Principles and Practice of Clinical Research
A GLOBAL JOURNAL OF CLINICAL RESEARCH

Open-Label Placebo Effects in Reducing Depression Severity Among Adults with Major Depressive Disorder: A Systematic Review and Meta-Analysis of Early Evidence

Janet Mercedes Arévalo-Ipanaqué[1,#], Maria Leticia Gobo Silva[2,#], Ricardo Romero-Vucovich[3,#], Gianluca Conte[4], Matheus Gerhard Rosenfeld[5], Sergio José Fausto Girón[6], Lisa Gonzalez[7], Jordy Batista[8], Sofia Ines Volpi[9], Maximiliano Espinosa[10], Daniela Lorena Alvarado Rodriguez[11], Salvatore Minnella[12], Tadeus Lopez Ulloa[13], Nadine Aranis Seguic[14], Yazmin Heredia Allegretti[15], Johan Antonio De La Rosa Laureano[16], Cristian Camilo Bermudez[17], Letícia Watanabe Koyama[18], Maria Eduarda Nabhan de Souza[18], Somaya Althobaiti[19], Karen Gamonal[20], Folasade Mary Adeyemi[21], Luis Pedro Lorenzana[22], Aderibigbe Adeniyi Sheriff[23], Peter Benedek[24], Emilio Vilorio-Alvarez[25], Wajih Jawhar[26,27,28*]

[1] Universidad Peruana Cayetano Heredia, Facultad de Enfermería, Lima, Perú; [2] Reference Center for Sarcomas and Bone Tumors / Department of Radiation Oncology, A.C.Camargo Cancer Center, Sao Paulo, Brazil; [3] Stroke Clinic, Instituto Nacional de Neurología y Neurocirugía, Mexico City, Mexico; [4] Clinical Research Service, IRCCS Policlinico San Donato, San Donato Milanese, Italy; [5] Hospital das Clinicas, Sao Paulo University, Sao Paulo, Brazil; [6] Faculty of Medicine, Universidad Francisco Marroquín, Guatemala City, Guatemala; [7] ONE HealthIT LLC, New Rochelle, NY, USA; [8] Pontificia Universidad Católica Madre y Maestra, Santiago, Dominican Republic; [9] Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA; [10] Departamento de Traumatología Infantil, Clínica Alemana Santiago, Santiago, Chile; [11] Clinica Alemana, Santiago, Chile; [12] Evidence Based Slow Musculoskeletal Medicine Approach Institute (EBaSMMAI), Lecce, Italy; [13] Two Oceans in Health, Santo Domingo, Dominican Republic; [14] Neuromodulation Center and Center for Clinical Research Learning, Spaulding Rehabilitation Hospital and Massachusetts General Hospital, Harvard Medical School, Boston, MA; [15] Sinai Hospital of Baltimore; [16] Clínica Corominas, Santiago, Dominican Republic; [17] Neurology department, Hospital Universitario Santa Clara, Bogotá DC, Colombia; [18] Faculty of Medicine São Leopoldo Mandic, São Paulo, Brazil; [19] Clinical Pharmacy Department Taif University; [20] Cardiology Resident, Universidad Ricardo Palma – Clínica Good Hope, Lima, Peru; [21] Radiology Department, Obafermi Awolowo University Teaching Hospitals Complex, Ile-Ife, Nigeria; [22] Faculty of Medicine, Universidad Francisco Marroquín, Guatemala City, Guatemala; [23] Division of Oncology, African Medical Center of Excellence, Abuja, Nigeria; [24] Gedeon Richter Plc, Budapest, Hungary; [25] CEDIMAT, Santo Domingo, Dominican Republic; [26] Division of Experimental Medicine, Department of Medicine, McGill University, Montreal, Quebec, Canada; [27] Cancer Research Program, The Research Institute of the McGill University Health Center, Montreal, Quebec, Canada; [28] Child Health and Human Development Program, The Research Institute of the McGill University Health Center, Montreal, Quebec, Canada.

Abstract

Background: *Open-label placebo (OLP) represents an intervention where patients knowingly receive inert treatments associated with clinical benefits. This approach has demonstrated efficacy across multiple medical conditions. Meanwhile, major depressive disorder (MDD) affects 280 million people worldwide, yet three-quarters receive no treatment, urging non-pharmacological approaches. Despite this high prevalence, there is scarce data on the potential of OLP intervention for MDD. Our objective is to determine whether OLP can meaningfully reduce depression scores in adults with MDD.*

Methods: *Following Cochrane methodology and PRISMA guidelines, six databases were systematically reviewed including studies from up to May 5, 2025. Inclusion criteria encompassed randomized controlled trials (RCTs) examining OLP interventions in adults with MDD, measuring depression severity using validated scales. Studies were independently screened by multiple reviewers. Risk of bias was assessed using RoB2. A meta-analysis was conducted using post-intervention mean scores and standard deviations.*

