure 4A). There was significant heterogeneity between these trials, with Q statistic of 25.2 and I^2 of 76.2% ($df = 6$, $P < .001$). The funnel plot of standard errors by effect size estimates was broadly symmetrical (Supplementary Figure 4), the fail-safe N could not be calculated because it was not relevant to the observed lack of significant difference, and the Egger intercept had a 1-tailed P value of .42, indicating a low probability of publication bias (see Supplementary Appendix 2).

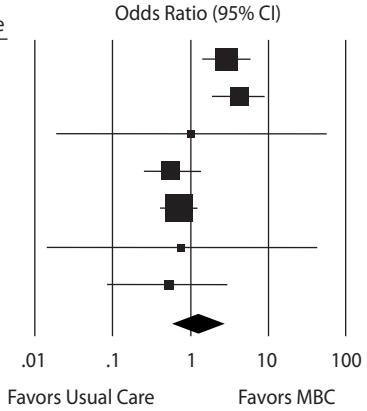
Three studies (1,042 participants) reported on medication adherence,[19,38,40] defined as the proportion of participants who continued their medication by study end. Medication adherence data were taken at the end of the trial for 2 studies[19,38] and 6 months following the end of the trial for 1 study.[39] Results using random-effects analysis revealed that medication adherence was more likely in the MBC group (OR 1.68; 95% CI, 1.22–2.30; $P = .001$) (Figure 4B). There was no significant heterogeneity between these trials, with Q statistic of 0.30 and I^2 of 0% ($df = 1$, $P = .862$). The funnel plot of standard errors by effect size estimates was broadly symmetrical (Supplementary Figure 5), the fail-safe N was 6, and the Egger intercept had a 1-tailed P

Figure 4. Forest Plots Displaying Meta-Analyses of (A) Odds Ratios for Attrition Rates for MBC Versus Usual Care and (B) Odds Ratios for Medication Adherence for MBC Versus Usual Care

A. Attrition

	MBC		Usual Care				
Study (first author, year)	Events	Total	Events	Total	Weight	Odds Ratio (95% CI)	*P* Value
Adli, 2017[42]	111	266	16	84	27.2%	3.04 (1.68–5.53)	.000
Bauer, 2009[25]	33	74	12	74	16.4%	4.16 (1.93–8.98)	.000
Chen, 2016[41]	1	54	1	54	0.6%	1.00 (0.02–51.33)	.000
Guo, 2015[38]	17	61	22	59	16.4%	0.65 (0.30–1.40)	.818
Wikberg, 2017[39]	37	125	48	133	35.6%	0.75 (0.44–1.26)	.268
Yeung, 2012[37]	1	364	1	278	0.6%	0.76 (0.02–38.59)	.893
Zhao, 2015[40]	2	60	4	60	3.2%	0.48 (0.09–2.74)	.411
Total (random)	202	1004	104	742		1.32 (0.61–2.87)	.485

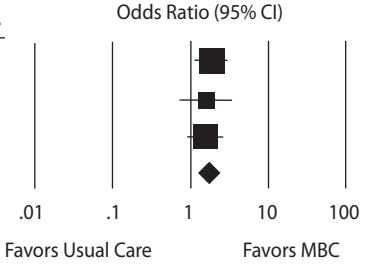
Heterogeneity $Q = 25.2$, $P = .000$, $I^2 = 76.2\%$



B. Medication Adherence

	MBC		Usual Care				
Study (first author, year)	Events	Total	Events	Total	Weight	Odds Ratio (95% CI)	*P* Value
Chang, 2014[19]	NA	380	NA	284	44.7%	1.85 (1.95–2.97)	.011
Guo, 2015[38]	44	61	37	59	17.0%	1.54 (0.71–3.32)	.272
Wikberg, 2017[39]	86	125	78	133	36.8%	1.56 (0.93–2.60)	.091
Total (random)	NA	566	NA	476		1.68 (1.22–2.30)	.001

Heterogeneity $Q = .30$, $P = .862$, $I^2 = 0\%$



Abbreviations: MBC = measurement-based care, NA = not available.

value of .48, suggesting a low probability of publication bias (see Supplementary Appendix 2). Overall unadjusted rates of medication adherence for the 3 studies were 76.1% for the MBC conditions and 63.9% for the comparison conditions.

There was an insufficient number of studies that reported quality of life and functional level measures for meta-analysis. Only 2 studies reported quality of life findings: Wikberg et al[39] found no significant differences in EuroQoL-5D (EQ-5D)[43] scores between the MBC and treatment-as-usual groups at the 12-month follow-up, while Zhao et al[40] found significant differences in scores on the World Health Organization Quality of Life assessment (WHOQOL-BREF)[44] in favor of MBC at 12-month follow-up. No studies reported functional outcomes, so the category was not further assessed.

DISCUSSION

Our objective was to determine the efficacy of MBC for depression from a meta-analysis of RCTs. We identified 8 articles reporting on 7 studies of MBC with pharmacologic treatment that met the inclusion criteria. The findings from our overall meta-analysis showed no significant difference between MBC and comparison groups in clinical response rates based on clinician- and patient-rated depression scales. The nonsignificant finding may be a result of Type II error, as only 3 studies were pooled for the clinical response outcome, which may have led to nonrejection of a false null hypothesis. In contrast, MBC was associated with a significant increase in clinical remission rates and decrease in endpoint depression severity, with clinically relevant effect size, compared to comparison conditions. Clinical remission is an important outcome because it is associated with lower risk of relapse[45,46] and better long-term outcomes[47] compared with achieving clinical response without remission. Additionally, MBC was associated with significantly greater medication adherence. Moreover, there was no significant difference in attrition rates between the two groups. Overall, these results support the benefit of MBC for improving outcomes in patients with depressive disorders treated with pharmacotherapy. However, a caution for these results is the high risk of bias in the majority of studies (5 of 7), although both trials with low overall risk of bias, Guo et al[38] and Zhao et al,[40] found positive effects in favor of MBC.

