medicina

MDPI

Systematic Review

Effects of Exercise-Based Interventions on Depressive Symptoms in Adults with Lung Cancer: A Systematic Review and Meta-Analysis of Randomized Controlled Trials

Mesut Süleymanoğulları [1] and Zarife Pancar [2,*]

[1] Department of Physical Education and Sports Teaching, Faculty of Sport Sciences, Ağrı İbrahim Çeçen University, Ağrı 04100, Türkiye; msuleymanogullari@agri.edu.tr
[2] Department of Physical Education and Sports, Faculty of Sports Science, Gaziantep University, Gaziantep 27350, Türkiye
* Correspondence: z_pancar@hotmail.com

Abstract

Background and Objectives: Exercise is increasingly used in supportive care for lung cancer, but its effects on depressive symptoms remain uncertain. This systematic review and meta-analysis synthesized randomized evidence. *Materials and Methods*: Four electronic databases were searched for reports published from January 2000 to July 2026. Eligible studies enrolled adults with lung cancer and compared a structured exercise-based intervention with usual care or an eligible active rehabilitation comparator. One depression result per independent trial was selected using documented rules for assessment time, outcome format, and multiple intervention arms. A random-effects model used REML with Hartung–Knapp inference. *Results*: Seventeen studies met the eligibility criteria; 15 independent trials (1041 participants) contributed to the primary synthesis. The pooled effect favored exercise (Hedges' g = −0.75, 95% CI −1.18 to −0.32; $p = 0.002$), but heterogeneity was considerable ($\tau^2 = 0.4582$; $\tau = 0.6769$; $I^2 = 87.1\%$; $H^2 = 7.76$; $Q(14) = 74.11$, $p < 0.001$) and the 95% prediction interval crossed the null (−2.27 to 0.76). The direction remained favorable in all sensitivity analyses, although exclusion of two influential studies attenuated the pooled effect to −0.49 and reduced I^2 to approximately 0%. Exploratory subgroup and study-level moderator analyses did not identify a stable explanation for heterogeneity. Certainty was very low after downgrading for risk of bias, inconsistency, and indirectness. *Conclusions*: Exercise-based interventions may reduce depressive symptom scores in some adults with lung cancer, but confidence in the magnitude and transferability of the average effect is limited. The evidence does not establish efficacy for major depressive disorder or identify an optimal modality, dose, or delivery format.

Keywords: lung cancer; depression; exercise oncology; psycho-oncology; randomized controlled trials; meta-analysis

Check for updates

Academic Editors: Silvia Varela and Fiorenzo Moscatelli

Received: 25 July 2026
Revised: 7 September 2026
Accepted: 18 September 2026
Published: 19 September 2026

Copyright: © 2026 by the authors. Published by MDPI on behalf of the Lithuanian University of Health Sciences. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Lung cancer was the most frequently diagnosed cancer worldwide in 2022 and remained the leading cause of cancer-related death, accounting for approximately 2.5 million new cases and 1.8 million deaths [1]. For many patients, however, the burden of the disease extends beyond tumor-related symptoms and treatment adverse effects. Psychological distress may arise at diagnosis and persist during treatment, recovery, or advanced disease.

Longitudinal evidence from patients with lung cancer indicates that depressive symptoms may develop or persist after diagnosis and are associated with poorer quality of life, greater symptom burden, lower perceived social support, and shorter survival [2]. Depressive symptoms are also common among adults with cancer, although prevalence estimates vary according to the population studied and the method of assessment [3]. These symptoms rarely have a single cause. Dyspnea, fatigue, pain, loss of physical function, treatment-related adverse effects, social isolation, and lung cancer-related stigma may all contribute [2,4]. Oncology guidelines therefore recommend regular assessment and management according to symptom severity and individual clinical needs [4,5]. Delivering this care can nevertheless be difficult, particularly when depressive and cancer-related somatic symptoms overlap, access to psycho-oncology services is limited, or patients already face a substantial treatment burden.

Exercise is relevant in this context because it may address physical functioning and psychological well-being within the same supportive-care program. International guidelines recommend appropriately prescribed aerobic and resistance exercise during and after cancer treatment when it is clinically suitable [6,7]. Meta-analyses conducted in cancer survivors have reported reductions in depressive symptoms, while evidence from the broader depression literature also supports a potential role for exercise [8–10]. These findings cannot be transferred directly to people with lung cancer. Respiratory limitation, cachexia, comorbidities, performance status, and anticancer treatment may affect both the ability to participate in exercise and the response to an exercise program.

This distinction is important because trials conducted in lung cancer represent neither a single clinical population nor a uniform intervention. Participants have been recruited before and after surgery, during chemotherapy or radiotherapy, with advanced disease, and in survivorship settings. Most trials did not require clinically significant depressive symptoms at enrollment, and depression was commonly assessed as a secondary outcome. The evidence may therefore reflect both improvement in existing symptoms and prevention of symptom-worsening and does not establish treatment efficacy for major depressive disorder. Interventions have ranged from supervised aerobic and resistance training to walking, breathing-based and mind–body exercise, home rehabilitation, and digitally supported multicomponent programs. Comparator conditions have ranged from usual care to education, self-management, and active rehabilitation. Intervention intensity, depression instruments, and assessment time points have also differed. Previous lung cancer-specific reviews have reported potentially favorable psychological outcomes but have consistently described a small and clinically heterogeneous evidence base [11–14]. Placing these studies on a common statistical scale may allow an average effect to be estimated, but it does not make their populations, interventions, comparator conditions, or clinical settings interchangeable.

Against this background, we conducted a systematic review and meta-analysis of randomized controlled trials evaluating structured exercise-based interventions in adults with lung cancer. The primary objective was to estimate the average effect on post-intervention depressive symptoms and quantify variation across trials. We also examined the stability of the pooled estimate and explored whether exercise category, depression instrument, intervention duration, and selected study-level participant characteristics contributed to between-study heterogeneity. Because the evidence base was small and clinically diverse, moderator analyses were treated as exploratory and hypothesis-generating.

2. Materials and Methods

2.1. Registration

This systematic review and meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [15], and the study selection process is shown in Figure 1. As this study was a systematic review and meta-analysis based exclusively on data from previously published studies, no new human participants were recruited, no interventions were administered by the authors, and no identifiable individual-level data were collected. Therefore, institutional ethics committee approval and informed consent were not required for the present study. Ethical approval and informed consent procedures for the individual studies included in this review were the responsibility of the respective original investigators and institutions.

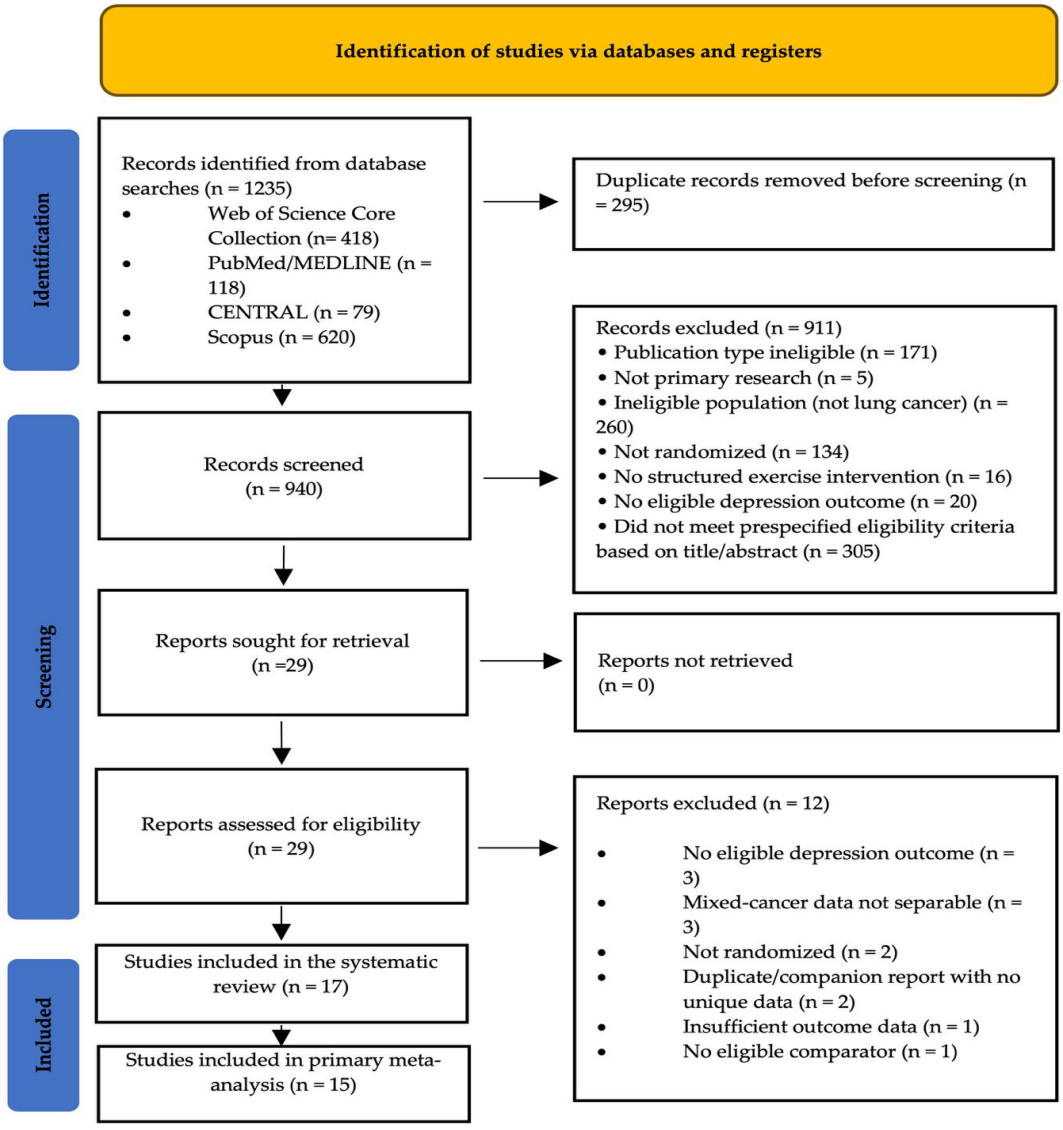


Figure 1. PRISMA 2020 flow diagram of the study selection process.

The review was registered in the Open Science Framework Registries on 17 July 2026 (registration version 1; https://osf.io/vuc87/ (accessed on 17 July 2026)). Database searching, screening, data extraction, risk-of-bias assessment, and statistical analysis had been completed before registration. Because registration was retrospective, the OSF record did not function as a prospective protocol.

Analytical differences between registration version 1 and the final report were limited to analyses that could not be implemented reliably or were added for exploratory purposes. Stage- and treatment-based subgroup analyses were not performed because the available categories were sparse, overlapping, and incompletely reported. Fail-safe N was not reported because it was considered to have limited interpretability in the presence of substantial between-study heterogeneity. Additional study-level moderator analyses were reported as exploratory and hypothesis-generating.

The registration record stated that screening, data extraction, and risk-of-bias assessment were conducted independently by two reviewers (M.S. and Z.P.). Title-and-abstract screening and full-text assessment were conducted independently by the same reviewers, as described in Section 2.4. Data extraction was performed by one reviewer (M.S.), who checked all entries against the source reports. Risk-of-bias assessment was performed by two reviewers (M.S. and Z.P.). Independent duplicate data extraction was not performed. The procedures implemented at each stage of the review are reported in Sections 2.4–2.7.