Results: *From 1,886 screened records, 3 randomized controlled trials (n = 118; mean age 46.8 years; 78% female) met the inclusion. Trials were conducted in the USA, Israel, and Austria between 2012 and 2022, with sample sizes ranging from 20–60 participants. All compared open-label placebo (OLP) to treatment-as-usual or waitlist controls, using validated depression scales (HAM-D-17, QIDS-SR, BDI-II). Individual trial findings were mixed: one trial found significant symptom reduction with OLP plus cognitive-behavioral therapy, though the effect was small; two trials reported no overall between-group differences, although one showed improvement in patients with early-onset depression. A meta-analysis of two small trials at 4 weeks suggested a modest reduction in depressive symptoms favoring OLP (Hedges's $g = -0.43$, 95% CI -0.83 to -0.04), although both individual trials were underpowered and did not reach statistical significance. Given the limited number of studies and high risk of bias, these findings should be interpreted as preliminary. Risk of bias was rated as some concerns in one study and high in two, primarily due to outcome measurement and missing data.*

Conclusions: *Current evidence from three small RCTs suggests that open-label placebo may be associated with modest short-term symptom improvement in adults with MDD; however, the evidence base is limited, heterogeneous, and at risk of bias. These findings should be considered preliminary and hypothesis-generating, and larger, methodologically rigorous trials are required before clinical recommendations can be made.*

Introduction

Major depressive disorder (MDD) is a common mental illness characterized by persistent sadness and loss of interest or pleasure in previously enjoyed activities. MDD affects about 5% of adults worldwide, approximately 280 million people, and more than 75% of individuals in low- and middle-income countries do not receive treatment (WHO, 2023).

Consequently, MDD is a complex public health problem that affects patients, their families, and their social networks. Beyond its impact on quality of life and productivity, MDD is also associated with an elevated risk of premature death, largely due to suicide and comorbid medical conditions (Herrman et al., 2019; Marwaha et al., 2023).

Psychotherapy and pharmacotherapy with antidepressants and selective serotonin reuptake inhibitors (SSRIs) have long been considered the standard of care for this condition (Simon et al., 2024). However, patients taking antidepressants often experience adverse effects, ranging from gastrointestinal problems, sexual dysfunction, and weight change to increased suicidal thoughts and serotonin syndrome. These side effects can significantly impact treatment adherence and efficacy (Braund et al., 2021). Therefore, developing alternative treatments are critically needed for MDD patients.

Recently, open-label placebo (OLP) has shown positive results in people with clinical disorders and non-clinical symptoms (Schaefer et al., 2023). Randomized control trials (RCTs) demonstrated that OLP can reduce anxiety and acute stress (Schaefer & Enge, 2024; Schaefer et al., 2021); either independently or in conjunction with depressive symptoms (Guevarra et al., 2024). Additionally, other studies reported that patients suffering from a range of disease such as irritable bowel syndrome, depression, allergic rhinitis, back pain, attention-deficit/hyperactivity disorder, induced pruritus, cancer-related fatigue, and menopausal hot flashes can benefit from OLP as an intervention. (Antonelli & Donelli, 2019; Charlesworth et al., 2017; von Wernsdorff et al., 2021).

Despite growing evidence, to date no systematic reviews have comprehensively assessed OLP in the context. MDD, even though it may represent a promising strategy to mitigate the risks of antidepressant overprescription (Krockow et al., 2023).

Given the current limitations of standard care, especially its prevalent adverse effects, and the need to offer a low-risk alternative, we sought to evaluate and synthesize the existing evidence on using OLP to reduce depression severity in patients with MDD. Addressing this critical knowledge gap will provide evidence-based guidance for clinical practice, policy development, and potential future research directions.

Materials and Methods

A systematic review of RCTs was conducted following the Cochrane Handbook for Systematic Reviews of Interventions methodology (Higgins et al., 2024) and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines (Page et al., 2021). The review protocol was developed internally a priori to minimize bias and ensure methodological rigor throughout the process. Moreover, it was not registered in any international systematic review registry. The systematic literature search commenced on April 10, 2025. Comprehensive searches, study selection, data extraction, and quality assessment were conducted until June 15th, 2025, and were performed by multiple independent reviewers to enhance reliability and reduce selection bias.

Search strategy

Following our research question, search strategies were prepared using MeSH terms (open label placebo, non-deceptive placebo, depressive disorder major, major depressive disorder, depressive symptoms), standardized terms, and Boolean operators for the databases Cochrane, PsycNet, PMC, ClinicalTrials.gov, Scopus and PubMed (supplementary table 1).

Inclusion and exclusion criteria

We included RCTs involving adults ($\geq$18 years) diagnosed with MDD, treated with OLP, and reporting depression severity using validated scales. Articles had to be published in English, with full-text available online. We excluded studies involving bipolar disorder, psychotic symptoms, double-blind designs, and study protocols.