MBC has shown benefits in the management of a range of psychiatric disorders.[6,9,10] Despite the benefits of MBC described in implementation and non-randomized studies, a 2016 Cochrane review of 17 RCTs[30] concluded that there was insufficient evidence to support the routine use of PROs due to low quality studies with high risk of bias. However, the evidence then was limited to use of multidomain PRO

measures instead of more specific measures from a single domain such as depression. In addition, a broad range of mental health disorders were included in populations with multiple comorbidities. Interestingly, most of the included studies focused on psychotherapy as a treatment modality, whereas none of the included RCTs in our current review used psychotherapy. Studies of MBC with psychotherapy have focused on broader groups of diagnoses and hence did not meet the depressive disorder diagnosis criterion for inclusion in our meta-analysis.

MBC may be particularly relevant to antidepressant prescribing, in which simple algorithms can guide stepwise changes to the dosage or medications depending on measurement outcome. In this regard, favorable outcomes with MBC may in part be due to increased adherence to medication. Our results support findings from other studies, such as the Combining Medications to Enhance Depression Outcomes (CO-MED) study,[48] which incorporated MBC into treatment delivery and found higher medication adherence among adults with chronic or recurrent MDD compared to rates reported by prior studies that did not use MBC. On the other hand, a meta-analysis of psychotherapy studies[49] found that MBC with clinical decision support tools offering specific treatment recommendations (analogous to medication algorithms) were also associated with larger positive effect sizes.

Unfortunately, there were too few studies that included quality of life and functional outcomes to synthesize. The 2 studies with quality of life measures found discrepant findings, with MBC showing significant benefit on the WHOQOL-BREF[40] but no significant improvement on the EQ-5D.[39] The EQ-5D has fewer items than the WHOQOL-BREF (6 items versus 26 items, respectively) and may be less responsive to change for mental health conditions.

A limitation of our meta-analysis was the heterogeneity in methodologies of the included studies. Currently, there is no consensus on the most effective measures or the optimal frequency of measurement for MBC. The MBC measures used in the included RCTs varied across all studies, with frequency of MBC assessment ranging from once a week to once a month. Similarly, there was substantial variation among studies in the nature, fidelity, and reporting of other elements of MBC, such as reviewing scores with patients, how scores were used to make clinical decisions, and the degree of shared decision making between clinicians and patients. For example, the type and complexity of medication algorithms used in the studies varied considerably: 4 studies did not use a treatment algorithm, Guo et al[38] used a simple medication dosing algorithm, and Adli et al[42] and Bauer et al[25] used a complex algorithm that included medication and somatic treatments such as electroconvulsive therapy. Other potential sources of heterogeneity in the studies included variability in depression severity, treatment setting, outcome assessments, duration of treatment, and country of study.

We anticipated heterogeneity in the studies and hence a priori set a random-effects model for all outcomes. We did find significant statistical heterogeneity in the efficacy outcomes, with I^2 ranging from 75.2% to 92.5%, regarded as considerable heterogeneity. The sensitivity analyses for potential methodology sources of study heterogeneity (studies of antidepressants only, studies in psychiatric settings, studies conducted outside China) were limited by small sample sizes, and the results were not consistent. There was suggestion that the studies of antidepressant-only treatment and studies conducted outside of China had less heterogeneity for clinical response and remission; these sensitivity analyses still showed increased ORs in favor of MBC. Future research should attempt to standardize the methodologies to reduce heterogeneity in outcomes.

MBC may also be more effective in specific patient subgroups, but we were unable to conduct subgroup analyses for the included studies because of lack of segregation of outcomes. Previous meta-analyses evaluating MBC in patients undergoing psychotherapy[18,50] showed that the effect of MBC was greater in treatment nonresponders compared to those who showed an initial improvement. The reasoning is that MBC provides information on change or lack of change in patient outcomes that cannot be reliably assessed by clinical judgment alone. Hence, future studies should examine the effects of MBC for treatment-resistant depression.

In summary, the findings from this systematic review and meta-analysis of RCTs support the use of MBC in the management of depression, particularly with pharmacotherapy. MBC allows clinicians to individualize therapeutic decisions based on up-to-date information regarding patient outcomes. MBC may also improve medication adherence. A limitation is that, because we did not identify any eligible psychotherapy studies, our results do not address efficacy of MBC for psychotherapy alone or psychotherapy combined with pharmacotherapy. Future RCTs of MBC should examine the effects of MBC for psychotherapy and for depression subgroups (including treatment-resistant depression), and optimization of algorithm-guided MBC. MBC studies should include outcome measures assessing functioning and quality of life to complement standard symptom measures. Further investigation is also necessary to standardize the type and frequency of routine outcome measures used for MBC in depression management.

Submitted: April 12, 2021; accepted June 28, 2021.
Published online: September 28, 2021.

Disclosure of off-label usage: The authors have determined that, to the best of their knowledge, no investigational information about pharmaceutical agents or device therapies that is outside US Food and Drug Administration–approved labeling has been presented in this article.