2.2. Eligibility Criteria

Eligibility was defined using PICOS (Population, Intervention, Comparator, Outcomes, and Study Design). Eligible participants were adults (≥18 years) with a confirmed diagnosis of lung cancer, irrespective of histological subtype, disease stage, treatment phase, or clinical setting. Eligible interventions comprised structured exercise or exercise-based rehabilitation delivered either alone or as a substantive component of a multicomponent rehabilitation program. These included aerobic, resistance, combined, pulmonary rehabilitation, breathing-based, mind–body, walking, digitally delivered, and other structured exercise interventions. Exercise was considered a substantive intervention component when it was scheduled, repeated, and prescribed in terms of content and dose rather than limited to general physical-activity advice. Eligible comparators included usual care; education, attention, or self-management controls; non-exercise control conditions; and active rehabilitation comparators when the intervention arm delivered an additional, more intensive, or substantively different structured exercise or rehabilitation program. Eligible outcomes were depressive symptoms assessed using validated instruments or separate extractable depression subscales.

Eligible studies were parallel-group randomized controlled trials published as full-text, peer-reviewed articles in English between January 2000 and July 2026. Studies were excluded if participants did not have lung cancer, treatment allocation was not randomized, the intervention did not include structured exercise, depression outcomes were unavailable or not separately extractable, or the publication was not an original peer-reviewed research article. Multi-arm randomized trials were eligible and were handled according to the documented data-synthesis procedures.

2.3. Search Strategy

Electronic searches were conducted in PubMed/MEDLINE, Scopus, Web of Science Core Collection, and the Cochrane Central Register of Controlled Trials (CENTRAL) on 12 July 2026. Searches were limited to English-language records published from 1 January 2000 to 12 July 2026. The searches retrieved 118 records from PubMed/MEDLINE, 620 from Scopus, 418 from Web of Science Core Collection, and 79 from CENTRAL, resulting in a total of 1235 records before deduplication. Supplementary Materials Table S1 provides the complete database-specific search strategies used, including controlled vocabulary, field restrictions, applied limits, search dates, and the number of records retrieved from each database before deduplication.

2.4. Study Selection

Database records were standardized and deduplicated using DOI, PMID, and compatible bibliographic fields. Two reviewers (M.S. and Z.P.) independently screened titles and abstracts against the eligibility criteria. Reports considered potentially eligible by either reviewer proceeded to full-text assessment. The same reviewers independently assessed the full-text reports, recorded one principal reason for each exclusion, and resolved disagreements through discussion and consensus. Multiple reports originating from the same randomized trial were linked using a unique study identifier to ensure that each trial contributed only one independent primary dataset to the quantitative synthesis.

2.5. Data Extraction

One reviewer (M.S.) extracted study characteristics and outcome data into structured forms and checked all entries against the corresponding source reports during construction of the analysis-ready dataset. A completed independent duplicate data-extraction procedure was not performed. Randomized, baseline-demographic, analyzed, and depression-outcome populations were recorded separately. Extracted fields included participant characteristics, intervention and comparator characteristics, depression instrument, assessment time, analysis population, summary statistics, attrition, adherence, and funding information when reported.

All eligible arm-level depression results were retained. Published endpoint means and SDs were preferred, followed by validly derived endpoints, adjusted endpoints with compatible SE-to-SD conversion, and change scores when no compatible endpoint was available. Median conversions used the Luo mean estimator and the Shi or Wan SD estimator, as appropriate to the reported data [16–18]. SD was calculated as SE $\times \sqrt{N}$ only when the corresponding group-specific sample size was available. One independent result was then selected from each trial. Compatible eligible intervention arms sharing one control group were combined using Cochrane continuous-outcome formulas, with the control group included only once; otherwise, the contrast that most directly isolated the additional structured exercise component was selected [19].

2.6. Outcome Definition

The selected outcome was a validated depression scale or separately extractable depression subscale. No quantitative trial reported more than one competing eligible depression instrument, so no instrument-priority choice was required. The assessment closest to intervention completion was selected, preferring immediate post-intervention data. Published unadjusted endpoints were preferred; valid derived or adjusted endpoints followed, and change scores were used only when no compatible endpoint was available. Adjusted and unadjusted estimates from the same participants were not entered together. Longer follow-up results were retained separately. Higher scores consistently indicated more severe depressive symptoms across all included instruments; therefore, negative standardized effect sizes favored exercise. Longer-term follow-up outcomes were retained for qualitative reporting but were not included in the primary meta-analysis. Studies reporting only the total Hospital Anxiety and Depression Scale (HADS) score, without a separately extractable Hospital Anxiety and Depression Scale–Depression (HADS-D) subscale, were excluded from the quantitative synthesis but retained in the qualitative review.

2.7. Risk of Bias and Certainty of Evidence

Risk of bias for the selected depression result was assessed by two reviewers (M.S. and Z.P.) using the Cochrane RoB 2 tool [20]. Disagreements were resolved through discussion and consensus. Signaling-question responses, supporting evidence, domain-level

judgments, and overall judgments were recorded for each included study. GRADE was applied to the primary outcome, with each downgrade linked to the corresponding body of evidence [21].

2.8. Statistical Analysis

All analyses were conducted in R (version 4.4.1; R Foundation for Statistical Computing, Vienna, Austria) using the metafor package [22]. For each independent comparison, Hedges' g and its sampling variance were calculated with the intervention group entered first and the comparator group second. Because higher scores indicated more severe depressive symptoms, negative effect estimates favored the intervention. Unusually large effect estimates were checked against the source data, and no study was excluded solely because of effect magnitude. A random-effects model was used because the true intervention effects were expected to vary across studies. The primary meta-analysis used restricted maximum likelihood (REML) estimation with Hartung–Knapp inference [23]. Statistical significance was defined using a two-sided α of 0.05. The pooled effect estimate, 95% confidence interval, 95% prediction interval, τ^2, τ, I^2, H^2, and Cochran's Q statistic were reported. Following Higgins et al. [24], I^2 was interpreted as an index of between-study inconsistency rather than methodological quality. Robustness analyses examined the exclusion of the high-risk study, influential studies, studies requiring converted or summary statistics, studies requiring the combination of multiple intervention arms, and studies without directly reported endpoint means, SDs, and group sizes. Leave-one-out analyses and alternative random-effects model specifications were examined separately. Subgroup analyses and univariable study-level meta-regressions were exploratory, were not adjusted for multiple testing, and were interpreted as hypothesis-generating. Small-study effects were assessed using funnel-plot inspection, Egger's regression test, Begg's rank-correlation test, and trim-and-fill analysis; these findings were interpreted cautiously because the number of trials was small and between-study heterogeneity was substantial [25–27]. An additional sensitivity analysis restricted the synthesis to exercise-dominant comparisons. Interventions were classified as exercise-dominant when their principal therapeutic content comprised scheduled, repeated, and dose-specified aerobic, resistance, walking, breathing, mind–body, or pulmonary rehabilitation exercise; digital delivery or monitoring alone did not result in exclusion. Programs in which education, psychological support, nutrition, or general self-management constituted coequal therapeutic components were exclude.

3. Results

3.1. Study Selection Results

Twenty-nine full-text reports were assessed for eligibility. Seventeen unique randomized studies met the eligibility criteria, whereas 12 reports were excluded (Figure 1). Fifteen independent trials contributed one depression effect each to the primary quantitative synthesis. Liu et al. (2024) [28] was retained in the qualitative synthesis because the reported distribution could not be converted using the applicable conversion methods. Tenconi et al. [29] was retained in the qualitative synthesis because only the total HADS score, without a separately extractable HADS-D subscale, was reported.

3.2. Study Characteristics

The 17 eligible studies evaluated conventional exercise or pulmonary rehabilitation, mind–body or breathing-based exercise, and digital, behavioral, or multicomponent exercise rehabilitation across perioperative, active-treatment, advanced-disease, and survivorship settings (Table 1). Table 1 distinguishes randomized, baseline-demographic, analyzed, and selected depression-analysis populations because these denominators differed in sev-

eral reports. The 15 quantitative trials contributed 543 intervention and 498 comparator participants (total N = 1041). Cheung et al. [30] and Wang et al. [31] were multi-arm trials. The loaded-deep-breathing arm was compared with routine care, and the PERMA-containing arm was excluded from this analysis. For quantitative synthesis, eligible intervention arms were combined within each study and the shared comparator group was counted only once, thereby retaining one statistically independent primary comparison per trial. Comparator conditions ranged from usual care, education or attention control, and self-management to active rehabilitation. In trials using pulmonary rehabilitation or chest physiotherapy as active comparators, the effect estimates represent comparative or incremental effects relative to active rehabilitation rather than exercise versus no exercise.

3.3. Risk of Bias

All 17 studies underwent result-specific RoB 2 assessment (Figure 2). Concerns most often involved self-reported depression in an unblinded comparison, missing outcome data, or selection of the reported result.

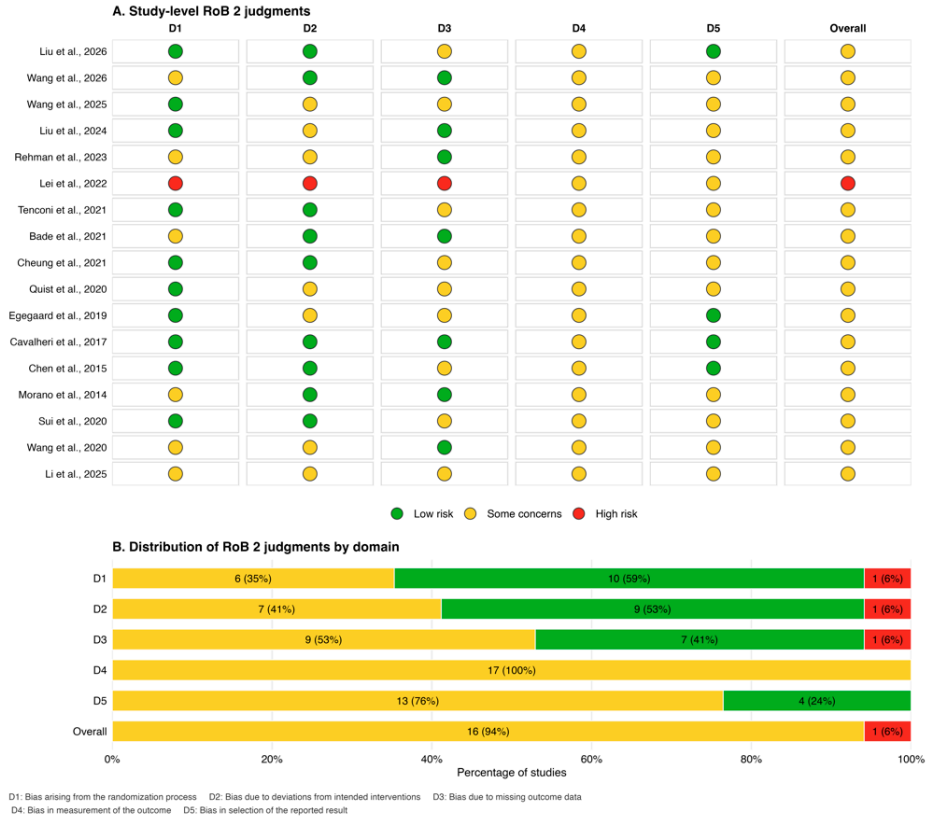


Figure 2. Risk-of-bias assessment using the Cochrane RoB 2 tool [28–44].

Table 1. Characteristics of studies included in the systematic review.