Selection of studies and data extraction

All authors were involved in the review process. Two reviewers conducted the initial search for articles. A third reviewer intervened if there was a disagreement regarding eligibility, and

*Corresponding author: Wajih.jawhar@mail.mcgill.ca
# Authors have contributed equally to this work.
Received: July 9, 2025 **Accepted:** February 16,2026
Published: April 26, 2026
Editor: Felipe Fregni
Reviewers: Inia Perez, Juan Moncada, Sibin Marian, Augusto J. Mendes, Ygor Junqueira
Keywords: open-label placebo, major depressive disorder, randomized clinical trials, diagnostic criteria, outcome measurement scales
DOI: https://doi.org/10.21801/ppcrj.2025.114.3

a consensus was reached. The reviewers independently screened the titles and abstracts after removing duplicates in the Covidence web-based platform (Kellermeyer et al., 2018). Subsequently, full texts were reviewed if a paper was considered relevant based on the inclusion and exclusion criteria.

Extracted items included:

- Study characteristics: author, year, country, design, funding source.
- Population: inclusion/exclusion criteria, baseline demographics (age, gender), diagnostic criteria for major depressive disorder (MDD), depression severity.
- Intervention details: group allocation, type and comparator, route of administration, frequency, duration, dosage.
- Study setting and recruitment methods, where available.
- Outcomes: depression scales used, direction of scoring, range, baseline and 4-week follow-up means and standard deviations.

For meta-analysis, the results of means, sample size, and SD were extracted to calculate the standardized mean difference (SMD) and its 95% CI.

Data synthesis

Selected clinical trials were summarized in Microsoft Excel tables. Given the small number of eligible trials, the quantitative synthesis was conducted for exploratory purposes. For the meta-analysis, a Random-effects model was used to combine the standardized mean differences between studies. Hedges's g method was used to estimate effect sizes, and 95% confidence intervals were calculated. Cochran's Q and I^2, and Tau-squared statistics were used to assess the heterogeneity of the included studies. A $I^2 > 50\%$ or $P < 0.05$ indicated statistically significant heterogeneity. All statistical analyses and forest plots were performed using STATA MP version 19 statistical software (StataCorp, 2025)

Risk of bias assessment

Three independent reviewers (ME, MLGS and SJFG) assessed the risk of bias of the included studies using the Cochrane Risk-of-Bias 2 tool (Sterne et al., 2019). Following Cochrane's recommendations, a traffic light plot was created manually using Excel to visually summarize the risk-of-bias assessments. After completing independent assessments, the reviewers convened to discuss any discrepancies, and consensus was reached through a joint meeting to ensure uniform judgments across all domains.

Results

The search strategy identified a total of 1,886 studies. One duplicate was identified manually, and 49 additional duplicates were detected by Covidence automatic detection tool, resulting in 1,836 studies screened by title and abstract. Of these, 1,812 records were excluded based on the inclusion criteria. A total of 24 studies were selected for full-text review. Twenty-one of these were excluded for the following reasons: wrong patient population (n = 12), wrong study design (n = 5), wrong intervention (n = 2), congress abstract (n = 1), and trial closed (n = 1). No studies were found to be ongoing at the time of the review. Ultimately, three RCTs (Kelley et al., 2012; Nitzan et al., 2020; Schienle et al., 2022) met all inclusion criteria and were included in this systematic review. The study selection process is detailed in the PRISMA flow diagram (Figure 1).

Description of the studies

Table 1 summarizes the key characteristics of the three RCTs included in this review. The studies were published between 2012 and 2022 and were conducted in the USA, Israel and Austria. All trials utilized randomized controlled designs comparing OLP to either treatment-as-usual (TAU) or waitlist controls. Kelley et al. and Schienle et. al used a parallel-group design to assess primary outcomes, while Nitzan et al. used a crossover design after four weeks of parallel-group design. Sample sizes across the studies ranged from 20 to 60 participants.

Population

The three studies enrolled a total of 118 participants (mean age 46.8 $\pm$ 14.2 years; females 77.8%) diagnosed with nonpsychotic MDD. Diagnostic criteria varied depending on the study (DSM-IV, ICD-10, DSM-V). Most study participants exhibited moderate depression scores, based on the different scales used and only the study by Schienle et al., included patients with severe depression (9/60 participants; 15%). No significant differences in baseline characteristics between groups were reported.

Intervention characteristics and effects / Exposure/Control

The OLP interventions varied in form, delivery, and context, ranging from oral capsules to