Financial disclosure: **Dr Michalak** has received funding to support patient education initiatives from Otsuka. **Dr Yatham** has been on speaker/advisory boards for, or has received research grants from Alkermes, AstraZeneca, Bristol-Myers Squibb, the Canadian Network for Mood and Anxiety Treatments (CANMAT), the Canadian Institutes of Health Research (CIHR), Sumitomo Dainippon Pharma (DSP), Eli Lilly, GlaxoSmithKline, Janssen, the Michael Smith Foundation for Health Research, Pfizer, Servier, Sunovion, and the Stanley Foundation. **Dr Chen** has received honoraria for ad hoc speaking for DSP, Eli Lilly, Pfizer, Lundbeck, Janssen, and Sunovion. **Dr Lam** has received

honoraria for ad hoc speaking or advising/consulting, or received research funds, from Allergan, Asia-Pacific Economic Cooperation, BC Leading Edge Foundation, CIHR, CANMAT, Healthy Minds Canada, Janssen, Lundbeck, Lundbeck Institute, Michael Smith Foundation for Health Research, Mitacs, Myriad Neuroscience, Ontario Brain Institute, Otsuka, Pfizer, Unity Health, and VGH-UBC Hospital Foundation. **Ms Zhu**, **Ms Hong**, **Ms Liu**, **Dr T. Yang**, **Dr X. Yang**, **Dr X. Wang**, **Dr Murphy**, and **Dr Z. Wang** have no disclosures.

Funding/support: This study was supported by the Enhanced Measurement-Based Care Effectiveness for Depression (EMBED) study, funded by the CIHR and the National Natural Sciences Foundation of China (81761128032) through the Global Alliance on Chronic Disease [CIHR Team Grant: GACD Mental Health, GAC-154985; Agreement #01709-000]. Ms Zhu was supported by a Summer Student Research Program (SSRP) award from the UBC Faculty of Medicine. Ms Hong was supported by a Mach-Gaensslen Foundation of Canada Award through the UBC Faculty of Medicine SSRP.

Role of the sponsor: The sponsors had no role in the design, conduct, interpretation, or publication of this study.

Supplementary material: Available at PSYCHIATRIST.COM.