Study	Sample Size	Age, Years	Female, *n*/N (%)	Cancer Type	Stage	Clinical Phase	Intervention	Comparator	Exercise Category	Duration	Depression Measure/Role	Assessment/Synthesis
Liu et al., 2026 [32]	Randomized 104 (52/52); depression results 77 (38/39)	IG: 60.23 ± 10.56; CG: 61.17 ± 9.57	IG: 26/52 (50.0%); CG: 28/52 (53.8%)	Lung cancer (NSCLC/SCLC)	III–IV	Advanced disease; active or oncological treatment	Hybrid home-based pulmonary rehabilitation	Usual care/education attention control with no prescribed structured exercise	Multicomponent exercise rehabilitation	8 weeks	HADS-D/primary	Week 8; quantitative
Wang et al., 2026 [33]	Randomized 57 (27/30); depression result 50 (25/25)	IG: 54.40 ± 10.87; CG: 58.40 ± 10.44	IG: 14/25 (56.0%); CG: 12/25 (48.0%)	Postoperative NSCLC	Not reported	Postoperative	HIM-guided rehabilitation	Standard care	Digital or telerehabilitation	4 weeks	SDS/secondary	1 month after discharge; quantitative
Wang et al., 2025 [31]	Randomized 102 (34/34/34); depression result 96 (65/31)	IG: 53.32 ± 9.29; CG: 57.84 ± 10.01	IG: 43/65 (66.2%); CG: 25/31 (80.6%)	NSCLC undergoing thoracoscopic lobectomy	I–III (predominantly stage I)	Perioperative	PERMA plus loaded deep breathing; Loaded deep breathing	Routine care	Breathing or respiratory muscle exercise	Admission–discharge (variable)	HADS-D/primary	Discharge; quantitative
Liu et al., 2024 [28]	Randomized 102 (34/34/34); analyzed 99 (33/34/32); narrative synthesis	CIG: 61.06 ± 8.30; MBG: 61.50 ± 9.09; CG: 60.63 ± 8.28	CIG: 18/33 (54.5%); MBG: 16/34 (47.1%); CG: 16/32 (50.0%)	NSCLC undergoing surgery	Carcinoma in situ–III	Perioperative	Mindful breathing plus diary guidance; Mindful breathing	Routine care	Mind–body exercise	Admission–discharge (variable)	HADS-D/secondary	Discharge; narrative synthesis (median and IQR reported)
Rehman et al., 2023 [34]	Randomized 40 (20/20); depression result 40 (20/20)	IG: 48.1 ± 4.0; CG: 48.3 ± 3.8	IG: 9/20 (45.0%); CG: 7/20 (35.0%)	NSCLC	I–II	During chemotherapy	Pulmonary rehabilitation plus aerobic training	Pulmonary rehabilitation	Aerobic exercise added to pulmonary rehabilitation	4	HADS-D/NR	Post-test; quantitative
Lei et al., 2022 [35]	Randomized: 52 (26/26); depression result 37 (19/18)	IG: 56.04 ± 11.67; CG: 58.03 ± 7.71	IG: 9/26 (34.6%); CG: 10/26 (38.5%)	Lung cancer/NSCLC	I–IIIB	Definitive treatment completed more than 1 week previously	Baduanjin plus elastic-band training	Control	Qigong or Baduanjin	8	SDS/primary	Post-intervention; quantitative
Tenconi et al., 2021 [29]	Randomized 140 (70/70); completed 6-month assessment 85 (45/40); narrative synthesis	IG: 66.00 ± 10.61; CG: 67.74 ± 10.84	IG 32/70 (45.7%); CG 22/70 (31.4%)	Clinical stage I–II NSCLC scheduled for resection	Clinical stage I–II	Perioperative	Standard care plus perioperative pulmonary rehabilitation	Standard care	Pulmonary rehabilitation	14–21 days preoperatively plus 8 weeks postoperatively	Total HADS/secondary	6 months; narrative synthesis (HADS-D not separately reported)
Bade et al., 2021 [36]	Randomized 40 (20/20)	IG: 66.55 ± 7.28; CG: 63.20 ± 9.80	IG: 12/20 (60.0%); CG: 18/20 (90.0%)	Stage III/IV NSCLC	IIIA–IV	Active treatment (85%) or post-treatment (15%)	Home-based physical activity	Usual care	Walking	12	PHQ-9/not explicitly designated	Week 12; quantitative (mixed-effects model estimates)

Table 1. *Cont.*

Study	Sample Size	Age, Years	Female, *n*/N (%)	Cancer Type	Stage	Clinical Phase	Intervention	Comparator	Exercise Category	Duration	Depression Measure/Role	Assessment/Synthesis
Cheung et al., 2021 [30]	Randomized 30 (19/11)	IG: 61.05 ± 9.76; CG: 58.36 ± 9.32	IG: 8/19 (42.1%); CG: 6/11 (54.5%)	Stage IIIB–IV NSCLC	IIIB–IV	Advanced disease; mixed treatment status	Aerobic exercise; Tai Chi	Self-management control with written exercise guidelines	Aerobic/strengthening exercise and Tai Chi	12	HADS-D/secondary	Post-intervention; quantitative
Quist et al., 2020 [37]	Randomized 216 (108/108)	IG: 65.2 ± 8.2; CG: 63.5 ± 8.7	IG: 55/108 (50.0%); CG: 56/108 (51.9%)	Advanced inoperable lung cancer (NSCLC/SCLC)	Advanced/inoperable	During chemotherapy	Supervised group exercise	Usual care	Combined aerobic and resistance	12	HADS-D/secondary	12 weeks; quantitative (multiple-imputation analysis)
Sui et al., 2020 [38]	Randomized 200 (100/100)	IG: 61.37 ± 11.21; CG: 62.35 ± 9.98	IG: 20/100 (20.0%); CG: 16/100 (16.0%)	NSCLC after surgical resection	I–III	Postoperative; intervention initiated 4–8 weeks after surgery	WeChat education and rehabilitation program	One-time education and rehabilitation guidance plus usual care	Digital or telerehabilitation	52	HADS-D/NR	12 months; quantitative
Wang et al., 2020 [39]	Randomized/result 65 (31/34)	IG: 59 (52–62); CG: 55.5 (46.75–63.25), median (IQR)	IG: 20 (64.5%); CG: 23 (67.6%)	NSCLC receiving surgery	Not reported	Perioperative	Breathing exercises	Routine care	Breathing or respiratory muscle exercise	NR	HADS-D/secondary	Discharge; quantitative
Egegaard et al., 2019 [40]	Randomized 15 (8/7); depression result 12 (7/5)	IG: 64 ± 5.8; CG: 65 ± 4.7	IG: 5/8 (62.5%); CG: 5/7 (71.4%)	Locally advanced NSCLC	locally advanced	During chemoradiotherapy	Aerobic interval training before radiotherapy	Control	Aerobic	7	HADS-D/secondary	7 weeks; quantitative
Cavalheri et al., 2017 [41]	Randomized 17 (9/8)	IG: 66 ± 10; CG: 68 ± 9	IG: 6/9 (66.7%); CG: 6/8 (75.0%)	NSCLC after curative-intent treatment	I–IIIA	6–10 weeks after lobectomy or 4–8 weeks after adjuvant chemotherapy	Supervised combined aerobic and resistance training	Usual activities plus weekly attention-control telephone calls	Combined aerobic and resistance	8	HADS-D/secondary	8 weeks; quantitative
Chen et al., 2015 [42]	Randomized 116 (58/58); depression result 101 (50/51)	IG: 64.76 ± 11.28; CG: 63.57 ± 10.54	IG: 32/58 (55.2%); CG: 30/58 (51.7%)	Lung cancer	I–IV; unknown in 12 participants	Mixed treatment statuses	Home-based walking	Usual care	Walking	12	HADS-D/primary	3 months; quantitative

Table 1. *Cont.*

Study	Sample Size	Age, Years	Female, *n*/N (%)	Cancer Type	Stage	Clinical Phase	Intervention	Comparator	Exercise Category	Duration	Depression Measure/Role	Assessment/Synthesis
Morano et al., 2014 [43]	Randomized 24 (12/12)	IG: 65 ± 8; CG: 69 ± 7	IG: 8/12 (66.7%); CG: 7/12(58.3%)	NSCLC awaiting resection with pre-existing inflammatory pulmonary disease	I–IIIA	Preoperative	Pulmonary rehabilitation	Chest physical therapy	Pulmonary rehabilitation	4	HADS-D/NR	1 month; quantitative
Li et al., 2025 [44]	Randomized 47 (25/22); depression result 36 (20/16)	IG: 49.50 ± 11.66; CG: 53.50 ± 10.71	IG: 19/22 (86.4%); CG: 12/18 (66.7%)	Early-stage NSCLC survivors	early stage	Survivorship	Digital cardiac telerehabilitation	Usual care	Digital or telerehabilitation	20	HADS-D/secondary	5 months; quantitative

Notes: Values are presented as mean ± SD unless otherwise indicated. Duration is reported in weeks unless otherwise specified. Sample-size entries distinguish randomized, baseline-demographic, analyzed, outcome-specific, and synthesis populations, as applicable. Female participants are reported as *n*/N (%) using the denominator reported for baseline characteristics, which may differ from the corresponding depression-analysis population. For Wang et al. (2025) and Cheung et al. (2021) [30,31], eligible active intervention arms were pooled for quantitative synthesis; pooled age, SD, and female counts are presented. For Wang et al. (2026) and Li et al. (2025) [33,44], baseline age and sex values refer to study completers included in the published baseline tables. Rehman et al. (2023) and Morano et al. (2014) [34,43] used active rehabilitation comparators; their effect estimates therefore represent comparative or incremental effects relative to active rehabilitation rather than exercise versus no exercise. Abbreviations: CG, comparator group; CIG, combined intervention group; HADS, Hospital Anxiety and Depression Scale; HADS-D, Hospital Anxiety and Depression Scale–Depression subscale; HIM, Hospital–Internet–Mobile; IG, intervention group; IQR, interquartile range; MBG, mindful breathing group; NR, not reported; NSCLC, non-small cell lung cancer; PERMA, Positive Emotion, Engagement, Relationships, Meaning, and Accomplishment; PHQ-9, Patient Health Questionnaire-9; SCLC, small cell lung cancer; SD, standard deviation; SDS, Self-Rating Depression Scale.

3.4. Primary Meta-Analysis

The primary synthesis included 15 independent trials and 1041 participants (Figure 3). Negative values favored exercise. REML with Hartung–Knapp inference yielded Hedges' g = −0.7518 (95% CI −1.1815 to −0.3221; t(14) = −3.7525; p = 0.00214). Heterogeneity was substantial: τ^2 = 0.4582, τ = 0.6769, I^2 = 87.12%, H^2 = 7.7619, and Q(14) = 74.1081 ($p = 3.45 \times 10^{-10}$). The 95% prediction interval was −2.2658 to 0.7623. Influence diagnostics identified Rehman et al. [34] as influential and Wang et al. [31] as potentially influential.