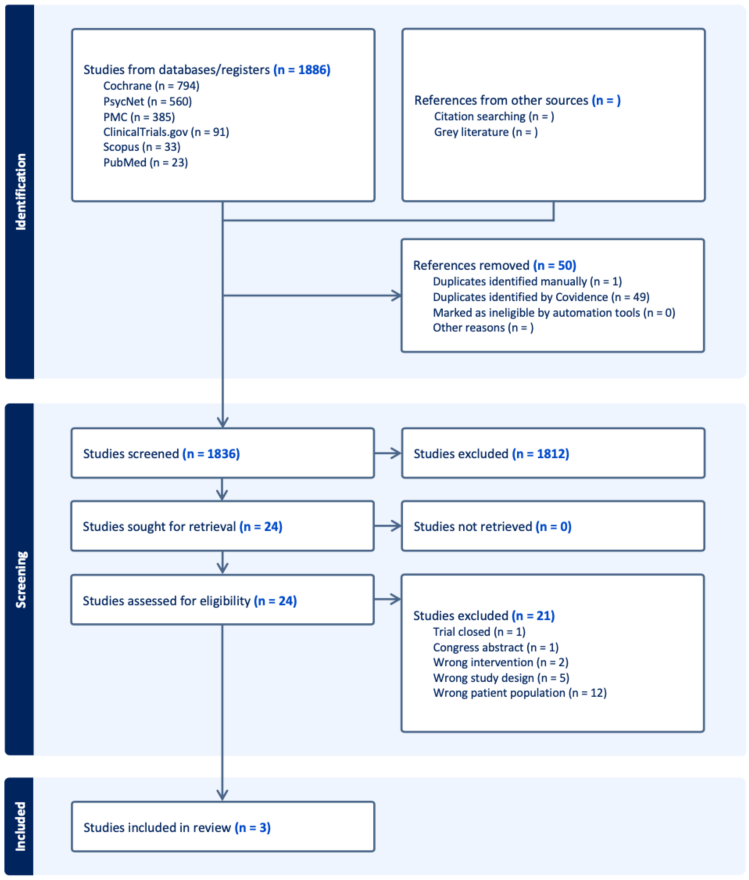

Figure 1: *PRISMA flowchart illustrating the identification, screening, and inclusion of studies in the systematic review.*

Author, year	Country	Study design	N	Age years (Mean ± SD)	Sex (% Female)	Diagnostic criteria	Depression severity	Intervention	Control	Outcome measures	Follow up
Kelley, 2012	USA	Parallel	20	38.8 ± 12.6	70%	DSM-IV	Moderately severe depression HAM-D-17: 18 ± 4.9	OLP: 2 pills BID for 4 weeks	Waitlist for 2 weeks followed by OLP for 4 weeks	HAM-D-17, QIDS, SDQ	2 weeks 4 weeks
Nitzan, 2020	Israel	Parallel: Crossover	38	OLP: 48.17 ± 16.86 TAU: 51.65 ± 17.68	OLP: 72% TAU: 57%	DSM-V	Mild to Moderate	OLP: 2 capsules BID for 8 weeks (or 4 weeks for TAU group after crossover)	TAU for 4 weeks followed by OLP for another 4 weeks	CGI, HARS, SES, ASEC, QIDS	4 weeks 8 weeks
Schienle, 2022	Austria	Parallel	60	CBT: 48.9 ± 11.8 CBT + OLP: 46 ± 11.7	83%	ICD-10	Mild CBT: 10 CBT+OLP: 13 Moderate CBT: 15 CBT+OLP:13 Severe CBT: 5 CBT+OLP: 4	CBT + OLP: 4-week CBT outpatient program with 0.15mL of open-label placebo (oral drops of sunflower oil)	CBT	BDI-II, relaxation quantity and quality, perceived overall effectiveness of the OLP, placebo usage and clinician rating (7-point Likert scale)	4 weeks 3 months

Abbreviations: OLP: Open-Label Placebo; TAU: Treatment as Usual; CBT: Cognitive-behavioral therapy; HAM-D: Hamilton Scale for Depression; QIDS: Quick Inventory of Depressive Symptoms; SDQ: Symptoms of Depression Questionnaire; CGI: Clinical Global Impression; HARS: Hamilton Anxiety Rating Scale; SES: Stanford Expectancy Scale; ASEC: Antidepressant Side Effects Checklist; BDI: Beck Depression Inventory.

Table 1: *Participant characteristics of the included studies (n = 3).*

placebo oil, but in all cases, patients were explicitly informed that they were receiving an inert substance. In the two studies conducted by Kelley et al. and Nitzan et al., patients were instructed to take two OLP pills, twice daily. The third study by Schienle et. al used sunflower oil in a labeled dropper bottle as their OLP and instructed participants to take three oral drops daily prior to home-based relaxation training.

Outcomes

The primary outcomes across the three studies were reduction in depressive symptoms, assessed using standardized tools such as the Hamilton Depression Rating Scale (HAM-D-17), the Beck Depression Inventory-II (BDI-II), and the Quick Inventory of Depressive Symptomatology (QIDS-SR16). These measures were used to evaluate changes in symptom severity over time and between treatment groups.

Secondary outcomes included adherence to therapeutic activities (e.g., frequency of relaxation practice), subjective ratings of relaxation quality, perceived effectiveness of the placebo, placebo intake quantity, clinician-rated global improvement, and dropout rates, and other correlative relationships. Nitzan et al., also explored subgroup effects based on age and onset of depression, non-geriatric adults (<65y) with early onset of the disorder (<50 years).