REFERENCES

1. Ferrari AJ, Charlson FJ, Norman RE, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med*. 2013;10(11):e1001547.
2. Richards D. Prevalence and clinical course of depression: a review. *Clin Psychol Rev*. 2011;31(7):1117–1125.
3. Kupfer DJ, Frank E, Phillips ML. Major depressive disorder: new clinical, neurobiological, and treatment perspectives. *Lancet*. 2012;379(9820):1045–1055.
4. Eaton WW, Shao H, Nestadt G, et al. Population-based study of first onset and chronicity in major depressive disorder. *Arch Gen Psychiatry*. 2008;65(5):513–520.
5. Steinert C, Hofmann M, Kruse J, et al. The prospective long-term course of adult depression in general practice and the community: a systematic literature review. *J Affect Disord*. 2014;152–154(1):65–75.
6. Fortney JC, Unützer J, Wrenn G, et al. A tipping point for measurement-based care. *Psychiatr Serv*. 2017;68(2):179–188.
7. Hatfield D, McCullough L, Frantz SHB, et al. Do we know when our clients get worse? an investigation of therapists' ability to detect negative client change. *Clin Psychol Psychother*. 2010;17(1):25–32..
8. Henke RM, Zaslavsky AM, McGuire TG, et al. Clinical inertia in depression treatment. *Med Care*. 2009;47(9):959–967.
9. Knaup C, Koesters M, Schoefer D, et al. Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis. *Br J Psychiatry*. 2009;195(1):15–22.
10. Brodey BB, Cuffel B, McCulloch J, et al. The acceptability and effectiveness of patient-reported assessments and feedback in a managed behavioral healthcare setting. *Am J Manag Care*. 2005;11(12):774–780.
11. Trivedi MH, Rush AJ, Wisniewski SR, et al; STAR*D Study Team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry*. 2006;163(1):28–40.
12. Hong RH, Murphy JK, Michalak EE, et al. Implementing measurement-based care for depression: practical solutions for psychiatrists and primary care physicians. *Neuropsychiatr Dis Treat*. 2021;17:79–90.
13. Gelenberg AJ, Freeman MP, Markowitz JC, et al; Work Group on Major Depressive Disorder. *Practice Guideline for the Treatment of Patients With Major Depressive Disorder*. 3rd ed. Arlington, VA: American Psychiatric Association; 2010.
14. Lam RW, McIntosh D, Wang J, et al; CANMAT Depression Work Group. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder: section 1, disease burden and principles of care. *Can J Psychiatry*. 2016;61(9):510–523.
15. Depression in adults: recognition and management (Clinical Guideline 90). National Institute for Health and Care Excellence. Published 2009. Accessed March 7, 2021. https://www.nice.org.uk/guidance/cg90
16. Crismon ML, Trivedi M, Pigott TA, et al. The Texas Medication Algorithm Project: report of the Texas Consensus Conference Panel on medication treatment of major depressive disorder. *J Clin Psychiatry*. 1999;60(3):142–156.
17. Valenstein M, Adler DA, Berlant J, et al. Implementing standardized assessments in clinical care: now's the time. *Psychiatr Serv*. 2009;60(10):1372–1375.
18. Lambert MJ, Whipple JL, Hawkins EJ, et al. Is it time for clinicians to routinely track patient outcome? a meta-analysis. *Clin Psychol Sci Pract*. 2003;10(3):288–301.
19. Chang TE, Jing Y, Yeung AS, et al. Depression monitoring and patient behavior in the Clinical Outcomes in MEasurement-Based Treatment (COMET) trial. *Psychiatr Serv*. 2014;65(8):1058–1061.
20. Jeon SW, Han C, Ko YH, et al. Measurement-based treatment of residual symptoms using clinically useful depression outcome scale: Korean validation study. *Clin Psychopharmacol Neurosci*. 2017;15(1):28–34.
21. Trivedi MH. How can measurement-based care help improve treatment outcomes for major depressive disorder in primary care? *J Clin Psychiatry*. 2020;81(2):UT17042BR2C.
22. Lewis CC, Boyd M, Puspitasari A, et al. Implementing measurement-based care in behavioral health: a review. *JAMA Psychiatry*. 2019;76(3):324–335.
23. Van Wert MJ, Malik M, Memel B, et al. Provider perceived barriers and facilitators to integrating routine outcome monitoring into practice in an urban community psychiatry clinic: a mixed-methods quality improvement project. *J Eval Clin Pract*. 2021;27(4):767–775.
24. Trivedi MH, Rush AJ, Crismon ML, et al. Clinical results for patients with major depressive disorder in the Texas Medication Algorithm Project. *Arch Gen Psychiatry*. 2004;61(7):669–680.
25. Bauer M, Pfennig A, Linden M, et al. Efficacy of an algorithm-guided treatment compared with treatment as usual: a randomized, controlled study of inpatients with depression. *J Clin Psychopharmacol*. 2009;29(4):327–333.
26. Zimmerman M, McGlinchey JB. Why don't psychiatrists use scales to measure outcome when treating depressed patients? *J Clin Psychiatry*. 2008;69(12):1916–1919.
27. Hatfield DR, Ogles BM. Why some clinicians use outcome measures and others do not. *Adm Policy Ment Health*. 2007;34(3):283–291.
28. Jensen-Doss A, Haimes EMB, Smith AM, et al. Monitoring treatment progress and providing feedback is viewed favorably but rarely used in practice. *Adm Policy Ment Health*. 2018;45(1):48–61.
29. Morris DW, Trivedi MH. Measurement-based care for unipolar depression. *Curr Psychiatry Rep*. 2011;13(6):446–458.
30. Kendrick T, El-Gohary M, Stuart B, et al. Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. *Cochrane Database Syst Rev*. 2016;7(7):CD011119.
31. Higgins J, Thomas J, Chandler J, et al, eds. *Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020)*. Cochrane Database of Systematic Reviews. Published 2020. https://www.training.cochrane.org/handbook
32. Higgins JPT, Savović J, Page MJ, et al. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, et al, eds. *Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated February 2020)*. Cochrane Database of Systematic Reviews. Published 2020. https://www.training.cochrane.org/handbook
33. Higgins JPT, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–560.
34. Rosenthal R. The file drawer problem" and tolerance for null results. *Psychol Bull*. 1979;86(3):638–641.
35. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–634.
36. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56(2):455–463.
37. Yeung AS, Jing Y, Brenneman SK, et al. Clinical Outcomes in Measurement-based Treatment (Comet): a trial of depression monitoring and feedback to primary care physicians. *Depress Anxiety*. 2012;29(10):865–873.
38. Guo T, Xiang YT, Xiao L, et al. Measurement-based care versus standard care for major depression: a randomized controlled trial with blind raters. *Am J Psychiatry*. 2015;172(10):1004–1013.
39. Wikberg C, Westman J, Petersson E-L, et al. Use of a self-rating scale to monitor depression severity in recurrent GP consultations in primary care - does it really make a difference? a randomised controlled study. *BMC Fam Pract*. 2017;18(1):6.
40. Zhao S, Qin G, Chen J, et al. A study on the effect of measurement-based care using Patient Health Questionnaire Depression Scale (PHQ-9) on patients with depression. *Zhejiang Prev Med*. 2015;28(5):469–472.
41. Chen P, Yang X. A study on the effect of measurement-based care on patients with depression in clinical practice. *Psychol Dr*. 2016;22(10):1007–8231.
42. Adli M, Wiethoff K, Baghai TC, et al. How effective is algorithm-guided treatment for depressed inpatients? results from the randomized controlled multicenter German algorithm project 3 trial. *Int J Neuropsychopharmacol*. 2017;20(9):721–730.
43. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16(3):199–208.
44. WHOQOL Group. Development of the WHOQOL: rationale and current status. *Int J Ment Health*. 1994;23(3):24–56.
45. Lin EHB, Katon WJ, VonKorff M, et al. Relapse of depression in primary care: rate and clinical predictors. *Arch Fam Med*. 1998;7(5):443–449.
46. Paykel ES, Ramana R, Cooper Z, et al. Residual symptoms after partial remission: an

important outcome in depression. *Psychol Med*. 1995;25(6):1171–1180.
47. Keller MB. Past, present, and future directions for defining optimal treatment outcome in depression: remission and beyond. *JAMA*. 2003;289(23):3152–3160.
48. Warden D, Trivedi MH, Carmody T, et al. Adherence to antidepressant combinations and monotherapy for major depressive disorder: a CO-MED report of measurement-based care. *J Psychiatr Pract*. 2014;20(2):118–132.
49. Lambert MJ, Whipple JL, Kleinstäuber M. Collecting and delivering progress feedback: a meta-analysis of routine outcome monitoring. *Psychotherapy (Chic)*. 2018;55(4):520–537.
50. Shimokawa K, Lambert MJ, Smart DW. Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *J Consult Clin Psychol*. 2010;78(3):298–311.

See supplementary material for this article at PSYCHIATRIST.COM.

POSTTEST

To obtain credit, go to PSYCHIATRIST.COM to take this Posttest and complete the Evaluation. A $10 processing fee is required.