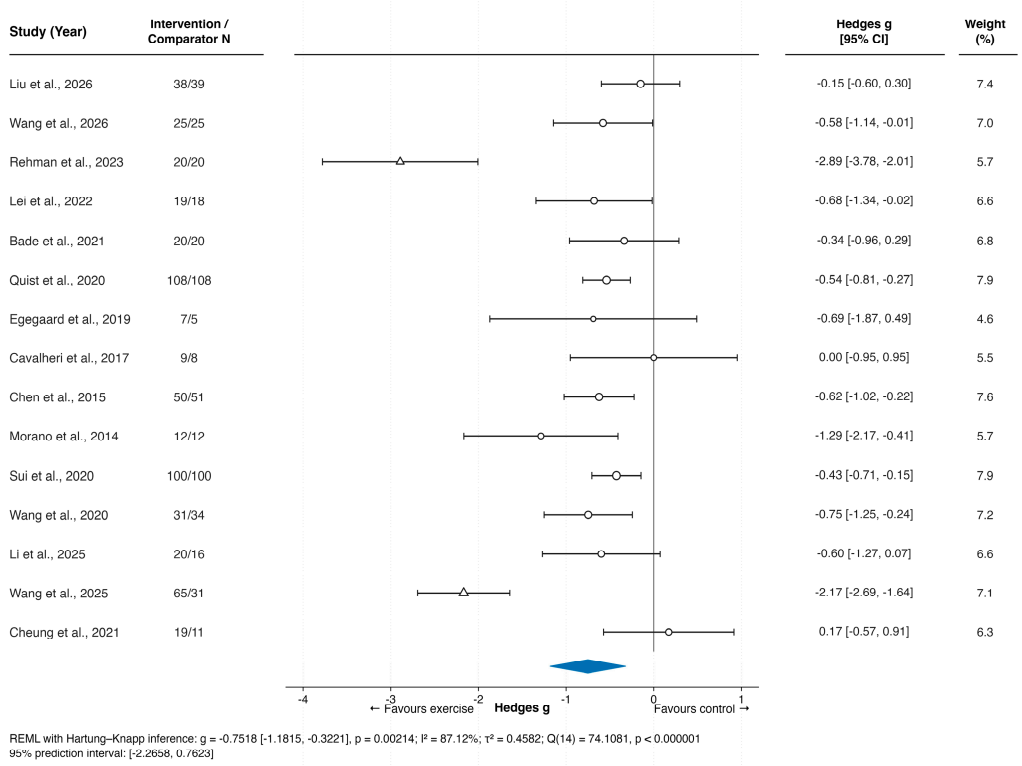


Figure 3. Primary random-effects meta-analysis [30–44]. Notes: Triangles indicate studies identified as influential in the influence diagnostics; circles indicate the remaining studies.

3.5. Sensitivity Analyses

Sensitivity analyses retained the same effect-size coding and REML/Hartung–Knapp method used in the primary analysis (Figure 4). Excluding Lei et al. [35], the study judged to be at high risk of bias, yielded g = −0.7584 (95% CI, −1.2244 to −0.2923; I^2 = 88.48%; k = 14). Excluding Rehman et al. [34] yielded g = −0.6235 (95% CI, −0.9507 to −0.2964; I^2 = 77.95%), whereas excluding Wang et al. [31] yielded g = −0.6291 (95% CI, −1.0115 to −0.2468; I^2 = 79.14%). Joint exclusion of Rehman et al. [34] and Wang et al. [31] attenuated the pooled effect to g = −0.4900 (95% CI, −0.6380 to −0.3419) and reduced heterogeneity to approximately zero ($I^2 \approx 0\%$; k = 13). Source-data verification for both influential trials confirmed the extracted group sizes, HADS-D scale units, and baseline and post-intervention summary statistics; no transcription or unit errors or other documented reasons for exclusion were identified. Both trials were therefore retained in the primary synthesis and examined in influence and sensitivity analyses. The analysis restricted to trials with directly reported endpoint means, SDs, and group sizes included 10 trials and yielded g = −0.7520 (95% CI, −1.3042 to −0.1997; I^2 = 83.70%). Other conversion- and derivation-based sensitivity analyses are presented in Figure 4.

Excluding the two median-converted trials yielded g = −0.6238 (95% CI, −1.0471 to −0.2006; I^2 = 81.78%; k = 13), whereas excluding trials with SE-derived SDs yielded

g = −0.8514 (95% CI, −1.3230 to −0.3797; I^2 = 88.03%; k = 13). Exclusion of the trial requiring the documented equal-allocation group-size rule yielded g = −0.7716 (95% CI, −1.2393 to −0.3038; I^2 = 86.34%; k = 14). Thus, the favorable average direction was retained across data-transformation sensitivity analyses, although substantial heterogeneity generally persisted.

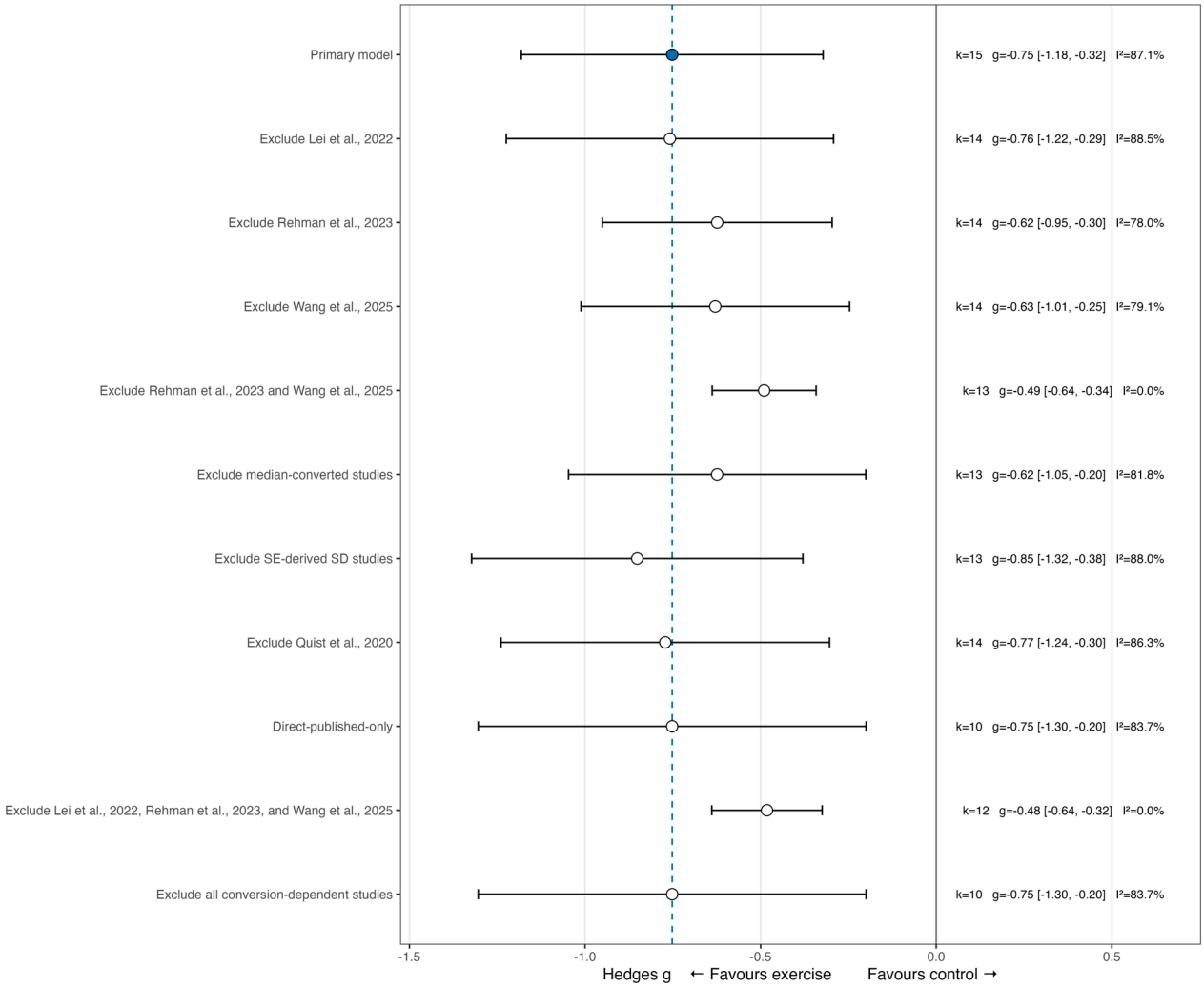


Figure 4. Sensitivity analyses of the primary pooled effect [31,34,35,37]. Notes: Circles represent pooled Hedges' g estimates and horizontal lines represent 95% confidence intervals. The blue dashed line indicates the primary pooled estimate. Abbreviations: CI, confidence interval; *k*, number of trials; I^2, percentage of variability attributable to between-study inconsistency; SD, standard deviation; SE, standard error.

In the exercise-dominant sensitivity analysis, the HIM program [33] and WeChat program [38] were excluded, and the loaded-deep-breathing-only arm from Wang et al. [31] was compared with routine care. Thirteen independent comparisons involving 757 participants yielded g = −0.7703 (95% CI −1.2525 to −0.2880; p = 0.0045; τ^2 = 0.4766; I^2 = 85.35%). The 95% prediction interval crossed the null (−2.3498 to 0.8093). Restricting the synthesis to exercise-dominant comparisons therefore retained a favorable average direction but did not resolve heterogeneity or uncertainty in transferability.

Leave-one-out analyses yielded pooled estimates ranging from g = −0.8129 to −0.6235 (Figure 5). Omission of any single trial changed neither the favorable direction nor the statistical significance of the pooled estimate. Nevertheless, stability of direction should not be interpreted as stability of effect magnitude or heterogeneity. In particular,

the joint exclusion of Rehman et al. [34] and Wang et al. [31] attenuated the pooled effect from g = −0.75 to −0.49 and reduced I^2 from 87.12% to approximately 0%. The magnitude and heterogeneity of the primary estimate were therefore sensitive to a small number of clinically and analytically distinctive trials.

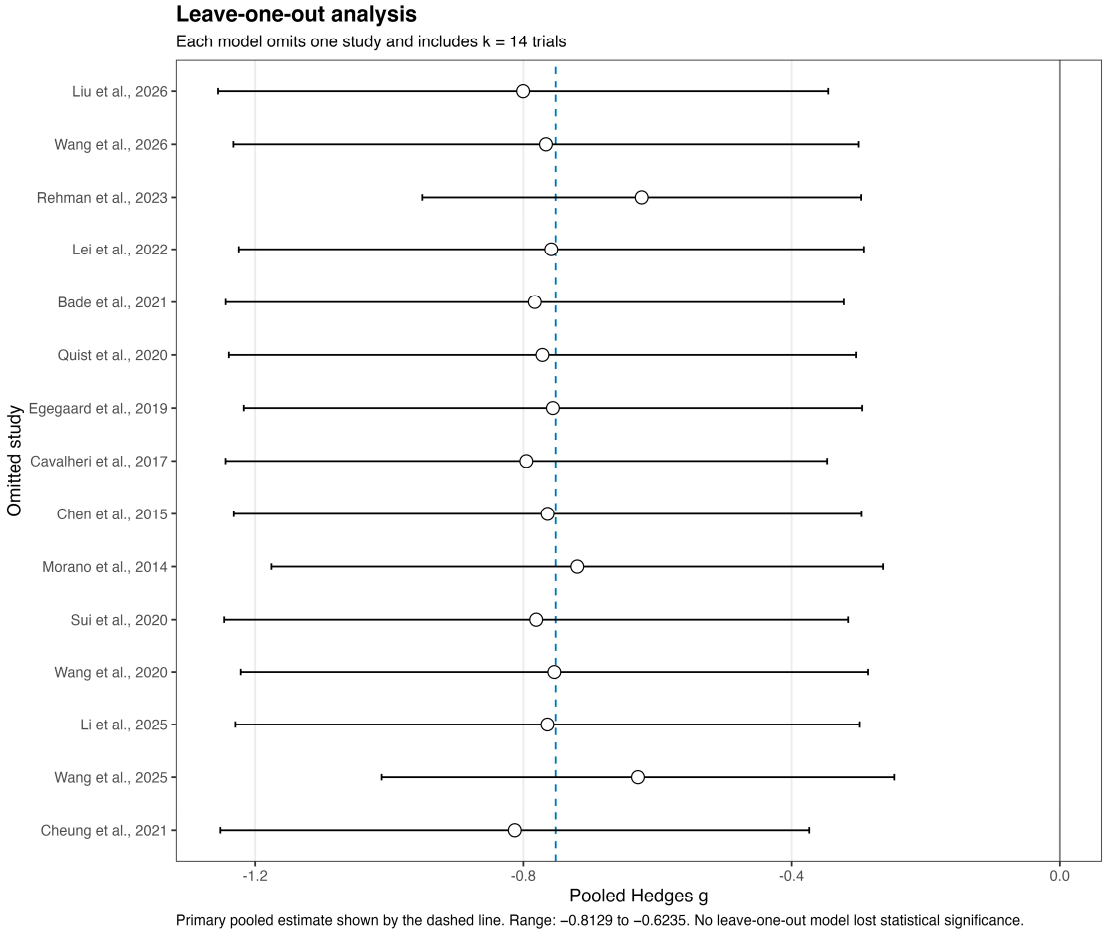


Figure 5. Leave-one-out sensitivity analysis [30–44]. Notes: Each row presents the pooled Hedges' g estimate after omission of the indicated study; all models included 14 trials. Circles represent pooled estimates and horizontal lines represent 95% confidence intervals. The blue dashed line indicates the primary pooled estimate. Negative values favor exercise. No leave-one-out model changed the favorable direction or statistical significance of the pooled estimate. Abbreviations: CI, confidence interval; k, number of trials.