Main findings

The three reviewed RCTs were not fully concordant in their conclusions, as their results diverged based on the specific analytical approach employed and the context studied (Table 2). Of these studies, Schienle et al. demonstrated that OLP interventions were associated with a statistically significant reduction in depressive symptoms over time. The reduction in depressive score was measured using the self-reported (BDI-II) and analyzed with a mixed-model Analysis of Variance (ANOVA), focusing on the group-by-time interaction (group*time).

Despite this statistical significance, the research team reported that these findings were not clinically meaningful, as BDI-II scores did not show a substantial difference between the intervention groups (CBT+OLP versus CBT) following the 3-week program. Furthermore, although most participants (83%) experienced a shift to a lower depression category (mild, moderate, or severe BDI-II scores) by the end of the program. There were no significant differences observed between the assigned groups regarding the number of patients within each severity category. The number of patients who experienced changes in their severity groups, and those who did not, also did not significantly vary depending on the assigned intervention arm.

Interestingly, the OLP group exhibited higher adherence to adjunctive therapeutic activities, such as relaxation exercises, when compared to the control groups. Nevertheless, this increased adherence did not translate into significant time-by-group interactions concerning the frequency or quality of relaxation reported.

Studies conducted by Kelley et al. and Nitzan et al. were unable to conclude statistically significant differences between OLP and control groups at their respective primary assessment endpoints (T0-T1 at 2 weeks for Kelley et al.; T0-T1-T2 at 8 weeks for Nitzan et al.) using HAM-D-17 and self-reported QIDS-SR16 as the main outcome, respectively. However, Nitzan et al., conducted a subgroup analysis using repeated measures ANOVA, which focused specifically on non-geriatric patients with early-onset depressive disorder. It concluded that for this subgroup, the rate or magnitude of improvement over the course of the treatment was significantly different between the groups, with the OLP arm demonstrating a greater reduction in symptoms, indicated by a significant Time*group interaction.

Lastly, although dropout rates and acceptance of OLP varied across studies, the use of OLP was associated with a higher rate of dropout, particularly among participants who perceived the placebo as ineffective in the Schienle et al. study.

Meta-Analysis

A meta-analysis was conducted using post-intervention mean scores and standard deviations at 4 weeks, as reported by two RCTs (Nitzan et al., 2020; Schienle et al., 2022) (Figure 2). Results from Kelley et al., were excluded from the quantitative synthesis as it reported outcomes only at 2 weeks post-intervention, precluding comparability with the other trials. The pooled analysis yielded a standardized mean difference (Hedges's g) of -0.43 (95% CI: -0.83 to -0.04, $p = .03$) in favor of open-label placebo. Heterogeneity across studies was negligible ($\tau^2 = 0.00$, $I^2 = 0.00\%$, $p = .93$). Neither individual study reached statistical significance (Nitzan et al.: $g = -0.41$ [95% CI: -1.04 to 0.22]; Schienle et al.: $g = -0.45$ [95% CI: -0.95 to 0.06]), although both studies showed effect estimates favoring OLP.

Risk of bias assessment

Author, year	N/Arm	Dropout	Primary Outcome Measurement	Baseline (Mean ± SD)	Symptom Improvement at 2 Weeks	Post-Intervention – 4 Weeks (Mean ± SD)
Kelley, 2012	OLP: 11 Waitlist: 9	25%	HAM-D-17 QIDS SDQ	18 ± 4.94 14.85 ± 2.68 146.94 ± 19.71	Waitlist: -0.67 ± 4 / OLP: 1.64 ± 4.52 Waitlist: -0.22 ± 2.44 / OLP: 2.27 ± 3.88 Waitlist: 1.38 ± 10.77 / OLP: 3.7 ± 18.98	14.75 ± 6.61 12.1 ± 4.34 137.06 ± 27.5
Nitzan, 2020	OLP: 18 TAU: 20	OLP: 4 TAU: 2	QIDS	All participants OLP: 11.28 ± 5.14 TAU: 11.6 ± 4.47 Non-geriatric adults OLP: 13.08 ± 5.32 TAU: 11.62 ± 4.21	NR	All participants OLP: 9.33 ± 4.97 TAU: 11.15 ± 3.65 Non-geriatric adults OLP: 10 ± 5.71 TAU: 11.38 ± 3.99
Schienle, 2022	CBT: 30 CBT + OLP: 30	CBT: 4 CBT + OLP: 9	BDI-II	CBT: 22.3 ± 4.8 CBT+OLP: 22.2 ± 5.49	NR	CBT: 14.77 ± 5.76 CBT+OLP: 11.9 ± 4.95

Abbreviations: OLP: Open-Label Placebo; TAU: Treatment as Usual; CBT: Cognitive-behavioral therapy; HAM-D: Hamilton Scale for Depression; QIDS: Quick Inventory of Depressive Symptoms; SDQ: Symptoms of Depression Questionnaire; BDI: Beck Depression Inventory.