1. **Which of the following strategies is a key element of measurement-based care?**
   a. Obtaining neuropsychological testing for patients with cognitive dysfunction
   b. Listing each patient's negative cognitions and behaviors
   c. Using standardized outcome scales in routine monitoring to adjust a patient's treatment
   d. Using a standardized diagnostic interview for every patient and noting how replies differ
2. **This meta-analysis found that, compared with usual care, measurement-based care for depression resulted in all of the following outcomes *except*:**
   a. Greater remission rates
   b. Greater response rates
   c. Lower endpoint severity
   d. Better adherence to medications
3. **You are treating Alyssa for a major depressive episode with an antidepressant. She has been taking the medication for 8 weeks and says that she is feeling much better with no troublesome side effects. With respect to the results from this article, what action would be indicated at this point?**
   a. Check her score on the Patient Health Questionnaire.
   b. Get a serum drug level to ensure that she is taking a therapeutic dose of the antidepressant.
   c. Get collateral information from her husband to ensure that her functioning is improved.
   d. Refer her to a psychologist for cognitive-behavioral therapy.

THE JOURNAL OF
CLINICAL PSYCHIATRY
THE OFFICIAL JOURNAL OF THE AMERICAN SOCIETY OF CLINICAL PSYCHOPHARMACOLOGY

Supplementary Material

Article Title: The Efficacy of Measurement-Based Care for Depressive Disorders: Systematic Review and Meta-Analysis of Randomized Controlled Trials

Author(s): Maria Zhu, MSc; Ran Ha Hong, BSc; Tao Yang, MD, PhD; Xiaorui Yang, MD; Xing Wang, MD; Jing Liu, MHA; Jill K. Murphy, PhD; Erin E. Michalak, PhD; Zuowei Wang, MD, PhD; Lakshmi N. Yatham, MBBS, MBA(Exec); Jun Chen, MD, PhD; and Raymond W. Lam, MD

DOI Number: https://doi.org/10.4088/JCP.21r14034

List of Supplementary Material for the article

1. Appendix 1 Search Strategy
2. Figure 1 Funnel plot for clinical response
3. Figure 2 Funnel plot for clinical remission
4. Figure 3 Funnel plot for standardized mean difference
5. Figure 4 Funnel plot for attrition
6. Figure 5 Funnel plot for medication adherence
7. Appendix 2 Publication Bias Reports
8. Table 1 Post Hoc Sensitivity Analyses Exploring Potential Sources of Heterogeneity

Disclaimer

This Supplementary Material has been provided by the author(s) as an enhancement to the published article. It has been approved by peer review; however, it has undergone neither editing nor formatting by in-house editorial staff. The material is presented in the manner supplied by the author.

Supplementary Material

Appendix 1. Search Strategy

MEDLINE

Publication Dates: Jan. 1, 1946-July 1, 2020

exp Patient Reported Outcome Measures/ OR *"Outcome Assessment (Health Care)"/ OR measurement-based care.ti,ab,kw. OR feedback informed treatment.ti,ab,kw. OR patient reported OR outcome measure*.ti,ab,kw. OR continuous assessment.ti,ab,kw. OR monitoring treatment progress.ti,ab,kw. OR routine outcome monitoring.ti,ab,kw. OR progress monitoring.ti,ab,kw. OR patient level feedback.ti,ab,kw. OR patient focused research.ti,ab,kw. OR outcome* assessment.ti,ab,kw. OR outcome-based care.ti,ab,kw. OR collaborative care.ti,ab,kw. OR care management.ti,ab,kw. OR measurement feedback system.ti,ab,kw. OR measurement-based.ti,ab,kw. OR self-rating.ti,ab,kw. OR self report.ti,ab,kw.

AND

exp Antidepressive Agents/ OR antidepress*.ti,ab,kw. OR exp psychotherapy/ OR psychotherapy.ti,ab,kw.

AND

depressive disorder/ or exp depressive disorder, major/ or exp depressive disorder, treatment-resistant/ or exp dysthymic disorder/ OR *Depression/ OR depress*.ti,ab,kw.

AND

randomized controlled trial.pt. OR RCT.ti,ab,kw. OR random*.ti,ab,kw.

RESULT: 1802

Embase

Publication Dates: Jan. 1, 1980-July 1, 2020

exp patient-reported outcome/ OR exp outcome assessment/ OR measurement-based care.ti,ab,kw. OR feedback informed treatment.ti,ab,kw. OR patient reported outcome measure*.ti,ab,kw. OR continuous assessment.ti,ab,kw. OR monitoring treatment progress.ti,ab,kw. OR routine outcome monitoring.ti,ab,kw. OR progress monitoring.ti,ab,kw. OR patient level feedback.ti,ab,kw. OR patient focused research.ti,ab,kw. OR outcome* assessment.ti,ab,kw. OR outcome-based care.ti,ab,kw. OR collaborative care.ti,ab,kw. OR care management.ti,ab,kw. OR measurement feedback system.ti,ab,kw. OR measurement-based.ti,ab,kw. OR self-rating.ti,ab,kw. OR self report.ti,ab,kw.

AND

exp antidepressant agent/ OR antidepress*.ti,ab,kw. OR exp psychotherapy/ OR psychotherapy.ti,ab,kw.

AND

depression/ OR depress.ti,ab,kw.

AND

exp randomized controlled trial/ OR (random* and control*).ti,ab,kw. OR RCT.ti,ab,kw.