3.6. Exercise Subgroup Analysis

The exploratory exercise-category subgroup analysis included 14 trials. Cheung et al. [30] was excluded from this analysis because its combined intervention group comprised both aerobic exercise and Tai Chi and therefore could not be assigned exclusively to one exercise category; the study remained included in the primary meta-analysis (Figure 6). Seven trials classified as conventional exercise or pulmonary rehabilitation [34,36,37,40–43] yielded a pooled Hedges' g of −0.8866 (95% CI, −1.7498 to −0.0234; $p = 0.0457$; $\tau^2 = 0.6807$; $I^2 = 88.70\%$). Three trials classified as mind–body or breathing-based exercise [31,35,39] yielded g = −1.2074 (95% CI, −3.3008 to 0.8860; $p = 0.1312$; $\tau^2 = 0.6318$; $I^2 = 88.58\%$). Four trials classified as digital, behavioral, or multicomponent rehabilitation [32,33,38,44] yielded g = −0.4034 (95% CI, −0.6759 to −0.1309; $p = 0.0181$; $\tau^2 = 0$; $I^2 = 0.00\%$). The omnibus test showed no statistically significant differences between exercise categories (QM (2) = 0.9990; $p = 0.3994$). Exercise category explained little of the estimated between-

study heterogeneity (pseudo-R^2 = 2.91%; residual τ^2 = 0.4257; residual I^2 = 85.20%). These exploratory between-study comparisons do not provide evidence that any exercise category is superior to another.

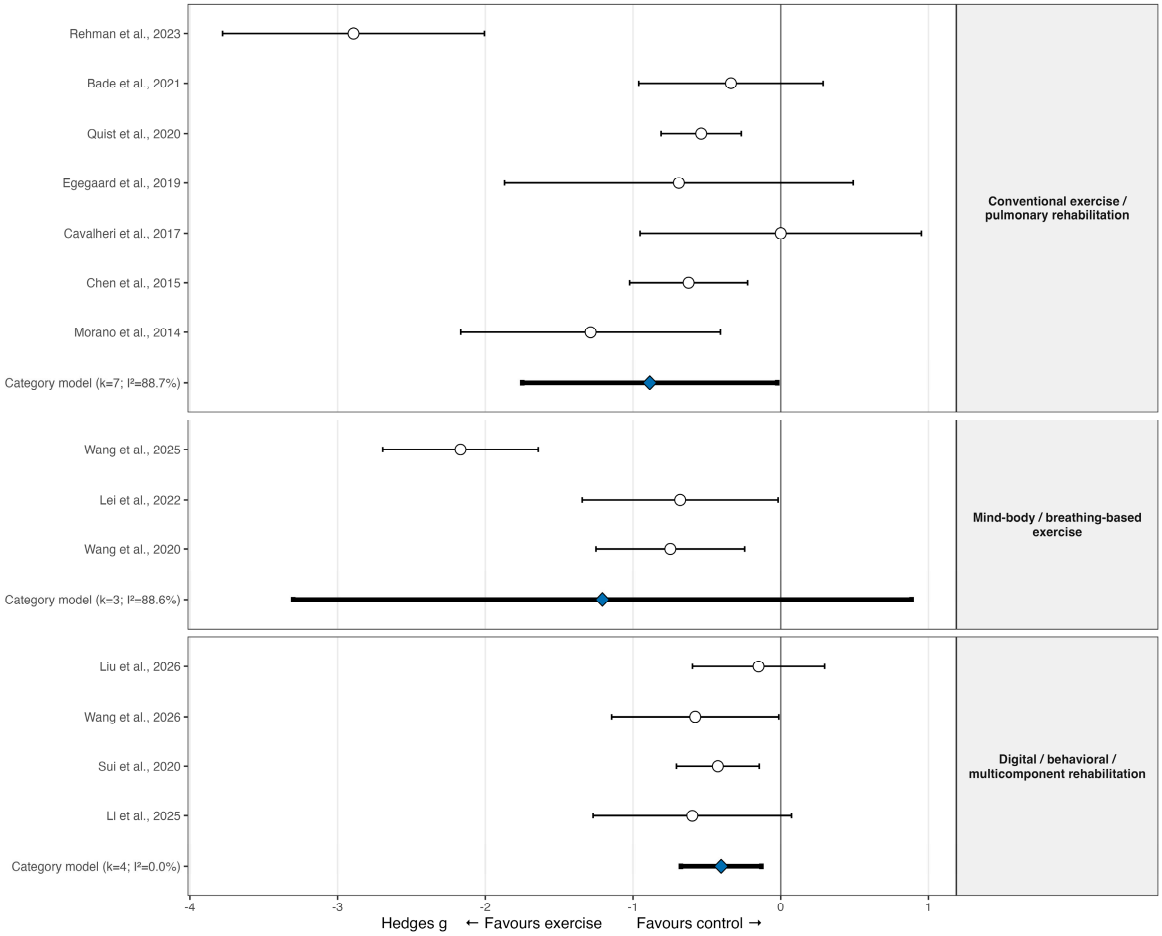


Figure 6. Exploratory exercise-category subgroup analysis [31–44]. Notes: Circles represent study-specific Hedges' g estimates and horizontal lines represent 95% confidence intervals. Diamonds represent pooled category estimates. Negative values favor exercise. Cheung et al. [30] remained in the primary meta-analysis but was excluded from this subgroup analysis because its combined intervention group included both aerobic exercise and Tai Chi. Abbreviations: CI, confidence interval; k, number of trials; I^2, percentage of variability attributable to between-study inconsistency.

3.7. Depression-Instrument Subgroup Analysis

The exploratory depression-instrument subgroup analysis (Figure 7) included all 15 independent trials. Twelve trials used the Hospital Anxiety and Depression Scale–Depression subscale (HADS-D), two used the Self-Rating Depression Scale (SDS), and one used the Patient Health Questionnaire-9 (PHQ-9). The pooled effect estimate for the HADS-D subgroup was g = −0.8129 (95% CI, −1.3665 to −0.2594; p = 0.00798; k = 12), whereas the estimate for the SDS subgroup was g = −0.6219 (95% CI, −1.2619 to 0.0181; p = 0.05145; k = 2). The single PHQ-9 trial yielded g = −0.3371 (95% CI, −0.9613 to 0.2871; p = 0.28978) and was presented descriptively rather than as a pooled subgroup estimate. The omnibus test showed no statistically significant differences between depression instruments (QM(2) = 0.1796; p = 0.8378). Residual heterogeneity remained substantial (τ^2 = 0.5461; I^2 = 89.29%), and the instrument moderator explained none of the estimated between-study heterogeneity (pseudo-R^2 = 0.00%). Although the HADS-D subgroup estimate excluded the null, the nonsignificant omnibus test provides no evidence that intervention effects differed according to the depression instrument used. Because the SDS

subgroup contained only two trials and PHQ-9 was represented by a single trial, these instrument-specific findings should be interpreted cautiously.

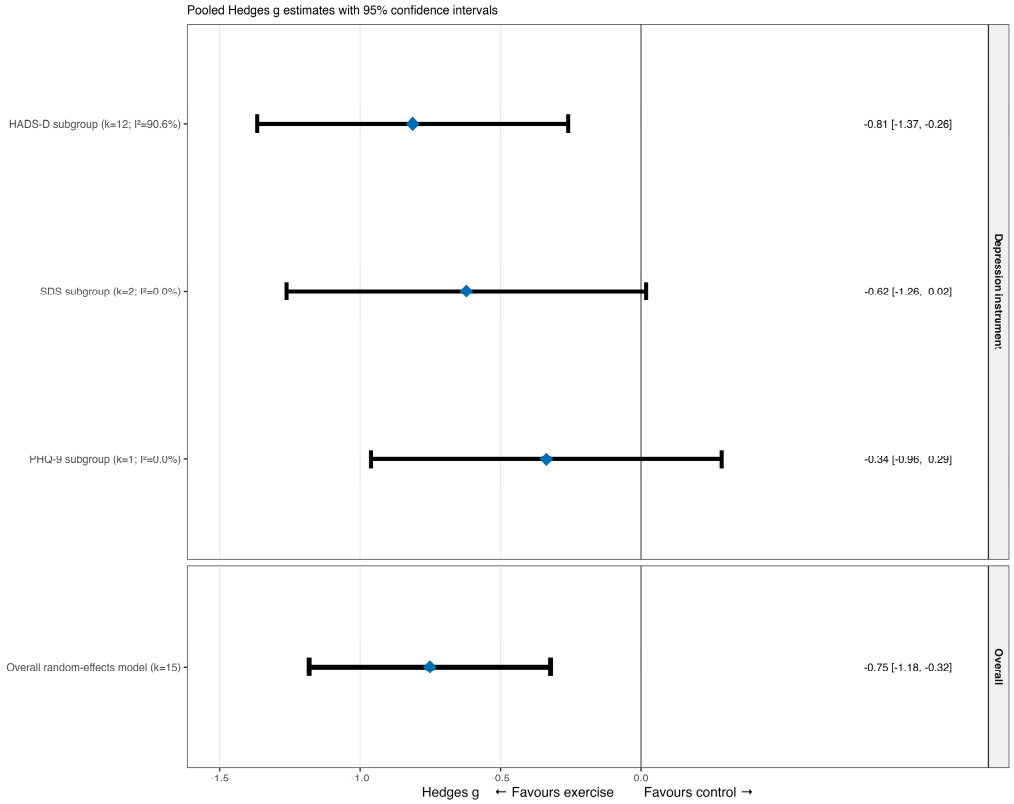


Figure 7. Exploratory subgroup analysis according to depression instrument. Notes: Diamonds represent subgroup and overall estimates with 95% confidence intervals. Negative Hedges' g values favor exercise. The PHQ-9 result represents a single-study descriptive estimate. Instrument-specific estimates should be interpreted cautiously because the SDS and PHQ-9 categories contained few trials.