Table 2: *Results of the included studies.*

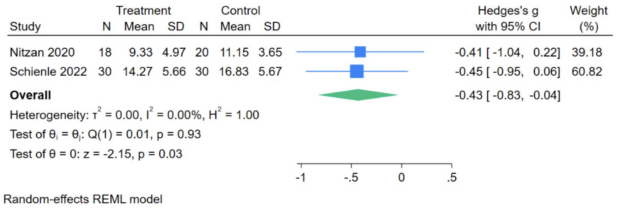

Figure 2: *Forest plot of the meta-analysis comparing OLP versus control on depression severity at 4 weeks post-intervention.*

As summarized in figure 3, all three studies raised some concerns regarding bias arising from the randomization process (Domain 1). This was because none of the three studies described the method of allocation, merely referring to their studies as randomized controlled trials, without providing further details on randomization techniques used.

It is important to interpret certain domains in the specific context of open-label placebo (OLP) trials. By design, OLP interventions are unblinded; therefore, domains related to deviations from intended interventions (Domain 2) and measurement of the outcome (Domain 4) require nuanced consideration. However, Kelley et al., and Schienle et al., still presented some concerns related to Domain 2, because caregivers, personnel delivering the intervention, and/or psychologists were aware of participants' assignments.

Similarly, Nitzan et al. and Schienle et al. were judged to be at high risk in the measurement of the outcome (Domain 4) because outcome assessors were aware of participants' treatment allocations.

Schienle et al. showed a high risk of bias in Domain 3 because of a substantial differential dropout rate—27% in the CBT + OLP group compared to 7% in the CBT-only group at follow-up.

Pertaining to domain 5, Kelley et al. raised some concerns as no statistical analysis plan was mentioned in the trial registration or protocol, and multiple outcome measures (HAM-D, QIDS, and SDQ) were used, allowing flexibility in the emphasis of reported results. Finally, although Nitzan et al., used QIDS to assess depression severity, their clinical trial registry (NCT02666989) specified intention to use HAM-D-17 (Hamilton, 1960) as a readout (outcome) to be analyzed. Surprisingly analysis using QIDS scores was conducted and reported. This clear shift in the methodology used was alarming and indicated potential bias leading to high concerns for domain 5.

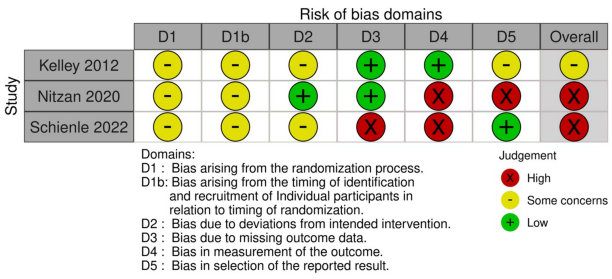

Figure 3: *RoB2 assessment of included articles (n=3).*

Discussion

This systematic review summarizes evidence from three RCTs evaluating OLP for MDD in adults, which collectively suggest modest symptom improvements compared with TAU or control conditions, though not all results reached statistical significance. While previous research has studied the potential benefits of OLP for a variety of clinical conditions including chronic low back pain (Flavio-Reis et al., 2025), cancer-related fatigue (Hoenemeyer et al., 2021), and irritable bowel syndrome (Charlesworth et al., 2017), to our knowledge, this is the first systematic review to focus exclusively on the use of OLP for depression.

Despite an extensive search strategy and a meticulous screening process conducted through Covidence, the evidence base remains scarce. Only three eligible studies met the inclusion criteria, underscoring the novelty and underexplored nature of OLP as a treatment for MDD. Nevertheless, the studies available provide initial insight into the feasibility, safety, and potential efficacy of this approach. A key limitation of the review is relying exclusively on results in the English language. This potentially entails losing studies conducted in other languages.

There was substantial heterogeneity among the included RCTs, particularly with respect to the diagnostic criteria used to define the study population and the assessment of primary outcomes. The varying diagnostic tests employed across these studies have direct implications for the comparability of their findings. The shift in diagnostic criteria from the pilot study by Kelley et al. suggests that subsequent studies may have included a more heterogeneous MDD population. DSM-IV criteria were often considered more stringent for MDD, emphasizing core depressive symptoms. In contrast, DSM-V and ICD-10, include broader presentations such as prolonged grief or anxiety features. This may lead to the inclusion of individuals with less severe or more complex, co-morbid symptom profiles under the definition or spectrum of MDD. Consequently, the baseline patient populations in Schienle et al. (2022) and Nitzan et al. (2020), could on average be less severely ill or more heterogeneous in their symptom presentation compared to Kelley et al. (2012).

Similarly to inclusion criteria which were based on different diagnostic tests, outcome assessment suffered substantial heterogeneity. Kelley et al. (2012) and Schienle et al. (2022) incorporated a mix of clinician-rated and self-report instruments, while Nitzan et al. (2020) relied solely on the self-reported version of the QIDS. Although interviews were conducted at 0, 4 and 8 weeks, the lack of clinician-rated outcome measures may have increased the risk of measurement bias in their study.