RESULT: 4257

PsycINFO

Publication Dates: 1980-July 1, 2020

DE "Patient Reported Outcome Measures" OR TI “measurement-based care” OR TI “feedback informed treatment” OR TI “patient reported outcome measure*” OR TI “continuous assessment” OR TI “monitoring treatment progress” OR TI “routine outcome monitoring” OR TI “progress monitoring” OR TI “patient level feedback” OR TI “patient focused research” OR TI “outcome* assessment” OR TI “outcome-based care” OR TI “care management” OR TI “measurement feedback system” OR TI “measurement-based” OR TI “self-rating” OR TI “self report” OR AB “measurement-based care” OR AB “feedback informed treatment” OR AB “patient reported outcome measure*” OR AB “continuous assessment” OR AB “monitoring treatment progress” OR AB “routine outcome monitoring” OR AB “progress monitoring” OR AB “patient level feedback” OR AB “patient focused research” OR AB “outcome* assessment” OR AB “outcome-based care” OR AB “care management” OR AB “measurement feedback system” OR AB “measurement-based” OR AB “self-rating” OR AB “self report” OR KW “measurement-based care” OR KW “feedback informed treatment” OR KW “patient reported outcome measure*” OR KW “continuous assessment” OR KW “monitoring treatment progress” OR KW “routine outcome monitoring” OR KW “progress monitoring” OR KW “patient level feedback” OR KW “patient focused research” OR KW “outcome* assessment” OR KW “outcome-based care” OR KW “care management” OR KW “measurement feedback system” OR KW “measurement-based” OR KW “self-rating” OR KW “self report”

AND

TI "depress*" OR AB "depress*" OR KW "depress*" OR (DE "Major Depression")) AND (DE "Major Depression" OR DE "Depression (Emotion)" OR DE "Treatment Resistant Depression" OR DE "Late Life Depression" OR DE "Long-term Depression (Neuronal)")

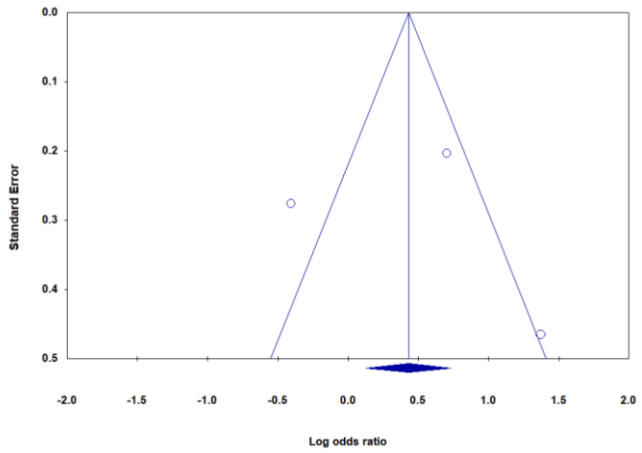
AND
(DE "Antidepressant Drugs" OR DE "Bupropion" OR DE "Citalopram" OR DE "Fluoxetine" OR DE "Fluvoxamine" OR DE "Iproniazid" OR DE "Isocarboxazid" OR DE "Lithium Carbonate" OR DE "Methylphenidate" OR DE "Mianserin" OR DE "Moclobemide" OR DE "Molindone" OR DE "Nefazodone" OR DE "Nialamide" OR DE "Nomifensine" OR DE "Paroxetine" OR DE "Phenelzine" OR DE "Pheniprazine" OR DE "Pipradrol" OR DE "Serotonin Norepinephrine Reuptake Inhibitors" OR DE "Sertraline" OR DE "Sulpiride" OR DE "Tranylcypromine" OR DE "Trazodone" OR DE "Tricyclic Antidepressant Drugs" OR DE "Venlafaxine" OR DE "Zimeldine") OR DE "Antidepressant Drugs" OR TI "antidepress*" OR AB "antidepress*" OR KW "antidepress*" OR DE “psychotherapy” OR TI "psychotherapy" OR AB "psychotherapy*" OR KW "psychotherapy" OR TI “therapy” OR AB “therapy” OR KW “therapy”
AND
DE "Randomized Controlled Trials" OR TI "random*" OR AB "random*" OR KW "random*" OR TI "RCT" OR AB "RCT" OR KW "RCT"
RESULT: 712

ClinicalTrials.gov
Publication Dates: first posted on or before July 1, 2020
“depression” OR “depressive disorder”
AND
“measurement-based” OR “patient reported outcome measure” OR “feedback informed treatment” OR “continuous assessment” OR “routine outcome monitoring” OR “monitoring treatment progress” OR “progress monitoring” OR “patient level feedback”
RESULT: 148

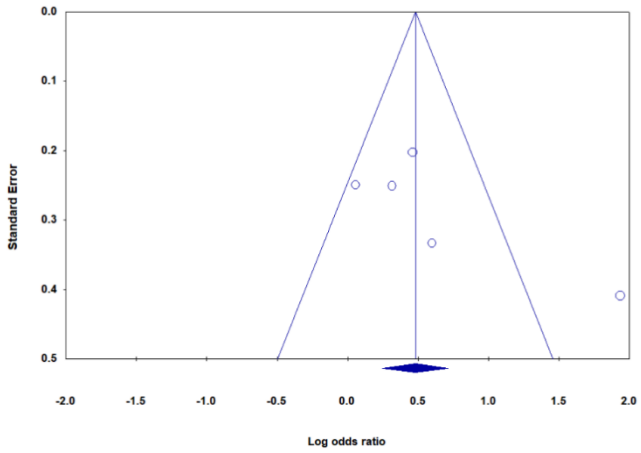
CNKI
(AB='基于评估的治疗'+'患者报告的结果评估'+'结果评估'+'反馈信息处理'+'患者报告结果评估'+'连续评估'+'监测治疗进展'+'常规结果监测'+'进度监测'+'患者反馈'+'患者焦点研究'+'结果评估'+'结局评估'+'基于结果的治疗'+'协作治疗'+'合作治疗'+'治疗管理'+'评估反馈系统'+'基于评估的'+'自评') AND (AB='抗抑郁药'+'抗抑郁'+'心理治疗') AND (AB='抑郁症'+'难治性抑郁症'+'情绪障碍'+'恶劣心境障碍'+'抑郁') AND (AB='随机对照试验'+'RCT'+'随机')
(AB='Measurement-based Care'+'Patient Reported Outcome Measures'+'Outcome Assessment'+'Feedback informed treatment'+'Patient reported'+'continuous assessment'+' monitoring treatment progress'+' routine outcome monitoring '+' progress monitoring'+' patient level feedback '+' patient focused research '+'outcome assessment'+'outcome measure'+'outcome-based care'+'collaborative care'+'collaborative treatment'+'care management'+'measurement feedback system'+'measurement-based'+'self-rating') AND (AB='antidepressant '+'antidepress'+'psychotherapy') AND (AB='depressive disorder'+'treatment resistant depression'+'mood disorders'+' dysthymic disorder '+'depression') AND (AB=' randomized controlled trial '+'RCT'+'random')
Result: 888
Hand search using “depression” + “measurement-based care”
Result: 5