3.8. Small-Study-Effect Diagnostics

Small-study-effect diagnostics included all 15 trials contributing to the primary synthesis (Figure 8). The Egger-type regression did not indicate statistically significant funnel-plot asymmetry (coefficient = −1.1540, SE = 1.6925; t(13) = −0.6818; $p = 0.5073$), and the Begg–Mazumdar rank-correlation test was also nonsignificant (Kendall's $\tau = -0.1619$; $p = 0.4351$) [25–27]. Exploratory trim-and-fill analysis imputed no studies, and the pooled estimate remained unchanged (Hedges' g = −0.75, 95% CI −1.18 to −0.32). The Egger and rank-correlation tests remained nonsignificant after excluding Rehman et al. [34], Wang et al. [31], or both trials. These diagnostics did not provide clear statistical evidence of small-study effects. Nevertheless, their power and specificity were limited by the small number of trials and substantial between-study heterogeneity; therefore, these findings cannot establish the absence of publication bias.

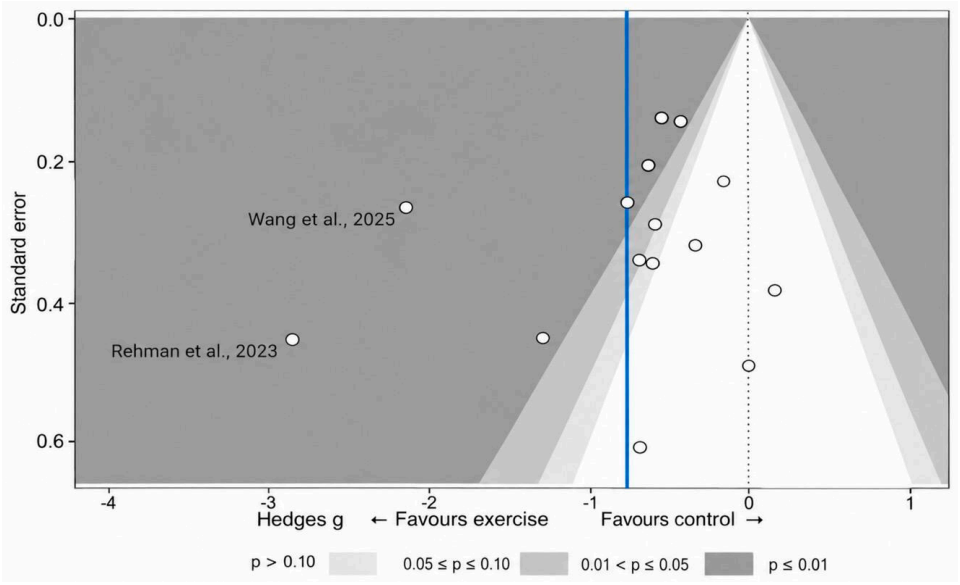


Figure 8. Contour-enhanced funnel plot of the 15 trials included in the primary meta-analysis [31,34]. Notes: Circles represent individual trial estimates plotted against their standard errors. The dotted vertical line indicates the null effect, and the blue vertical line indicates the primary pooled estimate. Shaded contours represent two-sided statistical-significance regions around the null. Negative Hedges' g values favor exercise. This diagnostic does not establish the presence or absence of publication bias. Abbreviation: SE, standard error.

3.9. Meta-Regression

Study-level mean age was available for 14 trials. Each additional year of mean age was associated with a 0.0814-unit increase in Hedges' g, corresponding to a smaller estimated benefit of exercise (95% CI, 0.0074 to 0.1554; $p = 0.0338$). Substantial residual heterogeneity remained ($\tau^2 = 0.3343$; $I^2 = 83.01\%$), and the association was sensitive to Rehman et al. [34]. This univariable ecological association was not adjusted for multiple testing and does not support patient-level treatment selection. Intervention duration was available for 13 trials and was not associated with effect magnitude ($\beta = 0.0113$ per week; 95% CI, −0.0219 to 0.0445; $p = 0.4689$).

Baseline depression means on the selected instrument were available for 12 quantitative trials. None of the 17 eligible studies required elevated depressive symptoms for enrollment. In an exploratory POMP-standardized study-level analysis, each 10-percentage-point increase in baseline depressive symptom burden was associated with a more negative effect estimate ($\beta = -0.409$; 95% CI, −0.633 to −0.186; $p = 0.0022$; k = 12). This association was sensitive to Rehman et al. [34]. Because these analyses were univariable, ecological, and unadjusted for multiple testing, they are presented as hypothesis-generating and do not establish patient-level effect modification.

3.10. GRADE Certainty

GRADE was applied to depressive symptoms assessed at the time point closest to intervention completion (Table 2). Randomized evidence started at high certainty and was downgraded by one level for risk of bias because no trial was judged to be at overall low risk and one was judged to be at high risk. Evidence was downgraded by one level for serious inconsistency because between-study heterogeneity was substantial and the 95% prediction interval crossed the null. Evidence was downgraded once more for serious indirectness because no trial required clinically elevated depressive symptoms for enrollment, depression was commonly a secondary outcome, several interventions were complex

or multicomponent, comparators included both usual care and active rehabilitation, and clinical settings ranged from surgery to advanced disease and survivorship.

Table 2. GRADE certainty of evidence for depressive symptoms.

Certainty Assessment							№ of Patients				
№ of Studies	Study Design	Risk of Bias	Inconsistency	Indirectness	Imprecision	Other Considerations	Exercise	Control	Standardized Effect (95% CI)	Certainty	Importance
15	Randomized controlled trials	Serious [a]	Serious [b]	Serious [c]	Not serious [d]	Publication bias not strongly suspected [e]	543	498	Hedges' g = −0.75 (95% CI −1.18 to −0.32)	⊕◯◯◯ Very low	Important

Notes: [a] Downgraded by one level for risk of bias because no trial was judged to be at overall low risk and one was judged to be at high risk. [b] Downgraded by one level for serious inconsistency because I^2 was 87.1% and the 95% prediction interval crossed the null. [c] Downgraded by one level for serious indirectness because participants were not selected for clinically elevated depression, depression was commonly a secondary outcome, interventions and comparators varied, and settings spanned the care continuum. [d] No separate downgrade for imprecision was applied because the confidence interval excluded both the null and an interpretive small-effect threshold of $|g| = 0.20$, although it permitted materially different magnitudes of benefit. This threshold is not a validated patient-level minimal important difference. [e] Publication bias was not downgraded because Egger's regression, rank correlation, and trim-and-fill did not provide clear evidence of small-study effects; however, these diagnostics had limited power because only 15 heterogeneous trials were available. Abbreviations: CI, confidence interval; GRADE, Grading of Recommendations Assessment, Development and Evaluation; I^2, percentage of variability attributable to between-study inconsistency.

Imprecision was not downgraded separately. No validated cross-instrument minimal important difference was available; therefore, an absolute standardized effect of 0.20 was used as an interpretive small-effect threshold. The Hartung–Knapp confidence interval (g = −1.18 to −0.32) excluded both the null and effects smaller than this threshold, although it encompassed materially different magnitudes of benefit. The threshold should be understood as an interpretive benchmark rather than a validated patient-level minimal important difference. The prediction interval was considered under inconsistency and was not counted again under imprecision. Publication bias was not downgraded, although this judgment remained uncertain because the available diagnostics had limited power with only 15 heterogeneous trials. The overall certainty of evidence was very low.

4. Discussion

4.1. Principal Findings

This review included 17 eligible randomized studies; 15 independent trials involving 1041 participants contributed to the primary synthesis. The average effect favored exercise (g = −0.75, 95% CI −1.18 to −0.32), but the prediction interval crossed the null and the magnitude of the pooled effect was sensitive to two influential trials. Excluding Rehman et al. [34] and Wang et al. [31] attenuated the pooled effect to g = −0.49 and reduced I^2 from 87.1% to approximately 0%. The pooled estimates therefore indicated a favorable average direction but not a uniform or precisely transferable effect. Certainty was very low after downgrading for risk of bias, inconsistency, and indirectness; 16 trials had some concerns and one was judged to be at high risk of bias, with none at overall low risk.

4.2. Evidence Evolution

Evidence concerning exercise and depressive symptoms in adults with lung cancer has evolved from feasibility-focused reviews to broader quantitative syntheses. Earlier reviews suggested possible psychological benefit but emphasized small trial numbers, heterogeneous interventions, and inconsistent reporting of depression outcomes [12,14,45].

Subsequent meta-analyses reported beneficial pooled effects, including a standardized mean difference of −0.60 across nine randomized trials in Lu et al. [46] and an effect of −0.55 for depressive symptoms in the broader psychological-distress synthesis by Tadsuan et al. [47]. In contrast, Hu et al. [48], which focused on advanced lung cancer, did not identify a statistically significant pooled effect on depressive symptoms. Differences in disease stage, treatment phase, eligibility criteria, intervention composition, outcome selection, and analytical methods likely account for part of this variation.

The present review extends earlier syntheses by retaining one independent primary effect per trial and reporting the prediction interval, influence analyses, risk of bias, and certainty alongside the pooled mean. These analyses show that a favorable average estimate coexists with substantial variation in magnitude and transferability. Evidence from other cancer populations and major depressive disorder is supportive but cannot be directly extrapolated to lung cancer [8–10].

4.3. Understanding Heterogeneity

The two influential trials differed from most of the dataset. Rehman et al. [34] enrolled 40 patients receiving chemotherapy and compared pulmonary rehabilitation plus aerobic training with pulmonary rehabilitation alone over four weeks. Baseline HADS-D means were 12.75 and 13.02; directly reported post-test means were 4.10 and 9.42 ($g = -2.89$). Wang et al. [31] enrolled a perioperative stage 0–III surgical population in a three-arm trial. The primary comparison combined PERMA plus loaded deep breathing and loaded deep breathing alone against routine care; discharge HADS-D medians were converted and active arms combined ($g = -2.17$). Both were rated as having some concerns, but Rehman had limited randomization and analysis reporting, whereas Wang used a per-protocol population, had losses, and required conversion and arm combination. Their extreme estimates may reflect baseline burden, short assessment, small samples, intervention contrast, and analytical assumptions.

The depression data from both trials were rechecked against the original reports, including group sizes, HADS-D scale units, and baseline and post-intervention summary statistics; no transcription or unit errors were identified. Baseline HADS-D values were closely balanced between groups in Rehman et al. (12.75 ± 2.20 vs. 13.02 ± 2.00) and across the three arms in Wang et al. (medians 4 [IQR 3–4.25], 3 [IQR 3–4], and 4 [IQR 3–4]), making baseline imbalance an unlikely explanation for the extreme effects. Rehman's estimate was calculated from directly reported post-test means and SDs, whereas Wang's estimate required conversion of discharge medians and combination of the two active arms. The large standardized effects therefore arose principally from substantial short-term between-group separation relative to the reported or derived within-group dispersion. Small sample sizes reduced precision, while the conversion and arm-combination procedures used for Wang et al. introduced additional analytical uncertainty.

Broad subgroup categories did not explain heterogeneity. Baseline depression means were available for 12 quantitative trials, and no study selected participants because of clinically elevated depression. Rehman et al. [34] was the only trial whose group means were clearly above a commonly used HADS-D screening range, which does not establish major depression. An exploratory baseline-severity meta-regression suggested larger effects with greater baseline burden, but it was ecological, based on 12 trials, and sensitive to Rehman et al.; it cannot separate treatment from prevention effects at patient level.

4.4. Clinical Interpretation and Sources of Variability

The pooled estimate represents an average effect across clinically and methodologically diverse trials rather than the effect of a single standardized intervention. Participants

differed in disease stage, treatment phase, surgical status, functional capacity, and baseline symptom burden. Intervention content, duration, supervision, and delivery also varied, while comparators ranged from usual care and education to self-management and active rehabilitation. Standardization permitted statistical synthesis but did not make these populations, interventions, or contrasts clinically equivalent.