While both clinician-rated and self-reported scales are generally sensitive in detecting treatment effects in clinical trials, they capture distinct features of depression. Clinician-based rating scales, such as the HAM-D-17 offer an objective, expert assessment of observable signs, often aligning with a medical model of illness and focusing on criteria relevant for diagnosis and treatment from a professional standpoint. Conversely, self-report scales, including the BDI-II and QIDS-SR16, capture patients' subjective experiences. This encompasses internal states, cognitive distortions, and personal distress that may not be readily apparent to an observer.

Hence, it is important to point out that the choice of primary outcome measure profoundly shapes the narrative of OLP efficacy. Given that OLP effects are hypothesized to involve psychological mechanisms such as expectation and conditioning, the patient's subjective experience, as captured by self-report measures, might serve as a more direct or sensitive indicator of the impact of OLP. However, clinician ratings are often perceived as more robust indicators of clinical significance. For instance, Schienle et al. (2022) observed a statistically significant reduction in BDI-II scores yet concluded that the practical value of OLP could not be demonstrated. Hence, these findings from Schienle et al. (2022) do not suggest robust causal evidence for the efficacy of OLP in decreasing depression symptoms.

Although the analysis by Nitzan et al. did not draw out significant differences between OLP and TAU groups, subgroup analysis on non-geriatric participants with early onset highlights an important variable that could potentially influence a patient's response to OLP interventions and prompts further investigation. The results of the meta-analysis indicate that OLP may offer a modest but consistent reduction in depression severity relative to control conditions. The absence of heterogeneity suggests that the effect was stable across studies. Going forward, the adoption of standardized, validated outcome measures, preferably including both clinician-rated and self-report assessments, would enhance methodological consistency and interpretability across trials.

While the use of RCT designs across all three studies strengthens internal validity by reducing confounding, several limitations warrant caution. The meta-analysis included only two small trials, and all three studies enrolled modest sample sizes (20–60 participants), limiting statistical power and the reliability of heterogeneity estimates, and reducing generalizability. Notably, one was a pilot study focused on feasibility rather than efficacy, further constraining the robustness of the overall findings.

Risk of bias was also a concern, as assessed using

the Cochrane Risk of Bias 2 (RoB2) tool. Two of the three studies, Nitzan et al. (2020) and Schienle et al. (2022), were rated as having a high overall risk of bias, while Kelley et al. (2012) raised some concerns. The most frequent sources of potential bias across studies were related to the randomization process, deviations from intended interventions, missing outcome data, and, most notably, the measurement of outcomes. In two studies, the latter domain was rated as high risk, largely due to reliance on unblinded, self-reported measures of depressive symptoms. While this may appear as a significant methodological limitation, it is important to acknowledge that the unique nature of OLP trials complicates the interpretation of these bias ratings. OLP interventions are inherently unblinded and rely on participants' awareness and engagement with the treatment rationale.

Based on the three available randomized controlled trials, there is currently insufficient evidence to conclude that OLP reduces depressive symptoms in adults with MDD. Importantly, the quantitative synthesis was based on only two small trials, limiting the stability of effect estimates and precluding meaningful assessment of publication bias or between-study heterogeneity. Two of the three included studies were judged to have a high risk of bias, and one was designed primarily as a pilot feasibility trial, further restricting confidence in the pooled findings. Although all studies reported trends toward symptom improvement—particularly when OLP was combined with usual care or behavioral interventions—these findings were derived from small, heterogeneous samples that limit reliability and generalizability. The study populations were predominantly female and included mostly patients with mild to moderate depression, leaving uncertainty about OLP's applicability to men or those with more severe or comorbid conditions. Despite these limitations, OLP remains an ethically transparent, low-risk intervention worth further investigation. Future studies should recruit larger and more diverse samples, standardize OLP delivery, assess both clinician- and self-rated outcomes, and evaluate the durability of effects. Developing assessment frameworks better suited to non-deceptive interventions may also strengthen methodological rigor in this emerging field.

Supplementary Materials

Supplementary Table 1: Comprehensive search strategies by database.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

References

Antonelli, M., & Donelli, D. (2019). Reinterpreting homoeopathy in the light of placebo effects to manage patients who seek homoeopathic care: A systematic review. Health Soc Care Community, 27(4), 824-847. https://doi.org/10.1111/hsc.12681

Braund, T. A., Tillman, G., Palmer, D. M., Gordon, E., Rush, A. J., & Harris, A. W. F. (2021). Antidepressant side effects and their impact on treatment outcome in people with major depressive disorder: an iSPOT-D report. Transl Psychiatry, 11(1), 417. https://doi.org/10.1038/s41398-021-01533-1

Charlesworth, J. E. G., Petkovic, G., Kelley, J. M., Hunter, M., Onakpoya, I., Roberts, N., Miller, F. G., & Howick, J. (2017). Effects of placebos without deception compared with no treatment: A systematic review and meta-analysis. J Evid Based Med, 10(2), 97-107. https://doi.org/10.1111/jebm.12251