Wanfang Database
(全部:(主题：'基于评估的治疗"+"患者报告的结果评估"+"结果评估"+"反馈信息处理"+"患者报告结果评估"+"连续评估"+"监测治疗进展"+"常规结果监测"+"进度监测"+"患者反馈"+"患者焦点研究"+"结果评估"+"结局评估"+"基于结果的治疗"+"协作治疗"+"合作治疗"+"治疗管理"+"评估反馈系统"+"基于评估的"+"自评") AND (主题："抗抑郁药"+"抗抑郁"+"心理治疗") AND (主题：'抑郁症"+"难治性抑郁症"+"情绪障碍"+"恶劣心境障碍"+"抑郁") AND (主题：'随机对照试验"+"RCT"+"随机"))*Date:-2019
(ALL:(Topic：'Measurement-based Care'+'Patient Reported Outcome Measures'+'Outcome Assessment'+'Feedback informed treatment'+'Patient reported'+'continuous assessment'+' monitoring treatment progress'+' routine outcome monitoring '+' progress monitoring'+' patient level feedback '+' patient focused research '+'outcome assessment'+'outcome measure'+'outcome-based care'+'collaborative care'+'collaborative treatment'+'care management'+'measurement feedback system'+'measurement-based'+'self-rating') AND (Topic: 'antidepressant '+'antidepress'+'psychotherapy') AND (Topic: 'depressive disorder'+'treatment resistant depression'+'mood disorders'+' dysthymic disorder '+'depression') AND (Topic: ' randomized controlled trial '+'RCT'+'random'))*Date:-2019
Result: 1010
Hand search using “depression” + “measurement-based care”
Result: 8

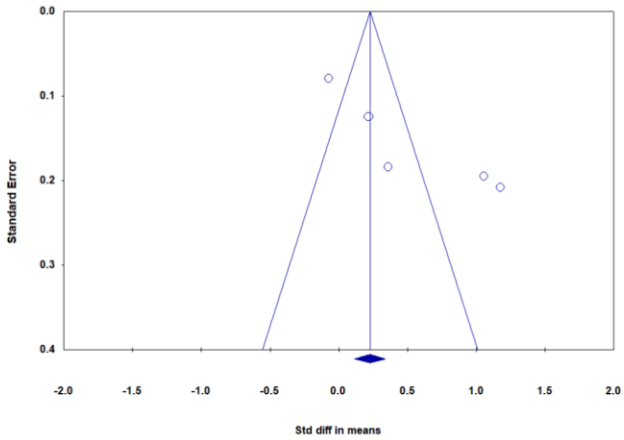
Supplementary Figure 1. Funnel plot for clinical response.



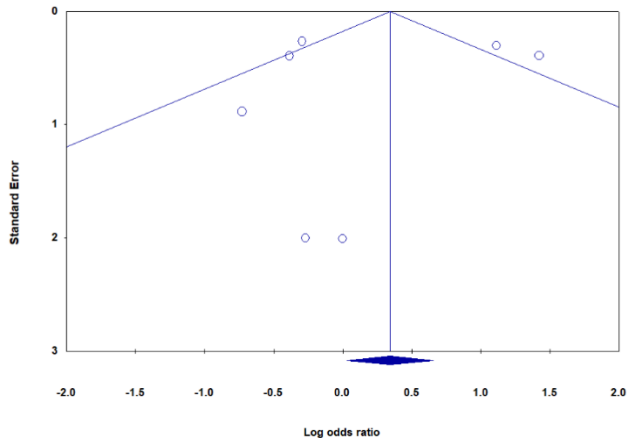
Supplementary Figure 2. Funnel plot for clinical remission.



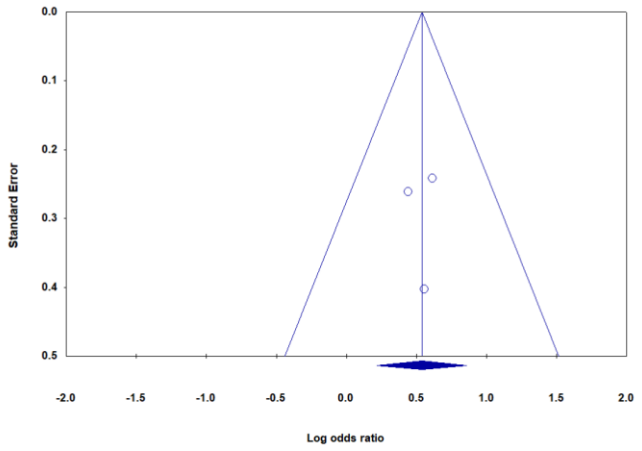
Supplementary Figure 3. Funnel plot for standardized mean difference.



Supplementary Figure 4. Funnel plot for attrition.



Supplementary Figure 5. Funnel plot for medication adherence.