The 95% prediction interval included the null, and exclusion of two influential trials attenuated the pooled effect and removed most of the estimated heterogeneity. These findings limit confidence in the magnitude and transferability of the average effect and do not identify the patients, intervention components, or clinical settings most likely to benefit.

None of the included trials evaluated causal mediators. The review therefore cannot determine whether observed changes arose from exercise-specific effects, nonspecific behavioral support, concurrent changes in physical symptoms, or other components of multicomponent care.

The exercise-dominant sensitivity analysis yielded an average estimate similar to the primary model, indicating that the broader exercise-containing programs did not account for the favorable average direction. However, substantial residual heterogeneity and a prediction interval crossing the null persisted, so this restriction did not establish a uniform exercise-specific effect.

4.5. Certainty of Evidence and Clinical Implications

Certainty was very low after downgrading for risk of bias, inconsistency, and indirectness. None of the included trials were judged to be at overall low risk of bias; the interventions and comparator conditions were heterogeneous, and the 95% prediction interval included the possibility of no benefit in a future setting. Confidence in both the magnitude and generalizability of the pooled effect is therefore limited.

The findings do not establish exercise as a treatment for major depressive disorders and do not support a universal exercise prescription for adults with lung cancer. The evidence is also insufficient to recommend a particular modality, intensity, duration, or delivery format for reducing depressive symptoms. Exercise-based interventions may be considered only as potential individualized adjuncts to supportive care when clinically appropriate, feasible, and safe.

Exercise should not replace depression screening, diagnostic assessment, psychotherapy, pharmacological treatment, or specialist psycho-oncology care when these are indicated. Further adequately powered randomized trials are needed to determine whether effects differ according to baseline depressive symptom severity, disease and treatment characteristics, intervention content, and comparator intensity.

4.6. Strengths

This review focused specifically on randomized evidence concerning depressive symptoms in adults with lung cancer and maintained one statistically independent primary effect per trial. An explicit hierarchy was applied to outcome, intervention-arm, and time-point selection, and multi-arm trials were handled without double counting shared comparator groups. Published and derived values were distinguished, and median-based conversions, SE-to-SD transformations, combined-arm estimates, and externally derived group-size allocations were examined in robustness analyses.

The primary synthesis used REML random-effects estimation with Hartung–Knapp inference and reported a prediction interval in addition to conventional heterogeneity statistics. Primary, sensitivity, influence, moderator, and small-study-effect analyses were kept analytically distinct. Risk of bias was assessed for the selected depression result using RoB 2, and certainty of evidence was evaluated using GRADE. Together, these procedures

enhanced transparency, reproducibility, and the distinction between a favorable average effect and uncertainty in its magnitude and transferability.

4.7. Limitations

Only 15 independent randomized trials contributed to the primary meta-analysis, and many had small samples. Although all trials addressed the same broad review question, their populations were not clinically interchangeable. Participants differed in disease stage, treatment phase, surgical status, functional capacity, symptom burden, and baseline depressive symptom severity. The interventions, comparator conditions, depression instruments, and assessment times also varied. Standardized mean differences placed outcomes on a common statistical scale but did not remove these clinical differences. The pooled estimate should therefore be understood as an average across diverse settings rather than as a treatment effect expected in every patient population. This limitation is reinforced by the prediction interval crossing the null and by attenuation of the effect after exclusion of two influential trials.

Disease stage could be classified with sufficient specificity in only 7 of the 15 quantitative trials; categories often overlapped, and several reports combined patients across stages. Oncological treatment likewise ranged from surgery and perioperative care to chemotherapy, chemoradiotherapy, advanced-disease care, and survivorship, with treatment exposure incompletely reported. Formal stage- or treatment-stratified meta-analysis was therefore not performed because the resulting groups would have been sparse, non-exclusive, and vulnerable to misclassification. These variables remain plausible sources of effect modification.

Comparators included usual care, education, self-management, chest physiotherapy, and pulmonary rehabilitation. Active-comparator estimates represent comparative or incremental effects relative to active rehabilitation rather than exercise versus no exercise.

Some effects required median conversion, SD reconstruction, or active-arm combination; raw and derived data are separated; and direct-endpoint sensitivity analysis is reported. One eligible study remained qualitative because conversion was not defensible and another reported only total HADS.

Subgroup and meta-regression analyses were exploratory, involved few studies, and were susceptible to ecological confounding and multiple testing. The age association was sensitive to an influential trial and should not guide clinical decision-making. Small-study-effect tests had limited power; nonsignificant results do not establish the absence of publication bias. The database searches included a depression concept, which may have reduced sensitivity because depressive symptoms were frequently secondary outcomes and may not have been reported in titles, abstracts, or indexing terms. A supplementary search without the depression concept was not conducted. Restriction to English-language publications and the absence of grey-literature and trial-registry searches may also have resulted in relevant evidence being missed.

Nine of the 15 quantitative trials were conducted in China, Hong Kong, or Taiwan. This geographic concentration may limit generalizability to other healthcare systems and cultural contexts, including settings in which intervention delivery, supportive-care pathways, and reporting of depressive symptoms may differ.

4.8. Future Research

Future randomized trials should be adequately powered for depressive symptoms and should prospectively specify the depression instrument, primary assessment time point, outcome format, analysis population, and missing-data strategy. Trials should report arm-specific randomized and analyzed sample sizes, means and SDs—or complete distributional summaries for skewed outcomes—at all time points. CONSORT-consistent

reporting of intervention adherence, contamination, comparator content, co-interventions, adverse events, and reasons for missing depression assessments is essential.

Eligibility based on clinically meaningful depressive symptoms, or stratification by baseline symptom severity, would help distinguish prevention from treatment effects and reduce floor effects. Trials should also report cancer histology and stage, active treatment, performance status, frailty, concomitant psychological or pharmacological treatment, functional capacity, and cardiopulmonary limitations. Longer-term assessments are required to determine whether benefits persist after intervention completion. Head-to-head, factorial, or component-based designs may help identify which intervention components contribute to benefit. Individual participant data meta-analysis could evaluate patient-level effect modification by baseline depressive symptom severity, disease stage, treatment, age, sex, frailty, performance status, and adherence without relying on ecological study-level comparisons. If a sufficiently connected evidence network emerges, component network meta-analysis may help distinguish the contributions of aerobic, resistance, breathing, mind–body, behavioral, and digital components within multicomponent programs. Future mechanistic and implementation studies should use longitudinal mediator assessment, objective activity or adherence measures where feasible, and standardized definitions of intervention content and adverse events. Development of a core outcome set for psychological outcomes in exercise-oncology trials would improve comparability and reduce selective outcome reporting.

5. Conclusions

Exercise-based interventions may reduce depressive symptom scores in some adults with lung cancer, but the magnitude of the average effect was not robust. Exclusion of two influential trials attenuated Hedges' g from −0.75 to −0.49 and reduced I^2 from 87.1% to approximately 0%. Together with a prediction interval crossing the null and very-low-certainty evidence, this finding limits the transferability of the pooled estimate across clinical settings. The evidence does not establish efficacy for major depressive disorder or identify an optimal modality, dose, or delivery format. Better-powered trials with clearly characterized populations, comparators, intervention components, and depression-specific outcomes are required.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/medicina62091805/s1, Table S1. Search strategies used across electronic databases.

Author Contributions: Conceptualization, M.S. and Z.P.; methodology, M.S.; software, M.S.; validation, M.S. and Z.P.; formal analysis, M.S.; investigation, M.S.; resources, M.S.; data curation, M.S.; writing—original draft preparation, M.S. and Z.P.; writing—review and editing, M.S. and Z.P.; visualization, M.S.; supervision, Z.P.; project administration, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The study-level extraction data, analysis-ready dataset, effect-size calculations, data-conversion and multi-arm combination records, RoB 2 assessments, GRADE evidence profile, and R code used for the primary, sensitivity, subgroup, meta-regression, and small-study-effect analyses are publicly available in the OSF project associated with the registration (VUC87): https://osf.io/eyp5v/ (accessed on 17 July 2026). Database-specific search strategies are provided in Supplementary Materials Table S1.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bray, F.; Laversanne, M.; Sung, H.; Ferlay, J.; Siegel, R.L.; Soerjomataram, I.; Jemal, A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2024**, *74*, 229–263. [CrossRef] [PubMed]
2. Sullivan, D.R.; Forsberg, C.W.; Ganzini, L.; Au, D.H.; Gould, M.K.; Provenzale, D.; Slatore, C.G. Longitudinal changes in depression symptoms and survival among patients with lung cancer: A national cohort assessment. *J. Clin. Oncol.* **2016**, *34*, 3984–3991. [CrossRef] [PubMed]
3. Walker, J.; Holm Hansen, C.; Martin, P.; Sawhney, A.; Thekkumpurath, P.; Beale, C.; Symeonides, S.; Wall, L.; Murray, G.; Sharpe, M. Prevalence of depression in adults with cancer: A systematic review. *Ann. Oncol.* **2013**, *24*, 895–900. [CrossRef] [PubMed]
4. Grassi, L.; Caruso, R.; Riba, M.B.; Lloyd-Williams, M.; Kissane, D.; Rodin, G.; McFarland, D.; Campos-Ródenas, R.; Zachariae, R.; Santini, D.; et al. Anxiety and depression in adult cancer patients: ESMO Clinical Practice Guideline. *ESMO Open* **2023**, *8*, 101155. [CrossRef] [PubMed]
5. Andersen, B.L.; Lacchetti, C.; Ashing, K.; Berek, J.S.; Berman, B.S.; Bolte, S.; Dizon, D.S.; Given, B.; Nekhlyudov, L.; Pirl, W.; et al. Management of anxiety and depression in adult survivors of cancer: ASCO guideline update. *J. Clin. Oncol.* **2023**, *41*, 3426–3453. [CrossRef] [PubMed]
6. Campbell, K.L.; Winters-Stone, K.M.; Wiskemann, J.; May, A.M.; Schwartz, A.L.; Courneya, K.S.; Zucker, D.S.; Matthews, C.E.; Ligibel, J.A.; Gerber, L.H.; et al. Exercise guidelines for cancer survivors: Consensus statement from international multidisciplinary roundtable. *Med. Sci. Sports Exerc.* **2019**, *51*, 2375–2390. [CrossRef] [PubMed]
7. Ligibel, J.A.; Bohlke, K.; May, A.M.; Clinton, S.K.; Demark-Wahnefried, W.; Gilchrist, S.C.; Irwin, M.L.; Late, M.; Mansfield, S.; Marshall, T.F.; et al. Exercise, diet, and weight management during cancer treatment: ASCO guideline. *J. Clin. Oncol.* **2022**, *40*, 2491–2507. [CrossRef] [PubMed]
8. Brown, J.C.; Huedo-Medina, T.B.; Pescatello, L.S.; Ryan, S.M.; Pescatello, S.M.; Moker, E.; LaCroix, J.M.; Ferrer, R.A.; Johnson, B.T. The efficacy of exercise in reducing depressive symptoms among cancer survivors: A meta-analysis. *PLoS ONE* **2012**, *7*, e30955. [CrossRef] [PubMed]
9. Craft, L.L.; Vaniterson, E.H.; Helenowski, I.B.; Rademaker, A.W.; Courneya, K.S. Exercise effects on depressive symptoms in cancer survivors: A systematic review and meta-analysis. *Cancer Epidemiol. Biomark. Prev.* **2012**, *21*, 3–19. [CrossRef] [PubMed]
10. Noetel, M.; Sanders, T.; Gallardo-Gómez, D.; Taylor, P.; del Pozo Cruz, B.; van den Hoek, D.; Smith, J.J.; Mahoney, J.; Spathis, J.; Moresi, M.; et al. Effect of exercise for depression: Systematic review and network meta-analysis of randomised controlled trials. *BMJ* **2024**, *384*, e075847. Erratum in *BMJ* **2024**, *385*, q1024. https://doi.org/10.1136/bmj.q1024. [CrossRef] [PubMed]
11. Cavalheri, V.; Burtin, C.; Formico, V.R.; Nonoyama, M.L.; Jenkins, S.; Spruit, M.A.; Hill, K. Exercise training undertaken by people within 12 months of lung resection for non-small cell lung cancer. *Cochrane Database Syst. Rev.* **2019**, *6*, CD009955. [CrossRef] [PubMed]
12. Lee, J. Physiologic and psychologic adaptation to exercise interventions in lung cancer patients undergoing chemotherapy: A systematic review and meta-analysis of randomized controlled trials. *Support. Care Cancer* **2021**, *29*, 2863–2873. [CrossRef] [PubMed]
13. Ni, H.-J.; Pudasaini, B.; Yuan, X.-T.; Li, H.-F.; Shi, L.; Yuan, P. Exercise training for patients pre- and postsurgically treated for non-small cell lung cancer: A systematic review and meta-analysis. *Integr. Cancer Ther.* **2017**, *16*, 63–73. [CrossRef] [PubMed]
14. Singh, B.; Spence, R.; Steele, M.L.; Hayes, S.; Toohey, K. Exercise for individuals with lung cancer: A systematic review and meta-analysis of adverse events, feasibility, and effectiveness. *Semin. Oncol. Nurs.* **2020**, *36*, 151076. [CrossRef] [PubMed]
15. 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.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]
16. Luo, D.; Wan, X.; Liu, J.; Tong, T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Stat. Methods Med. Res.* **2018**, *27*, 1785–1805. [CrossRef] [PubMed]
17. Shi, J.; Luo, D.; Weng, H.; Zeng, X.-T.; Lin, L.; Chu, H.; Tong, T. Optimally estimating the sample standard deviation from the five-number summary. *Res. Synth. Methods* **2020**, *11*, 641–654. [CrossRef] [PubMed]
18. Wan, X.; Wang, W.; Liu, J.; Tong, T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med. Res. Methodol.* **2014**, *14*, 135. [CrossRef] [PubMed]
19. Higgins, J.P.T.; Thomas, J.; Chandler, J.; Cumpston, M.; Li, T.; Page, M.J.; Welch, V.A. (Eds.) *Cochrane Handbook for Systematic Reviews of Interventions*; Version 6.5; Cochrane: London, UK, 2024. Available online: https://www.cochrane.org/handbook (accessed on 25 July 2026).