Flavio-Reis, V. H. P., Pessoa-Goncalves, Y. M., Diaz, C. A. V., Lamoglia, A. S. A., Desiderio, C. S., & Oliveira, C. J. F. (2025). Open label placebo for chronic low back pain: a systematic review and meta-analysis of randomized controlled trials. Pain Manag, 15(3), 149-160. https://doi.org/10.1080/17581869.2025.2467025

Guevarra, D. A., Webster, C. T., Moros, J. N., Kross, E., & Moser, J. S. (2024). Remotely administered non-deceptive placebos reduce COVID-related stress, anxiety, and depression. Appl Psychol Health Well Being, 16(4), 2204-2224. https://doi.org/10.1111/aphw.12583

Herrman, H., Kieling, C., McGorry, P., Horton, R., Sargent, J., & Patel, V. (2019). Reducing the global burden of depression: a Lancet-World Psychiatric Association Commission. Lancet, 393(10189), e42-e43. https://doi.org/10.1016/S0140-6736(18)32408-5

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2024). Cochrane handbook for systematic reviews of interventions (Version 6.4). Cochrane. https://training.cochrane.org/handbook.

Hoenemeyer, T. W., Baidwan, N. K., Hall, K., Kaptchuk, T. J., Fontaine, K. R., & Mehta, T. S. (2021). An Exploratory Analysis of the Association Between Catechol-O-Methyltransferase and Response to a Randomized Open-Label Placebo Treatment for Cancer-Related Fatigue. Front Psychiatry, 12, 684556. https://doi.org/10.3389/fpsyt.2021.684556

Kellermeyer, L., Harnke, B., & Knight, S. (2018). Covidence and Rayyan. Journal of the Medical Library Association, 106(4), 580-583. https://doi.org/10.5195/JMLA.2018.513

Kelley, J. M., Kaptchuk, T. J., Cusin, C., Lipkin, S., & Fava, M. (2012). Open-label placebo for major depressive disorder: a pilot randomized controlled trial. Psychother Psychosom, 81(5), 312-314. https://doi.org/10.1159/000337053

Krockow, E. M., Emerson, T., Youssef, E., Scott, S., & Tromans, S. (2023). Evidencing general acceptability of open-label placebo use for tackling overtreatment in primary care: a mixed methods study. BMC Med, 21(1), 362. https://doi.org/10.1186/s12916-023-03074-4

Marwaha, S., Palmer, E., Suppes, T., Cons, E., Young, A. H., & Upthegrove, R. (2023). Novel and emerging treatments for major depression. Lancet, 401(10371), 141-153. https://doi.org/10.1016/S0140-6736(22)02080-3

Nitzan, U., Carmeli, G., Chalamish, Y., Braw, Y., Kirsch, I., Shefet, D., Krieger, I., Mendlovic, S., Bloch, Y., & Lichtenberg, P. (2020). Open-Label placebo for the treatment of unipolar depression: Results from a randomized controlled trial. J Affect Disord, 276, 707-710. https://doi.org/10.1016/j.jad.2020.07.077

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Schaefer, M., & Enge, S. (2024). Open-label placebos enhance test performance and reduce anxiety in learner drivers: a randomized controlled trial. Sci Rep, 14(1), 6684. https://doi.org/10.1038/s41598-024-56600-6

Schaefer, M., Hellmann-Regen, J., & Enge, S. (2021). Effects of Open-Label Placebos on State Anxiety and Glucocorticoid Stress Responses. Brain Sci, 11(4). https://doi.org/10.3390/brainsci11040508

Schaefer, M., Kuhnel, A., Schweitzer, F., Enge, S., & Gartner, M. (2023). Neural underpinnings of open-label placebo effects in emotional distress. Neuropsychopharmacology, 48(3), 560-566. https://doi.org/10.1038/s41386-022-01501-3

Schienle, A., & Jurinec, N. (2022). Open-label placebos as adjunctive therapy for patients with depression. Contemp Clin Trials Commun, 28, 100948. https://doi.org/10.1016/j.conctc.2022.100948

Simon, G. E., Moise, N., & Mohr, D. C. (2024). Management of Depression in Adults: A Review. JAMA, 332(2), 141-152. https://doi.org/10.1001/jama.2024.5756

StataCorp. (2025). Stata Statistical Software: Release 19 [Computer software]. StataCorp LLC. Stata.

Sterne, J. A. C., Savovic, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H. Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernan, M. A., Hopewell, S., Hrobjartsson, A., Junqueira, D. R., Juni, P., Kirkham, J. J., Lasserson, T., Li, T., ... Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ, 366, l4898. https://doi.org/10.1136/bmj.l4898

von Wernsdorff, M., Loef, M., Tuschen-Caffier, B., & Schmidt, S. (2021). Effects of open-label placebos in clinical trials: a systematic review and meta-analysis. Sci Rep, 11(1), 3855. https://doi.org/10.1038/s41598-021-83148-6

World Health Organization. (2023). Depressive disorder (depression). WHO. https://www.who.int/news-room/fact-sheets/detail/depression