Appendix 2. PUBLICATION BIAS REPORTS

RESPONSE

Classic fail-safe N

This meta analysis incorporates data from 3 studies, which yield a z-value of 2.85133 and corresponding 2-tailed p-value of 0.00435. The fail-safe N is 4. This means that we would need to locate and include 4 'null' studies in order for the combined 2-tailed p-value to exceed 0.050.

Egger's Test of the Intercept

In this case the intercept (B0) is 1.25855, 95% confidence interval (-91.48020, 93.99730), with t=0.17243, df=1. The 1-tailed p-value (recommended) is 0.44565, and the 2-tailed p-value is 0.89129.

REMISSION

Classic fail-safe N

This meta analysis incorporates data from 5 studies, which yield a z-value of 4.61903 and corresponding 2-tailed p-value of 0.00000. The fail-safe N is 23. This means that we would need to locate and include 23 'null' studies in order for the combined 2-tailed p-value to exceed 0.050.

Egger's Test of the Intercept

In this case the intercept (B0) is 5.68818, 95% confidence interval (-4.19552, 15.57189), with t=1.83154, df=3. The 1-tailed p-value (recommended) is 0.08221, and the 2-tailed p-value is 0.16442.

STANDARDIZED MEAN DIFFERENCE

Classic fail-safe N

This meta analysis incorporates data from 5 studies, which yield a z-value of 6.19746 and corresponding 2-tailed p-value of 0.00000. The fail-safe N is 45. This means that we would need to locate and include 45 'null' studies in order for the combined 2-tailed p-value to exceed 0.050.

Egger's Test of the Intercept

In this case the intercept (B0) is 8.10452, 95% confidence interval (2.25827, 13.95077), with t=4.41175, df=3. The 1-tailed p-value (recommended) is 0.01080, and the 2-tailed p-value is 0.02161.

Duval and Tweedie's Trim and Fill

The program is looking for missing studies based on a random effects model, and is looking for missing studies only to the left side of the mean effect (these parameters are set by the user). Using these parameters the method suggests that no studies are missing. Under the random effects model the point estimate and 95% confidence interval for the combined studies is 0.52513 (0.06300, 0.98727). Using Trim and Fill these values are unchanged.

ALL-CAUSE DISCONTINUATION

Classic fail-safe N

This meta analysis incorporates data from 7 studies, which yield a z-value of 1.60477 and corresponding 2-tailed p-value of 0.10855. Since the combined result is not statistically significant, the Fail-Safe N (which addresses the concern that the observed significance may be spurious) is not relevant.

Egger's Test of the Intercept

In this case the intercept (B0) is -0.35304, 95% confidence interval (-4.51414, 3.80806), with t=0.21810, df=5. The 1-tailed p-value (recommended) is 0.41799, and the 2-tailed p-value is 0.83598.

MEDICATION ADHERENCE

Classic fail-safe

This meta analysis incorporates data from 3 studies, which yield a z-value of 3.24099 and corresponding 2-tailed p-value of 0.00119. The fail-safe N is 6. This means that we would need to locate and include 6 'null' studies in order for the combined 2-tailed p-value to exceed 0.050.

Egger's Test of the Intercept

In this case the intercept (B0) is -0.07954, 95% confidence interval (-17.92785, 17.76877), with t=0.05662, df=1. The 1-tailed p-value (recommended) is 0.48200, and the 2-tailed p-value is 0.96399.

Supplementary Table 1. Post Hoc Sensitivity Analyses Exploring Potential Sources of Heterogeneity

Response

Adli removed (non-algorithm studies remain)

Model		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	2	2.249	1.560	3.241	4.347	0.000	1.730	1	0.188	42.212	0.094	0.315	0.099	0.307
Random	2	2.475	1.355	4.518	2.950	0.003								

Guo removed (non-Chinese studies remain)

Model		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	2	1.368	0.992	1.886	1.912	0.056	10.422	1	0.001	90.405	0.554	0.867	0.751	0.744
Random	2	1.180	0.399	3.489	0.299	0.765								

Yeung removed (psychiatric setting studies remain)

Model		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	2	1.060	0.666	1.688	0.245	0.806	10.769	1	0.001	90.714	1.429	2.228	4.962	1.195
Random	2	1.559	0.274	8.861	0.501	0.616								

Remission

Non-algorithm studies only

Model		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	3	1.670	1.252	2.227	3.492	0.000	15.541	2	0.000	87.131	0.497	0.609	0.371	0.705
Random	3	2.124	0.896	5.036	1.711	0.087								

Non-Chinese studies only

Model		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	4	1.415	1.110	1.804	2.805	0.005	2.278	3	0.517	0.000	0.000	0.052	0.003	0.000
Random	4	1.415	1.110	1.804	2.805	0.005								

Psychiatric settings studies only

Model		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	3	2.044	1.436	2.910	3.971	0.000	11.555	2	0.003	82.691	0.506	0.626	0.392	0.711
Random	3	2.478	1.019	6.027	2.002	0.045								

Standardized Mean Difference

Note: All 3 Chinese studies involved psychiatric clinic patients, while both non-Chinese studies involved primary care patients, so the results are the same

Psychiatric settings OR Chinese studies (Guo, Chen, Zhao remain)

Model		Effect size and 95% confidence interval					Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	3	0.831	0.113	0.013	0.610	1.052	7.380	0.000	10.566	2	0.005	81.071	0.164	0.202	0.041	0.405
Random	3	0.858	0.260	0.067	0.349	1.367	3.306	0.001								

Primary care studies OR non-Chinese studies (Wikberg & Yeung remain)

Model		Effect size and 95% confidence interval					Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	2	0.011	0.067	0.005	-0.120	0.143	0.167	0.867	3.894	1	0.048	74.323	0.032	0.060	0.004	0.178
Random	2	0.057	0.145	0.021	-0.228	0.342	0.393	0.695								