20. Sterne, J.A.C.; Savović, J.; Page, M.J.; Elbers, R.G.; Blencowe, N.S.; Boutron, I.; Cates, C.J.; Cheng, H.-Y.; Corbett, M.S.; Eldridge, S.M.; et al. RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **2019**, *366*, l4898. [CrossRef] [PubMed]
21. Guyatt, G.H.; Oxman, A.D.; Vist, G.E.; Kunz, R.; Falck-Ytter, Y.; Alonso-Coello, P.; Schünemann, H.J. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* **2008**, *336*, 924–926. [CrossRef] [PubMed]
22. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **2010**, *36*, 1–48. [CrossRef]
23. Knapp, G.; Hartung, J. Improved tests for a random effects meta-regression with a single covariate. *Stat. Med.* **2003**, *22*, 2693–2710. [CrossRef] [PubMed]
24. Higgins, J.P.T.; Thompson, S.G.; Deeks, J.J.; Altman, D.G. Measuring inconsistency in meta-analyses. *BMJ* **2003**, *327*, 557–560. [CrossRef] [PubMed]
25. Begg, C.B.; Mazumdar, M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* **1994**, *50*, 1088–1101. [CrossRef]
26. 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*, 455–463. [CrossRef] [PubMed]
27. Egger, M.; Davey Smith, G.; Schneider, M.; Minder, C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* **1997**, *315*, 629–634. [CrossRef] [PubMed]
28. Liu, Y.-Y.; Lu, H.-B.; Li, Y.-J.; Wang, Z.-Y.; Liu, Y.-H.; Qiao, S.; Xie, J. Effects of mindful breathing training combined with diary-based rehabilitation guidance in lung cancer patients undergoing surgery: A randomized controlled trial. *Complement. Ther. Clin. Pract.* **2024**, *55*, 101849. [CrossRef] [PubMed]
29. Tenconi, S.; Mainini, C.; Rapicetta, C.; Braglia, L.; Galeone, C.; Cavuto, S.; Merlo, D.F.; Costi, S.; Paci, M.; Piro, R.; et al. Rehabilitation for lung cancer patients undergoing surgery: Results of the PUREAIR randomized trial. *Eur. J. Phys. Rehabil. Med.* **2021**, *57*, 1002–1011. [CrossRef] [PubMed]
30. Cheung, D.S.T.; Takemura, N.; Lam, T.C.; Ho, J.C.M.; Deng, W.; Smith, R.; Yan, Y.; Lee, A.W.M.; Lin, C.C. Feasibility of aerobic exercise and Tai-Chi interventions in advanced lung cancer patients: A randomized controlled trial. *Integr. Cancer Ther.* **2021**, *20*, 15347354211033352. [CrossRef] [PubMed]
31. Wang, Z.; Yu, L.; Song, C.; Cao, H.; Yang, Z.; Qiao, S.; Li, X.; Xie, J. Effects of loaded deep breathing training combined with PERMA model in lung cancer patients undergoing surgery: A randomized controlled trial. *Support. Care Cancer* **2025**, *33*, 574. [CrossRef] [PubMed]
32. Liu, W.-L.; Chien, J.-Y.; Lu, Y.-Y.; Liu, K.-F. Nurse-supported hybrid home-based pulmonary rehabilitation improves psychological distress, quality of life, and functional performance in advanced lung cancer: A randomized controlled trial. *Eur. J. Oncol. Nurs.* **2026**, *81*, 103130. [CrossRef] [PubMed]
33. Wang, L.; Liu, C.; Li, H.; Qin, L.; Shen, W.; Zheng, H.; Yang, D.; Huang, L.; Wang, Q. Effects of a rehabilitation program based on home intelligent management platform guided among patient after lung cancer surgery: A randomized controlled trial. *Support. Care Cancer* **2026**, *34*, 379. [CrossRef] [PubMed]
34. Rehman, M.; Ahmad, U.; Waseem, M.; Ali, B.; Tariq, M.I. Effects of exercise training in patients with lung cancer during chemotherapy treatment. *Malays. J. Med. Sci.* **2023**, *30*, 141–152. [CrossRef] [PubMed]
35. Lei, J.; Yang, J.; Dong, L.; Xu, J.; Chen, J.; Hou, X.; Bai, Z. An exercise prescription for patients with lung cancer improves the quality of life, depression, and anxiety. *Front. Public Health* **2022**, *10*, 1050471. [CrossRef] [PubMed]
36. Bade, B.C.; Gan, G.; Li, F.; Lu, L.; Tanoue, L.; Silvestri, G.A.; Irwin, M.L. Randomized trial of physical activity on quality of life and lung cancer biomarkers in patients with advanced-stage lung cancer: A pilot study. *BMC Cancer* **2021**, *21*, 352. Erratum in *BMC Cancer* **2024**, *24*, 1409. https://doi.org/10.1186/s12885-024-13183-9. [CrossRef] [PubMed]
37. Quist, M.; Langer, S.W.; Lillelund, C.; Winther, L.; Laursen, J.H.; Christensen, K.B.; Rørth, M.; Adamsen, L. Effects of an exercise intervention for patients with advanced inoperable lung cancer undergoing chemotherapy: A randomized clinical trial. *Lung Cancer* **2020**, *145*, 76–82. [CrossRef] [PubMed]
38. Sui, Y.; Wang, T.; Wang, X. The impact of WeChat app-based education and rehabilitation program on anxiety, depression, quality of life, loss of follow-up and survival in non-small cell lung cancer patients who underwent surgical resection. *Eur. J. Oncol. Nurs.* **2020**, *45*, 101707. [CrossRef] [PubMed]
39. Wang, Y.-Q.; Cao, H.-P.; Liu, X.; Yang, Z.; Yin, Y.-Y.; Ma, R.-C.; Xie, J. Effect of breathing exercises in patients with non-small cell lung cancer receiving surgical treatment: A randomized controlled trial. *Eur. J. Integr. Med.* **2020**, *38*, 101175. [CrossRef]
40. Egegaard, T.; Rohold, J.; Lillelund, C.; Persson, G.; Quist, M. Pre-radiotherapy daily exercise training in non-small cell lung cancer: A feasibility study. *Rep. Pract. Oncol. Radiother.* **2019**, *24*, 375–382. [CrossRef] [PubMed]
41. Cavalheri, V.; Jenkins, S.; Cecins, N.; Gain, K.; Phillips, M.J.; Sanders, L.H.; Hill, K. Exercise training for people following curative-intent treatment for non-small cell lung cancer: A randomized controlled trial. *Braz. J. Phys. Ther.* **2017**, *21*, 58–68. [CrossRef] [PubMed]

42. Chen, H.-M.; Tsai, C.-M.; Wu, Y.-C.; Lin, K.-C.; Lin, C.-C. Randomised controlled trial on the effectiveness of home-based walking exercise on anxiety, depression and cancer-related symptoms in patients with lung cancer. *Br. J. Cancer* **2015**, *112*, 438–445. [CrossRef] [PubMed]
43. Morano, M.T.A.P.; Mesquita, R.; da Silva, G.P.F.; Araújo, A.S.; Pinto, J.M.S.; Neto, A.G.; Viana, C.M.S.; de Moraes Filho, M.O.; Pereira, E.D.B. Comparison of the effects of pulmonary rehabilitation with chest physical therapy on the levels of fibrinogen and albumin in patients with lung cancer awaiting lung resection: A randomized clinical trial. *BMC Pulm. Med.* **2014**, *14*, 121. [CrossRef] [PubMed]
44. Li, G.; Zhou, X.; Deng, J.; Wang, J.; Ai, P.; Zeng, J.; Ma, X.; Liao, H. Digital therapeutics-based cardio-oncology rehabilitation for lung cancer survivors: Randomized controlled trial. *JMIR mHealth uHealth* **2025**, *13*, e60115. [CrossRef] [PubMed]
45. Henshall, C.L.; Allin, L.; Aveyard, H. A systematic review and narrative synthesis to explore the effectiveness of exercise-based interventions in improving fatigue, dyspnea, and depression in lung cancer survivors. *Cancer Nurs.* **2019**, *42*, 295–306. [CrossRef] [PubMed]
46. Lu, Y.; Bai, X.; Pan, C. Impact of exercise interventions on quality of life and depression in lung cancer patients: A systematic review and meta-analysis. *Int. J. Psychiatry Med.* **2024**, *59*, 199–217. [CrossRef] [PubMed]
47. Tadsuan, J.; Lai, Y.-H.; Lee, Y.-H.; Chen, M.-R. The effectiveness of exercise interventions on psychological distress in patients with lung cancer: A systematic review and meta-analysis. *J. Cancer Surviv.* **2026**, *20*, 910–926. [CrossRef] [PubMed]
48. Hu, Y.; Gu, S.; Bu, Z.; Liu, Z.; Dong, J.; Shi, J.; Xu, Y. Effect of exercise for patients with advanced lung cancer and cancer-related fatigue: A systematic review and meta-analysis. *J. Sport Health Sci.* **2025**, *14*, 101017. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.